

UNIVERSITÉ DE NANTES
FACULTÉ DES SCIENCES ET DES TECHNIQUES

ÉCOLE DOCTORALE
SCIENCES POUR L'INGÉNIEUR, GÉOSCIENCES, ARCHITECTURE (SPIGA)

Année 2013

Une approche bayésienne pour estimer
les propriétés physiques
dans la zone de transition
à partir des ondes de surface

THÈSE DE DOCTORAT
Discipline : Sciences de la Terre
Spécialité : Géophysique

*Présentée
et soutenue publiquement par*

Mélanie DRILLEAU

Le jeudi 30 mai 2013, devant le jury ci-dessous

Rapporteurs	M. Éric DEBAYLE, directeur de recherche, Laboratoire de Géologie de Lyon M. Dominique GIBERT, professeur, Institut de Physique du Globe de Paris
Examineurs	M. Jean-Paul MONTAGNER, professeur, Université Paris 7 – Institut de Physique du Globe de Paris Mme Helle PEDERSEN, physicienne CNAP, Institut des Sciences de la Terre, Université Joseph Fourier, Grenoble
Membre invité	M. Olivier VERHOEVEN, maître de conférences, Laboratoire de Planétologie et de Géodynamique de Nantes

Directeur de thèse : M. Antoine Mocquet, professeur, Laboratoire de Planétologie et de Géodynamique de Nantes

*Co-encadrant de thèse : M. Éric Beucler, maître de conférences, Laboratoire de Planétologie et de Géodynamique de
Nantes*

Remerciements

Je voudrais d'abord remercier Antoine Mocquet, mon directeur de thèse, et Éric Beucler, mon co-encadrant, responsables de m'avoir transmis l'envie d'étudier la zone de transition dès le stage de M2. Merci à tous les deux de m'avoir donné la chance de faire ce travail et de m'avoir fait confiance pendant ces 3 ans et demi. J'ai beaucoup appris à vos côtés, et je suis sûre que cela me servira dans mon travail futur, qu'il soit dans la recherche ou ailleurs. Je voudrais également remercier Oliver Verhoeven, qui bien que ne faisant pas partie de l'ANR TAM qui a financé ce travail, s'est particulièrement impliqué dans l'encadrement de la thèse. Au-delà de sa grande intégrité scientifique, et malgré son chic pour me gâcher régulièrement mes week-ends, sa bonne humeur et sa gentillesse m'ont beaucoup aidé à aller de l'avant. Promis, un jour je vous dirai lequel de vous trois est le plus pointilleux... J'adresse mes plus sincères remerciements à Guy Moebs, responsable scientifique du Centre de Calcul Intensif des Pays de la Loire, qui m'a fait découvrir les secrets de l'optimisation et de la parallélisation des codes. Merci d'avoir pris de ton temps et de m'avoir "mis un pied à l'étrier", tu as été d'un grand secours.

Je tiens à remercier Jean-Paul Montagner, responsable de l'ANR TAM, pour toutes les discussions que nous avons eues ensemble, notamment sur l'anisotropie sismique. Ce fut pour moi un grand plaisir qu'il ait accepté d'être le président de mon jury de thèse. Je voudrais remercier aussi Gaël Burgos, avec qui les discussions m'ont permis de voir mon travail sous un autre angle. Merci pour ton aide, tes conseils et surtout ton amitié. Je remercie Pierre Vacher, pour avoir répondu à mes questions en ce qui concerne la routine de minéralogie. Merci également à Attilio Rivoldini pour les discussions que nous avons eues sur les méthodes de Monte Carlo. J'ai été très impressionnée par son expérience et sa rigueur. Je tiens à remercier également les équipes du CCIPL (Centre de Calcul Intensif des Pays de la Loire) et du CINES (Centre Informatique National de l'Enseignement supérieur), qui se sont toujours montrées disponibles pour répondre à mes questions. Pendant ma thèse, j'ai eu l'opportunité de faire plusieurs formations à l'IDRIS (Institut du Développement et des Ressources en Informatique Scientifique). Merci à l'ensemble des intervenants, qui proposent des cours de très grande qualité.

Je suis particulièrement honorée qu'Éric Debayle et Dominique Gibert aient accepté d'être rapporteurs de ma thèse. Vos remarques et critiques m'ont toutes paru justi-

fiées et ont permis de soulever de nouvelles perspectives de travail. Même si elle n’a pas pu être présente le jour de ma soutenance, je remercie Helle Perderson d’avoir accepté de juger mon travail. J’espère que nous aurons l’occasion d’en discuter de vive voix.

Dans ce laboratoire pluridisciplinaire, j’ai pu échanger avec d’autres chercheurs qui ont pu apporter un point de vue différent sur mon travail. Je voudrais particulièrement remercier Jonathan Besserer, Olivia Golle, Yann Capdeville, Gaël Choblet, Matthieu Lefevre, Philippe Cance et Olivier Grasset. Tous m’ont permis d’une façon ou d’une autre d’évoluer dans ma démarche scientifique.

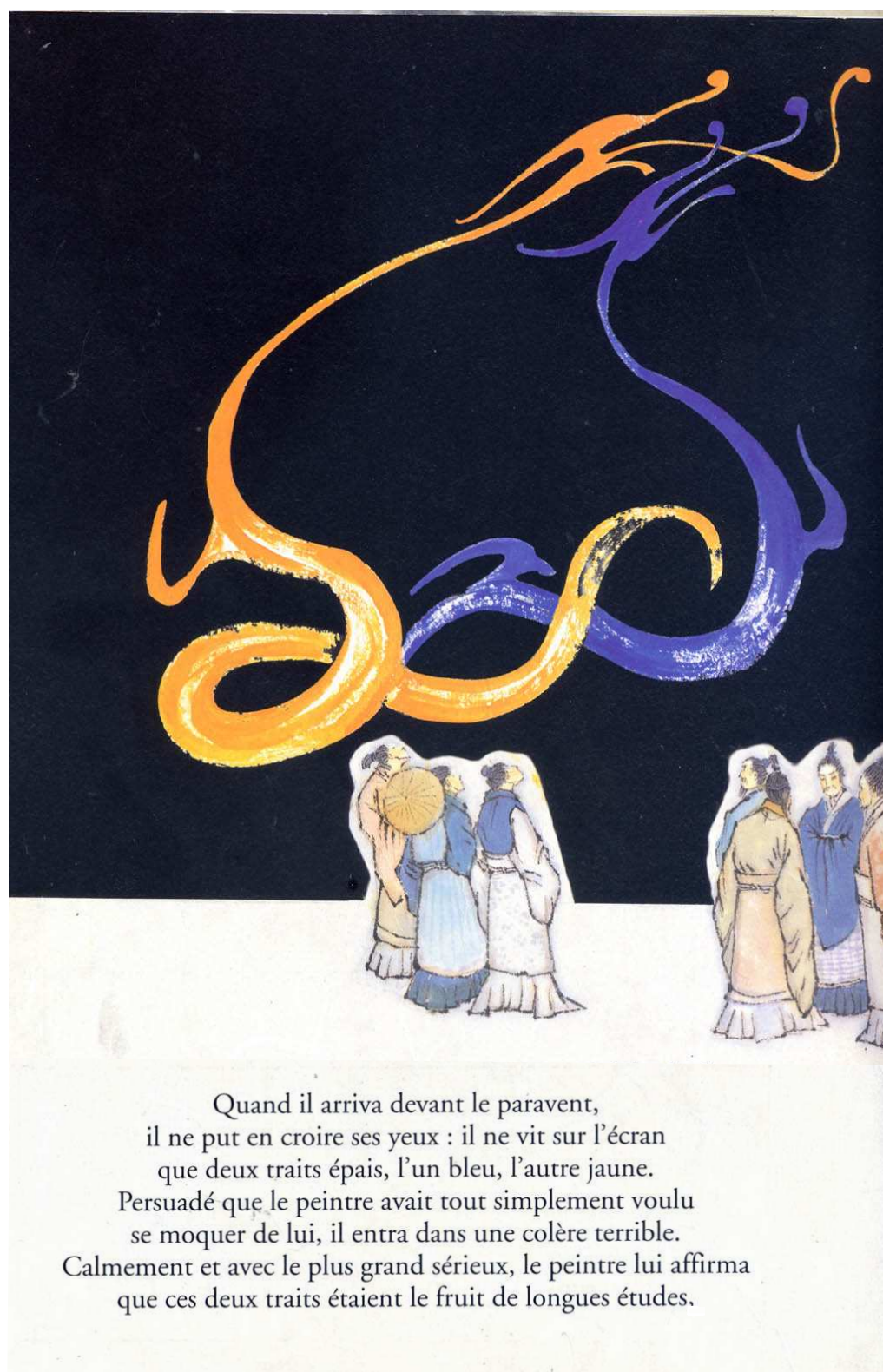
Parmi les permanents du laboratoire, je voudrais remercier : Éric Boeuf, toujours présent et adorable pour résoudre le moindre problème informatique ; Marie-Claire Bréhier et Stéphanie Beaunay pour la gestion des déplacements en France et à l’étranger ; Laurent Lenta pour son intervention avec les problèmes d’ampoules qui grésillent dans mon bureau ! Plus généralement, je remercie l’ensemble des permanents du LPGN pour leur accueil.

Un merci spécial à toutes mes co-bureaux qui ont fait un séjour dans le-bureau-27-au-bout-du-couloir-à-droite, loin de l’imprimante, de la photocopieuse et de la machine à café. Merci à Zeineb Kassouk, pour m’avoir motivée au quotidien. Merci à Susan Conway, la coach de touch-rugby. Merci à Lucile Bezacier, qui dort maintenant près de son synchrotron. Merci à Colleen Milbury, pour toutes nos discussions. Ces années n’auraient pas été les mêmes sans vous. Merci également aux deux petits nouveaux, Ianis Gaudot et Diana Saturnino, qui m’ont supporté pendant les derniers mois. Je vous souhaite une très belle réussite pour votre travail de thèse.

Évidemment, je voudrais remercier les quatre autres docteurs de la promotion 2012-2013, avec qui j’ai partagé ces quatre dernières années. Merci à youhou-Thomas Cornet, à mes côtés depuis le M2, dont la thèse n’a (heureusement) pas eu raison de sa bonne humeur. Merci à Erwin Dehouck, qui a récemment décidé de s’expatrier au pays des burgers : attention, tu vas avoir droit à la pesée à ton retour ! Merci à Mélanie Segard, avec qui je partage des souvenirs particuliers de l’AGU. Merci à Patrick Thollot, pour m’avoir suggéré de manger autre chose que des pâtes chinoises en fin de thèse.

Quelques pensées en vrac pour des (ex-)nantais : Nadia Marounina, Théodolina Lopez, Adriana Oancea, Olivier Bollangier, Gaëlle Plissart, Axel Lefevre, Benjamin Guillaume, Joana Oliveira, Marine Gourronc, Thibault Fougeroux, Erell Leocat, Cécile Taffin, Déborah Chavrit, Aurore de Bigault De Granrut, Farzaneh Kazemipour...

Je voudrais également remercier quelques proches qui m’ont toujours apporté leur soutien. Merci à Izma, Julien, Thomas, Jeanne, Matthieu et Elliot. Un merci spécial à Pauline (mon DJ ! !), sur qui je peux toujours compter. Enfin un TRÈS TRÈS TRÈS grand merci à Fred et à ma famille pour leur soutien sans faille, qui s’exprime bien au-delà de ce travail de thèse.



Quand il arriva devant le paravent,
il ne put en croire ses yeux : il ne vit sur l'écran
que deux traits épais, l'un bleu, l'autre jaune.
Persuadé que le peintre avait tout simplement voulu
se moquer de lui, il entra dans une colère terrible.
Calmement et avec le plus grand sérieux, le peintre lui affirma
que ces deux traits étaient le fruit de longues études.

D'après l'album, "Dragon bleu, dragon jaune", de Ré et Philippe Soupault,
et Li Zhong Yao.

Table des matières

Remerciements	i
Introduction générale	1
1 Cadre de l'étude	5
1.1 La zone de transition	5
1.1.1 Les discontinuités sismiques	6
1.1.2 Variations spatiales des discontinuités	9
1.1.3 Géodynamique de la zone de transition	13
1.1.4 Réconcilier les données sismologiques et les données de laboratoire	18
1.1.5 Anisotropie sismique	21
1.1.6 Hypothèses de travail	24
1.2 Tomographie sismique	25
1.2.1 Des modèles 3D complexes	26
1.2.2 Le problème des couches superficielles	28
1.2.3 La méthode des gradients	30
1.2.4 L'approche probabiliste en sismologie	31
1.2.5 Interprétation des modèles sismologiques en termes de température et de composition	35
1.3 Objet de cette étude et démarche adoptée	37
2 Problème direct	39
2.1 Introduction	39
2.2 Paramètres sismiques	40
2.2.1 Équation d'état	41
2.2.1.1 Chemin thermodynamique des minéraux	42
2.2.1.2 Chauffage isobare	44
2.2.1.3 Compression adiabatique	45
2.2.2 Du minéral à la roche	47

2.2.3	Diagrammes de phases	49
2.3	Anisotropie	53
2.3.1	Le tenseur élastique	53
2.3.2	Anisotropie radiale	54
2.4	Vitesses de phase	57
2.4.1	Les oscillations libres	57
2.4.2	Solutions du problème aux valeurs propres	59
2.4.3	La dispersion des ondes de surface	61
2.5	Paramétrisation des modèles	64
2.5.1	Différents types de paramétrisation	64
2.5.2	Les courbes de Bézier cubiques	67
2.6	Ce qu'il faut retenir	70
3	Inversion probabiliste	71
3.1	Introduction	71
3.1.1	Problème inverse	71
3.1.2	Stratégie adoptée	72
3.2	Inférence bayésienne	73
3.2.1	Notion de densité de probabilité marginale	73
3.2.2	Théorème de Bayes	73
3.2.3	Distribution <i>a posteriori</i> des paramètres	75
3.3	Méthode de Monte Carlo par chaînes de Markov (McMC)	77
3.3.1	Chaînes de Markov	77
3.3.1.1	Principe	77
3.3.1.2	Définitions et propriétés	77
3.3.2	L'algorithme de Metropolis-Hastings	80
3.3.3	Échantillonnage de l'espace des paramètres	81
3.4	Ce qu'il faut retenir	83
4	Implémentation et application pratique à la zone de transition	85
4.1	Introduction	85
4.2	Conditions aux limites	86
4.2.1	La croûte et le manteau supérieur	86
4.2.2	Conditions de raccord	87
4.2.3	Priors sur les paramètres de l'inversion	88
4.3	L'algorithme	89
4.3.1	Les principales étapes	89
4.3.2	Générer de nouveaux modèles tout au long d'une chaîne de Markov	91
4.3.3	Construction des courbes de Bézier de classe C^1	91

4.3.4	Optimisation de l'algorithme	92
4.3.4.1	Avantage en temps de calcul	93
4.3.4.2	Avantage en performance	94
4.3.5	Atteindre la période stationnaire	95
4.3.6	Une procédure en trois étapes	97
4.3.7	Temps de calcul	99
4.4	Choix particuliers pour l'étude de la zone de transition	100
4.4.1	Contraintes <i>a priori</i>	100
4.4.1.1	Prior sur la température	100
4.4.1.2	Prior sur l'anisotropie	104
4.4.1.3	Prior sur les paramètres sismiques	105
4.4.1.4	Distributions de probabilité <i>a priori</i>	106
4.4.2	Modalités de tirage au sort de nouveaux modèles	109
4.4.2.1	Écart-type des fonctions gaussiennes	109
4.4.2.2	Échantillonnage continu/discontinu	113
4.4.2.3	Tirage au sort des paramètres	114
4.4.3	La fonction coût (misfit)	116
4.4.3.1	Normes L_1 et L_2	116
4.4.3.2	Influence de la discrétisation des profils de vitesses	119
4.5	Ce qu'il faut retenir	122
5	Inversion de données synthétiques	123
5.1	Introduction	123
5.2	Les modèles utilisés	124
5.3	Les différents types de représentation	127
5.4	Résultats	131
5.4.1	Température	131
5.4.2	Anisotropie radiale	135
5.4.3	Vitesse de cisaillement dans la croûte et le manteau supérieur	140
5.4.4	Corrélations entre les paramètres	143
5.4.4.1	Paramètres primaires	143
5.4.4.2	Paramètres secondaires	143
5.5	Tests supplémentaires	146
5.5.1	Influence de la taille des incertitudes sur les données	146
5.5.2	Résultats avec une norme L_1	148
5.5.3	Profondeurs des paramètres fixées	149
5.5.4	Influence du lissage des profils en température	154
5.6	Ce qu'il faut retenir	158

6	Inversion de données réelles	159
6.1	Introduction	159
6.2	Les données	160
6.3	Résultats	163
6.3.1	Vitesse de cisaillement dans la croûte et le manteau supérieur . .	163
6.3.2	Température	163
6.3.3	Anisotropie radiale	167
6.3.4	Corrélation entre les paramètres primaires	170
6.3.5	Les distributions en V_S , V_P isotropes et ρ dans la zone de transition	173
6.4	Influence de la valeur de la pente de Clapeyron à 660 km	178
6.5	Déterminer la distribution de température à partir des vitesses de cisaillement	181
6.5.1	Des vitesses de phase vers la distribution en V_S	182
6.5.2	De V_S vers la distribution en température	186
6.6	Ce qu'il faut retenir	191
7	Conclusions et perspectives	193
7.1	Synthèse	193
7.2	Avantages et critiques de la méthode	195
7.3	Applications potentielles de ce travail	196
7.3.1	En utilisant l'algorithme en l'état	196
7.3.2	En modifiant l'algorithme	198
7.4	Améliorations potentielles de l'algorithme	200
	Annexes	203
A	Notations	203
B	Submitted article	207
B.1	Introduction	209
B.2	Forward problem	210
B.2.1	Physical parameters	211
B.2.1.1	Isotropic parameters	211
B.2.1.2	Anisotropy	211
B.2.2	Model parameterization	212
B.3	McMC method	213
B.4	Practical implementation	216
B.4.1	Boundary conditions	216

B.4.2	Prior information	216
B.4.3	The MCMC algorithm within a three-step exploration	217
B.5	Inversion of synthetic data	219
B.5.1	Input models	219
B.5.2	Characterization of posterior results	220
B.5.3	Results	221
B.5.3.1	Temperature	221
B.5.3.2	Radial anisotropy	222
B.5.3.3	V_S in the uppermost mantle	223
B.6	Real data	224
B.6.1	Data and <i>a priori</i> conditions	224
B.6.2	Results	224
B.6.2.1	Temperature distribution	224
B.6.2.2	Radial anisotropic structure	225
B.6.2.3	Seismic velocities and density	226
B.6.3	Effect of the Clapeyron slope at 660 km depth on the temperature distribution	227
B.7	Conclusion	228
B.8	Figures and tables	230
Références bibliographiques		245

Introduction générale

C'est un fait : il n'existe pas d'observation directe de la Terre au-delà d'une douzaine de kilomètres de profondeur¹. Hormis dans les productions hollywoodiennes, il n'est malheureusement pas possible à l'heure actuelle de creuser un puits suffisamment profond, ni d'envoyer une sonde (habitée ou non) suffisamment loin vers le centre de la Terre pour observer directement ses propriétés physiques. Moins spectaculaires, mais d'un intérêt primordial, seules des mesures indirectes sont possibles.

La sismologie est l'une des rares disciplines qui puisse fournir des informations sur la structure interne du globe. En effet, les ondes sismiques générées par les tremblements de terre véhiculent des informations sur les matériaux traversés le long de leur trajet. Ces informations sont extraites des sismogrammes à l'aide de théories physiques et mathématiques. Elles constituent des observations indirectes de la Terre profonde, et permettent d'établir des modèles du manteau terrestre en s'appuyant sur diverses hypothèses. Les premiers modèles sismologiques globaux se sont focalisés sur la structure moyenne radiale de la Terre. Puis des modèles 3D sont apparus, obtenus grâce à des théories de plus en plus complexes, qui quantifient les perturbations de vitesses par rapport à ces modèles unidimensionnels. Parmi les différentes enveloppes accessibles par la tomographie sismique, la zone de transition *s. l.*, située entre 350 et 1000 km de profondeur, suscite un intérêt particulier. Localisée au cœur du manteau terrestre, celle-ci est le siège d'échanges de matière entre les manteaux supérieur et inférieur et comprend de nombreuses transformations de phase minéralogiques. Il s'agit d'une région extrêmement hétérogène, à la fois d'un point de vue thermique et chimique, et dont l'étude constitue une des clés pour comprendre les processus de convection. Bien que la résolution des modèles sismiques 3D soit en constante augmentation depuis ces trois dernières décennies, les différents modèles s'accordent seulement aux grandes longueurs d'ondes, et les divergences s'amplifient au fur et à mesure que l'on s'enfonce dans le manteau terrestre.

La connaissance des variations des vitesses sismiques dans le manteau n'est pourtant pas une fin en soi. Effectivement, encore faut-il être capable d'expliquer quels sont les paramètres physiques qui décrivent la dynamique de la zone de transition, et ceux

¹Il s'agit du forage effectué dans la péninsule de Kola, en Russie, réalisé de 1970 à 1989.

qui sont responsables des hétérogénéités de vitesses. Ces hétérogénéités sont-elles d'origine thermique ? Minéralogique ? Ou bien chimique ? Pour comprendre la dynamique du manteau terrestre, les géophysiciens ont développé de nombreuses techniques. Alors que certains d'entre eux s'enferment dans des laboratoires expérimentaux pour reproduire les conditions thermodynamiques qui règnent à l'intérieur du globe, et ainsi synthétiser les phases minérales profondes, d'autres se retranchent derrière un ordinateur pour modéliser les variations de vitesses sismiques dans la Terre à partir des données sismologiques. Ces dernières années, un axe de recherche actif en géophysique est de confronter les données issues de la physique des minéraux avec les données sismologiques. Les nouvelles avancées sur la connaissance de la zone de transition se situent dans ce contexte pluridisciplinaire.

Cependant, la non-unicité des solutions intrinsèque à ce genre d'approche est parfois négligée, les équipes proposant souvent un seul modèle de Terre. Notre planète est évidemment unique, mais compte-tenu des incertitudes sur les données géophysiques, il n'existe pas une unique configuration de paramètres capable de rendre compte des données. C'est pourquoi des théories linéarisées, fortement dépendantes d'un modèle de référence, sont souvent utilisées. Aujourd'hui, de plus en plus d'études cherchent non pas à produire un modèle unique, mais à prendre en compte la non-linéarité entre les paramètres et les données en décrivant l'espace des paramètres qui constituent ce modèle en termes de probabilités, selon leur adéquation aux données. De telles techniques d'exploration (puisque'il ne s'agit pas d'inversions au sens classique du terme), où la connaissance de nouveaux modèles est éclairée par la connaissance de tous les modèles échantillonnés précédemment, sont novatrices et appliquées depuis peu en sismologie.

Le but de ce travail de thèse est d'utiliser les théories de physique des matériaux pour expliquer les vitesses de phase des ondes de surface en termes de distributions 1D de température et d'anisotropie radiale dans la zone de transition. La non-linéarité du problème est prise en compte en employant une méthode de recherche dans l'espace des modèles.

Le manuscrit est divisé en sept chapitres, dont le contenu est esquissé ci-dessous :

- Le **chapitre 1**, qui définit le cadre de l'étude, est divisé en deux grandes parties. La première partie est une synthèse bibliographique concise des connaissances sur la zone de transition. Elle met l'accent sur les principaux facteurs responsables des anomalies de vitesses sismiques dans cette région. La deuxième partie est focalisée sur les méthodes tomographiques à l'heure actuelle, en mettant en parallèle les méthodes déterministes, sur lesquelles sont basés la plupart des modèles tomographiques 3D, et les méthodes probabilistes, plus novatrices en sismologie, qui prennent en compte la non-unicité des solutions. Ce chapitre s'achève en décrivant les objectifs de cette thèse ainsi que la démarche adoptée pour les remplir, qui s'inscrit dans un cadre pluridisciplinaire.

- Le **chapitre 2** est consacré à la description du problème direct, c'est-à-dire le passage d'un modèle en température et en anisotropie radiale, aux vitesses de phase correspondantes. Il décrit comment à partir des diagrammes de phases et d'une équation d'état, il est possible de déterminer les propriétés sismiques d'une roche mantellique, pour une composition minéralogique et un profil de température donné. Le calcul des vitesses de phase a lieu par sommations de modes propres, en incorporant l'anisotropie radiale. La paramétrisation particulière des modèles en courbes de Bézier est également détaillée.
- Le **chapitre 3** concerne la méthode d'inversion qui a été appliquée aux vitesses de phase. Il s'agit d'une inversion probabiliste de type Monte Carlo en utilisant les propriétés des chaînes de Markov. Ce chapitre présente les grands traits mathématiques de la méthode.
- Le **chapitre 4** traite de l'implémentation de la procédure d'inversion. Cette dernière consiste en une détermination jointe des paramètres d'un modèle 1D en température et en anisotropie radiale dans la zone de transition, et de vitesse des ondes S dans la croûte et le manteau supérieur. La méthode d'inversion choisie (détaillée dans le chapitre 3) s'avère très générale, c'est pourquoi les choix particuliers qui ont été effectués pour l'étude de la zone de transition sont décrits en détail.
- Le **chapitre 5** est dédié à la validation de l'algorithme sur des données synthétiques, en laissant un maximum de liberté sur les paramètres de l'inversion. Une attention spéciale est portée sur la manière d'analyser les résultats, qui ne correspondent pas à un modèle unique, mais à une distribution de modèles. Les effets de différents facteurs sur les résultats de l'inversion sont également illustrés.
- L'application de la méthode sur des vitesses de phase réelles pour un trajet situé au niveau de l'Océan Pacifique, est exposée dans le **chapitre 6**. Les résultats obtenus amènent à une première compréhension des distributions de la température, de l'anisotropie radiale et des vitesses sismiques dans la zone de transition. L'essentiel de la discussion est focalisé sur ce que les ondes de surface et la méthode sont capables de détecter.
- Une synthèse de ce travail de thèse est présentée dans le **chapitre 7**. Les principaux résultats sont résumés et des applications potentielles de l'algorithme développé dans ce travail sont proposées.
- En **Annexe** est présenté un article² qui décrit l'ensemble de la méthode, ainsi que des applications à des données synthétiques et réelles.

²Drilleau, M., *et al.*, A Bayesian approach to infer radial models of temperature and anisotropy in the transition zone from surface wave dispersion curves (soumis à GJI).

Chapitre 1

Cadre de l'étude

1.1 La zone de transition

Dans cette première partie, l'état des connaissances sur la zone de transition est brièvement présenté. Au premier ordre, la structure sismique radiale de cette région est assez bien connue, mais il existe des variations latérales importantes selon les régions du globe. La figure 1.1 est un schéma qui synthétise les principaux facteurs responsables des anomalies de vitesses sismiques dans la zone de transition. Tout au long de ce chapitre, on y fera régulièrement référence. Certains facteurs sont bien identifiés tandis que d'autres sont moins bien contraints (teneur en eau, présence de fusion partielle, ...). Si l'on examine la situation à l'heure actuelle, on constate que les travaux récents visent à étudier le manteau dans son ensemble, en prenant en compte les effets de la température, de la minéralogie, de l'anisotropie, ainsi que le contenu en eau, et se placent dans le cadre d'études pluridisciplinaires. Une des perspectives actuelles des géophysiciens est de confronter les données provenant d'études très différentes, les données sismologiques et les données issues des expériences de laboratoire sur les minéraux du manteau, et d'élaborer des théories intégrant ces différentes données.

Dans cette première section, nous allons détailler les caractéristiques essentielles de la zone de transition présentées sur la figure 1.1. La zone de transition est caractérisée par un certain nombre de discontinuités et de gradients sismiques liés aux différentes transitions de phase et transformations minéralogiques des minéraux du manteau. Les profondeurs des deux principales discontinuités varient radialement et latéralement selon les conditions de pression, de température et les variations chimiques. Ces discontinuités jouent un rôle important au niveau géodynamique, en tant que barrières aux courants ascendants et descendants, ce qui implique des modèles de convection complexes. Malgré une amélioration croissante de la précision des données de laboratoire issues des expériences à hautes pressions et hautes températures, ces mesures ne permettent pas d'évaluer

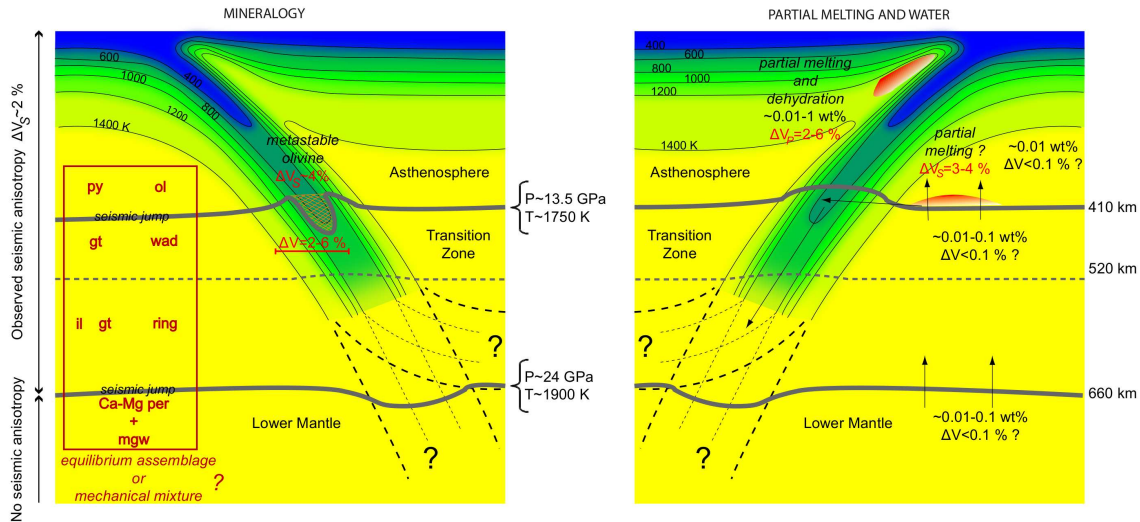


FIG. 1.1 : Figure schématique synthétisant les questions qui se posent à propos des variations des vitesses sismiques dans la zone de transition. Les variations de vitesses produites par différents facteurs sont indiquées en %. Au niveau de la minéralogie (à gauche), les expériences hautes pressions ont montré que les principales discontinuités sismiques à 410 et 660 km de profondeur sont liées aux transformations de l'olivine (e. g. Engdahl & Flinn, 1969; Adams, 1971). Les pyroxènes et les grenats subissent également des transformations, mais de façon plus graduelle avec la profondeur (e. g. Ito, 1977; Liu, 1977). De l'olivine métastable peut être présente à l'intérieur des plaques plongeantes (e. g. Vacher et al., 1999; Kawakatsu & Yoshioka, 2011). Ces plaques montrent des anomalies de vitesses positives et certaines arrivent à franchir la barrière endothermique que constitue la discontinuité à 660 km de profondeur, tandis que d'autres stagnent dans la zone de transition (e. g. van der Hilst et al., 1991; Fukao et al., 1992). La zone de transition est considérée comme étant anisotrope d'un point de vue sismique (Montagner, 1998), à la fois à l'échelle globale et régionale. Le contenu en eau (à droite) peut avoir un impact non négligeable sur l'orientation des cristaux, l'épaisseur des discontinuités et sur la fusion partielle (Karato, 2011).

précisément les contributions relatives de la température et de la composition minéralogique sur les vitesses sismiques. De plus, l'interprétation des profils sismiques dans la zone de transition s'avère particulièrement complexe car celle-ci est considérée comme étant anisotrope d'un point de vue sismique.

1.1.1 Les discontinuités sismiques

La zone de transition du manteau est définie comme étant la région délimitée par deux grandes discontinuités sismiques à 410 et 660 km de profondeur, à l'intérieur de laquelle la masse volumique et les vitesses sismiques augmentent brutalement (voir tableau 1.1). La limite supérieure est associée à la transition de phase de l'olivine $(\text{Mg,Fe})_2\text{SiO}_4$

profondeur (km)	$\Delta V_S/V_S$ (%)	$\Delta V_P/V_P$ (%)
410	3.6	2.6
660	6.7	4.7

TAB. 1.1 : Sauts de vitesses sismiques à 410 et 660 km de profondeur dans le modèle PREM ([Dziewonski & Anderson, 1981](#)).

en wadsleyite et la limite inférieure correspond principalement à la transformation minéralogique de la ringwoodite en pérovskite $(\text{Mg,Fe})\text{SiO}_3$ et magnésio-wüstite $(\text{Mg,Fe})\text{O}$ ([Katsura & Ito, 1989](#); [Ita & Stixrude, 1992](#)). La wadsleyite et la ringwoodite sont des polymorphes de l'olivine, c'est-à-dire qu'ils présentent la même formule chimique, mais qu'ils possèdent une structure cristalline plus compacte¹. À partir d'études pétrologiques, [Birch \(1952\)](#) a été le premier à formuler l'hypothèse que l'évolution des roches du manteau en fonction de la profondeur ne devait pas correspondre à une simple compression d'un matériau homogène, mais que les minéraux silicatés devaient se transformer en des phases plus denses, liées à des arrangements plus compacts. Les prédictions de [Birch \(1952\)](#) n'ont été confirmées par la sismologie que 15 ans plus tard (*e. g.* [Niazi & Anderson, 1965](#); [Ibrahim & Nuttli, 1967](#); [Julian & Anderson, 1968](#)). Les températures à 410 et 660 km de profondeur sont respectivement égales à 1750 ± 100 K et 1900 ± 150 K ([Ito & Takahashi, 1989](#)), et les pressions auxquelles ont lieu les transformations correspondent à environ 13.5 GPa et 24 GPa ([Akaogi et al., 1989](#); [Katsura & Ito, 1989](#)). Les premières cartographies à l'échelle globale des discontinuités à 410 et 660 km de profondeur ont été réalisées par [Shearer \(1991\)](#) et [Shearer & Masters \(1992\)](#) à partir des différences de temps de trajet entre les ondes SS et leurs précurseurs (voir section 1.1.2). Ces cartes soulignent le caractère global de ces deux interfaces.

Une transformation supplémentaire dans le système olivine a lieu au milieu de la zone de transition : il s'agit de la transition de phase wadsleyite $\rightarrow$ ringwoodite, qui se produit théoriquement à 520 km de profondeur. Les études sismiques ont d'abord révélé cette discontinuité uniquement sous les continents (*e. g.* [Gu et al., 1998](#)) et des études globales montrent que les relativement faibles contrastes d'impédance de part et d'autre de cette discontinuité s'y traduisent par une moindre réflexion des ondes sismiques que sur les discontinuités à 410 et 660 km de profondeur ([Flanagan & Shearer, 1998](#)). De plus, contrairement à la transition de phase à 410 km qui s'étale sur seulement 5 à 10 km d'épaisseur suivant les conditions de température, la wadsleyite et la ringwoodite peuvent coexister sur 10 à 50 km d'épaisseur ([Shearer, 2000](#)). À cause de cet intervalle de transition très large en profondeur, la discontinuité à 520 km serait sismiquement moins visible car elle ne correspond pas à un changement rapide des propriétés sismiques. De plus,

¹La structure cristalline de l'olivine est orthorhombique, celle de la wadsleyite est monoclinique et celle de la ringwoodite est cubique à structure spinelle.

cette discontinuité n'est pas répartie latéralement de manière homogène, à l'échelle globale (Deuss & Woodhouse, 2001). Elle n'est d'ailleurs pas prise en compte dans certains modèles de Terre de référence (par exemple le modèle AK135 de Kennett et al., 1995).

Les variations latérales des vitesses sismiques observées à différentes échelles sont une des manifestations de l'hétérogénéité du manteau. L'olivine et ses polymorphes comptent pour 40 à 60 % de la pétrologie du manteau total, selon le modèle pétrologique considéré. Le reste est composé de pyroxènes et de grenat, qui eux aussi subissent des transformations. L'effet des transitions de phase sur les discontinuités sismiques est évalué à partir de calculs thermodynamiques. Ces calculs confirment l'importance de considérer les minéraux de la composante non-olivine pour faire le lien entre les images tomographiques, la profondeur des discontinuités (Vacher et al., 1998) et l'origine thermique ou bien compositionnelle des anomalies (Cammarrano et al., 2003; Stixrude & Lithgow-Bertelloni, 2005a, 2010). Une synthèse sur les différentes transitions de phase et transformations minéralogiques qui ont lieu dans la zone de transition, ainsi que leurs profondeurs de référence, est montrée dans le tableau 1.2.

transitions de phase et transformations minéralogiques	profondeur de référence (km)	pente de Clapeyron (MPa/K)	référence
olivine → wadsleyite	410	+3.1 +4.0 +3.0	Akaogi et al. (2007) Katsura et al. (2004) Bina & Helffrich (1994)
wadsleyite → ringwoodite	520	+5.2 +6.9	Akaogi et al. (2007) Suzuki et al. (2000)
ringwoodite → pérovskite + magnésioiwüstite	660	-2.6 -1.3 -0.4 à -2.0 -3.0 -2.8	Akaogi et al. (2007) Fei et al. (2004) Katsura et al. (2003) Akaogi & Ito (1993) Ito & Takahashi (1989)
Cpx → Ca-pérovskite + grenat	540-560	+4.0	Saikia et al. (2008)
<i>région chaude :</i> grenat → pérovskite	660	+1.3	Hirose (2002)
<i>région froide :</i> grenat → ilménite	660	+4.0	Wang et al. (2004)
ilménite → pérovskite	660	-3.1	Fei et al. (2004)

TAB. 1.2 : Transitions de phase et transformations minéralogiques ayant lieu dans la zone de transition. Les valeurs des pentes de Clapeyron (notion définie dans la section 1.1.2) associées sont indiquées.

La présence de pyroxènes dans le manteau supérieur conduit à une transformation supplémentaire vers 520 km de profondeur, liée à la dissociation de clinopyroxène en Ca-pérovskite et grenat. Bien que cette transition ait lieu sur un intervalle de 40 à 60 km de profondeur, la majorité de la dissociation a lieu dans les 10 à 20 premiers kilomètres (Saikia et al., 2008). Dans les régions « froides » (par exemple au niveau des plaques plongeantes), trois transformations peuvent se produire : grenat $\rightarrow$ ilménite, puis ilménite $\rightarrow$ pérovskite, suivie par la transformation minéralogique ringwoodite $\rightarrow$ pérovskite + magnésioiwüstite. Ces deux dernières discontinuités sont proches en profondeur et peuvent apparaître comme une unique discontinuité (*e. g.* Weidner & Wang, 1998). Dans les régions « chaudes » (par exemple près des points chauds) a lieu la transition grenat $\rightarrow$ pérovskite. Dans une région « chaude » du manteau avec des grenats alumineux, la pérovskite n'est pas stable et la ringwoodite se transforme en grenat (et non en pérovskite) et magnésioiwüstite, suivie par la transformation de grenat en pérovskite alumineuse (Vacher et al., 1998; Weidner & Wang, 1998). À la base de la zone de transition, la présence d'aluminium à l'intérieur des grenats conduit à des discontinuités multiples qui dépendent à la fois de la température et de la composition du manteau (Weidner & Wang, 1998, 2000; Fei et al., 2004; Wang et al., 2004).

Les propriétés physiques d'une roche correspondent à une moyenne pondérée des propriétés individuelles de chacun des minéraux qui la composent. Les changements de phase des pyroxènes et des grenats ont lieu de façon plus graduelle dans le manteau que pour les polymorphes de l'olivine. Ils tendent à modifier les gradients des vitesses sismiques en fonction de la profondeur plutôt qu'à former des discontinuités discrètes. L'interprétation des profils de vitesse est donc complexe à cause des différentes transformations qui ont lieu dans les systèmes olivine et pyroxènes-grenat, car elles se réalisent à des intervalles de profondeurs qui parfois se superposent (tableau 1.2).

1.1.2 Variations spatiales des discontinuités

Les profils sismiques sont d'autant plus difficiles à interpréter que la largeur de l'intervalle de pression sur laquelle une transformation minéralogique se réalise dépend de la température et de la chimie du milieu. La figure 1.2 montre comme exemple le diagramme de phase de l'olivine dans le système fayalite-forstérite pour différentes températures. Ce sont les conditions de pression, de température et les variations chimiques qui définissent la profondeur des transformations. Les réactions ne se produisent pas instantanément, ce qui se traduit par une « épaisseur » des discontinuités terrestres. Les profondeurs et les épaisseurs des transformations vont donc varier latéralement si des variations latérales de ces paramètres existent. C'est l'analyse du contenu fréquentiel des ondes converties à travers les interfaces qui permet d'étudier ces phénomènes en détail. Les fréquences utilisées

sont en général entre 0.2 et 2.0 Hz. Il est généralement considéré que les ondes de surface possèdent des fréquences trop basses pour être affectées par ces phénomènes.

La pente de Clapeyron Γ relie la pression et la température aux changements de volume qui se produisent lors d'un changement d'état. Elle est définie pour une relation univariante selon l'équation :

$$\Gamma = \frac{dP}{dT} = \frac{\Delta S}{\Delta V}, \quad (1.1)$$

où P et T représentent respectivement la pression et la température, S l'entropie et V le volume. La pente de Clapeyron correspond à la trajectoire de la limite dans l'espace pression-température qui sépare les produits et les réactifs lorsqu'ils ont des compositions fixées. Les deux transformations principales à 410 et 660 km de profondeur ont des pentes de Clapeyron de signes opposés (voir tableau 1.2 et figure 1.3), ce qui signifie qu'en présence d'un même événement géodynamique, les deux discontinuités à 410 et 660 km de profondeur ont un comportement anticorrélé. D'après le tableau 1.2, il est clair que les valeurs obtenues pour les pentes de Clapeyron, en particulier pour celle de la transformation de ringwoodite en pérovskite et magnésiowüstite, sont sensiblement différentes selon les auteurs, et ce pour deux raisons. D'une part, les études expérimentales ne sont pas toutes effectuées dans les mêmes conditions (température de la transition légèrement différente, teneur en fer de l'olivine originelle variable, ...). D'autre part, les études théo-

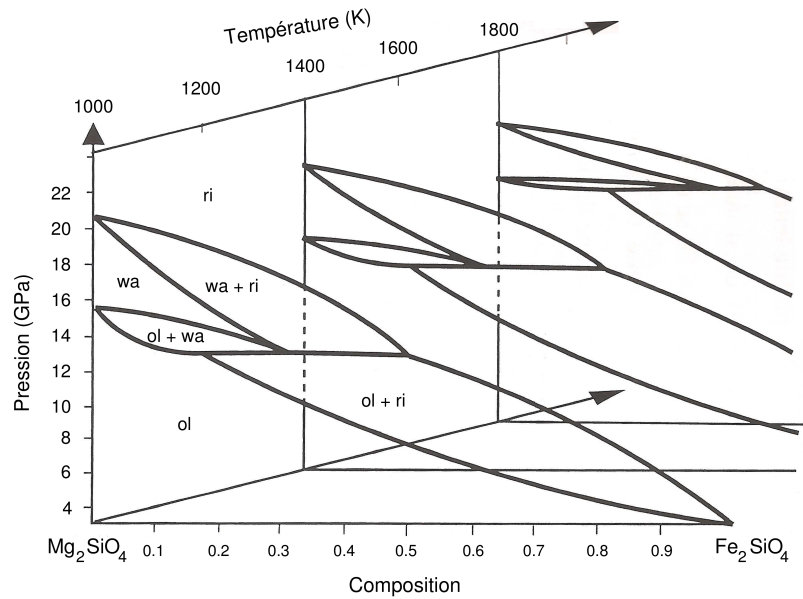


FIG. 1.2 : Domaines de transformations des polymorphes de l'olivine en fonction de la pression, de la température et de la proportion de fayalite Fe_2SiO_4 (d'après [Castillo, 2001](#), modifié). Les diagrammes de phase ont été construits d'après les valeurs de [Ito & Takahashi \(1989\)](#) et [Bina & Helffrich \(1994\)](#). Abréviations : ol : olivine, wad : wadsleyite, ri : ringwoodite.

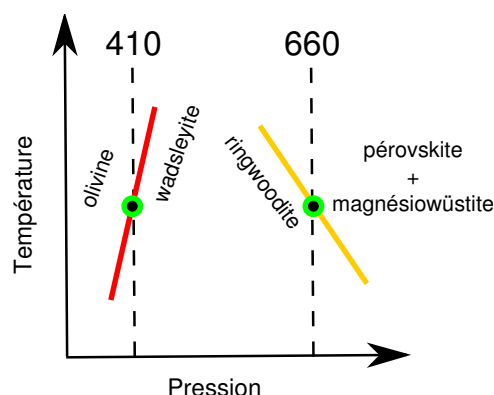


FIG. 1.3 : Représentation schématisée des pentes de Clapeyron de la transition de phase olivine $\rightarrow$ wadsleyite et de la transformation minéralogique ringwoodite $\rightarrow$ pérovskite + magnésio-wüstite.

riques n'utilisent pas forcément les mêmes équations d'état. Ces différentes estimations ont d'importantes conséquences, notamment sur les modèles numériques de convection qui prennent en compte les transitions de phase (voir section 1.1.3), et servent à l'interprétation en termes de température et de composition des ondulations des discontinuités. De manière générale, ce sont des valeurs situées autour de $\Gamma_{660} \sim 2.8 \text{ MPa/K}$ qui reviennent le plus souvent dans les publications.

Théoriquement, la zone de transition devrait s'épaissir dans les régions froides (par exemple lorsqu'une plaque plongeante traverse la zone de transition comme sur la figure 1.1), et s'affiner au niveau des régions chaudes, comme les panaches. À cause de leur importante dépendance par rapport à la température, les changements de phases minérales peuvent avoir lieu à plusieurs dizaines de kilomètres de différence par rapport à leurs positions moyennes respectives. La pente de Clapeyron de la discontinuité à 410 km prédit que la discontinuité devrait être élevée au niveau des zones de subduction et abaissée au niveau des points chauds. Un comportement inverse est attendu pour la discontinuité à 660 km (figure 1.1). Cette hypothèse peut être testée en cherchant une corrélation entre les topographies des discontinuités et les vitesses sismiques, en considérant que les variations de vitesses sont uniquement liées à la température.

Les précurseurs des ondes *SS* et *PP* sont les ondes les plus employées pour étudier les caractéristiques des discontinuités de la zone de transition à l'échelle globale. Les ondes *SS* (ou *PP*) sont des ondes de volumes télésismiques qui sont réfléchies à la surface de la Terre. Leur sensibilité n'est pas située directement sous la station, mais plutôt en un point situé entre la source sismique et la station (figure 1.4a). Leurs précurseurs² voyagent selon des chemins semblables, mais ils sont réfléchis sur des discontinuités et

²Les précurseurs sont nommés *SdS* (ou *PdP*), où *d* correspond à la profondeur de la discontinuité. Par exemple, *S660S* est le précurseur des ondes *SS* réfléchi sous la discontinuité à 660 km.

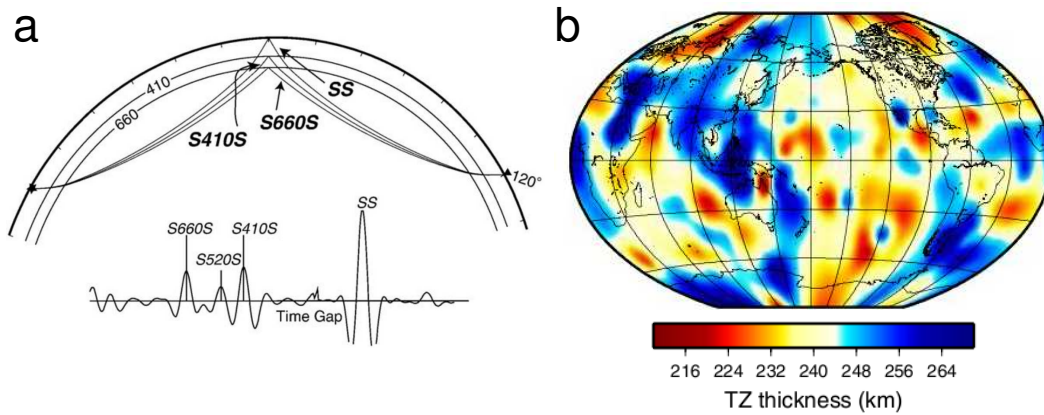


FIG. 1.4 : (a) En haut : géométrie des rais des précurseurs des ondes *SS*. En bas : exemple de sismogramme montrant les positions des arrivées des ondes *SS* et des ondes réfléchies sur les discontinuités de la zone de transition *S410S*, *S520S* et *S660S* (d'après [Houser & Williams, 2010](#)). (b) Mesures de l'épaisseur de la zone de transition en utilisant les précurseurs des ondes *SS* (d'après [Deuss, 2009](#)).

arrivent avant l'onde *SS* (ou *PP*) sur le sismogramme (figure 1.4a). Puisque les chemins empruntés par les ondes *SS* (ou *PP*) et leurs précurseurs sont très similaires, on considère que le temps différentiel entre l'arrivée des ondes *SS* (ou *PP*) est uniquement sensible aux structures entre la profondeur de la discontinuité et la surface, et que les autres hétérogénéités ont modifié l'onde principale de la même manière que ses précurseurs. Pour une synthèse sur les différentes techniques d'observations et d'analyse des précurseurs, le lecteur pourra se référer à [Deuss \(2009\)](#).

Pour évaluer la topographie des discontinuités, la différence entre les temps d'arrivée des précurseurs des différentes discontinuités et le temps d'arrivée de l'onde *SS* ou *PP* est donc mesurée. Les mouvements de la discontinuité à 410 km ne semblent pas toujours répondre aux perturbations thermiques du manteau ([Flanagan & Shearer, 1998](#); [Gu et al., 1998](#); [Houser et al., 2008](#)). [Chambers et al. \(2005\)](#) ont montré qu'une corrélation entre les vitesses sismiques et la température existe seulement pour de très grandes longueurs d'onde, mais qu'elle disparaît aux courtes longueurs d'onde. Ainsi, à longue période les variations de la topographie et des vitesses sismiques seraient liées au même facteur, c'est-à-dire la température, mais à plus courte période des hétérogénéités chimiques devraient jouer un rôle non négligeable. [Gossler & Kind \(1996\)](#) et [Gu et al. \(1998\)](#) suggèrent que la discontinuité à 410 km s'élève sous les continents. Cependant, d'autres études ne confirment pas cette corrélation ([Flanagan & Shearer, 1999](#); [Chambers et al., 2005](#)). Contrairement à la discontinuité à 410 km de profondeur, la topographie de la discontinuité à 660 km montre une corrélation avec les vitesses sismiques et s'enfonce au niveau des zones de subduction ([Flanagan & Shearer, 1998](#); [Gu et al., 1998](#); [Houser et al., 2008](#)).

Les mesures des temps d'arrivée des ondes *SS* (ou *PP*) et de leurs précurseurs doivent être corrigées des variations 3D de la structure de la croûte et du manteau supérieur, ce qui peut potentiellement affecter les résultats. Une mesure plus robuste est la détermination de l'épaisseur de la zone de transition, qui correspond à la différence de temps de trajet entre les phases réfléchies sur les discontinuités à 410 et 660 km. Cette mesure dépend uniquement de la structure entre ces deux profondeurs. Les études globales de l'épaisseur de la zone de transition (Gu et al., 2003; Houser et al., 2008; Lawrence & Shearer, 2008; Deuss, 2009) indiquent un épaississement au niveau des zones de subduction. La figure 1.4b montre qu'il est particulièrement important à l'ouest de l'Océan Pacifique. Cette tendance est confirmée par les études régionales dans lesquelles par exemple, une élévation de la discontinuité à 410 km de 60-70 km (Collier et al., 2001) et une déflexion de 57 km pour la discontinuité à 660 km (Tibi & Wiens, 2005) ont été remarquées pour la plaque plongeante d'Izu-Bonin, ce qui donne une épaisseur de la zone de transition d'environ 370 km.

Bien que la température joue un rôle primordial sur la topographie des principales discontinuités, l'amplitude des variations de leur localisation fluctue à l'échelle globale et indique qu'il ne s'agit pas de l'unique facteur de contrôle de la topographie. L'amplitude de ces variations peut en effet refléter dans les zones de subduction, en plus des différences thermiques, des différences compositionnelles, ce qui peut affecter à la fois la profondeur et l'épaisseur des discontinuités (*e. g.* Wood, 1995). Bien qu'elles soient moins bien contraintes que pour le système olivine, des transitions de phase exothermiques et endothermiques ont lieu dans le système pyroxènes-grenat (tableau 1.2). Les profondeurs des transitions de phase impliquant le grenat sont fortement dépendantes de la quantité de pyrope (c'est-à-dire la quantité d'aluminium du grenat), qui peut entraîner des variations de la profondeur des discontinuités de plus de 50 km (Weidner & Wang, 1998). Les variations de topographie des discontinuités minéralogiques sont susceptibles d'être liées aux effets cumulés de plusieurs réactions de pentes de Clapeyron de mêmes signes ou bien de signes opposés, ce qui peut conduire à des variations à des profondeurs éloignées de celles théoriquement attendues dans le cas où une composition d'olivine pure est considérée. Récemment, la présence d'olivine métastable a été décelée au niveau du sud-ouest du Japon (Kawakatsu & Yoshioka, 2011). La transformation olivine $\rightarrow$ wadsleyite peut être effectivement retardée à l'intérieur de la partie la plus froide de la plaque plongeante. Ainsi, l'olivine peut persister plus profondément qu'à 410 km de profondeur (figure 1.1).

1.1.3 Géodynamique de la zone de transition

Les discontinuités décrites précédemment jouent un rôle de premier plan dans la dynamique interne de la planète, en tant qu'obstacles aux courants convectifs ascendants et descendants.

Au fur et à mesure que les plaques plongeantes s'enfoncent dans le manteau au niveau des zones de subduction, des matériaux sont déposés à différentes profondeurs. Ces matériaux sont ensuite dispersés à travers le manteau par l'intermédiaire des mouvements de matière (*e. g.* [Christensen & Hofmann, 1994](#); [Nakagawa et al., 2009](#)). Progressivement, un équilibre thermique se met en place pour réduire les différences de rhéologie et de masse volumique. Ainsi, les matériaux déposés se mélangent au manteau environnant. Bien que les études sismologiques indiquent que la composante à grande longueur d'onde des hétérogénéités du manteau est généralement dominante (*p. e.* [Becker & Boschi, 2002](#)), ces processus sont responsables de la production d'hétérogénéités à toutes les échelles, radiales comme latérales. Des indices sismologiques et chimiques de la présence de ces hétérogénéités ont d'ailleurs été mis en évidence (*e. g.* [Helffrich & Wood, 2001](#); [Tackley et al., 2005](#)). Par exemple, des études utilisant les propriétés des ondes diffractées (*e. g.* [Hedlin et al., 1997](#)) ont détecté des corps dans le manteau inférieur dont la taille est d'environ 10 km. Étant donnée leur taille, il s'agirait plutôt d'hétérogénéités chimiques que thermiques, vestiges de plaques plongeantes. Un autre exemple est l'existence de basaltes qui montrent dans leur signature isotopique des évidences de la contribution d'une croûte océanique ayant subi un processus de subduction ([Hofmann, 1997](#)).

Les hétérogénéités du manteau induites par le processus de subduction sont ainsi présentes à tous les niveaux. Elles se rencontrent à des échelles variées, qui représentent les différents stades du processus de mélange : de l'ordre du millier de kilomètres pour une plaque plongeante ou bien pour la longueur d'onde de la topographie des discontinuités de la zone de transition, à la dizaine de kilomètres pour des morceaux de plaques plongeantes ([Helffrich, 2006](#)), et jusqu'à l'échelle microscopique (chimique) ([Hofmann, 1997](#)). Malgré ce modèle pour la production et la maintenance d'hétérogénéités latérales dans le manteau, et en particulier dans la zone de transition, la nature de ces hétérogénéités aux différentes échelles observées n'est pas clairement connue. Plusieurs questions se posent : Sont-elles d'origine compositionnelle ou thermique ? Ou bien une combinaison des deux ? Et quelle est leur relation avec le mode de convection du manteau ? Une des clés pour répondre à ces questions se trouve en partie dans les images obtenues par la tomographie sismique.

Historiquement, ce sont les images tomographiques régionales produites au début des années 1990 qui ont donné de nouveaux éléments de réflexion sur les modalités des processus de convection. L'étude de [Okino et al. \(1989\)](#) a été la première à détecter, à partir d'une analyse des temps de trajet, que les plaques plongeantes pouvaient présenter un comportement horizontal dans le manteau. Puis [van der Hilst et al. \(1991\)](#) ont montré, par l'intermédiaire d'une tomographie des ondes *P*, que les plaques plongeantes sous les arcs du Japon et d'Izu-Bonin subissent une flexion à 660 km de profondeur, tandis que sous les Mariannes et l'arc des Kouriles, elles pénètrent dans le manteau inférieur.

Par la suite, de nombreuses études tomographiques ont confirmé que les plaques plongeantes ont des comportements différents à l'intérieur du manteau (Fukao et al., 1992, 2001; Grand, 1994; Bijwaard et al., 1998; Kárason & van der Hilst, 2000; Li et al., 2008). Sur la figure 1.5, trois comportements peuvent être distingués. Certaines plaques stagnent à la base de la zone de transition (coupes BB' et CC'), d'autres subissent une flexion dans la zone de transition avant de pénétrer dans le manteau inférieur (coupe EE'), et d'autres encore plongent directement dans le manteau inférieur (coupes AA', DD' et FF'). À la

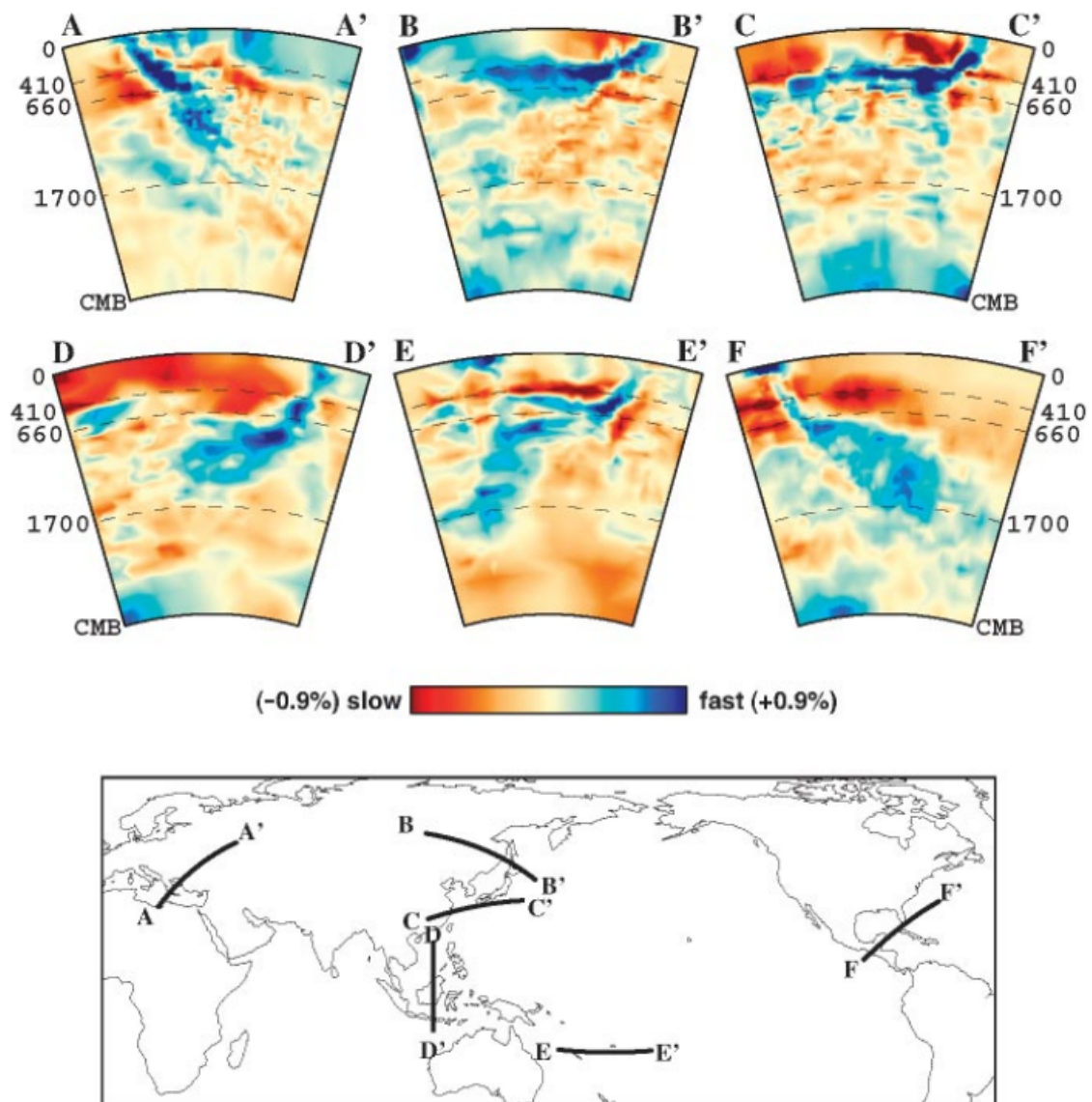


FIG. 1.5 : Exemples de sections verticales de vitesse des ondes P dans le manteau au niveau de plusieurs zones de subduction (d'après Kárason & van der Hilst, 2000). Elles révèlent la présence d'anomalies rapides des ondes P associées aux plaques plongeantes.

suite de ces observations, deux modèles extrêmes de convection ont été proposés : un modèle à une couche qui affecterait l'ensemble du manteau et un modèle de convection à deux couches (Hofmann, 1997; Turcotte et al., 2000). Les arguments en faveur d'un système de convection à deux couches proviennent d'études géochimiques (Hofmann, 1997). Le manteau supérieur est appauvri en éléments traces incompatibles, comparé à la composition moyenne théorique d'une Terre primitive, ce qui signifie que des matériaux complémentaires primitifs doivent exister ailleurs. De plus, il existe des basaltes de compositions isotopiques distinctes, ce qui implique la présence de réservoirs différents dans le manteau. En revanche, les observations sismologiques indiquent que les plaques plongeantes peuvent atteindre le manteau inférieur, ce qui est plutôt en faveur d'un scénario de convection globale. Que la convection soit à une ou deux couches dans le manteau, il existe certainement des interactions entre les manteaux supérieur et inférieur. Ces dernières années, on attribue une complexité plus importante au style de convection dans le manteau que ces deux modèles extrêmes. En effet, des modèles alternatifs ont été proposés en faveur d'un système de convection partiellement stratifié (*e. g.* Tackley, 2008).

Les modèles numériques de convection qui prennent en compte les transformations minéralogiques, et qui introduisent les valeurs des pentes de Clapeyron dans leurs simulations, ont largement alimenté le débat sur une éventuelle stratification de la convection mantellique. La discontinuité à 660 km de profondeur est l'une des structures les plus importantes de l'intérieur de la Terre. Les modèles numériques de convection du manteau montrent effectivement que la présence d'une transition de phase endothermique stabilise une convection stratifiée (*e. g.* Sotin & Parmentier, 1989), excepté lors de courtes périodes d'avalanches dans le manteau inférieur (*e. g.* Tackley et al., 1984; Tackley, 2008). Cependant, ces modèles sont dépendants des valeurs des pentes de Clapeyron utilisées. Ils sont confrontés à une revue à la baisse de la pente de Clapeyron de la transformation minéralogique ringwoodite $\rightarrow$ pérovskite + magnésio-wüstite (voir tableau 1.2). Des expériences récentes utilisant la diffraction des rayons X ont trouvé des valeurs comprises entre -0.4 et -1.3 MPa/K (Katsura et al., 2003; Fei et al., 2004; Litasov et al., 2005), ce qui est plus faible que les valeurs de -3.0 à -2.0 MPa/K déterminées par des études plus anciennes (Ito & Takahashi, 1989; Bina & Helffrich, 1994; Irifune et al., 1998).

Deschamps & Tackley (2009) ont réalisé des simulations de convection thermo-chimique en se focalisant sur les résultats qui montrent une bonne adéquation avec le modèle de tomographie sismique de Trampert et al. (2004). Ces auteurs en ont déduit un modèle de convection hybride, dont une représentation schématique est montrée sur la figure 1.6. Ils ont constaté que la transformation minéralogique endothermique à 660 km de profondeur empêche de façon significative les échanges de matière et aide à conserver des hétérogénéités de composition importantes dans le manteau inférieur. Une stratification compositionnelle globale est maintenue de chaque côté de la discontinuité. Les

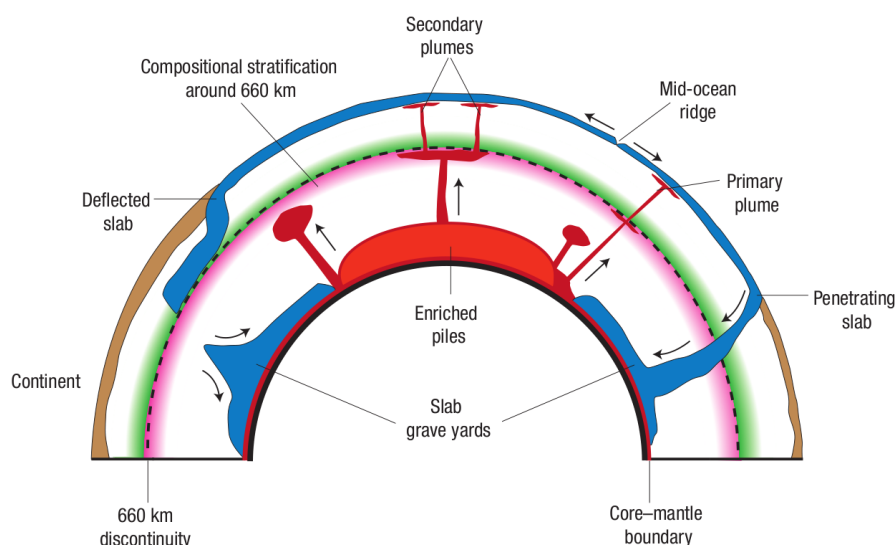


FIG. 1.6 : Modèle de convection alternatif de [Tackley \(2008\)](#).

modèles de convection qui fournissent le meilleur accord avec le modèle sismologique possèdent des valeurs de pente de Clapeyron comprises entre -1.5 et -3.0 MPa/K. La pénétration de plaques plongeantes dans le manteau inférieur serait épisodique dans le temps et l'espace. Dans ce type de modèle, les avalanches à travers la discontinuité à 660 km de profondeur sont liées aux ondulations de l'interface sous l'effet des variations de température. Elles ont lieu quand les anomalies de densité accumulées au-dessus de cet interface sont suffisamment importantes pour compenser les forces de bouillonnement³. Les anomalies de densité les plus efficaces pour pénétrer la discontinuité à 660 km de profondeur sont caractérisées par de grandes longueurs d'onde horizontales. Les panaches sont au contraire stoppés par la discontinuité, et des panaches secondaires, plus petits, sont générés au-dessus de la zone de transition (figure 1.6). Le manteau profond contiendrait des piles de matériaux primitifs et subductés. Bien que ce modèle de convection hybride soit basé sur des modélisations numériques qui contiennent de nombreuses simplifications (par exemple, les transformations du grenat n'ont pas été prises en compte par [Deschamps & Tackley, 2009](#)), il est certainement plus réaliste que les deux modèles extrêmes de convection proposés il y a 30 ans, et décrit le manteau comme étant une région extrêmement complexe, hétérogène à toutes les échelles, spatiales et temporelles.

³La force de « bouillonnement » est la poussée d'Archimède résultant des différences de masse volumique liées à une différence de température.

1.1.4 Réconcilier les données sismologiques et les données de laboratoire

La physique des matériaux terrestres a pour but de comprendre la composition interne de la Terre aussi bien d'un point de vue statique (quels sont les minéraux présents à une profondeur donnée ?) que d'un point de vue dynamique (comment les transitions de phase influent-elles sur la circulation de matière ?). Pour répondre à ces questions, des études expérimentales fournissent les données thermiques et élastiques sur les minéraux, et des études théoriques permettent d'extrapoler ces données à des pressions et des températures inaccessibles expérimentalement.

Depuis les années 1980, deux modèles sont classiquement utilisés pour décrire la minéralogie du manteau. Le premier est le modèle pyrolitique ([Ringwood, 1975](#)). Ce modèle décrit une roche théorique, la *pyrolite*, qui permet d'expliquer l'extraction des liquides basaltiques et la formation d'un résidu péridotitique. Le deuxième modèle, la *piclogite* ([Bass & Anderson, 1984](#)), est caractérisée par une proportion plus importante de grenat et de pyroxènes, et moins d'olivine. Dans la suite de cette thèse, les termes *pyrolite* et *piclogite* feront référence aux compositions listées dans le tableau 1.3. Le modèle piclogitique original possédait 16 % d'olivine en volume ([Bass & Anderson, 1984](#)). En se basant sur des estimations de l'élasticité du grenat à haute pression, la fraction volumique d'olivine dans la piclogite a été re-évaluée à 40 % ([Duffy & Anderson, 1989](#)). Ce modèle a été proposé dans le cadre d'une évolution du manteau qui conduit à une différenciation chimique du manteau supérieur par rapport à la zone de transition (voir section 1.1.3). Il y a une vingtaine d'années, [Ita & Stixrude \(1992\)](#) ont montré que les deux modèles

mol (%)	pyrolite	piclogite	basalte	harzburgite
SiO ₂	38.71	41.94	51.75	36.07
MgO	49.85	42.32	14.94	56.51
FeO	6.17	5.29	7.06	6.07
CaO	2.94	8.67	13.88	0.81
Al ₂ O ₃	2.22	1.78	10.19	0.53
Na ₂ O	0.11	0.00	2.18	0.00
fraction volumique (%)	pyrolite	piclogite		
olivine	61	40		
orthopyroxène	15	22		
clinopyroxène	10	8		
grenat	14	30		

TAB. 1.3 : Compositions chimiques et minéralogiques d'une pyrolite (valeurs de [Irfune, 1987](#)), d'une piclogite (valeurs de [Bass & Anderson, 1984](#)), d'un basalte (valeurs de [Workman & Hart, 2005](#)) et d'une harzburgite (valeurs de [Baker & Beckett, 1999](#)).

s'ajustaient bien aux modèles 1D de référence comme le PREM (Dziewonski & Anderson, 1981) ou AK135 (Kennett et al., 1995). En particulier, plusieurs études (*e. g.* Weidner, 1985; Ita & Stixrude, 1992) ont constaté qu'un modèle pyrolitique avec un gradient adiabatique et une température potentielle⁴ de 1300 °C pouvait expliquer les modèles de référence au premier ordre.

Plus récemment, Cammarano et al. (2005a,b) ont déterminé que, compte-tenu des incertitudes actuelles sur les modules élastiques des minéraux du manteau, le comportement sismique d'un tel modèle n'est pas capable d'expliquer les modèles sismiques de référence (figure 1.7). En effet, le profil adiabatique à 1300 °C possède un gradient de vitesse plus faible dans la zone de transition, ainsi qu'un saut de vitesse plus faible à 660 km de profondeur. Au-dessous de cette discontinuité, les vitesses sismiques sont également plus faibles. Les transformations de phases à une composition fixée ne paraissent pas suffisantes pour rendre compte des phénomènes sismiques.

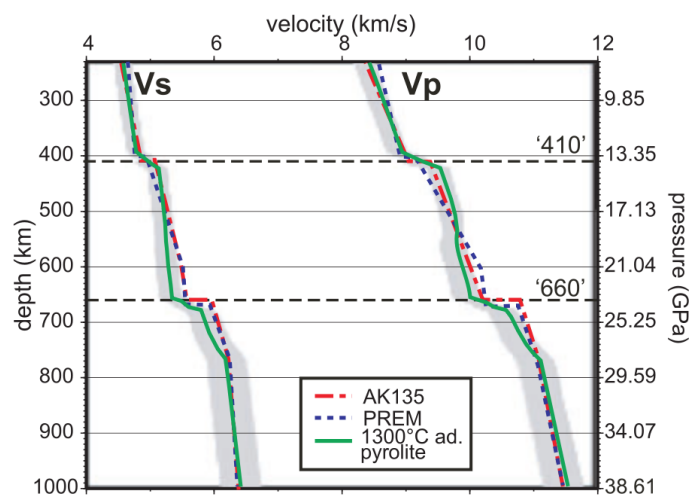


FIG. 1.7 : Comparaison entre le profil adiabatique pyrolitique à 1300 °C avec les modèles de référence AK135 (Kennett et al., 1995) et PREM (Dziewonski & Anderson, 1981). Les bandes grises indiquent les incertitudes sur le modèle pyrolitique adiabatique (d'après Cobden et al., 2008).

Cammarano & Romanowicz (2007) et Cobden et al. (2008) ont proposé que la composition globale moyenne du manteau se modifie avec la profondeur pour rendre compte des données sismiques, avec une région au-dessus de 660 km enrichie en matériaux basaltiques, et une région au-dessous de 660 km enrichie en harzburgite appauvrie. Une telle stratification est prédite par les calculs dynamiques qui prennent en compte la dépendance de la localisation des transitions de phase par rapport à la composition (Nakagawa et al., 2009).

⁴La température potentielle d'un matériau correspond à la température que celui-ci aurait s'il était détendu adiabatiquement à un niveau de pression standard (ici la surface de la Terre).

Jusqu'à présent, seules des compositions à l'équilibre chimique ont été évoquées. [Xu et al. \(2008\)](#) ont considéré deux modèles de composition du manteau composés de basalte et d'harzburgite. Il ont d'abord étudié un modèle à l'équilibre, où les composants basalte et harzburgite sont équilibrés chimiquement, et un mélange mécanique, qui représente le scénario extrême dans lequel le manteau pyrolitique a réalisé une complète différenciation en basalte et en harzburgite. Dans ce dernier modèle, les propriétés de la roche globale sont calculées en moyennant les propriétés des minéraux des composants basalte et harzburgite. En faisant l'hypothèse d'une composition mantellique pyrolitique, [Xu et al. \(2008\)](#) ont trouvé, sur la base d'une comparaison qualitative, que le mélange mécanique est plus proche des modèles de référence que l'assemblage à l'équilibre. D'un point de vue pétrologique, un modèle totalement différencié pour le manteau terrestre est difficilement concevable. En effet, un tel modèle n'est pas compatible avec le volcanisme basaltique au niveau des rides médio-océaniques. Cependant, étant donné les faibles vitesses des mouvements de convection ([Kellogg et al., 2002](#)) et la faible diffusion chimique dans le manteau solide ([Farber et al., 1994](#)), il est possible qu'une homogénéisation complète entre le basalte subducté et l'harzburgite ne soit pas réalisée, dans certaines régions du manteau.

En plus des hétérogénéités chimiques, un facteur qui pourrait également affecter la valeur des gradients de vitesses sismiques est le contenu en eau. Les constituants majeurs de la zone de transition, la wadsleyite et la ringwoodite, ont la capacité d'accommoder de l'eau jusqu'à 3 % en poids ([Inoue et al., 1995](#); [Kohlstedt et al., 1996](#)), à des pressions correspondant aux profondeurs de la zone de transition, alors que les manteaux supérieur et inférieur possèdent une faible capacité de stockage (figure 1.1). Cette capacité tend à diminuer avec l'augmentation de la température, mais [Smyth et al. \(2003\)](#) ont estimé qu'un contenu en eau d'environ 1 % en poids pouvait s'intégrer dans la structure de la wadsleyite et de la ringwoodite à 1300 °C. La zone de transition constitue donc un réservoir d'eau potentiel. Bien que la quantité maximale d'eau dans la zone de transition soit relativement bien contrainte par les données expérimentales sur la solubilité de l'eau dans les minéraux du manteau, la quantité d'eau supposée demeure très controversée. De petites quantités d'eau peuvent affecter de façon significative les propriétés des minéraux du manteau, entraînant un abaissement du solidus et une diminution des vitesses sismiques ([Karato, 2011](#)). Cependant, les études de [Jacobsen & Smyth \(2006\)](#), [Jacobsen et al. \(2008\)](#) et [Mao et al. \(2008\)](#) montrent que l'influence de l'eau sur les minéraux du manteau est mineure pour des quantités d'eau inférieure à 0.1 %. L'influence de l'eau reste faible sur les vitesses sismiques et est inférieure à celle des variations de composition ou de la température ([Karato, 2011](#)). La présence d'eau pourrait expliquer l'amplitude de la topographie à 660 km de profondeur, en bonne adéquation avec une faible pente de Clapeyron ([Litasov et al., 2005](#)), bien que des études supplémentaires soient nécessaires pour clarifier le lien

entre les expériences de laboratoire et les calculs théoriques. De la fusion partielle dans le coin du manteau (*e. g.* Zhao, 2001), liée à la déshydratation de la plaque plongeante, et près de la discontinuité à 410 km de profondeur (Revenaugh & Sipkin, 1994; Bercovici & Karato, 2003; Jasbinsek & Dueker, 2007; Tauzin et al., 2010) pourrait réduire localement les vitesses sismiques (voir figure 1.1). De plus, l'eau transporte des éléments dissous dans les fluides, ce qui peut également modifier la localisation des limites de phases ainsi que l'épaisseur des transitions (*e. g.* Wood, 1995).

1.1.5 Anisotropie sismique

Les variations de vitesses des ondes sismiques sont liées non seulement aux anomalies de température ou de composition chimique, mais aussi à l'anisotropie sismique. La cartographie des déplacements de manière dans le manteau est basée sur l'interprétation de l'anisotropie sismique observée en utilisant une relation entre l'orientation cristallographique préférentielle (appelée CPO ou LPO pour *Crystallographic/Lattice Preferred Orientation*) des minéraux du manteau et la géométrie des flux (Mainprice, 2007; Karato, 2008). Lors d'un processus de déformation, les cristaux individuels qui composent les roches mantelliques se déforment selon des directions en accord avec le mouvement. L'écoulement convectif n'est possible que parce qu'il existe des défauts dans les structures cristallines des minéraux. Le fluage des dislocations des minéraux entraîne la formation de réseaux cristallins orientés selon des directions préférentielles. Ainsi, l'anisotropie sismique est une conséquence directe de la convection et fournit l'histoire intégrée dans le temps de la déformation. Une structure anisotrope peut également être produite par une orientation préférentielle des minéraux (appelée SPO pour *Shape Preferred Orientation*), qui est associée à un alignement de matériaux avec des propriétés isotropes distinctes qui contrastent avec la matrice environnante, comme des poches de liquides ou des failles dans la croûte. De plus, une périodicité d'hétérogénéités isotropes peut conduire à une anisotropie extrinsèque. Dans ce cas, la longueur d'onde de l'onde sismique est plus importante que la longueur d'onde dominante des hétérogénéités. Ainsi, une succession de fines couches de matériaux avec des propriétés sismiques contrastées s'intercalant de façon périodique peut se comporter comme un milieu anisotrope homogène (*e. g.* Backus & Gilbert, 1962).

Il existe deux types d'anisotropie sismique qui peuvent être observées par l'intermédiaire des ondes dans le manteau. En considérant une cellule de convection simple, la figure 1.8 présente ces deux types d'anisotropie :

- l'*anisotropie radiale* ou *de polarisation* (ξ), qui fait référence à la différence de vitesse de propagation des ondes polarisées horizontalement (V_{SH}) et verticalement (V_{SV}). Ce phénomène n'existe que dans un milieu transversalement isotrope, c'est-à-dire dans un milieu qui contient un axe de symétrie vertical. Dans le cas où $V_{SV} > V_{SH}$, il s'agit

d'un mouvement vertical de matière. En revanche, si $V_{SV} < V_{SH}$, l'écoulement a lieu horizontalement.

- l'*anisotropie azimutale* ($\mathcal{G}$), qui correspond à la dépendance de la vitesse des ondes en fonction de l'azimut. Un écoulement horizontal induit une forte anisotropie azimutale, tandis que les régions de courants descendants ou ascendants sont plutôt caractérisées par de faibles variations azimutales.

Ces deux types d'anisotropie correspondent à deux manifestations du même phénomène. [Montagner & Nataf \(1986\)](#) ont élaboré une technique qui permet d'expliquer de façon simultanée ces deux formes d'anisotropie. Cette méthode peut être simplifiée en considérant que le milieu possède un axe de symétrie avec une certaine orientation. Dans ce cas, le nombre de paramètres anisotropes indépendants est réduit à 7 au lieu de 13 pour les ondes de surface ([Montagner & Nataf, 1988](#)).

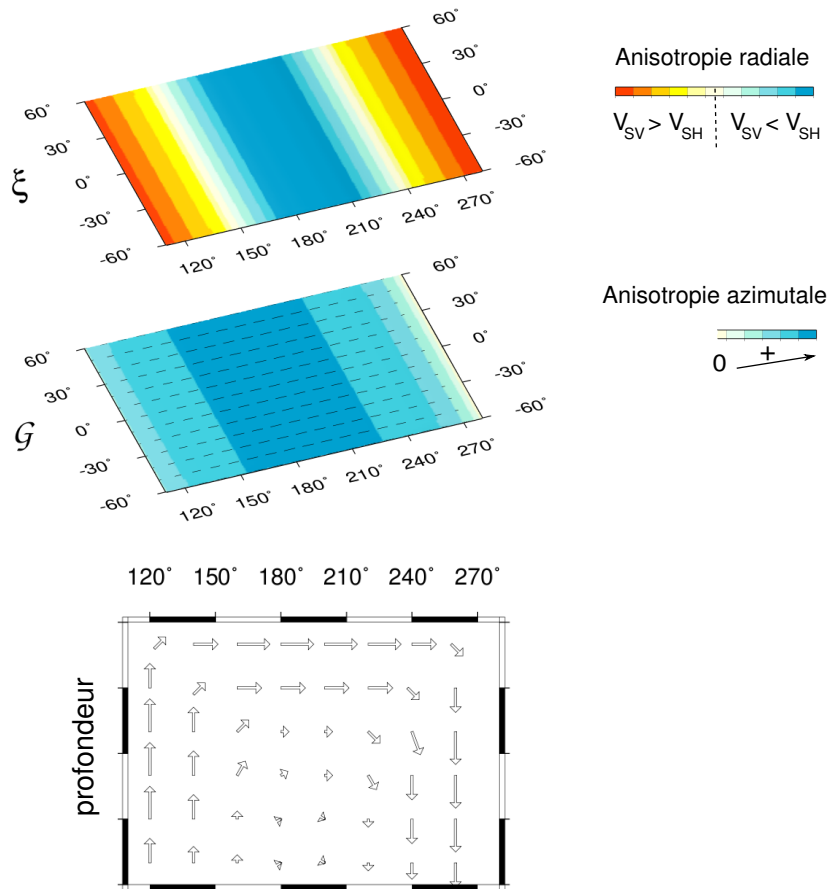


FIG. 1.8 : Schéma d'une cellule convective simple. Le mouvement circulaire de matière est indiqué par des flèches. L'influence de cet écoulement sur les paramètres d'anisotropies radiale (ξ) et azimutale ($\mathcal{G}$) est montrée. Les traits du paramètre $\mathcal{G}$ représentent la LPO d'un minéral anisotrope tel que l'olivine ou encore la wadsleyite (d'après [Montagner, 2002](#), modifié).

L'anisotropie sismique dans le manteau supérieur est interprétée en terme de LPO de l'olivine (Jung & Karato, 2001; Ohuchi et al., 2011) et de la serpentine (Katayama et al., 2009), ce qui indique que le style de convection dans le manteau supérieur est dominé par des flux horizontaux. Les directions de ces flux sont quasiment parallèles aux mouvements des plaques, que ce soit dans les zones de subduction (Katayama et al., 2009), ou même à l'échelle globale (Becker et al., 2003). D'après le tableau 1.4, on constate que l'olivine est l'un des minéraux les plus anisotropes. En considérant un seul cristal d'olivine, son anisotropie intrinsèque sur les ondes de cisaillement est de 18 %. Puis son anisotropie diminue tandis que ce minéral prend les formes compactes de la wadsleyite et de la ringwoodite. Des minéraux constitutifs de la zone de transition, c'est la wadsleyite qui montre l'anisotropie intrinsèque la plus importante (tableau 1.4). L'anisotropie sismique est donc un facteur important à prendre en compte lors d'études tomographiques pour ne pas imager des structures erronées. Cependant, l'anisotropie sismique sera observée seulement si la déformation dans la zone de transition conduit à un développement cohérent d'orientations préférentielles des cristaux aux échelles échantillonnées par les ondes sismiques. Ainsi, il est difficile de déterminer si l'effet est lié à de petites structures de l'ordre du centimètre ou bien à des structures de plusieurs dizaines de kilomètres. De plus, les orientations préférentielles des cristaux peuvent résulter en des motifs qui diffèrent de façon significative de ceux d'un cristal unique. Des structures stratifiées ou des inclusions orientées de matériaux avec des propriétés élastiques contrastées peuvent également produire de l'anisotropie sismique.

Minéral	$\Delta V_P/V_P$ (%)	$\Delta V_S/V_S$ (%)	références
olivine	25	18	Isaak (1992)
wadsleyite	16	17	Sawamoto et al. (1984)
ringwoodite	2.3-4.7	4.8-10.3	Kiefer et al. (1997); Sinogeikin et al. (1998)
pérovskite	7.6-13.7	15.4-33	Yeganeh-Haeri et al. (1989); Sinogeikin et al. (2004)
magnésiowüstite	11	21.5	Isaak et al. (1989)
orthopyroxène	12-15.1	11-15	Frisillo & Barsch (1972); Weidner et al. (1978); Webb & Jackson (1993)
clinopyroxène	29	20-24	Collins & Brown (1998)
majorite	1.8	9.1	Pacalo & Weidner (1997)
ilménite	21.1	36.4	Weidner & Ito (1985)

TAB. 1.4 : Données d'anisotropie intrinsèque de vitesses des ondes P et S pour les minéraux du manteau. Les références sont tirées de Mainprice (2007).

L'anisotropie sismique dans la zone de transition a été mise en évidence à l'échelle globale par une analyse jointe des temps de trajet des ondes de volume et des fréquences des oscillations libres (Montagner & Kennett, 1996). Les données peuvent être réconciliées par une faible anisotropie radiale dans la zone de transition où $V_{SV} < V_{SH}$, ce qui est en bonne corrélation avec l'accumulation de phases présentant une faible anisotropie intrinsèque comme la majorite et la ringwoodite. Plus récemment, Visser et al. (2008b) ont reporté la présence d'une anisotropie radiale $V_{SV} > V_{SH}$, ce qui est en opposition avec les résultats de Montagner & Kennett (1996). Une anisotropie azimutale a été mise en évidence par Trampert & van Heijst (2002) en utilisant les ondes de surface. À l'échelle régionale, Chen & Brudzinski (2003) ont détecté de l'anisotropie au niveau de la zone de subduction des Tonga-Fiji, qu'ils ont attribué à la présence d'une plaque plongeante. Fouch & Fisher (1996) ont reporté certaines régions isolées d'anisotropie en haut de la zone de transition (entre 410 et 520 km) sous l'arc des Kouriles, qu'ils attribuent à l'alignement préférentiel des cristaux de wadsleyite. D'autres études régionales n'ont pas mis en évidence la présence d'anisotropie au milieu du manteau (*e. g.* Fischer & Wiens, 1996).

À l'intérieur du manteau, l'anisotropie ne semble pas distribuée de façon uniforme. Elle pourrait être concentrée au niveau des discontinuités du manteau convectif, où la déformation est plus importante (Montagner, 2007). Cependant, les mécanismes d'orientation des cristaux et des minéraux dans la zone de transition demeurent encore mal compris. De plus, l'influence de composants mineurs tels que l'eau pourraient affecter profondément les orientations des cristaux (Jung & Karato, 2001). D'après l'ensemble de ces études, la localisation de l'anisotropie dans la zone de transition ainsi que son amplitude restent peu contraintes, et son interprétation s'avère complexe.

1.1.6 Hypothèses de travail

À travers les paragraphes précédents, nous avons vu que la zone de transition est une région extrêmement hétérogène, et ce à toutes les échelles. Au premier ordre, ce sont d'abord la température puis la minéralogie qui produisent l'impact le plus important sur les vitesses sismiques et la masse volumique. D'autres facteurs tels que l'anisotropie, le contenu en eau ou encore la fusion partielle peuvent également affecter les propriétés sismiques, mais de manière plus locale.

La démarche que nous avons choisie d'adopter dans cette thèse, afin d'obtenir des éléments de réponse sur les principaux facteurs responsables des anomalies de vitesses sismiques dans la zone de transition, se situe dans la continuité des travaux de Cammarano et al. (2003) et Cammarano et al. (2005a,b). L'approche est de considérer un modèle minéralogique, puis de déterminer ses différentes phases en fonction de la profondeur et de la température, et enfin de calculer les propriétés sismiques de l'assemblage minéralogique à partir des données de physique des minéraux obtenues à hautes pressions et hautes

températures. Les discontinuités sismiques sont directement corrélées à la température et aux proportions des différentes phases minéralogiques. L'objectif est ensuite de comparer les propriétés sismiques ainsi obtenues à celles des modèles globaux.

Cependant, peu de moyens sont à notre disposition pour observer la zone de transition. Les ondes de volume permettent d'obtenir un haut degré de résolution avec les phases réfléchies ou diffractées sur les interfaces, mais les points d'acquisition sismiques sont mal répartis à la surface de la Terre, ce qui résulte en une résolution très inhomogène. Au contraire, les ondes de surface permettent de couvrir efficacement l'ensemble du globe. Nous avons choisi de travailler avec ces dernières, et plus précisément avec leurs courbes de dispersion, intégrées le long d'un trajet source-station. Les ondes de surface possèdent un contenu fréquentiel à basse fréquence, et il est généralement considéré qu'elles ne sont que peu affectées par la position et les épaisseurs des discontinuités. Contraindre la composition minéralogique dans la zone de transition à partir des ondes de surface uniquement paraît difficile, c'est pourquoi notre travail est principalement focalisé sur les profils de température et d'anisotropie radiale capables d'expliquer les vitesses de phase des ondes de surface, pour une minéralogie fixée. La nature des équations qui relient les paramètres aux données (voir chapitre 2) est extrêmement non-linéaire. Dans un tel cas, nous allons voir que la méthode employée pour rechercher les paramètres dans l'espace des modèles est cruciale.

1.2 Tomographie sismique

La tomographie sismique est la méthode la plus employée pour obtenir de l'information sur la structure interne et la dynamique de la Terre. Il s'agit d'une technique qui exploite les informations contenues dans les enregistrements sismiques pour contraindre des modèles sismiques 2D ou 3D de l'intérieur de la Terre. Cette procédure consiste à trouver la solution d'un problème inverse (voir section 3.1.1) pour obtenir un modèle sismique hétérogène compatible avec les observables. De façon plus formelle, une relation $\mathbf{d} = A(\mathbf{p})$ entre les données sismiques $\mathbf{d}$ et la structure sismique $\mathbf{p}$ est employée, ce qui signifie que pour un modèle donné $\mathbf{p}$, les données $\mathbf{d}$ peuvent être prédites par l'intermédiaire des lois physiques A . Ainsi, la tomographie sismique consiste à trouver $\mathbf{p}$ tel que $\mathbf{d}$ explique les données observées $\mathbf{d}_{\text{obs}}$. L'écart entre les données testées et observées est quantifié par une fonction coût (appelée *misfit*, en anglais). Dans la plupart des cas, $\mathbf{d}$ et $\mathbf{p}$ sont des vecteurs de grande dimension.

Les images sismiques révèlent des structures à des résolutions radiales et latérales qui sont en constante progression. De nombreuses approches pour résoudre le problème inverse, c'est-à-dire la description de la structure interne en utilisant des données sismologiques, ont été implémentées (e. g. [Montagner & Tanimoto, 1990](#); [van der Hilst et al.,](#)

1997; Fukao et al., 2001; Romanowicz, 2003). Bien que les modèles de Terre soient de plus en plus nombreux, une comparaison quantitative est difficile à cause du manque d'information sur les incertitudes associées.

1.2.1 Des modèles 3D complexes

Les modèles sismiques 3D récents sont obtenus à partir de théories d'inversion sophistiquées basées sur le calcul des noyaux de sensibilité (*e. g.* Fichtner et al., 2009). Ils sont les produits de méthodes numériques exactes pour la propagation des ondes sismiques. Pour une synthèse sur ce type d'inversion novatrice, appelée en anglais *adjoint tomography*, le lecteur pourra se référer par exemple à Liu & Gu (2012). Ces méthodes permettent d'augmenter la résolution latérale des modèles 3D lissés utilisés comme modèles de départ de l'inversion (Ritsema et al., 1999; Mégnin & Romanowicz, 2000). Cependant, ils sont intrinsèquement liés à de petites perturbations au premier ordre à partir de modèles de référence homogènes latéralement comme le PREM (Dziewonski & Anderson, 1981) ou AK135 (Kennett et al., 1995). D'un point de vue géodynamique, dans des régions telles que la zone de transition, nous avons vu que le degré d'hétérogénéité induit par les transformations minéralogiques, les mouvements convectifs, les matériaux ascendants et descendants, et l'anisotropie, peut être considéré comme étant suffisamment important pour que le concept d'un modèle sismique 1D de référence soit remis en question.

Les ondes de volume, principalement sensibles aux contrastes d'impédance, permettent d'obtenir un haut degré de résolution. Cependant, les études qui les utilisent souffrent d'une couverture des données très inégale, liée à la distribution irrégulière des séismes et des stations d'acquisition. Dans la majorité des cas, les ondes arrivent presque verticalement à une station. Les continents sont donc particulièrement bien échantillonnés comparés aux deux tiers restants de la surface terrestre recouverts par les océans. Comparées aux ondes de volume, les ondes de surface présentent l'avantage de couvrir efficacement l'ensemble du globe à cause de leur mode de propagation (voir section 2.4). Cependant, elles ne peuvent pas sonder le manteau profond avec une grande résolution, et ont des difficultés à contraindre la structure de la croûte. De manière à obtenir une meilleure résolution en profondeur dans la zone de transition, il est indispensable de prendre en compte les informations véhiculées par les modes harmoniques (voir section 2.4.3).

En général, il existe un bon accord entre les différents modèles tomographiques globaux au niveau de la lithosphère (figure 1.9). En revanche, c'est dans la zone de transition que les incertitudes sont les plus importantes. Ces incertitudes montrent clairement que la zone de transition reste peu contrainte par les modèles tomographiques, et qu'il s'agit d'un manque critique pour comprendre les mouvements de matière dans le manteau et les transformations minéralogiques.

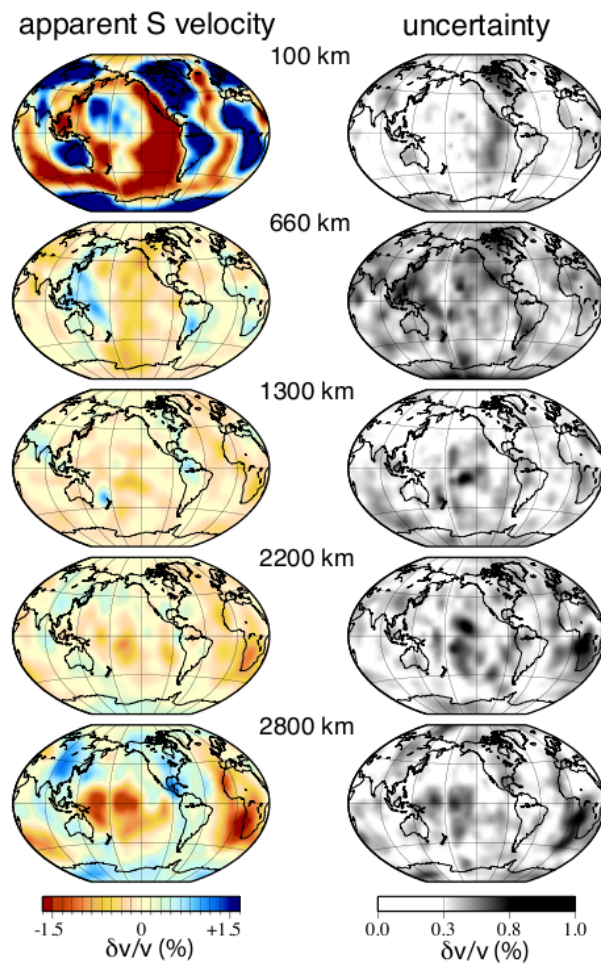


FIG. 1.9 : Moyenne des perturbations de vitesses de cisaillement et incertitudes associées à différentes profondeurs. Cette figure a été réalisée à partir de 18 modèles tomographiques globaux publiés (d'après [Liu & Gu, 2012](#)).

La principale difficulté pour déterminer l'anisotropie en tomographie sismique est que le problème inverse est sous-déterminé. Par conséquent, les études qui incluent l'anisotropie utilisent un nombre limité de modules élastiques, au lieu d'utiliser les 21 constantes élastiques indépendantes requises pour décrire les anomalies anisotropes. De plus, bien que de nombreuses études (*e.g.* [Boschi & Dziewonski, 2000](#); [Trampert & Woodhouse, 2003](#); [Panning & Romanowicz, 2006](#); [Beghein et al., 2008](#); [Visser et al., 2008a,b](#)) indiquent des variations d'anisotropie radiale à travers le manteau (elles n'utilisent que 5 paramètres élastiques indépendants), elles ne s'accordent qu'aux grandes longueurs d'ondes ([Kustowski et al., 2008](#); [Panning et al., 2010](#)), même lorsque les modes harmoniques des ondes de surface sont utilisés pour sonder les profondeurs de la zone de transition. Ces résultats suggèrent que la présence ou l'absence d'anisotropie radiale dans la zone de transition ne fait pas l'objet d'un consensus.

Ces incertitudes importantes peuvent être expliquées par les différentes méthodes inverses et de régularisation employées, le choix de la paramétrisation ainsi que les jeux de données utilisés. Idéalement, on voudrait extraire de l'information à partir des données sismiques sans imposer de limitations sur la nature des variations spatiales. En réalité, la Terre contient des variations continues et discontinues des vitesses sismiques, et montre des structures hétérogènes à toutes les échelles. Ainsi, le choix de la paramétrisation restreint immédiatement la gamme de modèles possibles, ce qui peut être vu comme une forme de régularisation. Dans la plupart des cas, les modèles de Terre sont paramétrés avec des fonctions de base et des cellules 2D ou 3D dont la taille est par avance fixée. Le choix de la taille de la cellule est un compromis entre la résolution du modèle et son incertitude (Nolet, 2008). L'information obtenue en tomographie sismique dépend fortement de la localisation des sources (aux limites de plaques) et de la position des récepteurs, qui ne sont pas distribués de façon uniforme à la surface du globe. Cela conduit à des régions qui présentent une bonne couverture de rais sismiques, et d'autres non. Pour résoudre ce problème, des techniques de lissage ainsi que l'agrandissement des cellules sont parfois nécessaires. Lorsque les cellules deviennent très petites, les incertitudes sur les paramètres du modèle deviennent importantes et la solution devient rapidement non unique. C'est pourquoi des contraintes de lissage et de régularisation sont souvent imposées pour obtenir un modèle unique, ce qui affecte l'estimation des incertitudes.

1.2.2 Le problème des couches superficielles

Une des difficultés pour estimer la structure sismique de la zone de transition et l'anisotropie radiale est liée à la croûte. En plus du manteau, les ondes de surface sont très sensibles à la structure crustale. Les effets de la croûte doivent donc être pris en compte, soit en incluant directement des paramètres pour décrire la croûte dans l'inversion, soit en réalisant des corrections en se basant sur un modèle crustal *a priori*. Cette dernière approche est souvent suivie en tomographie globale des ondes de surface et il a été montré que des corrections crustales pouvaient conduire à des erreurs importantes dans les modèles (Marone & Romanowicz, 2007; Ferreira et al., 2010). Notamment, les modèles d'anisotropie radiale sont particulièrement sensibles aux corrections crustales (Bozdog & Trampert, 2007).

Plutôt que d'employer un seul modèle de croûte pour réaliser les corrections crustales, Ferreira et al. (2010) ont utilisé différents modèles de croûte pour quantifier l'impact de la croûte lors de l'inversion de courbes de dispersion des ondes de Love et de Rayleigh du mode fondamental et jusqu'au sixième harmonique. Trois modèles ont été testés : les modèles CRUST2.0 (Bassin et al., 2000), 3SMAC (Nataf & Ricard, 1996) et le modèle de Meier et al. (2007), désigné par le nom CRUST07. CRUST2.0 est basé sur une importante compilation de données de sismique réfraction et d'épaisseurs de glaces et de

sédiments. Le modèle 3SMAC (Nataf & Ricard, 1996) est un modèle *a priori* basé sur des données de tectonique, du flux de chaleur et sur des connaissances géophysiques. CRUST07 consiste en une moyenne des vitesses de cisaillement et de la profondeur du Moho obtenue en utilisant une méthode probabiliste. La figure 1.10 présente les résultats de l'inversion pour l'anisotropie radiale à 100 et 600 km de profondeur. Bien que les caractéristiques à grande longueur d'onde soient en général compatibles d'un modèle à l'autre, il existe d'importantes différences entre les modèles. À 100 km de profondeur, les modèles sont particulièrement différents sous le Pacifique et l'Eurasie. Ces différences persistent jusqu'à 600 km de profondeur sous l'Eurasie. Ferreira et al. (2010) montrent que l'utilisation de corrections crustales à partir de différents modèles de croûte *a priori* peuvent conduire à des valeurs d'anisotropie radiale différentes jusque dans la zone de transition.

De récentes études sur le manteau supérieur ont essayé de quantifier les relations entre la structure crustale, qui est considérée comme étant une inconnue, et l'anisotropie radiale recherchée. Par exemple, Shapiro & Ritzwoller (2002) ont réalisé des inversions de Monte Carlo qui fournissent des estimations d'incertitudes sur les vitesses sismiques dans la croûte et le manteau supérieur. Cependant, ces auteurs imposent de fortes contraintes sur l'amplitude des variations par rapport au modèle *a priori*.

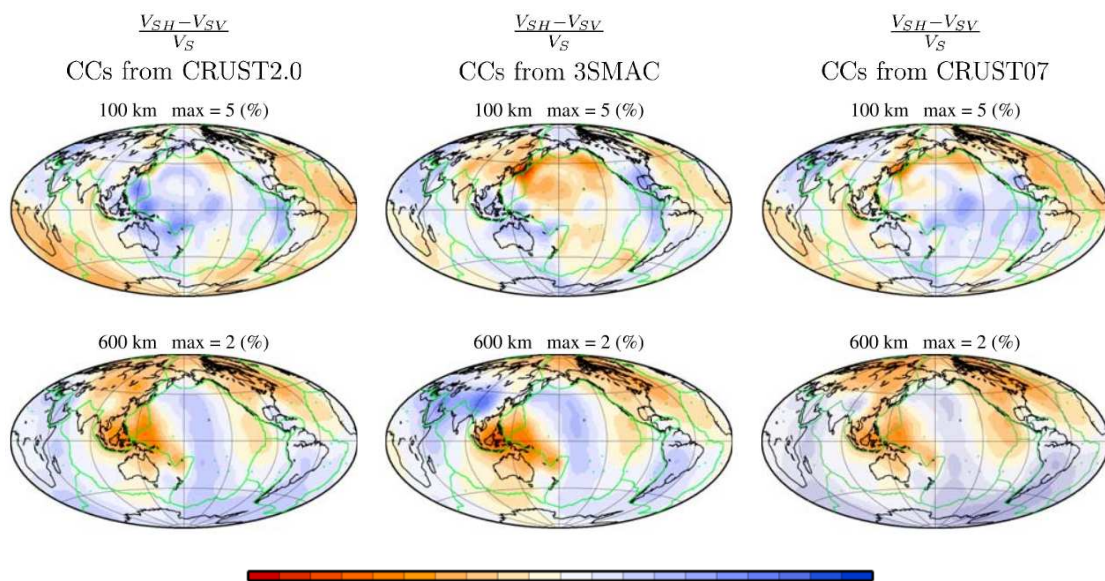


FIG. 1.10 : Variations d'anisotropie radiale par rapport au PREM obtenues pour des modèles de croûte différents à 100 et 600 km de profondeur (d'après Ferreira et al., 2010). L'anisotropie moyenne à chaque profondeur a été retirée. L'échelle de couleur est différente selon la profondeur considérée (un maximum de 5 et 2 %, respectivement, pour 100 et 600 km de profondeur).

1.2.3 La méthode des gradients

Les méthodes d'inversion qui utilisent la méthode des gradients pour obtenir une solution sont les plus employées en tomographie sismique. Partant d'un modèle initial, le calcul des dérivées partielles ($\partial \mathbf{A} / \partial \mathbf{p}$) permet de déterminer la direction de la plus grande pente de la fonction coût autour de ce point. $\partial \mathbf{A} / \partial \mathbf{p}$ décrit l'amplitude du changement des observables comparé au changement des paramètres du modèle. La principale hypothèse utilisée est de considérer qu'au voisinage des modèles expliquant les données, la forte non-linéarité du problème est suffisamment réduite pour pouvoir entrer dans le cadre d'une optimisation au sens des moindres carrés (Tarantola & Valette, 1982). La mesure de l'adéquation du modèle par rapport aux données est caractérisée par une fonction coût, qui contient un terme décrivant le résidu sur les données et un ou plusieurs termes de régularisation. Par exemple, une telle fonction peut s'écrire (Rawlinson et al., 2006) :

$$S(\mathbf{p}) = (\mathbf{A}(\mathbf{p}) - \mathbf{d}_{\text{obs}})^T \mathbf{C}_d^{-1} (\mathbf{A}(\mathbf{p}) - \mathbf{d}_{\text{obs}}) + \varepsilon (\mathbf{p} - \mathbf{p}_0)^T \mathbf{C}_p^{-1} (\mathbf{p} - \mathbf{p}_0) + \eta \mathbf{p}^T \mathbf{L}^T \mathbf{L} \mathbf{p}, \quad (1.2)$$

où $\mathbf{d}_{\text{obs}}$ sont les données réelles, $\mathbf{p}_0$ le modèle de référence, $\mathbf{A}(\mathbf{p})$ les données prédites, $\mathbf{C}_d$ et $\mathbf{C}_p$ sont les opérateurs de covariance⁵ sur les données et les paramètres, et $\mathbf{L}$ est un opérateur de lissage. ε est un facteur d'amortissement (le *damping*, en anglais), qui est un facteur de normalisation *ad hoc* appliqué sur les paramètres afin de contraindre l'amplitude de leurs variations, pour qu'ils restent relativement proches du modèle de départ. η est un facteur qui contrôle le degré de lissage du modèle (le *smoothing*, en anglais).

Il s'agit donc de déterminer un modèle qui permet au mieux d'expliquer les données sans pour autant s'écarter trop du point de départ de l'algorithme. Si la forme de la fonction coût est complexe, choisir un modèle de référence comme modèle de départ peut ainsi bloquer les recherches dans un minimum local et non dans le minimum global. Cela conduit directement à la constatation que la solution d'un tel problème est extrêmement dépendante du modèle de départ. La figure 1.11 montre les effets des facteurs ε et η sur le résultat de l'inversion. Quand un amortissement et un lissage faibles sont utilisés, la structure retrouvée contient de nombreux artefacts de petites longueurs d'onde. L'amortissement tend à diminuer l'amplitude des perturbations sans filtrer l'image, alors que le lissage joue essentiellement le rôle d'un filtre passe-bas. Selon les valeurs employées pour ces deux facteurs, le modèle trouvé est différent. Bien qu'il existe des techniques pour choisir des facteurs d'amortissement et de lissage optimaux (e. g. Lukas, 2008), une des conséquences de l'utilisation d'une telle méthode est que les informations *a priori* imposées sur le modèle prennent énormément de poids dans le processus d'inversion.

⁵Il s'agit de matrices qui gèrent les relations entre les différentes composantes des vecteurs $\mathbf{d}$ et $\mathbf{p}$. Les opérateurs de covariance sont introduits pour éviter une trop grande sous-détermination du problème et reflètent les différentes informations disponibles entre les composantes de ces deux vecteurs.

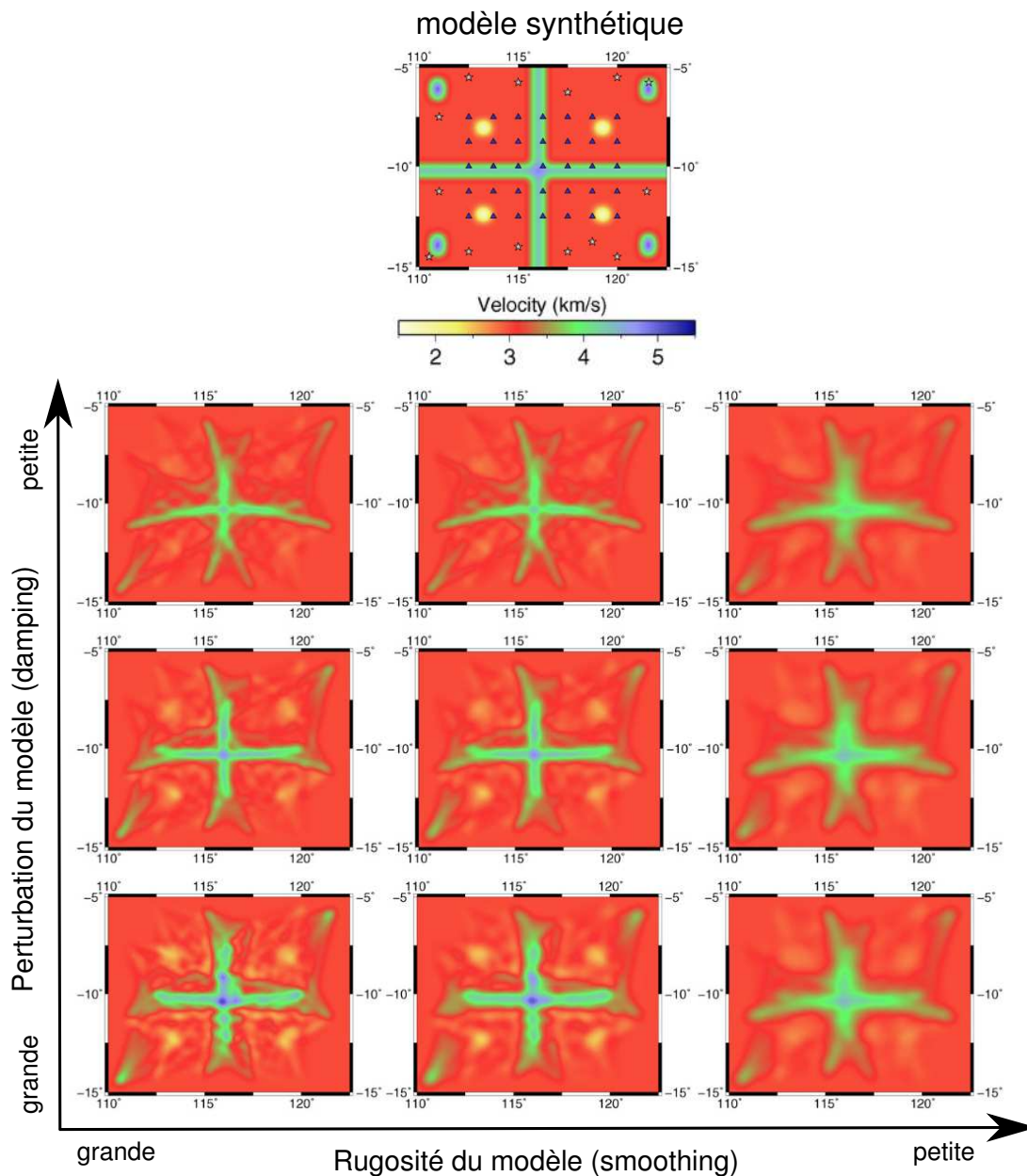


FIG. 1.11 : Illustration des effets du *smoothing* et du *damping* sur le résultat de l'inversion d'un modèle synthétique (d'après [Rawlinson et al., 2010](#), modifié). Sur le modèle synthétique, les étoiles grises correspondent aux sources sismiques et les triangles bleus aux récepteurs.

1.2.4 L'approche probabiliste en sismologie

La méthode d'inversion décrite précédemment est considérée comme étant « locale », car elle exploite l'information uniquement dans des régions limitées de l'espace des paramètres pour arriver à une solution. La limitation de ce type de méthode est qu'elle est fortement dépendante du modèle de départ, et qu'elle ne fournit pas une mesure robuste de l'incertitude des modèles.

La procédure d'inversion en tomographie sismique requiert l'ajustement des paramètres du modèle p pour satisfaire les observables d_{obs} . Ces paramètres sont sujets à différentes contraintes. Une des principales difficultés dans cette procédure est la non-unicité de la solution. La plupart des études en sismologie interprètent un modèle unique. De nombreuses techniques existent pour analyser la pertinence de la solution. En particulier, les tests « en damier »⁶ (*checkerboard tests*, en anglais) sont très répandus. Se focaliser sur retrouver un seul jeu de paramètres à travers un processus de minimisation est une approche *déterministe*. Mais comment cette approche peut-elle être justifiée puisque de nombreux modèles peuvent satisfaire les données ? L'appel à la régularisation a pour objectif de réduire considérablement l'espace des solutions, ce qui permet de sélectionner un modèle capable d'expliquer les données de façon plus directe. L'inconvénient est que des contraintes *ad hoc* sont souvent imposées. En plus de la non-unicité de la solution, la non-linéarité du problème inverse affecte également de nombreuses applications en tomographie sismique.

La plupart des techniques d'inversion sont basées sur des algorithmes qui minimisent une fonction coût entre les données réelles et testées. Les inversions linéarisées sont basées sur le calcul du gradient de cette fonction, par exemple en utilisant la méthode des gradients conjugués ou la méthode de la pente la plus forte (*steepest descent*, en anglais). Or si la non-linéarité entre les paramètres et les données est importante, la fonction coût peut devenir extrêmement complexe. Une question intéressante se pose alors : en quoi consiste un modèle qui satisfasse les données ? Une alternative est d'envisager une solution probabiliste basée sur la sélection d'un ensemble de modèles acceptés selon un critère de tolérance. Le principe de base est de tester aléatoirement un très grand nombre de modèles et de tester leur adéquation avec les données. Il s'agit des méthodes de Monte Carlo (Metropolis & Ulam, 1949), dans lesquelles des approches non-linéaires permettent de décrire l'espace des modèles de manière statistique. Ces méthodes sont qualifiées de *bayésiennes*, car elles sont basées sur un théorème central, appelé *théorème de Bayes* (voir section 3.2.2). Dans le cadre bayésien, l'information connue *a priori* sur les paramètres est combinée avec les données observées pour générer la distribution *a posteriori* des paramètres du modèle. Plus particulièrement, les méthodes de Monte Carlo par chaînes de Markov (désignées par l'abréviation MCMC pour *Markov chain Monte Carlo*, en anglais) réalisent une recherche guidée en échantillonnant l'espace des paramètres selon les probabilités *a posteriori*. Un avantage important de cette méthode est la complète indépendance du choix du modèle initial.

⁶Il s'agit d'un test de résolution qui consiste à retrouver un modèle synthétique constitué d'une alternance de régions lentes et rapides. Le résultat de l'inversion de ce modèle peut donner des informations sur les régions les moins bien résolues (par exemple les endroits où il y a peu de stations sismiques).

La figure 1.12a montre un exemple de fonction coût dans le cas d'un problème non-linéaire. Pour un tel problème, il existe de nombreux minima locaux, en plus du minimum global. Le minimum global de cette fonction correspond au meilleur modèle, et le localiser serait très compliqué avec des techniques basées sur des calculs de gradients, à moins de débiter l'inversion avec un modèle proche de la solution, c'est-à-dire dans la « vallée » de la solution globale. Il existe un réel risque de rester confiné dans un minimum local. En fait, un modèle initial qui est relativement proche de la solution est requis pour

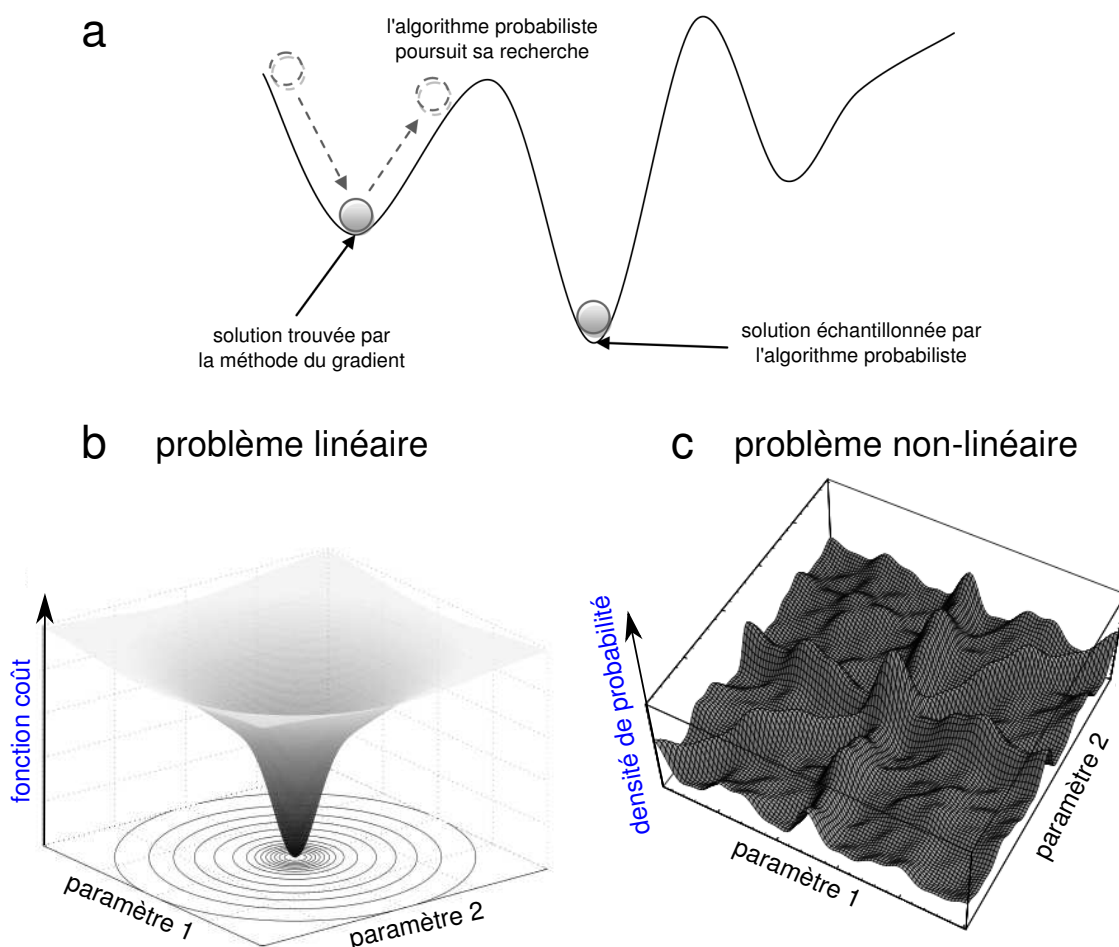


FIG. 1.12 : (a) Illustration de la forme d'une fonction coût. Dans le cas d'une méthode reposant sur le calcul du gradient, il existe un risque de rester confiné dans un minimum local. Un algorithme probabiliste de type Monte Carlo, par échantillonnage de nombreux modèles, permet d'accéder aux paramètres correspondant au minimum global de cette fonction. (b) Représentation de la fonction coût dans le cas d'un problème linéaire (d'après Sambridge & Mosegaard, 2002, modifié). (c) Représentation de la fonction de densité de probabilité dans le cas d'un problème non-linéaire (d'après Deng & Scales, 1999, modifié). Dans les deux cas, le modèle est constitué de deux paramètres.

justifier l'hypothèse de linéarité locale. Des exemples de solutions pour des problèmes linéaire et non-linéaire sont présentés sur la figure 1.12b et 1.12c. La figure 1.12b montre le cas d'un problème linéaire. Pour deux paramètres, il n'existe qu'une seule combinaison de valeurs possibles qui donnent un minimum global de la fonction coût. La figure 1.12c représente la fonction de densité de probabilité pour un problème non-linéaire, déterminée grâce à un algorithme probabiliste. En 3D, cette fonction donne l'impression qu'il s'agit d'une carte topographique, où l'on peut voir des « montagnes » et des « vallées ». Elle représente les régions de l'espace des paramètres échantillonnées un grand nombre de fois (les montagnes), et les régions peu échantillonnées (les vallées). Ce type de méthode est particulièrement intéressant car il donne une vision complète de la solution dans l'espace des paramètres.

Plutôt que d'échantillonner uniformément l'espace des paramètres, ce qui est très coûteux en temps de calcul, des méthodes d'inversion non-linéaires ont été adaptées aux problèmes en sismologie, telles que :

- *le recuit simulé* (e. g. [Weber, 2000](#)). Le principe de cette méthode est basé sur un processus très utilisé en métallurgie pour obtenir un alliage sans défaut appelé « le recuit ». Le but est d'atteindre un état d'équilibre thermodynamique. Cet état d'équilibre, où l'énergie est minimale, représente la solution optimale du problème. En pratique, un paramètre fictif de température est considéré. La température restera élevée pendant les premières itérations, pour pouvoir s'extraire des minima locaux de la fonction coût, puis elle sera à décroissance lente jusqu'à tendre vers zéro, où l'optimum global est atteint.
- *l'algorithme génétique* (e. g. [Bodin et al., 2012](#)). Ce type d'algorithme s'inspire de l'évolution biologique d'une espèce pour rechercher de nouveaux modèles à partir d'une « population d'individus » initiale générée aléatoirement. Cette population d'individus correspond à un modèle, par exemple un profil de vitesses sismiques. Ensuite l'algorithme va passer d'une génération (d'un modèle) à l'autre en appliquant des mécanismes d'évaluation, de sélection et de modification.
- *l'algorithme du réseau de neurones* (e. g. [Meier et al., 2007](#)). Son fonctionnement est très schématiquement inspiré du fonctionnement des neurones biologiques. Il met en œuvre un principe d'apprentissage par l'expérience (principe de l'induction). Par confrontation avec des situations particulières, ils en déduisent un système de décision qui dépend du nombre de cas d'apprentissages rencontrés et de leur complexité par rapport à la difficulté du problème à résoudre.

Ces algorithmes conviennent parfaitement pour les problèmes inverses non-linéaires. De plus, la description statistique de l'ensemble des modèles offre une vision plus objective de la résolution des modèles que les tests standards de résolution d'une inversion linéaire ([Sambridge & Mosegaard, 2002](#); [Fichtner & Trampert, 2011](#)).

L'approche bayésienne permet d'aller au-delà de la procédure classique qui consiste à retrouver un unique modèle donnant la meilleure adéquation possible avec les données, en fournissant une mesure quantitative probabiliste de la résolution du modèle, des incertitudes et de la non-unicité (*e. g.* [Mosegaard & Tarantola, 1995](#)). L'idée dominante d'une telle méthode est de combiner les informations contenues dans les données avec les informations *a priori*. Bien que ces méthodes, telles que la méthode MCMC soient très populaires en géophysique, leurs utilisations pour des études sismologiques et les interprétations des modèles en termes de température et de composition sont récentes. Le fait que les méthodes probabilistes soient peu employées en sismologie est principalement lié au grand nombre de paramètres à inverser et au temps de calcul important qu'elles requièrent (*e. g.* [Verhoeven et al., 2009](#); [Khan et al., 2009, 2011](#); [Hauser et al., 2011](#); [Bodin et al., 2012](#); [Mosca et al., 2012](#)).

Même en parallélisant la procédure d'inversion, des techniques de recherches entièrement non-linéaires sont limitées, du moins dans le contexte de la tomographie sismique, à une centaine d'inconnues au plus. Par exemple, [Khan et al. \(2011\)](#) ont réalisé une tomographie du manteau jusqu'à 1300 km de profondeur, selon une procédure entièrement non-linéaire avec 74 paramètres par région inversée. Pour arriver à un tel résultat, la Terre a été divisée en seulement 27 régions, et l'espace possible des paramètres a été restreint. Une alternative pratique serait d'utiliser une méthode hybride. Par exemple, [Rawlinson et al. \(2008\)](#) ont développé une méthode alternative basée sur la méthode des gradients, qui exploite l'information obtenue des solutions précédentes pour guider la recherche de nouveaux modèles. Cette technique est possible en ajoutant un terme d'évolution à la fonction coût qui crée un maximum local à chaque point dans l'espace des modèles occupé par toutes les solutions précédentes. Les nouveaux modèles évitent ainsi le voisinage des modèles générés précédemment, et un ensemble de solutions distinctes est produit.

La philosophie globale adoptée pour ce travail de thèse est de laisser un maximum de liberté aux paramètres des modèles, tout en préservant la réalité physique des profils de température et d'anisotropie radiale. C'est pourquoi nous avons choisi de réaliser une inversion probabiliste de type Monte Carlo en employant les propriétés des chaînes de Markov. L'objectif est d'évaluer si une seule ou plusieurs familles de modèles peuvent rendre compte de la complexité de la zone de transition.

1.2.5 Interprétation des modèles sismologiques en termes de température et de composition

Bien que la résolution des modèles sismiques soit en constante amélioration, les modèles sismiques *s. s.* ne donnent pas d'information sur le comportement dynamique du

manteau. Les modèles sismiques sont déjà des interprétations des données (temps de trajet, vitesses de phase, etc.). Ils ne sont pas uniques, et dépendent de la paramétrisation, du type, de la distribution, et de la qualité des données. De plus, ce n'est pas forcément parce qu'un modèle fournit une très bonne adéquation aux données qu'il correspond nécessairement à un modèle physique.

Un aspect fondamental de la recherche en géophysique est de pouvoir interpréter les informations sismologiques en termes de température et de composition du manteau. En effet, les modèles tomographiques fournissent seulement des images de l'intérieur de la Terre et non des informations fondamentales pour comprendre la physique du manteau, c'est-à-dire la température et la composition. Mais il n'existe pas de combinaison unique de ces variables à partir d'un modèle sismique particulier, puisque différents modèles de température et de composition minéralogique peuvent aboutir aux mêmes vitesses sismiques.

Deux approches sont généralement suivies. La première méthode consiste à extraire la température et la composition à partir de l'interprétation de modèles sismiques préexistants (Vacher et al., 1996, 1998; Goes et al., 2000; Trampert et al., 2004; Cammarano & Romanowicz, 2007). Dans un premier temps, un modèle sismique est construit pour expliquer les données de façon satisfaisante, puis le modèle est interprété en termes de température et de composition en se basant sur les propriétés élastiques et anélastiques des minéraux du manteau. Par exemple, l'analyse des changements relatifs des vitesses des ondes de cisaillement et de compression ($\partial \ln V_S / \partial \ln V_P$) est parfois utilisée pour déterminer les causes thermiques ou chimiques des variations sismiques (*e. g.* Deschamps & Trampert, 2003; Simons et al., 2009). La contribution relative de chaque paramètre (température ou composition) peut être écrite comme le produit d'une dérivée partielle des vitesses sismiques par rapport à ce paramètre. $\partial \ln V_S$ et $\partial \ln V_P$ sont considérés comme étant les données à inverser. Deschamps & Trampert (2003) ont montré pour le manteau inférieur que cette technique ne donne seulement qu'une information qualitative des variations chimiques.

La deuxième approche consiste à déterminer directement la température et la composition à partir des temps de trajet observés, des vitesses de phase des ondes de surface ou des fréquences propres (Cammarano et al., 2009; Khan et al., 2009, 2011; Ritsema et al., 2009). L'idée générale est d'utiliser des informations issues de la physique des minéraux (les modules élastiques des différents minéraux et leurs dérivées par rapport à la pression et à la température) en amont du calcul des vitesses sismiques, pour imposer une signification physique au modèle sismologique.

L'un des principaux avantages de cette deuxième approche est qu'elle connecte directement les incertitudes *a posteriori* sur les paramètres géodynamiques aux incertitudes sur les observations. Cependant, la nature non-linéaire du problème nécessite d'utiliser

des méthodes de type Monte Carlo, qui sont coûteuses en temps de calcul. Cette difficulté a déjà été mise en évidence en testant un nombre limité de modèles sélectionnés (Cammarrano et al., 2009; Ritsema et al., 2009), ou bien en explorant des espaces des paramètres relativement restreints (Khan et al., 2009, 2011).

1.3 Objet de cette étude et démarche adoptée

Cette première partie avait un double objectif. D'une part, les différentes questions qui se posent quant à l'origine des variations de vitesses dans la zone de transition ont été succinctement présentées. D'autre part, un point a été fait sur les méthodes de tomographie sismique qui permettent d'obtenir de l'information sur cette région.

Nous avons vu que la zone de transition est hétérogène à toutes les échelles à cause de la géodynamique du manteau. Bien que la principale source d'hétérogénéité soit d'origine thermique, la composition, l'anisotropie et le contenu en eau jouent également un rôle important. L'interprétation des modèles sismologiques en termes de température et de composition n'est donc pas unique. Déconvoluer leurs différents effets à partir des données sismologiques seules constitue un véritable défi scientifique. Les modèles tomographiques 3D sont basés sur des méthodes de plus en plus sophistiquées, mais ils montrent des résultats proches du modèle de départ de l'inversion et la non-unicité n'est pas prise en compte. Les modèles tomographiques globaux présentent d'importantes différences dans la zone de transition, preuve que cette région demeure mal connue. De plus, ce sont en général ces mêmes modèles tomographiques globaux qui sont utilisés pour dériver la température et la composition dans un deuxième temps.

Ces dernières années, de plus en plus d'études ont soulevé deux problèmes majeurs :

- Les modèles sismologiques ne suffisent pas pour décrire la géodynamique de la zone de transition et une interprétation en termes de paramètres physiques (température, composition) est nécessaire. Si on analyse la situation à l'heure actuelle, on constate que les résultats théoriques et expérimentaux concernant la physique des minéraux sont cruciaux pour donner un sens physique aux modèles sismologiques.
- Les problèmes directs employés sont en général fortement non-linéaires, et les méthodes de tomographie classiques ne permettent pas d'obtenir une vision d'ensemble de la solution dans l'espace des paramètres. C'est pourquoi des approches non-linéaires sont développées pour prendre en compte la non-unicité de la solution.

Ce sont ces deux problématiques qui sont au cœur de ce mémoire de thèse, dont l'objectif est d'apporter des éléments de réponse aux questionnements suivants :

- Quelle est la distribution des profils radiaux de température et d'anisotropie pour expliquer les courbes de dispersion des ondes de surface ?
- Quelle est la contribution de la composante anisotrope dans la zone de transition ?

Pour répondre à ces questions, une étude pluridisciplinaire a été nécessaire. J'ai développé une approche non-linéaire pour interpréter directement les courbes de dispersion des ondes de Love et de Rayleigh en termes de distributions 1D de température et d'anisotropie radiale, pour une composition fixée. La philosophie globale est de mettre peu d'information *a priori* sur les modèles afin d'échantillonner de grandes gammes de valeurs possibles des paramètres et d'éviter de réaliser trop de choix arbitraires qui pourraient influencer la forme de la solution. À partir de modèles de température et d'anisotropie choisis de façon aléatoire, les vitesses sismiques sont calculées en utilisant une équation d'état. L'avantage est que les vitesses des ondes P et S , et la masse volumique, sont naturellement liées. De plus, la taille et la localisation des discontinuités des propriétés physiques en fonction de la pression induisent des changements de phase minéralogiques, qui sont modélisés d'une manière physique réaliste, puisque leurs variations dépendent de la composition et des conditions physiques du modèle particulier considéré. Les courbes de dispersion synthétiques sont ensuite obtenues par sommations de modes propres et comparées aux données à travers une méthode de MCMC. Afin de générer une paramétrisation adaptative, selon la capacité de résolution des données, ainsi que de réduire le temps de calcul, des courbes de Bézier polynomiales sont choisies pour la paramétrisation.

Chapitre 2

Problème direct

2.1 Introduction

La prédiction ou le calcul des données selon des valeurs fixées des paramètres est appelé *problème direct*. Celui-ci incarne l'interprétation physique du phénomène observé, c'est-à-dire ici les vitesses de phase des ondes de surface, qui sont considérées comme étant les données. Dans ce chapitre sont détaillées la procédure et les hypothèses utilisées pour le calcul des vitesses de phase à partir d'un profil de température et de l'anisotropie radiale, qui sont les paramètres des modèles. La composition minéralogique est fixée. Pour estimer les vitesses de phase des ondes de surface en fonction de la fréquence (ou de la période), le problème direct $\mathbf{d} = A(\mathbf{p})$ doit être résolu, où $\mathbf{p}$ désigne les paramètres du modèle, $\mathbf{d}$ les données, et où A comprend l'ensemble des lois physiques qui connectent les paramètres du modèle aux données.

Les différentes étapes du problème direct sont synthétisées sur la figure 2.1. Le problème direct à proprement parler, représenté de façon schématique à l'intérieur du

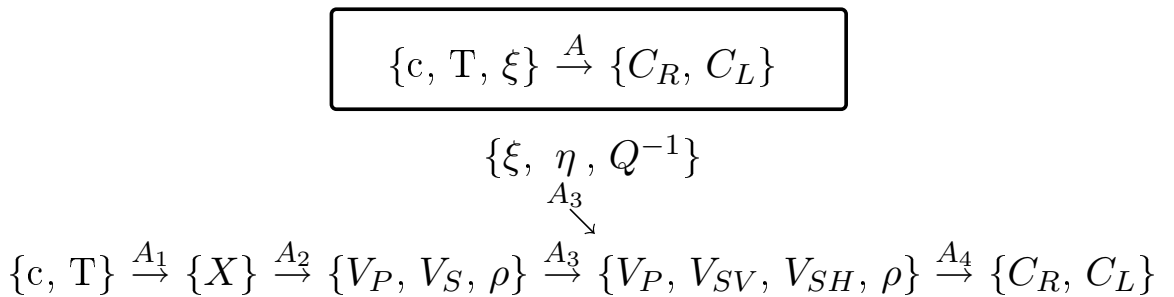


FIG. 2.1 : Illustration schématique du problème direct et des théories physiques employées ($A_1, A_2, \dots$).

cadre sur la figure 2.1, consiste à déterminer les vitesses de phase des ondes de Love et de Rayleigh (C_L et C_R) à partir de la composition minéralogique c , de la température T et de l'anisotropie radiale ξ . Pour résoudre ce problème, des étapes intermédiaires sont nécessaires et correspondent aux symboles A_i sur la figure 2.1 :

- A_1 comprend l'utilisation des diagrammes de phases pour calculer les fractions volumiques X des phases minérales correspondant aux conditions locales de pression et de température dans la zone de transition.
- L'étape A_2 permet de déterminer, grâce aux lois thermodynamiques, les modules élastiques des minéraux individuels à ces conditions de pression et de température pour en déduire ensuite les propriétés sismiques de la roche totale V_P , V_S et ρ , qui correspondent aux vitesses des ondes P et S et à la masse volumique. V_P , V_S et ρ sont considérés comme étant des paramètres secondaires, puisqu'ils dépendent directement des paramètres primaires du modèle que sont la température et la composition.
- Puis, les vitesses des ondes V_{SV} et V_{SH} , qui sont les vitesses des ondes S polarisées verticalement et horizontalement, sont déterminées par A_3 à partir de la vitesse des ondes S isotropes et du profil d'anisotropie radiale ξ . Cette étape requière l'introduction du paramètre anisotrope η (défini dans la section 2.3).
- Dans la dernière étape A_4 , les courbes de dispersion synthétiques des ondes de Love et de Rayleigh sont finalement calculées en prenant en compte un modèle d'atténuation Q^{-1} fixé.

Les sections suivantes définissent brièvement les différentes lois physiques employées dans le problème direct.

2.2 Paramètres sismiques

Une procédure classique pour calculer les valeurs de la masse volumique et des vitesses sismiques en profondeur est d'extrapoler leurs valeurs à pression et température ambiantes (conditions STP) en utilisant des lois thermodynamiques. Une présentation de ces procédures est donnée dans, par exemple, Jackson (1998) et Stixrude & Lithgow-Bertelloni (2010). Les valeurs des modules élastiques aux conditions standards (STP, $P = 10^5$ Pa, $T = 298$ K) et leurs dérivées par rapport à la température et à la pression sont obtenues grâce aux expériences de laboratoire. Il est possible d'extrapoler les valeurs des modules élastiques aux conditions de pression et de température de la zone de transition grâce aux équations d'état et aux valeurs de laboratoire aux conditions STP. En se donnant une composition chimique ou minéralogique du manteau, les valeurs de V_P , V_S et ρ pour l'assemblage minéralogique sont ensuite calculées à partir des valeurs des modules élastiques déterminées pour les conditions du manteau. L'interprétation des modèles sismologiques en termes de température et de composition est sensible à la fois aux détails de la méthodologie employée dans le problème direct, mais aussi aux incertitudes sur les paramètres thermoélastiques.

2.2.1 Équation d'état

Les valeurs de la masse volumique et des vitesses sismiques en fonction de la profondeur sont calculées en employant la méthode décrite dans [Vacher et al. \(1998\)](#) et [Verhoeven et al. \(2005\)](#), en utilisant les valeurs issues d'expériences de laboratoire compilées par [Cammarano et al. \(2003\)](#) et [Verhoeven et al. \(2005\)](#) (voir tableaux 2.1 et 2.2). Cette méthode repose sur l'équation d'état de Birch-Murnaghan au troisième ordre avec une correction de Grüneisen pour la température (*e. g.* [Poirier, 2000](#)).

L'effet de la teneur en fer sur les modules élastiques est pris en compte. Pour calculer la masse volumique ρ , le module de cisaillement G et le module d'incompressibilité K_S à une pression et à une température données pour chaque minéral (i), deux étapes sont nécessaires : un chauffage isobare du minéral est d'abord simulé, suivi par une compression adiabatique.

Minéral	ρ_0 (g/cm ³)	$\rho_{0,Fe}$ (g/cm ³)	K_{S_0} (GPa)	$K_{S_0,Fe}$ (GPa)	G_0 (GPa)	$G_{0,Fe}$ (GPa)
olivine	3.222	1.182	129 ± 1.3	0	81 ± 0.81	-31
wadsleyite	3.472	1.24	172 ± 1.8	0	112 ± 1.68	-40
ringwoodite	3.548	1.30	185 ± 1.9	35	120.4 ± 1.81	-28
Mg-pérovskite	4.107	1.07	263 ± 2.6	0	175 ± 3.50	0
Mg-wüstite	3.584	2.28	162 ± 1.6	-15	130 ± 2.60	-77
clinopyroxène BP	3.208	0.80	112 ± 2.2	-5	75 ± 1.50	10
clinopyroxène HP	3.292	0.82	112 ± 2.2	-5	75 ± 1.50	10
orthopyroxène	3.204	0.81	109 ± 2.2	-5	75 ± 1.50	10
Ca-pyroxène	3.277	0.38	105 ± 1	13	67 ± 1.34	-6
ilménite	3.810	1.1	212 ± 2.1	0	132 ± 2.64	-41
majorite	3.565	0.76	171 ± 1.8	15	92 ± 1.84	7

TAB. 2.1 : Modules élastiques des minéraux du manteau aux conditions STP et incertitudes associées. Les données sont d'après les compilations de [Cammarano et al. \(2003\)](#) et [Verhoeven et al. \(2005\)](#). ρ_0 , K_{S_0} et G_0 désignent respectivement la masse volumique, le module d'incompressibilité et le module de cisaillement. L'indice Fe indique la dépendance du module élastique par rapport à la teneur en fer. Les lettres BP et HP signifient Basse Pression et Haute Pression.

Minéral	$\left(\frac{\partial K_{S_0}}{\partial P}\right)_S$	$-\left(\frac{\partial K_{S_0}}{\partial T}\right)_S$ (GPa/K)	$\left(\frac{\partial G_0}{\partial P}\right)_S$	$-\left(\frac{\partial G_0}{\partial T}\right)_S$ (GPa/K)
olivine	4.2 ± 0.13	0.016 ± 0.0027	1.4 ± 0.098	0.014 ± 0.00238
wadsleyite	4.5 ± 0.23	0.016 ± 0.004	1.5 ± 0.150	0.014 ± 0.00350
ringwoodite	4.1 ± 0.41	0.024 ± 0.0048	1.3 ± 0.195	0.015 ± 0.00300
Mg-pérovskite	4.0 ± 0.2	0.015 ± 0.003	1.8 ± 0.270	0.029 ± 0.00725
Mg-wüstite	4.0 ± 0.4	0.019 ± 0.0038	2.3 ± 0.470	0.024 ± 0.00500
clinopyroxène BP	6.6 ± 1.4	0.012 ± 0.003	1.6 ± 0.240	0.012 ± 0.00300
clinopyroxène HP	6.6 ± 1.4	0.012 ± 0.003	1.6 ± 0.240	0.012 ± 0.00300
orthopyroxène	7.0 ± 1.4	0.012 ± 0.003	1.6 ± 0.240	0.012 ± 0.00300
Ca-pyroxène	$6.2 - 1.9X_{\text{Fe}}$ ± 0.62	0.013 ± 0.0032	1.7 ± 0.255	0.010 ± 0.00250
ilménite	5.6 ± 1.12	0.017 ± 0.0042	1.7 ± 0.255	0.017 ± 0.00425
majorite	4.4 ± 0.44	0.021 ± 0.0042	1.4 ± 0.210	0.010 ± 0.00250

TAB. 2.2 : Dérivées des modules élastiques des minéraux du manteau par rapport à la pression et à la température aux conditions STP et incertitudes associées. Les données sont d'après les compilations de [Cammarano et al. \(2003\)](#) et [Verhoeven et al. \(2005\)](#). X_{Fe} représente la fraction molaire de fer. Les lettres BP et HP signifient Basse Pression et Haute Pression.

2.2.1.1 Chemin thermodynamique des minéraux

Supposons que l'on veuille calculer les vitesses sismiques et la masse volumique pour une pyrolite ([Ringwood, 1975](#)) à 400 km de profondeur et à une température de 1760 K. Sur la figure 2.2 sont détaillés les chemins thermodynamiques de l'olivine et du grenat dans le diagramme pression-température. À partir du point A, aux conditions STP (P_0, T_0), on souhaite amener le minéral au point C aux conditions (P, T) connues, pour lesquelles on veut déterminer les valeurs des modules élastiques. Un chauffage isobare est simulé jusqu'à la température $\vartheta^{(i)}$ (point B). Puis une compression adiabatique permet de passer du point B au point C. Cette température est différente pour chacun des minéraux (i). Dans l'exemple de la figure 2.2, $\vartheta_{\text{ol}} = 1573$ K et $\vartheta_{\text{gt}} = 1605$ K. Le long du chemin AB, une température intermédiaire $\vartheta^{(i)}$ est calculée. Il s'agit de la température à pression standard au départ de laquelle une compression adiabatique mène le minéral (i) aux conditions locales de pression et de température qui règnent au point C. $\vartheta^{(i)}$ est déterminée de façon itérative. À chaque itération, les valeurs de $\vartheta^{(i)}$, $\rho^{(i)}$, $K_S^{(i)}$ et $G^{(i)}$ sont recalculées. Pour plus de simplicité de lecture, l'indice (i) ne sera pas écrit systématiquement dans la suite du chapitre.

La température ϑ peut être calculée à partir de l'intégration du paramètre de Grüneisen :

$$\gamma_{\text{th}} = \left(\frac{\partial \ln T}{\partial \ln \rho} \right)_S. \quad (2.1)$$

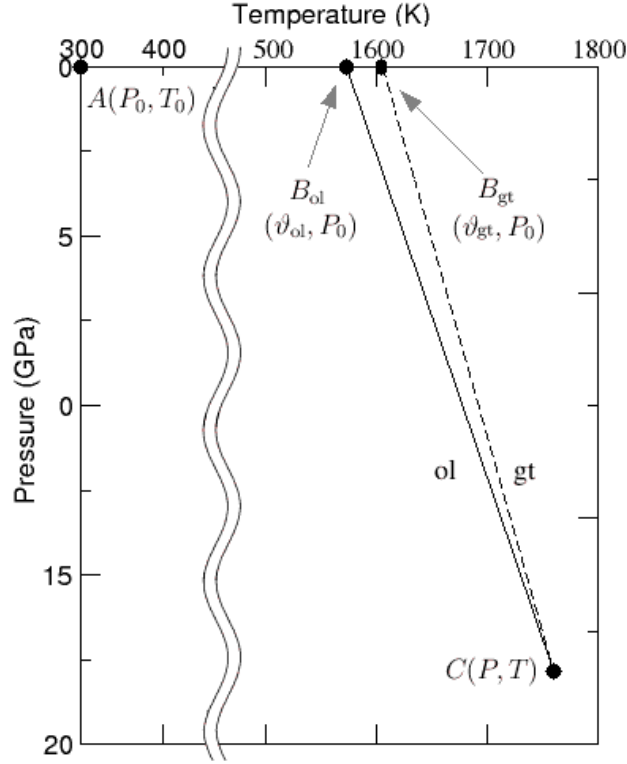


FIG. 2.2 : Exemple de chemins thermodynamiques pris par deux minéraux (olivine et grenat) jusqu'à une profondeur de 400 km et une température de 1760 K (adapté d'après Verhoeven et al., 2005). Les minéraux partent des conditions STP, qui sont matérialisées par le point A, et arrivent aux conditions locales de pression et de température (point C). Chaque minéral subit d'abord un chauffage isobare, qui l'amène du point A au point B, puis une compression adiabatique jusqu'au point C. La température ϑ au point B est différente pour chacun des minéraux.

L'indice S indique une entropie constante. La valeur du paramètre de Grüneisen dépend peu de la température, mais surtout de la pression (Anderson, 1988). D'après les résultats d'expériences de laboratoire, la relation empirique suivante est utilisée (Anderson, 1979) :

$$\gamma_{th}(P, T)\rho(P, T)^q = \text{constante}, \quad (2.2)$$

où q est une constante approximativement égale à 1. En intégrant l'équation (2.1) à entropie constante et en utilisant l'équation (2.2), la température T peut être décrite selon l'équation suivante :

$$T = \vartheta \exp \left\{ \frac{\gamma_{th0}}{q} \left[\frac{\rho_0}{\rho(P_0, \vartheta)} \right]^q \left[1 - \left(\frac{\rho(P_0, \vartheta)}{\rho(P, T)} \right)^q \right] \right\}. \quad (2.3)$$

Cette équation permet de calculer la température potentielle ϑ pour chaque minéral à l'aide d'une procédure itérative.

2.2.1.2 Chauffage isobare

Le chauffage isobare consiste à chauffer un minéral (i) à pression constante P_0 de la température T_0 à la température potentielle ϑ . Pour la plupart des minéraux, l'expansivité thermique $\alpha(P_0, T)$ peut s'écrire sous la forme d'une loi polynomiale :

$$\alpha(P_0, T) = a_0 + b_0 T - c_0 T^{-2} + d_0 T^{-1}. \quad (2.4)$$

Les valeurs des coefficients a_0 , b_0 , c_0 et d_0 , ainsi que celles du paramètre de Grüneisen, sont récapitulées pour chaque minéral dans le tableau 2.3. L'expansivité thermique est utilisée pour déterminer la masse volumique $\rho(P_0, T)$ dans l'équation (2.3) :

$$\left(\frac{\partial \rho}{\partial T} \right)_{P_0} = -\alpha(P_0, T) \rho(P_0, T). \quad (2.5)$$

En intégrant cette dernière équation à pression constante P_0 , on obtient l'expression de la masse volumique à la pression P_0 et à la température ϑ :

$$\rho(P_0, \vartheta) = \rho_0 \left(\frac{T_0}{\vartheta} \right)^{d_0} \exp \left[a_0 (T_0 - \vartheta) + \frac{b_0}{2} (T_0^2 - \vartheta^2) + c_0 \left(\frac{1}{T_0} - \frac{1}{\vartheta} \right) \right]. \quad (2.6)$$

La valeur du module d'incompressibilité adiabatique $K_S(P_0, \vartheta)$ est calculée à P_0 en utilisant la formule :

$$K_S(P_0, \vartheta) = K_{S_0} \left(\frac{\rho(P_0, \vartheta)}{\rho_0} \right)^{\delta_{S_0}}, \quad (2.7)$$

Minéral	a_0 (K ⁻¹)	b_0 (K ⁻²)	c_0 (K)	γ_{th_0}
olivine	2.832	0.758	0	1.14
wadsleyite	2.711	0.6885	0.5767	1.32
ringwoodite	1.872	0.421	0.6537	1.21
Mg-pérovskite	1.17	1.51	0	1.31
Mg-wüstite	3.0	1.2	0	1.45
clinopyroxène BP	2.86	0.72	0	1.05
clinopyroxène HP	2.86	0.72	0	1.05
orthopyroxène	2.86	0.72	0	1.05
Ca-pyroxène	2.32	1.88	0	1.06
ilménite	2.27	0.682	-0.385	1.38
majorite	2.08	1.43	0	1.17

TAB. 2.3 : Coefficients d'expansion thermique et paramètre de Grüneisen (γ_{th_0}) aux conditions STP. Les données sont d'après les compilations de [Vacher et al. \(1998\)](#) et de [Verhoeven et al. \(2005\)](#). La constante d_0 est égale à zéro excepté pour l'ilménite, pour laquelle $d_0 = -0,1808$.

où δ_{S_0} est le paramètre d'Anderson-Grüneisen, défini par

$$\delta_{S_0} = -\frac{1}{\alpha} \left(\frac{\partial \ln K_S}{\partial T} \right)_{P_0}. \quad (2.8)$$

Le paramètre d'Anderson-Grüneisen peut être considéré comme étant indépendant de la température (Anderson, 1988), c'est pourquoi il est calculé uniquement aux conditions STP pour chaque minéral. Dans le cas du module de cisaillement $G(P_0, \vartheta)$ à P_0 , on suppose une variation linéaire avec la température :

$$G(P_0, \vartheta) = G_0 + \frac{\partial G}{\partial T}(\vartheta - T_0). \quad (2.9)$$

2.2.1.3 Compression adiabatique

La seconde partie de la procédure consiste à déterminer les valeurs des modules élastiques aux conditions (P, T) . La théorie de la déformation finie est utilisée, où la déformation eulerienne ε d'un minéral pour une compression adiabatique est donnée par (Birch, 1952) :

$$\varepsilon = \frac{1}{2} \left\{ 1 - \left[\frac{\rho(P, T)}{\rho(P_0, \vartheta)} \right]^{(2/3)} \right\}, \quad (2.10)$$

et l'énergie élastique U est exprimée par une approximation polynomiale en déformation :

$$U = U_1\varepsilon + U_2\varepsilon^2 + U_3\varepsilon^3 + \dots \quad (2.11)$$

Davies & Dziewonski (1975) ont montré que l'équation (2.11) devrait inclure l'ordre 4 en déformation pour rendre compte des paramètres sismiques des modèles de référence dans le manteau inférieur. Puisque notre étude est restreinte à la zone de transition, une expansion d'ordre 3 en déformation est considérée. La pression peut être exprimée par l'équation d'état de Birch-Murnaghan au 3^{ème} ordre :

$$P(r) = -3K_S(1 - 2\varepsilon)^{5/2} \left[\varepsilon + \frac{3}{2} \left(4 - \frac{\partial K_S}{\partial P} \right) \varepsilon^2 \right]. \quad (2.12)$$

La valeur de la masse volumique $\rho(P, T)$ est obtenue en calculant l'unique racine réelle négative de cette dernière équation. Les modules élastiques à une profondeur z donnée s'obtiennent alors grâce aux expressions suivantes :

$$P(z) = -(1 - 2\varepsilon)^{5/2} \left(C_1\varepsilon + \frac{1}{2}C_2\varepsilon^2 + \frac{1}{6}C_3\varepsilon^3 \right) \quad (2.13)$$

$$K_S(z) + \frac{4}{3}G(z) = (1 - 2\varepsilon)^{5/2} \left(L_1 + L_2\varepsilon + \frac{1}{2}L_3\varepsilon^2 \right) \quad (2.14)$$

$$G(z) = (1 - 2\varepsilon)^{5/2} \left(M_1 + M_2\varepsilon + \frac{1}{2}M_3\varepsilon^2 \right). \quad (2.15)$$

Les coefficients des expressions (2.13)-(2.15) sont fonctions des modules élastiques et de leurs dérivées (*e. g.* Jackson, 1998; Trampert et al., 2001; Cammarano et al., 2003) :

$$C_1 = 3K_S \quad (2.16)$$

$$C_2 = 9K_S \left(4 - \frac{\partial K_S}{\partial P} \right) \quad (2.17)$$

$$C_3 = 27K_S \left[K_S \frac{\partial^2 K_S}{\partial^2 P} - \frac{\partial K_S}{\partial P} \left(7 - \frac{\partial K_S}{\partial P} \right) + \frac{149}{9} \right] \quad (2.18)$$

$$L_1 = K_S + \frac{4}{3}G \quad (2.19)$$

$$L_2 = 5L_1 - 3K_S \left(\frac{\partial K_S}{\partial P} + \frac{4}{3}G \right) \quad (2.20)$$

$$L_3 = 9K_S^2 \left(\frac{\partial^2 K_S}{\partial^2 P} + \frac{4}{3}G \right) - 3L_2 \left(\frac{\partial K_S}{\partial P} - 4 \right) + 5L_1 \left(3\frac{\partial K_S}{\partial P} - 5 \right) \quad (2.21)$$

$$M_1 = G \quad (2.22)$$

$$M_2 = 5M_1 - 3K_S \frac{\partial G}{\partial P} \quad (2.23)$$

$$M_3 = 9K_S^2 \frac{\partial^2 G}{\partial^2 P} - 3M_2 \left(\frac{\partial K_S}{\partial P} - 4 \right) + 5M_1 \left(3\frac{\partial K_S}{\partial P} - 5 \right). \quad (2.24)$$

On suppose que $\frac{\partial K_S}{\partial P}$ et $\frac{\partial G}{\partial P}$ sont indépendants de la température à pression ambiante :

$$\frac{\partial K_S}{\partial P}_{(P_0, \vartheta)} \approx \frac{\partial K_S}{\partial P}_{(P_0, T_0)} \quad (2.25)$$

$$\frac{\partial G}{\partial P}_{(P_0, \vartheta)} \approx \frac{\partial G}{\partial P}_{(P_0, T_0)}. \quad (2.26)$$

Jackson (1998) a montré que cette approximation est valide pour le module d'incompressibilité. En revanche, cela n'a pas encore été montré pour G .

Pour évaluer L_3 et M_3 , il est nécessaire de connaître les dérivées secondes des modules élastiques par rapport à la pression. Ces dérivées secondes ont été mesurées

avec de grandes incertitudes sur un petit nombre de minéraux, par exemple l'olivine (Zha et al., 1998), la magnésio-wüstite (Murakami et al., 2009) et l'orthopyroxène (Flesch et al., 1998). L'intérêt de tronquer l'équation de Birch-Murnaghan à l'ordre 3, est d'imposer implicitement que le coefficient d'ordre 3 dans l'expression (2.13) de la pression est nul, ce qui donne l'équation (2.18). La dérivée seconde de K_S par rapport à P est alors reliée à $\frac{\partial K_S}{\partial P}$. Le long d'un profil de température adiabatique, la relation mise en place par Stacey (1992) selon laquelle $G = AK_S + BP$, où A et B sont calculés à pression nulle, conduit à :

$$\frac{\partial^2 G}{\partial^2 P} = A \frac{\partial^2 K_S}{\partial^2 P} = 0,631 \frac{\partial^2 K_S}{\partial^2 P}. \quad (2.27)$$

Les mesures de laboratoire des dérivées secondes des modules élastiques sont ainsi remplacées par les valeurs issues des équations (2.18) et (2.27). Les constantes C_j , L_j et M_j sont ensuite calculées à partir des modules élastiques et de leurs dérivées par rapport à la pression isentropique.

En admettant que la pression à chaque profondeur donnée est connue (dans ce travail on prendra le profil de pression du PREM), la déformation eulerienne ε est déterminée à partir de l'équation (2.13) en utilisant une méthode numérique de recherche de racine de type Newton-Raphson. Les expressions (2.10), (2.14) et (2.15) permettent ensuite d'évaluer la masse volumique et les modules élastiques d'un minéral donné aux conditions (P, T) désirées. En comparant le résultat T de l'équation (2.3) avec la valeur de $T(r)$ recherchée, la valeur de la température ϑ est modifiée, et ainsi de suite jusqu'à ce que la différence entre T et $T(r)$ soit inférieure à 5 K.

2.2.2 Du minéral à la roche

La théorie exposée dans le paragraphe précédent permet de calculer les propriétés élastiques des minéraux pris individuellement. Or le manteau terrestre correspond à un assemblage de minéraux. Pour un modèle pétrologique donné, les proportions des différentes phases minéralogiques peuvent être estimées à chaque profondeur grâce aux diagrammes de phases. Connaissant ces proportions, la moyenne pondérée des modules élastiques et des masses volumiques est calculée pour estimer les vitesses sismiques et la masse volumique de la roche à une profondeur donnée.

La masse volumique de la roche est calculée par une moyenne arithmétique des masses volumiques des minéraux pondérée par leurs fractions volumiques. Pour les vitesses sismiques, le problème est plus complexe. En général, deux moyennes correspondant à des cas extrêmes sont calculées. La « vraie » moyenne recherchée appartient à cet intervalle. La moyenne la plus communément utilisée est celle de Voigt-Reuss-Hill (Hill, 1952), qui correspond à la moyenne arithmétique des moyennes de Voigt et de Reuss.

La moyenne de Voigt considère que la déformation est uniforme dans le composé multi-phasé :

$$K = \frac{1}{n} \sum_{i=1}^n X^{(i)} K^{(i)}, \quad (2.28)$$

tandis que la moyenne de Reuss considère une contrainte uniforme :

$$\frac{1}{K} = \frac{1}{n} \sum_{i=1}^n X^{(i)} \frac{1}{K^{(i)}}, \quad (2.29)$$

avec n le nombre de minéraux considérés et $X^{(i)}$ la fraction volumique du minéral (i) . Ces équations sont également valables pour G .

La moyenne d'Hashin-Shtrikman (Hashin & Shtrikman, 1962), obtenue à partir d'expressions variationnelles, fournit de meilleures approximations pour des constituants polyphasés. Celle-ci est basée sur le principe de minimisation de l'énergie potentielle. Cette méthode remplace un corps hétérogène par un corps de référence équivalent qui produit la même déformation. Watt et al. (1976) ont montré que, même pour des monocristaux, il existe des cas où la la moyenne de Voigt-Reuss-Hill se situe en-dehors de l'intervalle formé par les deux bornes de la moyenne d'Hashin-Shtrikman. L'erreur introduite dans le calcul de la moyenne est plus petite dans le cas de la moyenne d'Hashin-Shtrikman, c'est pourquoi nous la préférons à la moyenne de Voigt-Reuss-Hill.

Pour les modules d'incompressibilité K et de cisaillement G , les bornes inférieures K_{HS}^- et G_{HS}^- , et supérieures K_{HS}^+ et G_{HS}^+ , sont reliées à la plus petite et à la plus grande des n valeurs de K et G (Hashin & Shtrikman, 1963; Watt et al., 1976) :

$$K_{\text{HS}}^{-,+} = K_{1,n} + \frac{A_{1,n}}{1 + \alpha_{1,n} A_{1,n}} \quad (2.30)$$

$$G_{\text{HS}}^{-,+} = G_{1,n} + \frac{1}{2} \frac{B_{1,n}}{1 + \beta_{1,n} B_{1,n}}, \quad (2.31)$$

où $K_{\text{HS}}^- < K < K_{\text{HS}}^+$ et $G_{\text{HS}}^- < G < G_{\text{HS}}^+$, K et G étant les valeurs recherchées. Avant de calculer la moyenne, les modules élastiques des différents minéraux sont classés par ordre croissant. Le minéral possédant le plus petit module élastique est donc indiqué par l'indice 1 et celui avec le plus grand module élastique avec l'indice n . La borne inférieure est alors calculée avec le module élastique le plus petit et la borne supérieure avec celui le plus grand. Les coefficients $\alpha_{1,n}$, $\beta_{1,n}$, $A_{1,n}$ et $B_{1,n}$ sont estimés de la manière suivante :

$$\alpha_{1,n} = -\frac{3}{3K_{1,n} + 4G_{1,n}} \quad (2.32)$$

$$\beta_{1,n} = -\frac{3(K_{1,n} + 2G_{1,n})}{5G_{1,n}(3K_{1,n} + 4G_{1,n})} \quad (2.33)$$

$$A_{1,n} = \sum_{i=2,1}^{n,n-1} \frac{X^{(i)}}{(K^{(i)} - K_{1,n})^{-1} - \alpha_{1,n}} \quad (2.34)$$

$$B_{1,n} = \sum_{i=2,1}^{n,n-1} \frac{X^{(i)}}{\frac{1}{2}(G^{(i)} - G_{1,n})^{-1} - \beta_{1,n}}. \quad (2.35)$$

Une fois que la masse volumique et les modules élastiques à une pression et une température données sont déterminés, les vitesses sismiques de l'assemblage minéralogique peuvent être calculées par les équations :

$$V_S(P, T) = \sqrt{\frac{K_S(P, T)}{\rho(P, T)}}, \quad (2.36)$$

$$V_P(P, T) = \sqrt{\frac{K_S(P, T) + \frac{4}{3}G(P, T)}{\rho(P, T)}}. \quad (2.37)$$

2.2.3 Diagrammes de phases

Les minéraux mantelliques qui sont considérés dans ce travail sont : l'olivine, l'orthopyroxène, le clinopyroxène et le grenat. Au fur à mesure que la température et la pression augmentent, les minéraux du manteau subissent des transformations minéralogiques. Les fractions volumiques des nouvelles phases minérales sont calculées en utilisant les diagrammes de phases construits à partir des mesures de laboratoire. Deux sous-systèmes indépendants sont considérés : le diagramme de phases de l'olivine et de ses phases hautes pressions, et un diagramme de phases pour tous les autres composants. Les diagrammes employés dans ce travail sont décrits dans [Vacher et al. \(1998\)](#). Ces derniers ont été construits en utilisant une méthode développée par [Ita & Stixrude \(1992\)](#) qui minimise l'énergie libre de Gibbs. Elle incorpore des changements des rapports Fe/Mg dans les minéraux en fonction de la profondeur.

D'après le diagramme de phases représenté sur la figure [2.3](#), l'olivine (phase α) se transforme progressivement en wadsleyite (phase β) et en ringwoodite (phase γ). À partir de 16 GPa, seule la ringwoodite peut exister. À des pressions supérieures à 24 GPa la ringwoodite se transforme en pérovskite et en magnésio-wüstite. La localisation des transitions de phase de l'olivine et de ses phases de hautes pressions ne dépend pas seulement de la pression et de la température, mais aussi de la teneur en fer. Cette dernière est quantifiée par le rapport entre les atomes de fer ferreux, et la somme des atomes magnésiens et des atomes de fer ferreux :

$$\text{Fe}\# = \frac{n(\text{Fe}^{2+})}{n(\text{Mg}) + n(\text{Fe}^{2+})}. \quad (2.38)$$

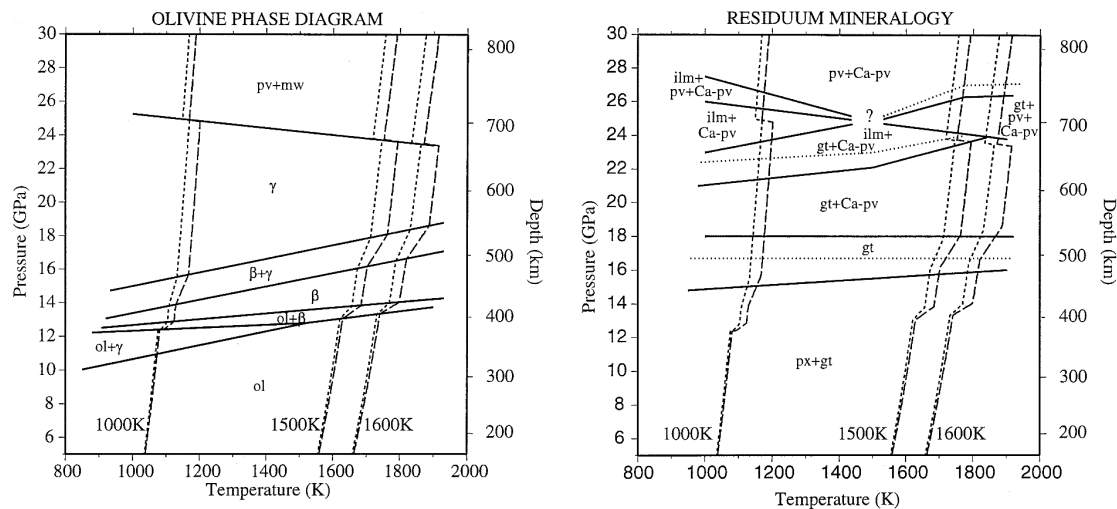


FIG. 2.3 : Diagrammes de phases de l'olivine ($\text{Mg}_{0.89}, \text{Fe}_{0.11})_2\text{SiO}_4$ (à gauche) et des pyroxènes-grenat (à droite) (Vacher et al., 1998). Les adiabats initiés à 1000, 1500 et 1600 K sont montrés pour la pyrolite (en pointillés longs) et la piclogite (en pointillés courts). Pour le diagramme de phases du sous-système pyroxènes-grenat, les traits pleins correspondent aux limites de phases pour une composition pyrolitique. Les pointillées représentent les limites de phases pour une composition piclogitique. ol : olivine, β : wadsleyite, γ : ringwoodite, pv : pérovskite magnésienne, mg : magnésiowüstite, px : pyroxènes, gt : grenat, Ca-pv : Ca-pérovskite, ilm : ilménite.

$\text{Fe}\#$ est appelé *iron number*, en anglais. Comme illustré sur la figure 1.2, qui présente le diagramme de phases de l'olivine dans le système fayalite-forstérite, les pressions auxquelles ont lieu les transformations diminuent lorsque la teneur en fer augmente.

Le sous-système des pyroxènes-grenat est plus complexe (figure 2.3). Les traits pleins et en pointillés correspondent respectivement aux limites de phases pour une composition pyrolitique et piclogitique (voir tableau 1.3). Puisque la piclogite est plus riche en Al_2O_3 que la pyrolite, les différents champs de stabilité de l'ilménite sont réduits pour une composition piclogitique par rapport à une composition pyrolitique. Pour la même raison, le champ de stabilité du grenat et de la pérovskite à haute température augmente lorsque le taux d' Al_2O_3 augmente (e. g. Irifune et al., 1996). Le constat important que l'on peut déduire de ces diagrammes, est que les transitions de phase principales dépendent fortement de la température. Comme le font remarquer Vacher et al. (1998), il existe dans le détail des incertitudes sur la localisation des limites de phases dans le système pyroxènes-grenat entre 22 et 24 GPa. Par exemple, la localisation de la transition du grenat en ilménite est incertaine. Puisqu'il est difficile de concilier les résultats issus des expériences menées à haute et basse températures, un point d'interrogation a été placé au niveau des pressions correspondant au centre du manteau.

Les formules chimiques des minéraux du manteau sont celles employées par [Verhoeven et al. \(2005\)](#) et sont détaillées dans le tableau 2.4. Les polymorphes de hautes pressions possèdent des formules chimiques identiques à celles des minéraux de basses pressions. Cependant, leurs propriétés thermoélastiques associées sont significativement différentes (voir tableaux 2.1, 2.2 et 2.3). Les propriétés des minéraux individuels dépendent explicitement de la teneur en fer :

$$\rho_0^{(i)}(P_0, T_0) = \rho^{(i)}(P_0, T_0) + y \rho_{0,Fe}^{(i)}(P_0, T_0) \quad (2.39)$$

$$K_{S_0}^{(i)}(P_0, T_0) = K_S^{(i)}(P_0, T_0) + y K_{S_0,Fe}^{(i)}(P_0, T_0) \quad (2.40)$$

$$G_0^{(i)}(P_0, T_0) = G^{(i)}(P_0, T_0) + y G_{0,Fe}^{(i)}(P_0, T_0), \quad (2.41)$$

où l'indice i indique la phase minérale et y la teneur en fer. En général, la fraction d'atomes de fer dans chaque minéral est fonction de la pression et de la température, mais également des autres minéraux présents à ces mêmes conditions thermodynamiques. Ainsi, le fer peut être réparti différemment entre les minéraux individuels. Dans cette étude, on considère que le manteau est chimiquement homogène et que la répartition du fer dans les minéraux individuels est identique. Pour les phases minérales pures, la teneur en fer est donc identique à $Fe\#$. À l'intérieur du domaine de pressions et de températures, lorsque deux phases coexistent pendant un changement de phase, la fraction d'atomes de fer pour chaque phase minérale est déterminée à partir des diagrammes de phases. Par exemple, dans le système olivine, Mg_2SiO_4 et Fe_2SiO_4 forment une solution solide. Lorsqu'une seule phase est présente, la teneur en fer est égale à $Fe\#$. Quand une transformation a lieu, la proportion relative de forstérite et de fayalite dans chaque phase est calculée par l'intermédiaire du diagramme de phases.

Minéral	formule chimique
olivine	$(Mg_{1-y} Fe_y)_2 SiO_4$
wadsleyite	$(Mg_{1-y} Fe_y)_2 SiO_4$
ringwoodite	$(Mg_{1-y} Fe_y)_2 SiO_4$
Mg-pérovskite	$(Mg_{1-y} Fe_y) SiO_3$
Mg-wüstite	$(Mg_{1-y} Fe_y) O$
clinopyroxène BP	$(Mg_{1-y} Fe_y) SiO_3$
clinopyroxène HP	$(Mg_{1-y} Fe_y) SiO_3$
orthopyroxène	$(Mg_{1-y} Fe_y) SiO_3$
Ca-pyroxène	$Ca (Mg_{1-y} Fe_y) Si_2O_6$
ilménite	$(Mg_{1-y} Fe_y) SiO_3$
majorite	$(Mg_{1-y} Fe_y) SiO_3$

TAB. 2.4 : Formules chimiques des minéraux du manteau. y correspond à la teneur en fer.

Lorsque la composition globale du manteau contient de l'aluminium, il est en général incorporé plutôt dans les pyroxènes, le grenat, l'ilménite et la pérovskite. Il est considéré dans cette étude que la teneur en aluminium est suffisamment faible pour que les phases alumineuses soient négligées.

L'effet de Verhoogen ([Jeanloz & Thompson, 1983](#)), c'est-à-dire le fait qu'un profil thermique adiabatique soit décalé lorsqu'il franchit une transition de phase, est pris en compte pour le système olivine. Les transitions de phases induisent une augmentation ou une diminution de température ΔT :

$$\Delta T = \frac{\mathcal{L}}{C_P} = \frac{\Gamma T \Delta V}{C_P}, \quad (2.42)$$

où $\mathcal{L}$ est la chaleur latente de la réaction, ΔV la variation de volume et Γ la pente de Clapeyron de la transition de phase, et où C_P est une capacité calorifique moyenne ([Turcotte & Schubert, 1982](#), p. 193). Sur la figure 2.3 sont représentés trois profils de température adiabatiques pour différents pieds de profils adiabatiques : 1000, 1500 et 1600 K. Le gradient adiabatique est décalé lorsqu'il franchit les limites de phases du système olivine. Il est décalé vers des températures plus importantes à 410 et 520 km de profondeur car les réactions sont exothermiques (voir tableau 1.2). En revanche, il se déplace vers des températures plus faibles à 660 km de profondeur, où la transformation minéralogique de ringwoodite en pérovskite et magnésiowüstite est endothermique. La variation en température correspondante est égale à 60 K et 40 K pour une pyrolite et une piclogite, respectivement, en considérant une pente de Clapeyron égale à -2 MPa/K. Évidemment cette variation dépend des valeurs choisies pour les pentes de Clapeyron, qui restent encore à affiner (voir tableau 1.2).

D'après les tableaux 2.1 et 2.2, bien que les incertitudes des données expérimentales sur les modules élastiques des minéraux de l'olivine et de ses polymorphes soient parmi les plus faibles, les incertitudes demeurent importantes pour les pyroxènes et le grenat. Parmi les rares travaux qui tiennent compte de ces incertitudes, [Cobden et al. \(2008\)](#) ont généré un ensemble de 10 000 modèles du manteau, en considérant une composition pyrolitique et une température potentielle de 1600 K, en faisant varier les valeurs des paramètres élastiques à l'intérieur des barres d'erreur décrites dans [Cammarano et al. \(2003\)](#). Ils ont ensuite comparé directement les données synthétiques issues de ces modèles avec des temps de trajet réels d'ondes P , S et SS . Malgré la prise en compte de la variabilité des paramètres élastiques, le gradient de vitesse entre 410 et 660 km de profondeur demeure inférieur à celui des modèles de référence. Il est clair que les incertitudes sur les paramètres des phases de l'olivine ont un effet plus important sur l'amplitude des sauts de vitesse à 410, 520 et 660 km de profondeur que l'effet Verhoogen. Dans le diagramme pyroxènes-grenat, les réactions s'étendent sur des gammes de profondeurs plus grandes que dans le diagramme de l'olivine et les valeurs des pentes de Clapeyron sont encore

moins bien contraintes que pour celles des transformations de l'olivine. Pour ces raisons, l'effet Verhoogen n'est pas pris en compte dans le diagramme pyroxènes-grenat.

2.3 Anisotropie

2.3.1 Le tenseur élastique

Contrairement au corps isotrope qui, indépendamment de l'orientation, se comporte de façon identique, l'anisotropie caractérise un milieu qui possède des propriétés différentes selon la direction considérée, abstraction faite du caractère homogène du milieu. La manifestation de l'anisotropie provient d'une sensibilité des ondes sismiques différente selon les coefficients du tenseur élastique anisotrope. Dans un référentiel fixe de coordonnées x_i ($i = 1, 2, 3$), le tenseur des contraintes σ_{ij} et le tenseur des déformations ϵ_{ij} sont reliés par le tenseur élastique C_{ijkl} selon la loi de Hooke généralisée (ou loi de l'élasticité linéaire) :

$$\sigma_{ij} = C_{ijkl} \epsilon_{ij}. \quad (2.43)$$

C_{ijkl} est un tenseur d'ordre 4 qui contient 81 composantes. Le tenseur des contraintes et celui des déformations étant symétriques, il ne reste plus que 36 coefficients indépendants. En raison des différentes symétries des tenseurs σ ($\sigma_{ij} = \sigma_{ji}$) et ϵ ($\epsilon_{ij} = \epsilon_{ji}$), ainsi qu'une condition sur l'énergie libre, le tenseur élastique devient également symétrique ($C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}$). Dans ce cas, il n'existe plus que 21 coefficients indépendants qui décrivent un milieu anisotrope. Pour simplifier l'écriture de la matrice C_{ijkl} , l'écriture matricielle de Voigt est employée. Le remplacement des indices s'exprime pour le tenseur élastique d'ordre 4 de la manière suivante :

$$C_{ijkl} \rightarrow C_{pq} \begin{cases} \text{si } i = j, & p = i \\ \text{si } i \neq j, & p = 9 - i - j \\ \text{si } k = l, & q = k \\ \text{si } k \neq l, & q = 9 - k - l \end{cases} \quad (2.44)$$

Ce dernier devient alors

$$C_{pq} = \begin{pmatrix} \mathbf{C_{11}} & \mathbf{C_{12}} & \mathbf{C_{13}} & \mathbf{C_{14}} & \mathbf{C_{15}} & \mathbf{C_{16}} \\ C_{21} & \mathbf{C_{22}} & \mathbf{C_{23}} & \mathbf{C_{24}} & \mathbf{C_{25}} & \mathbf{C_{26}} \\ C_{31} & C_{32} & \mathbf{C_{33}} & \mathbf{C_{34}} & \mathbf{C_{35}} & \mathbf{C_{36}} \\ C_{41} & C_{42} & C_{43} & \mathbf{C_{44}} & \mathbf{C_{45}} & \mathbf{C_{46}} \\ C_{51} & C_{52} & C_{53} & C_{54} & \mathbf{C_{55}} & \mathbf{C_{56}} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & \mathbf{C_{66}} \end{pmatrix}, \quad (2.45)$$

où les composantes indépendantes sont représentées en gras.

À partir de cette forme générale, il est possible de réduire encore le nombre de composantes indépendantes lorsque le milieu élastique présente une ou plusieurs symétrie(s) dans certaines directions. Les trois cas principaux sont :

1. un milieu élastique avec une symétrie monoclinique (13 composantes indépendantes),
2. un milieu transverse isotrope à axe hexagonal (5 composantes indépendantes),
3. un milieu complètement isotrope (2 composantes indépendantes, qui correspondent aux paramètres de Lamé).

L'anisotropie radiale est décrite par ce deuxième cas et c'est celui-ci que nous allons étudier.

2.3.2 Anisotropie radiale

Sous l'hypothèse d'anisotropie hexagonale, le nombre de coefficients élastiques indépendants est réduit à 5 au lieu de 21. Il s'agit des coefficients de Love A , C , F , L et N (Love, 1927). Le tenseur élastique peut s'écrire

$$C_{pq} = \begin{pmatrix} A & A - 2N & F & 0 & 0 & 0 \\ A - 2N & A & F & 0 & 0 & 0 \\ F & F & C & 0 & 0 & 0 \\ 0 & 0 & 0 & L & 0 & 0 \\ 0 & 0 & 0 & 0 & L & 0 \\ 0 & 0 & 0 & 0 & 0 & N \end{pmatrix}. \quad (2.46)$$

Ces coefficients de Love sont reliés aux vitesses sismiques par les relations :

$$A = \rho V_{PH}^2 \quad (2.47)$$

$$C = \rho V_{PV}^2 \quad (2.48)$$

$$L = \rho V_{SV}^2 \quad (2.49)$$

$$N = \rho V_{SH}^2 \quad (2.50)$$

$$F = \frac{\eta}{A - 2L}, \quad (2.51)$$

où V_{PH} et V_{PV} sont les vitesses des ondes P se propageant horizontalement et verticalement, et V_{SH} et V_{SV} sont les vitesses des ondes S polarisées horizontalement et verticalement. Le paramètre η décrit la dépendance de la vitesse de propagation d'une onde selon l'angle d'incidence.

À partir des moyennes de Voigt des modules de cisaillement et d'incompressibilité, il est possible d'exprimer V_P et V_S en fonction des vitesses anisotropes (Babuska & Cara, 1991). Les moyennes de Voigt de G et K dans un milieu radialement anisotrope s'écrivent (Babuska & Cara, 1991) :

$$K = \frac{1}{9}(C + 4A - 4N + 4F) \quad (2.52)$$

$$G = \frac{1}{15}(C + A + 6L + 5N - 2F). \quad (2.53)$$

Les vitesses isotropes sont définies en fonction des modules de cisaillement et d'incompressibilité de la manière suivante :

$$V_P^2 = \frac{K + \frac{4}{3}G}{\rho} \quad (2.54)$$

$$V_S^2 = \frac{K}{\rho}. \quad (2.55)$$

On peut substituer (2.51), (2.52) et (2.53) dans (2.54) et (2.55) pour obtenir :

$$\rho V_P^2 = \frac{1}{15} [3C + (8 + 4\eta)A + 8(1 - \eta)L] \quad (2.56)$$

$$\rho V_S^2 = \frac{1}{15} [C + (1 - 2\eta)A + (6 + 4\eta)L + 5N]. \quad (2.57)$$

η est égal à 1 dans un milieu isotrope. Les vitesses moyennes isotropes décrites dans les équations (2.56) et (2.57) dépendent des observables sismiques des équations (2.47)-(2.51), ainsi que de η . Cependant, si on suppose que l'anisotropie est faible, il est possible de faire l'hypothèse que $\eta \approx 1$. De plus, en considérant l'approximation que l'anisotropie des ondes P est bien inférieure à celle des ondes S , le terme $C - A$ peut être négligé au premier ordre par rapport au terme $10L + 5N$ dans l'équation (2.57). Ainsi, les moyennes de Voigt des vitesses isotropes peuvent être simplifiées par :

$$V_S^2 = \frac{2L + N}{3\rho} = \frac{2V_{SV}^2 + V_{SH}^2}{3} \quad (2.58)$$

$$V_P^2 = \frac{C + 4A}{5\rho} = \frac{V_{PV}^2 + 4V_{PH}^2}{5}, \quad (2.59)$$

de telle manière que la vitesse moyenne isotrope des ondes S dépende uniquement de V_{SV} et V_{SH} , et que la vitesse des ondes P soit uniquement liée à V_{PV} et V_{PH} . Cette simplification a déjà été utilisée pour diminuer le nombre de paramètres de l'inversion par Panning & Romanowicz (2006) ou encore Khan et al. (2009, 2011). Les paramètres

anisotropes ξ et ϕ décrivent l'anisotropie des ondes S et P , respectivement :

$$\xi = \frac{V_{SH}^2}{V_{SV}^2} \quad (2.60)$$

$$\phi = \frac{V_{PV}^2}{V_{PH}^2}. \quad (2.61)$$

Il est donc possible de déterminer les vitesses anisotropes en utilisant les équations (2.58) à (2.61) :

$$V_{SH} = V_S \sqrt{\frac{3\xi}{2+\xi}} \quad (2.62)$$

$$V_{SV} = V_S \sqrt{\frac{3}{2+\xi}} \quad (2.63)$$

$$V_{PH} = V_P \sqrt{\frac{5}{4+\phi}} \quad (2.64)$$

$$V_{PV} = V_P \sqrt{\frac{5\phi}{4+\phi}}. \quad (2.65)$$

Notons que cette formulation de l'anisotropie est totalement découplée des calculs thermodynamiques sur les propriétés des roches isotropes. Il est théoriquement possible de prendre en compte l'anisotropie dans les calculs thermodynamiques (Stixrude & Lithgow-Bertelloni, 2005b). Cependant, à notre connaissance, il n'existe pas encore de base de données sur les coefficients élastiques anisotropes pour l'ensemble des minéraux considérés.

Les ondes de Love et de Rayleigh étant peu sensibles à l'anisotropie des ondes P , seul le paramètre ξ est inversé. Si V_{SH} est plus grande que V_{SV} ($\xi > 1$), nous aurons tendance à interpréter cela comme un mouvement horizontal de matière, alors que l'inverse ($\xi < 1$) désignera plutôt des courants verticaux. Les ondes de surface ne sont également que faiblement dépendantes de η dans le manteau. Khan et al. (2009, 2011) ont évalué les deux paramètres ϕ et η selon une méthode bayésienne et ont montré qu'ils n'étaient que très peu contraints. η est peu souvent considéré comme étant un paramètre variable de l'inversion. Par exemple, Shapiro & Ritzwoller (2002) le fixent selon les valeurs du PREM. Gung et al. (2003), Panning & Romanowicz (2006) et Beghein (2010) utilisent une relation d'échelle selon laquelle $\partial \ln \eta = 2,5 \partial \ln \xi$. Cette relation est basée sur des calculs et résultats issus d'expériences de laboratoire (Montagner & Anderson, 1989). Afin d'être cohérent avec l'approximation selon laquelle l'anisotropie est faible et avec la simplification théorique employée, η est ici fixé à la valeur de 1.

2.4 Vitesses de phase

Les modèles sismiques radiaux qui correspondent à un profil de température, un profil d'anisotropie et à une composition minéralogique donnés sont ensuite employés pour calculer les courbes de dispersion des vitesses de phase des ondes de surface. Le code MINEOS, largement employé aujourd'hui pour calculer les fréquences et les fonctions propres, a été utilisé. Il est basé sur le travail de [Gilbert & Dziewonski \(1975\)](#), mais a été repris par de nombreux collaborateurs tels que [Woodhouse \(1988\)](#) et G. Masters. Le programme MINEOS est disponible sur le site du CIG (Computational Infrastructure for Geodynamics, <http://www.geodynamics.org/cig/software/mineos>). Dans cette section, nous rappelons brièvement l'essentiel de la théorie des modes propres. Pour plus de détails, le lecteur se référera par exemple à [Takeushi & Saito \(1972\)](#), [Aki & Richards \(1980\)](#), ou encore à [Dahlen & Tromp \(1998\)](#).

2.4.1 Les oscillations libres

Tout corps élastique de dimensions finies peut vibrer librement à des fréquences déterminées par sa forme et sa constitution. Le spectre de ses oscillations contient un nombre infini de fréquences discrètes, appelées *modes propres* ou *oscillations libres*. Le terme « libre » est important car il signifie que ces oscillations continuent même si la source d'excitation a cessé d'agir. Sous l'effet d'une impulsion initiale, une corde attachée à ses deux extrémités se met à osciller autour de sa position de repos. Seules certaines longueurs d'onde et certaines fréquences sont permises, déterminées par les propriétés physiques du corps et par les conditions aux limites. Deux approches mathématiques équivalentes sont possibles pour décrire ces vibrations : soit une approche modale, soit une approche d'ondes progressives. C'est la première que nous allons suivre. Les oscillations ont lieu à des fréquences différentes, appelées *fréquences propres* (ω_n). L'ordre $n = 0$ est qualifié de mode fondamental et tous les suivants ($n = 1, 2, \dots$) correspondent aux modes harmoniques. Le déplacement associé à chaque fréquence propre est appelé *fonction propre*. L'ordre radial n donne le nombre de nœuds en profondeur, c'est-à-dire le nombre de fois où la courbe représentant la fonction propre en fonction de la profondeur passe par zéro. Les fonctions et fréquences propres ne dépendent pas de la force extérieure responsable de l'excitation, mais de la nature physique et de la géométrie de l'objet considéré.

Les interférences entre les ondes produites par le séisme entraînent deux types d'oscillation de la sphère : les oscillations *sphéroïdales*, qui sont des oscillations pour lesquelles le rotationnel du déplacement autour de la verticale est nul, et les oscillations *toroïdales*, qui sont des oscillations de torsion pour lesquelles le déplacement radial et la divergence du déplacement sont nuls (c'est-à-dire qu'il n'y a pas de changement de

volume). D'un point de vue mathématique, les ondes de surface de Rayleigh et de Love sont équivalentes à ces oscillations libres. Leur propagation est schématisée sur la figure 2.4.

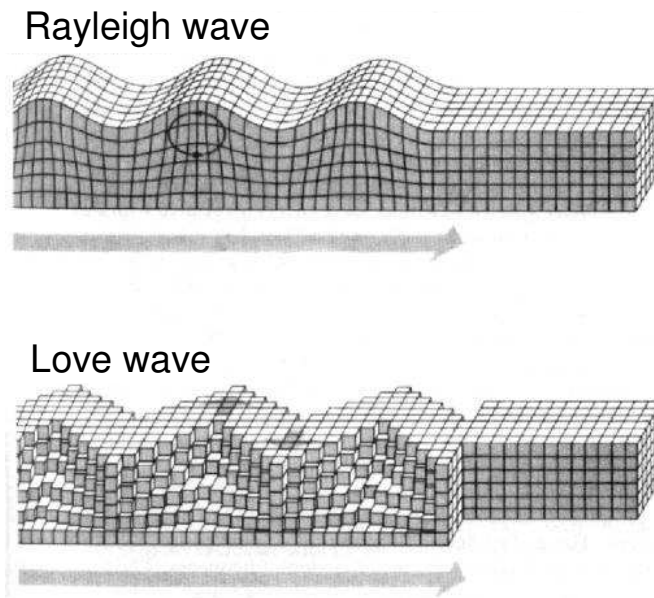


FIG. 2.4 : Mouvement des particules lors d'un passage d'une onde de Rayleigh et de Love (d'après [Lay & Wallace, 1995](#)).

- Les ondes de Rayleigh résultent des interférences constructives qui se produisent entre les ondes P et SV aux différentes interfaces de la Terre. Dans un milieu isotrope, elles ont un mouvement particulier elliptique dans le plan vertical et parallèle à la direction de propagation. Ces ondes se formalisent comme étant la somme des fonctions propres sphéroïdales ou des produits des fonctions propres sphéroïdales et toroïdales (suivant qu'il s'agit de la projection verticale ou horizontale du déplacement). Les ondes de Rayleigh sont observées sur les composantes longitudinale et verticale d'un sismogramme.
- Les ondes de Love sont le résultat d'interférences entre les ondes SH et correspondent à la somme des différentes fonctions propres toroïdales. Elles sont polarisées dans le plan horizontal perpendiculairement à la direction de propagation. Ces ondes sont visibles sur la composante transverse d'un sismogramme.

Les interférences entre ces deux types d'ondes, qui équivalent à des couplages entre modes, peuvent être provoquées par des hétérogénéités latérales, la rotation propre de la Terre et son ellipticité.

2.4.2 Solutions du problème aux valeurs propres

Comment calculer les modes propres de la Terre ? Supposons que par rapport à un état de référence en équilibre hydrostatique, une perturbation est introduite au niveau de la masse volumique et du champ de contraintes. Par rapport à cet état perturbé, la théorie des modes propres est une théorie linéaire. En effet, la réponse de la Terre est supposée instantanée et linéaire par rapport à l'excitation initiale et les équations du mouvement sont linéarisées.

Le modèle de Terre utilisé présente les propriétés suivantes. La Terre est à symétrie sphérique et elle est autogravitante. Les contraintes sont non nulles même à l'état d'équilibre. Le modèle est élastique, c'est-à-dire qu'il existe une relation linéaire entre les contraintes et les déformations. La Terre est formée de couches concentriques latéralement homogènes et elle peut être modélisée comme un milieu transverse isotrope à axe vertical. Pour chaque couche, le tenseur des contraintes est donc entièrement déterminé par les cinq modules élastiques A, C, F, L , et N .

L'équation linéarisée de conservation du moment cinétique met en jeu les forces inertielles, les forces élastiques liées à la déformation du corps et les forces gravitationnelles dues à la variation du potentiel gravitationnel. En employant le formalisme de [Aki & Richards \(1980\)](#), cette équation prend la forme pour un milieu isotrope :

$$\begin{aligned}
 \underbrace{-\rho_0 \omega^2 \mathbf{u}}_{\text{inertie}} &= \overbrace{\left(u_r \frac{d\rho_0}{dr} + \rho_0 \nabla \cdot \mathbf{u} \right) g_0 \hat{\mathbf{r}} + \rho_0 \nabla \kappa - \nabla(\rho_0 u_r g_0)}^{\text{gravitation}} \\
 &+ \underbrace{(\lambda + 2\mu) \nabla \nabla \cdot \mathbf{u} - \mu \nabla \times (\nabla \times \mathbf{u}) + \frac{d\lambda}{dr} (\nabla \cdot \mathbf{u}) \hat{\mathbf{r}}}_{\text{élasticité}} \\
 &+ 2 \frac{d\mu}{dr} \frac{\partial \mathbf{u}}{\partial r} + \frac{d\mu}{dr} \hat{\mathbf{r}} \times (\nabla \times \mathbf{u}),
 \end{aligned} \tag{2.66}$$

avec :

$\mathbf{u}$, le champ de déplacement et u_r sa composante radiale,

$\hat{\mathbf{r}}$, le rayon vecteur en coordonnées sphériques,

ρ_0 , la masse volumique non perturbée,

κ , la perturbation eulérienne du potentiel gravitationnel,

λ , le premier coefficient de Lamé,

μ , la rigidité,

g_0 , le champ de gravité non perturbé,

et ω , la pulsation.

Les solutions de cette équation sont les fonctions propres associées chacune à une fréquence propre. Dès lors, il faut considérer les conditions aux limites appropriées à notre planète, à savoir la continuité des vecteurs de déplacement et de contrainte à la surface, qui peut être également formulée comme étant l'annulation de la traction aux interfaces. L'équation du mouvement est résolue grâce à une décomposition en harmoniques sphériques vectorielles. Dans le système de coordonnées sphériques (r, Δ, ϕ) (voir figure 2.5), le vecteur du déplacement élastique $\mathbf{u}$ peut être représenté par l'intermédiaire de trois vecteurs $\mathbf{R}_l^m$, $\mathbf{S}_l^m$ et $\mathbf{T}_l^m$ qui définissent une base de vecteurs propres orthogonaux :

$$\begin{aligned} \mathbf{u}(r, \Delta, \phi, \omega) = & \sum_n \sum_l \sum_m [{}_n U_l(r) \mathbf{R}_l^m(\Delta, \phi) + {}_n V_l(r) \mathbf{S}_l^m(\Delta, \phi) \\ & + {}_n W_l(r) \mathbf{T}_l^m(\Delta, \phi)] \exp(-i_n \omega_l t), \end{aligned} \quad (2.67)$$

où les fonctions propres radiales ${}_n U_l$, ${}_n V_l$ et ${}_n W_l$ sont fonctions du rayon uniquement. Les nombres l et m correspondent aux indices des harmoniques sphériques qui donnent les nœuds en latitude et en longitude. La Terre étant supposée à symétrie sphérique, il y a dégénérescence de l'ordre azimutal m et les modes propres dépendent uniquement de l'ordre radial n et de l'ordre angulaire l . Pour un mode n fixé, l'ordre angulaire l détermine la pulsation ${}_n \omega_l$ considérée. Ainsi, pour chaque couple (n, l) , c'est-à-dire pour des pulsations discrètes ${}_n \omega_l$, il est possible de déterminer les fonctions propres de la Terre associées aux deux types d'oscillations toroïdales et sphéroïdales. Les vecteurs $\mathbf{R}_l^m$, $\mathbf{S}_l^m$ et $\mathbf{T}_l^m$ correspondent respectivement aux composantes radiale, sphéroïdale et toroïdale. Ils s'expriment en fonction des harmoniques sphériques de la façon suivante :

$$\mathbf{R}_l^m(\Delta, \phi) = \hat{\mathbf{r}} Y_l^m \quad (2.68)$$

$$\mathbf{S}_l^m(\Delta, \phi) = \frac{\nabla_s Y_l^m}{\sqrt{l(l+1)}} = \frac{1}{\sqrt{l(l+1)}} \left(\frac{\partial Y_l^m}{\partial \Delta} \hat{\Delta} + \frac{1}{\sin \Delta} \frac{\partial Y_l^m}{\partial \phi} \hat{\phi} \right) \quad (2.69)$$

$$\mathbf{T}_l^m(\Delta, \phi) = \frac{\hat{\mathbf{r}} \times \nabla_s Y_l^m}{\sqrt{l(l+1)}} = \frac{1}{\sqrt{l(l+1)}} \left(\frac{1}{\sin \Delta} \frac{\partial Y_l^m}{\partial \phi} \hat{\Delta} - \frac{\partial Y_l^m}{\partial \Delta} \hat{\phi} \right), \quad (2.70)$$

où ∇_s est le gradient surfacique et Y_l^m est le scalaire harmonique sphérique

$$Y_l^m(\Delta, \phi) = (-1)^m \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!} \right] P_l^m(\cos \Delta) e^{im\phi}. \quad (2.71)$$

Le tenseur de contrainte est développé de façon analogue.

En remplaçant $\mathbf{u}$ par son développement en harmoniques sphériques vectorielles dans l'équation du mouvement linéarisée (équation 2.67), on obtient pour chaque pulsation discrète ${}_n \omega_l$ deux systèmes d'équations différentielles. Pour les oscillations sphéroïdales le système est constitué de 6 équations qui permettent de déterminer les fonctions propres ${}_n U_l$, ${}_n V_l$, ${}_n R_l$, ${}_n S_l$, ${}_n K_l$ et ${}_n G_l$. Les fonctions propres ${}_n U_l$ et ${}_n V_l$ sont respective-

ment proportionnelles au déplacement radial et au déplacement horizontal. ${}_nR_l$ est liée à la contrainte radiale, ${}_nS_l$ aux contraintes cisailantes et ${}_nK_l$ au potentiel gravitationnel. Enfin, ${}_nG_l$ est liée au gradient de ce potentiel. La gravité n'affecte pas les modes toroïdaux car ceux-ci n'ont pas de composante radiale de déplacement et laissent donc la distribution de masse volumique inchangée. Ainsi, dans le cas des oscillations toroïdales, l'équation du mouvement est résolue en se ramenant à un système de deux équations différentielles grâce auxquelles il est possible de calculer les fonctions propres ${}_nW_l$ et ${}_nT_l$. ${}_nW_l$ est associée au déplacement horizontal et ${}_nT_l$ aux contraintes cisailantes. Les expressions des équations différentielles sont détaillées dans [Aki & Richards \(1980\)](#) avec le même formalisme.

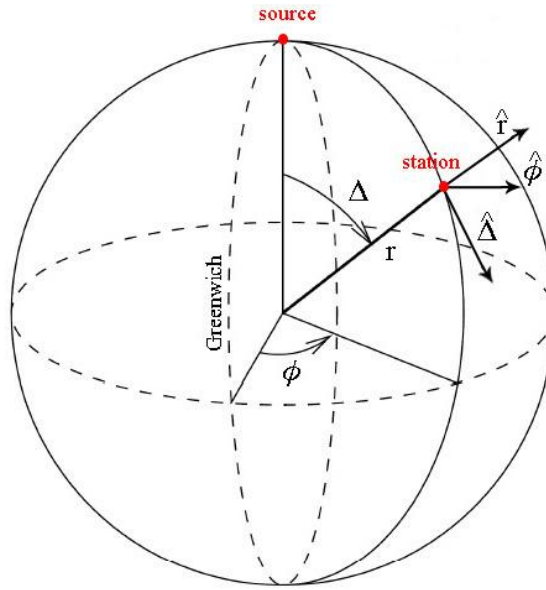


FIG. 2.5 : Coordonnées sphériques (r, Δ, ϕ) , avec l'origine située au centre de la Terre.

2.4.3 La dispersion des ondes de surface

Nous avons vu que la Terre, en tant qu'objet physique, possède plusieurs fréquences propres et qu'il existe un déplacement associé à chacune d'entre elles. Les figures 2.6 et 2.7 présentent les fonctions propres U et W , respectivement, pour les modes $n = 0$ à $n = 3$. Dans le cas d'une Terre latéralement homogène mais contenant des variations radiales, pour une profondeur donnée, ces fonctions propres sont identiques pour les tous les points de la coquille sphérique. En revanche, les figures 2.6 et 2.7 montrent qu'elles varient de façon différente avec la profondeur. Les ondes de surface sont des ondes dispersives, c'est-à-dire que leur vitesse de propagation apparente dépend de la fréquence. Deux fréquences propres différentes associées à une même branche ($n = 0$ ou $n = 1, 2, \dots$) auront des vitesses de propagation différentes.

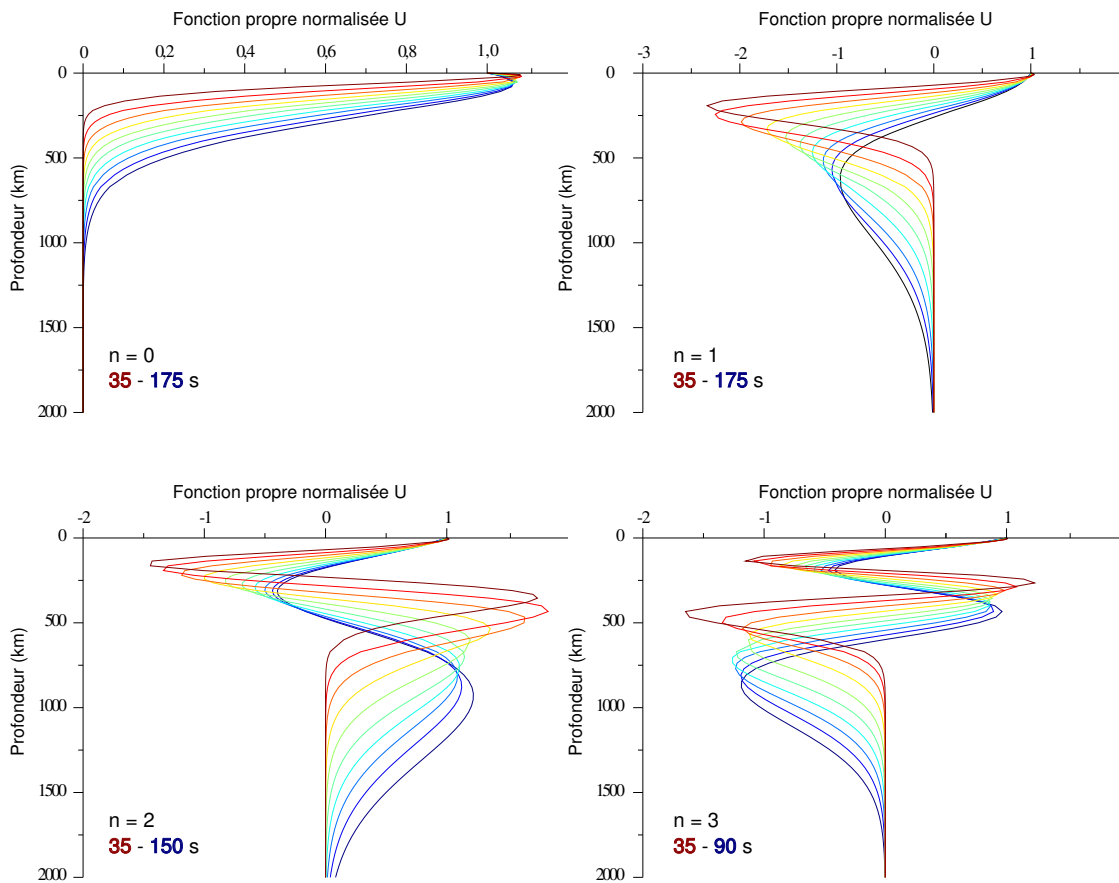


FIG. 2.6 : Courbes de déplacement radial des premiers modes sphéroïdaux en fonction de la profondeur. Les calculs ont été réalisés pour le PREM (Dziewonski & Anderson, 1981) en utilisant MINEOS. Les fonctions propres sont normalisées par rapport au déplacement en surface ($z = 0$). Les courbes du mode fondamental et du premier harmonique sont calculées toutes les 15 s alors que l'intervalle est de 9 s pour le mode 2 et de 5 s pour le mode 3. Du bleu au rouge, les périodes sont de plus en plus faibles et indiquent des ordres angulaires l de plus en plus petits.

Par définition, plus le mode est élevé, plus la courbe est oscillante. Effectivement, l'ordre radial n indique le nombre de nœuds radiaux, c'est-à-dire le nombre de fois où la fonction propre s'annule. Ces nœuds peuvent être considérés comme étant des nœuds de vibration. Plus la période est faible, plus la sensibilité à la Terre profonde diminue. Cette propriété est très intéressante car l'étude de la dispersion des ondes de surface donne accès à la structure moyenne des milieux traversés entre la surface et la profondeur maximale de pénétration de ces modes. Ainsi, un trajet d'onde de surface échantillonne mieux les milieux traversés qu'un trajet d'onde de volume, ces dernières étant sensibles à la structure le long d'un rai uniquement. La pénétration d'un mode augmente avec la période mais également avec l'ordre radial n , d'où l'intérêt d'utiliser plusieurs modes harmoniques et une large gamme de périodes pour obtenir de l'information sur la zone de transition. En

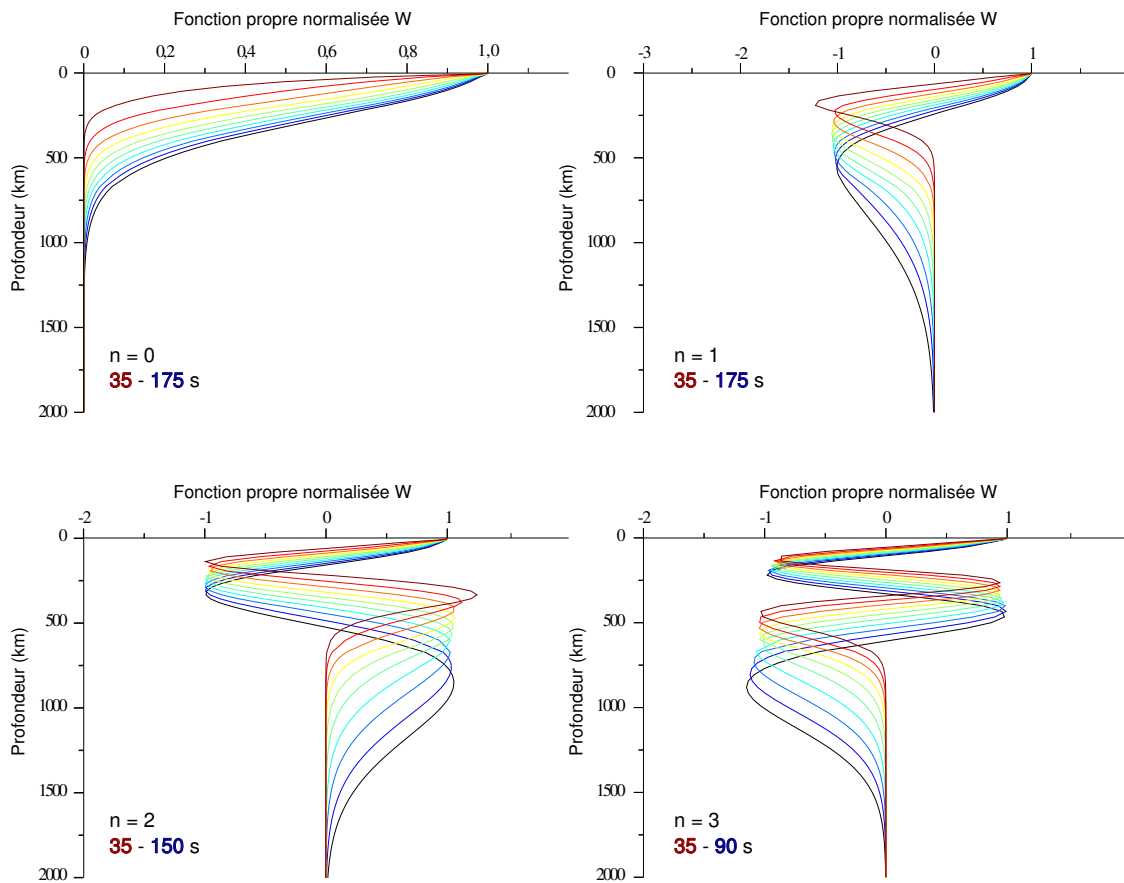


FIG. 2.7 : Courbes de déplacement horizontal des premiers modes toroïdaux en fonction de la profondeur. Les calculs ont été réalisés pour le PREM (Dziewonski & Anderson, 1981) en utilisant MINEOS. Les fonctions propres sont normalisées par rapport au déplacement en surface ($z = 0$). Les courbes du mode fondamental et du premier harmonique sont calculées toutes les 15 s alors que l'intervalle est de 9 s pour le mode 2 et de 5 s pour le mode 3. Du bleu au rouge, les périodes sont de plus en plus faibles et indiquent des ordres angulaires l de plus en plus petits.

effet, les différents modes sont complémentaires. Le mode fondamental permet de sonder le manteau jusqu'à environ 250 km de profondeur. Avec le premier harmonique, bien que la sensibilité des fonctions propres décroisse avec la période, il est théoriquement possible d'obtenir de l'information jusqu'à 600 km. Les modes 2 et 3 viennent combler la lacune de sensibilité des deux modes précédents, avec des maxima situés à la base de la zone de transition, vers 800 km. Prendre en compte les modes harmoniques pour étudier la zone de transition représente une source d'informations précieuse, mais le maximum de sensibilité pour tous les modes se situe dans les 500 premiers kilomètres de la Terre. Il faut donc s'attendre à ce que les paramètres de l'inversion soient de moins en moins bien contraints passée cette région.

Les sources sismiques génèrent des ondes qui occupent une plage fréquentielle continue. Chaque onde monochromatique possède sa propre vitesse, appelée *vitesse de phase*. La vitesse de phase C correspondant à la pulsation ω est décrite par l'équation suivante :

$$C = \frac{\omega}{k}, \quad (2.72)$$

où k représente le nombre d'onde. En raison de la distribution radiale des propriétés du milieu, les longues périodes, plus sensibles en profondeur, ont une vitesse de phase plus élevée que les courtes périodes. Cela signifie qu'en un point fixe de la Terre et pour un mode donné, les déplacements à longues périodes vont toujours précéder les déplacements à courtes périodes (figure 2.8). À partir du mode fondamental, les modes harmoniques présentent des vitesses de phase de valeurs croissantes. En fonction de la période (ou de la fréquence), les vitesses de phase d'un même mode constituent une *courbe de dispersion*. Ces dernières peuvent être extraites selon différentes méthodes à partir du sismogramme, qui constitue la donnée brute de l'observation sismologique. La plupart des techniques pour estimer les vitesses de phase sont basées sur la comparaison entre le sismogramme observé et un sismogramme synthétique, souvent calculé par sommations de modes propres.

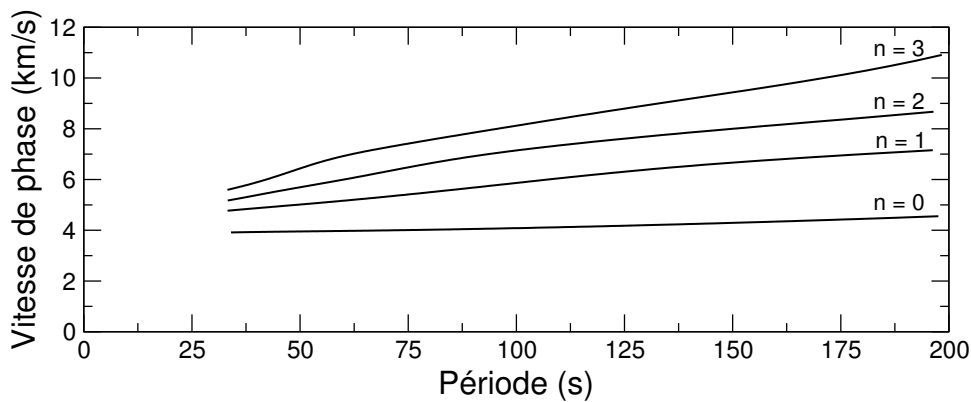


FIG. 2.8 : Vitesses de phase théoriques du mode fondamental et des trois premiers modes harmoniques des ondes de Rayleigh. Les calculs ont été réalisés pour le PREM (Dzie-wonski & Anderson, 1981) en utilisant MINEOS.

2.5 Paramétrisation des modèles

2.5.1 Différents types de paramétrisation

La paramétrisation est une contrainte importante sur le résultat de l'inversion. Dans le contexte de recherches de modèles entièrement non-linéaire, le nombre de paramètres

descriptifs des modèles est limité à une centaine, selon le problème posé (voir section 1.2.4). C'est pour cette raison qu'une paramétrisation en couches est souvent employée. Par exemple, l'étude de [Khan et al. \(2011\)](#), dont l'objectif est assez similaire au notre, à savoir évaluer les structures thermochimique et d'anisotropie radiale dans le manteau à partir des ondes de surface, ont utilisé des couches de tailles uniformes pour décrire la température et l'anisotropie radiale en fonction de la profondeur. La taille de ces couches est adaptée à la résolution des ondes de surface, qui diminue avec la profondeur. En particulier, l'épaisseur des couches est de 50 km entre la surface et 700 km, puis elle augmente à 100 km au-dessous de la zone de transition. Une paramétrisation en couches est également souvent employée pour décrire les vitesses sismiques des couches superficielles (*e. g.* [Shapiro & Ritzwoller, 2002](#); [Khan et al., 2009, 2011](#)), mais cette fois-ci en rajoutant un paramètre supplémentaire qui est la profondeur du Moho.

Avec de telles paramétrisations, le risque est de ne pas retrouver les variations des paramètres physiques, par exemple une discontinuité sismique, à la bonne profondeur, puisque les limites des couches sont fixées. Une approche originale appliquée par [Bodin et al. \(2012\)](#), afin de décrire V_S dans les 60 premiers kilomètres de la Terre, est d'utiliser des nuclei de Voronoï, dont les positions sont variables en profondeur et constituent les paramètres du modèle. La limite entre deux couches est équidistante entre deux nuclei adjacents. Cette approche permet de ne pas imposer la profondeur des discontinuités, mais la valeur de V_S à l'intérieur d'une même couche demeure constante.

Utiliser des paramètres qui varient selon une fonction avec la profondeur paraît plus approprié pour représenter les propriétés physiques de la Terre, et en particulier celles de la zone de transition. En effet, celle-ci comprend plusieurs gradients sismiques, en plus des principales discontinuités liées aux transformations minéralogiques (voir figure 1.7). De plus, cette technique présente l'avantage de limiter le nombre de paramètres. [Hauser et al. \(2011\)](#) ont paramétré les profils de vitesses sismiques des premiers 250 kilomètres de la Terre avec des points dont la profondeur est variable en profondeur. Des profils sismiques continus sont obtenus en interpolant linéairement les points. La transition lithosphère-asthénosphère est imposée sous la forme d'une discontinuité, mais sa profondeur est variable.

En tomographie sismique, les courbes B-splines, qui correspondent à des fonctions polynomiales par morceaux, sont couramment utilisées pour décrire les variations des paramètres sismiques en fonction de la profondeur (*e. g.* [Masters et al., 1996](#); [Ritsema et al., 2004](#); [Panning & Romanowicz, 2006](#); [Kustowski et al., 2008](#); [Nettles & Dziewonski, 2008](#)). Seules quelques études probabilistes les ont employées pour représenter les propriétés sismiques du manteau et de la croûte ([Shapiro & Ritzwoller, 2002](#); [Visser et al., 2008b](#); [Shen et al., 2013](#)). De manière générale, les courbes sont décrites par des points

de contrôle et un vecteur nœud¹. Les vecteurs construits avec les points de contrôle sont pondérés par des fonctions B-splines. Pour obtenir plus de détails sur les expressions des fonctions B-splines, le lecteur peut se référer par exemple à [Demengel & Pouget \(1998\)](#). Ces fonctions sont sommées avec les points de contrôle pour obtenir la courbe. L'inconvénient des B-splines est leur complexité. Il est difficile de gérer les points et les nœuds en même temps, c'est pourquoi en général, seuls les points de contrôle varient durant le processus d'inversion et les nœuds sont souvent fixés à une profondeur donnée. Des tests de résolution, notamment celui de [Backus & Gilbert \(1968\)](#), sont réalisés au préalable pour déterminer la profondeur et le nombre de fonctions de base. Les B-splines sont en général répartis de façon plus dense dans le manteau supérieur comparé à la zone de transition et au manteau inférieur, en fonction de la résolution des ondes de surface avec la profondeur (figure 2.9).

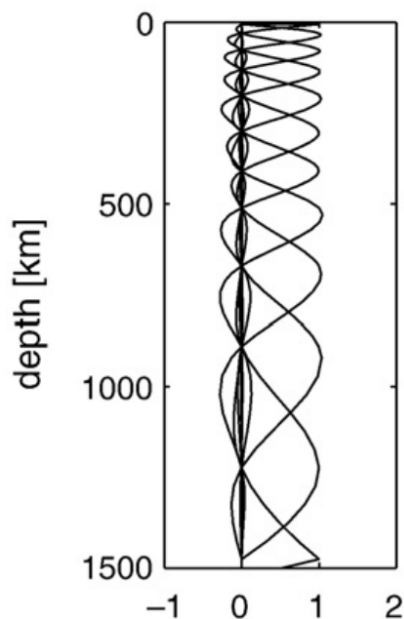


FIG. 2.9 : Représentation des 12 splines utilisées par [Visser et al. \(2008b\)](#) dans le cadre d'une inversion probabiliste de l'anisotropie radiale dans le manteau.

Dans notre cas, l'idéal serait de trouver une paramétrisation des profils qui requière peu de paramètres, dont la profondeur des variations de ces paramètres ne soit pas fixée *a priori*, et qui puisse produire des profils à la fois lisses et rugueux. L'ensemble de ces critères peut être rempli par les *courbes de Bézier cubiques*. Notons que celles-ci correspondent à des B-splines particulières.

¹Le vecteur nœud correspond à une suite non décroissante de nombres réels.

2.5.2 Les courbes de Bézier cubiques

Les profils de température et d'anisotropie dans la zone de transition, ainsi que le profil de vitesse de cisaillement dans le manteau supérieur et la croûte, sont paramétrés en fonction de la profondeur avec des courbes de Bézier cubiques de classe C^1 (Bézier, 1966, 1967). À l'origine, ces courbes ont été proposées en 1962 au sein du bureau d'étude du constructeur automobile Renault par l'ingénieur français Pierre Bézier, afin de concevoir les différents éléments d'une voiture. Aujourd'hui, elles sont devenues un des modèles de base dans de nombreux logiciels de conception assisté par ordinateur (C. A. O.), utilisés dans des domaines divers d'applications tels que : la mécanique, l'aéronautique, le design, etc.

Cette paramétrisation offre plusieurs avantages :

1. Avec peu de paramètres, les courbes de Bézier sont capables de décrire à la fois des variations lisses, comme par exemple des gradients de température, des solutions solides de minéraux, et des variations brutales, qui peuvent être associées à de fines couches limites thermiques ou bien à des transitions de phase minéralogiques.
2. Les paramètres des courbes n'ont pas besoin d'être discrétisés de façon régulière et il n'est pas nécessaire de faire des hypothèses sur les épaisseurs des couches ou bien sur la localisation des discontinuités.
3. Pendant le processus d'inversion, il est possible d'adapter le nombre de courbes ainsi que le degré de courbure nécessaires pour décrire une structure donnée.

La paramétrisation d'un profil de température, d'anisotropie ou de vitesse de cisaillement est schématisée sur la figure 2.10. Une courbe de Bézier cubique $B_j(t)$, dont l'indice j représente le numéro de la courbe composite selon la profondeur, est décrite de la manière suivante :

$$B_j(t) = \sum_{i=0}^3 b_{i,3}(t) \mathcal{P}_{ji}, \quad t \in [0, 1], j = [1, \dots, N], \quad (2.73)$$

où $\mathcal{P}_{ji}, i = [0, \dots, 3]$ sont les points de contrôle de la courbe et où les polynômes

$$b_{i,k}(t) = \binom{k}{i} t^i (1-t)^{k-i}, \quad i = [0, \dots, k] \quad (2.74)$$

sont les polynômes de Bernstein de degré k . D'une manière analogue au modèle des B-splines, les vecteurs construits avec les points de contrôle sont pondérés par les polynômes de Bernstein, lesquels remplacent les fonctions B-splines. Au niveau des extrémités $\mathcal{P}_{j0}$ et $\mathcal{P}_{j3}$, la courbure est définie par la norme des vecteurs tangents locaux $\overrightarrow{\mathcal{P}_{j0}\mathcal{P}_{j1}}$ et $\overrightarrow{\mathcal{P}_{j2}\mathcal{P}_{j3}}$, respectivement. Pour obtenir une suite de courbes de Bézier continue, sans rupture de

penne, les dérivées au-dessus et au-dessous d'un point commun à deux courbes de Bézier consécutives sont identiques. Il s'agit de courbes de classe C^1 , c'est-à-dire que la direction et la norme du vecteur dérivée sont continues.

L'idée des courbes de Bézier repose sur l'utilisation de points de contrôle et non de points d'interpolation. En général, la courbe ne passe ni par $\mathcal{P}_{j1}$ ni par $\mathcal{P}_{j2}$. Les points de définition $\mathcal{P}_{j1}$ et $\mathcal{P}_{j2}$ sont simplement présents pour donner une information de direction. Dans la suite du manuscrit, nous appellerons *points de Bézier* uniquement les points placés sur la courbe, qui correspondent aux paramètres de l'inversion (par exemple $\mathcal{P}_{j0}$ et $\mathcal{P}_{j3}$). Les autres points seront qualifiés de *points de définition* ou de *points intermédiaires* ($\mathcal{P}_{j1}$ et $\mathcal{P}_{j2}$). La manière dont ils sont calculés est détaillée dans la section 4.4.1.1.

Une caractéristique des courbes de Bézier est que le déplacement d'un point de contrôle sur une courbe de Bézier a des répercussions au-delà de la zone de localité du point de contrôle. C'est pour cette raison que des courbes composites sont utilisées, afin de ne provoquer qu'une variation locale de la courbe, tout en laissant la possibilité d'avoir des points d'inflexion. Comme illustré sur la figure 2.11a, lorsque la position d'un point

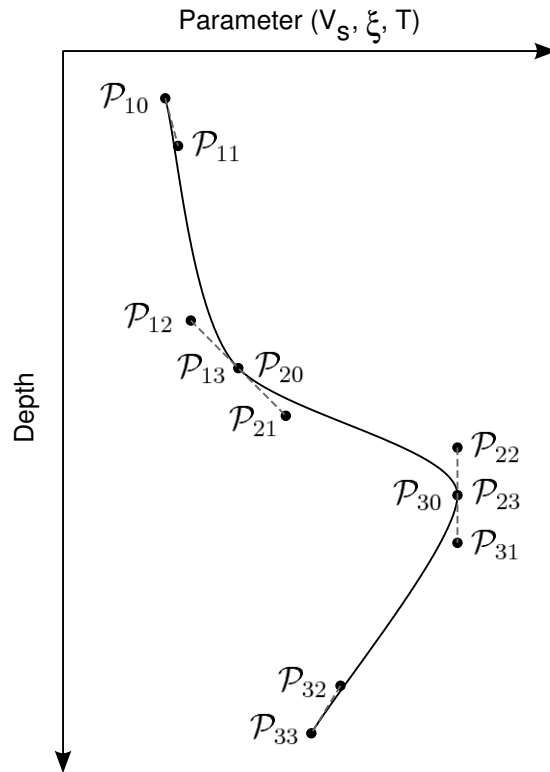


FIG. 2.10 : Représentation schématique de la paramétrisation des modèles en utilisant trois courbes de Bézier cubiques de classe C^1 . Quatre points de contrôle ($\mathcal{P}_{j0}$, $\mathcal{P}_{j1}$, $\mathcal{P}_{j2}$ et $\mathcal{P}_{j3}$) définissent chaque polynôme et ses dérivées. Pour obtenir une courbe totale continue, les dérivées sont identiques à chaque point commun à deux courbes de Bézier consécutives.

de Bézier est modifiée, seules les deux courbes de part et d'autre de ce point subissent une variation. Il est donc facile de modifier la courbe puisqu'il ne faut changer que les points de contrôle, qui sont relativement peu nombreux. La distance entre $\mathcal{P}_{j0}$ et $\mathcal{P}_{j1}$ détermine la « longueur » du déplacement dans la direction de $\mathcal{P}_{j1}$ avant de tourner vers $\mathcal{P}_{j3}$. Plus la norme des vecteurs tangents locaux est grande, plus le profil radial obtenu est lisse et courbé (figure 2.11b). À l'inverse, une norme proche de zéro revient à réaliser une interpolation linéaire. Nous verrons dans la section 4.4.1.1 que cette dernière propriété peut être exploitée pour ne pas avoir à fixer le degré de lissage des profils de température échantillonnés.

Les courbes de Bézier sont décrites selon la variable $t \in [0, 1]$, et non directement en fonction de la profondeur. L'équation (2.73) pour une courbe de Bézier cubique $B_j(t)$ donnée peut être développée de façon à s'écrire sous la forme d'une équation polynomiale du troisième degré, avec pour inconnue la valeur t . Cette équation est résolue grâce à la méthode de Cardan pour obtenir la valeur de la température, de l'anisotropie ou de la vitesse de cisaillement à une profondeur donnée.

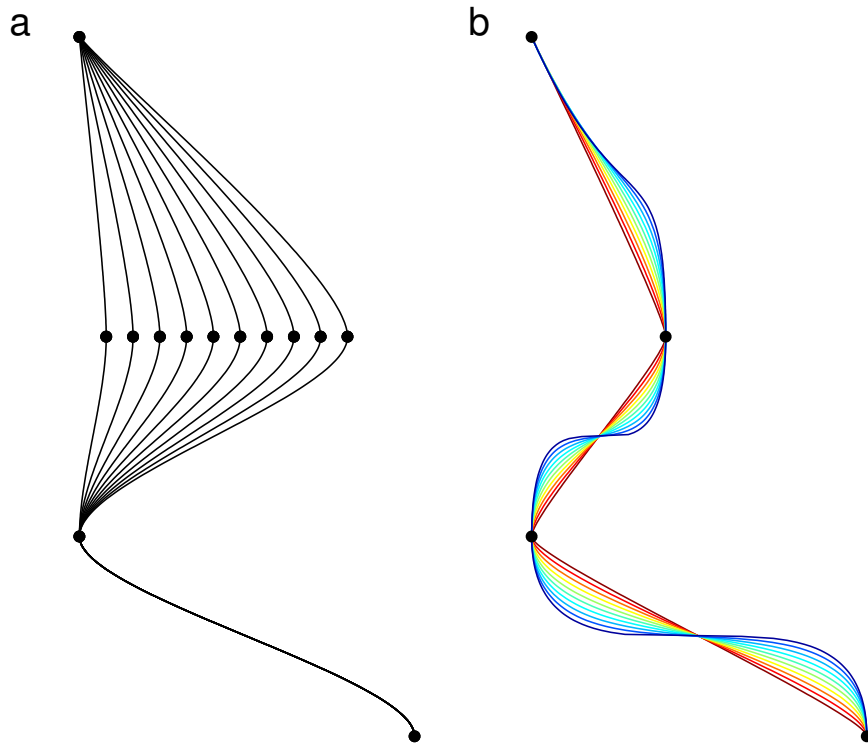


FIG. 2.11 : (a) Modification de la forme d'un profil selon la variation de la position d'un point de Bézier. Trois courbes cubiques sont représentées. Seules les deux courbes adjacentes sont modifiées. (b) Modification de la courbure du profil en fonction de la valeur de la norme des vecteurs tangents locaux. Du rouge au bleu, la norme est de plus en plus grande.

2.6 Ce qu'il faut retenir

Dans ce chapitre, les différentes étapes du problème direct ont été exposées. Celui-ci permet de calculer les vitesses de phase à partir des paramètres physiques que sont la température et l'anisotropie radiale, pour une composition minéralogique donnée. L'application de cette méthode pour déterminer la structure physique de la zone de transition est différente de celle traditionnellement employée en tomographie sismique, puisque les variations de V_S , V_P et ρ sont naturellement liées. Les ondes de surface sont alors sensibles aux variations des différents paramètres sismiques de façon simultanée. Les courbes de dispersion sont calculées par la méthode classique de sommations de modes propres jusqu'à l'ordre 3, ce qui assure une bonne sensibilité des ondes de surface à l'ensemble de la zone de transition. Les profils de température et d'anisotropie sont paramétrés grâce à des points de Bézier reliés par des courbes cubiques polynomiales. La paramétrisation en termes de points de Bézier permet de s'affranchir d'une structure en couches dont les épaisseurs seraient fixées *a priori*. En effet, les points de Bézier ont la possibilité de se placer directement aux endroits où il existe des variations importantes des propriétés physiques dans la zone de transition. Puisqu'il n'est pas possible de directement convertir des données sismiques en une unique structure physique, le problème inverse consiste à déterminer quels sont les modèles susceptibles d'expliquer les données, compte-tenu des incertitudes associées. Il s'agit de l'objet du chapitre suivant.

Chapitre 3

Inversion probabiliste

3.1 Introduction

3.1.1 Problème inverse

La plupart de nos connaissances sur l'intérieur de la Terre proviennent de l'analyse de données à la surface de la planète, les *observations*. Les structures internes majeures, telles que la croûte, le manteau, le noyau externe, la graine, et en particulier la zone de transition, ont été détectées par des mesures indirectes réalisées à la surface du globe. Le fait d'utiliser les observations d'un phénomène, comme par exemple le sismogramme issu d'un tremblement de terre, pour obtenir de l'information sur les propriétés intrinsèques à la Terre profonde, est appelé *inversion*.

En pratique, résoudre un problème inverse est délicat pour différentes raisons. L'une de ces raisons est la non-unicité. En effet, plusieurs valeurs des paramètres peuvent correspondre aux mêmes observations. De plus, l'espace des paramètres peut être extrêmement vaste, c'est pourquoi des procédures de régularisation sont souvent employées avant même de rechercher la solution (voir section 1.2.3). Le choix de la méthode de résolution d'un problème inverse particulier est lié à la nature du problème direct ainsi qu'à l'incertitude sur les données (ou bruit). Si le problème direct est linéaire et le bruit assimilé à une distribution gaussienne, le problème inverse peut être résolu grâce à la théorie des moindres carrés (e. g. [Tarantola & Valette, 1982](#)). Mais si le problème est non-linéaire et le bruit de distribution quelconque, les méthodes algébriques classiques ne peuvent plus être employées.

Les méthodes classiques d'inversion sont déterministes. Elles reposent sur le fait qu'il existe une valeur vraie des paramètres, et donc une solution unique au problème posé. L'inversion probabiliste est basée sur une tout autre vision du problème. Dans certains cas, en plus des erreurs de mesure sur les données, le modèle contient des erreurs de

modélisation, dues à une mauvaise connaissance du système physique. Dans ce cas, pour une valeur donnée des paramètres, on ne peut plus prédire une valeur exacte des observations mais plutôt une probabilité que les observations aient une valeur comprise dans un certain domaine. Puisqu'il existe des erreurs, les observations (et les paramètres correspondant à ces mêmes observations) sont considérées comme étant des variables aléatoires. Le résultat obtenu pour un jeu de paramètres n'est donc qu'une réalisation particulière des observations.

Une autre particularité importante de l'inversion probabiliste est qu'elle permet d'incorporer une information concernant les paramètres du modèle, indépendante des données. Cette information est fournie sous la forme d'une densité de probabilité sur les paramètres. Elle est appelée densité de probabilité *a priori*. Dans le cadre probabiliste, une inversion consiste donc à passer d'une densité de probabilité *a priori* sur les paramètres, qui représente l'état de connaissance des paramètres avant d'avoir pris en compte les informations véhiculées par les données, à une densité dite *a posteriori* qui tient compte des données. Ainsi, l'inversion peut être considérée comme une amélioration du degré de connaissance du système.

3.1.2 Stratégie adoptée

L'objectif de ce travail est de déterminer les paramètres qui gouvernent la structure sismique de la zone de transition, et ce à partir des courbes de dispersion des ondes de surface. Les paramètres sont les points de contrôle des courbes de Bézier cubiques, c'est-à-dire la température T et l'anisotropie ξ , ainsi que les profondeurs de ces points particuliers. Une exploration systématique s'avère extrêmement coûteuse en temps de calcul, et quasiment impossible à mettre en œuvre. C'est pourquoi une méthode alternative doit être envisagée. La démarche adoptée dans la suite de cette thèse a consisté à déterminer les différents modèles de température et d'anisotropie radiale capables de rendre compte des courbes de dispersion. Il s'agit d'une inversion de type Monte Carlo utilisant les propriétés des chaînes de Markov, à partir de laquelle les paramètres du modèle sont déduits par échantillonnage de la distribution de probabilité *a posteriori* de l'espace des paramètres, au moyen d'un algorithme de type Metropolis-Hastings (*e. g.* [Metropolis et al., 1953](#); [Hastings, 1970](#)). Cette méthode consiste à estimer quelles configurations des paramètres sont les plus probables compte-tenu des données observées.

Ce chapitre rappelle les principes fondamentaux de l'inversion bayésienne, basée sur la méthode de Monte Carlo par chaînes de Markov, qui est détaillée par exemple dans [Mosegaard & Tarantola \(1995\)](#), [Mosegaard \(1998\)](#) et [Tarantola \(2005\)](#). Cette méthode a déjà été employée pour contraindre la structure du manteau terrestre en utilisant des données géophysiques (*e. g.* [Khan et al., 2006, 2009](#); [Verhoeven et al., 2009](#); [Khan &](#)

[Shankland, 2012](#)). Avant de présenter l'algorithme d'échantillonnage, nous allons faire un bref rappel du formalisme mathématique concernant la méthode bayésienne.

3.2 Inférence bayésienne

3.2.1 Notion de densité de probabilité marginale

Pour une variable aléatoire réelle continue p , la probabilité que p prenne une valeur x dans l'intervalle $a \leq p \leq b$ est :

$$P(a \leq p \leq b) = \int_a^b P(p = x) dx, \quad (3.1)$$

où $P(p = x)$ est appelée densité de probabilité associée à la variable aléatoire p . Pour plusieurs variables aléatoires $p_1, p_2, \dots, p_m$, on peut généraliser cette expression en définissant une densité de probabilité jointe $P(p_1 = x_1, p_2 = x_2, \dots, p_m = x_m)$. À partir de cette densité de probabilité, il est possible de calculer la probabilité jointe que chaque variable aléatoire p_i se trouve dans un intervalle particulier $a_i \leq p_i \leq b_i$:

$$P(a_1 \leq p_1 \leq b_1 \text{ et } a_2 \leq p_2 \leq b_2 \text{ et } \dots a_m \leq p_m \leq b_m) = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_m}^{b_m} dx_m P(p_1 = x_1, p_2 = x_2, \dots, p_m = x_m). \quad (3.2)$$

À partir d'une densité de probabilité jointe, l'expression de la densité de probabilité d'une variable aléatoire particulière peut être obtenue en intégrant sur toutes les valeurs possibles des autres variables aléatoires :

$$P(p_i = x_i) = \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{i-1} \int_{-\infty}^{\infty} dx_{i+1} \dots \int_{-\infty}^{\infty} dx_m P(p_1 = x_1, p_2 = x_2, \dots, p_m = x_m). \quad (3.3)$$

Le fait d'intégrer sur certaines variables aléatoires est appelé *marginalisation* et la densité de probabilité $P(p_i = x_i)$ correspond à la *densité de probabilité marginale*.

3.2.2 Théorème de Bayes

Le passage de la densité *a priori* à la densité *a posteriori* est effectué par le *théorème de Bayes*. Ce théorème étant au cœur de l'inversion probabiliste, cette dernière est également appelée *inversion bayésienne* ou encore *inférence bayésienne*.

On note $\mathbf{p} = (p_1, p_2, \dots, p_m)$ et $\mathbf{d} = (d_1, d_2, \dots, d_N)$ les paramètres du modèle et les données sismologiques, respectivement. m est ici le nombre de paramètres et N le nombre de données. Les données sont contraintes par la relation

$$\mathbf{d} = A(\mathbf{p}), \quad (3.4)$$

où l'opérateur A représente le problème direct présenté dans le chapitre précédent, qui comprend les calculs thermodynamiques ainsi que le calcul des vitesses de phase synthétiques.

En général, les observations $\mathbf{d}$ diffèrent des prédictions $A(\mathbf{p})$ à cause des erreurs de mesure. Si la distribution de probabilité des erreurs de mesure est connue, il est possible de calculer la probabilité que les données $\mathbf{d}$ soient situées dans un certain domaine, sachant que les paramètres $\mathbf{p}$ ont une valeur donnée. Cette probabilité $P(\mathbf{d}|\mathbf{p})$ est appelée fonction de vraisemblance (ou *likelihood*, en anglais). Dans la suite du manuscrit, la vraisemblance est notée $L(\mathbf{d}|\mathbf{p})$.

La solution du problème inverse consiste à calculer la probabilité *a posteriori* $P(\mathbf{p}|\mathbf{d})$ que les paramètres soient dans une configuration $\mathbf{p}$ étant donnée que les données sont dans une configuration $\mathbf{d}$. L'espace des paramètres est échantillonné selon cette probabilité $P(\mathbf{p}|\mathbf{d})$. Le théorème de Bayes relie la distribution *a priori* $P(\mathbf{p})$ et la distribution *a posteriori* $P(\mathbf{p}|\mathbf{d})$ de la manière suivante :

$$P(\mathbf{p}|\mathbf{d}) = \frac{L(\mathbf{d}|\mathbf{p})P(\mathbf{p})}{\sum_{\mathbf{p} \in \mathcal{M}} L(\mathbf{d}|\mathbf{p})P(\mathbf{p})}, \quad (3.5)$$

$\mathcal{M}$ correspondant à l'ensemble des configurations de l'espace des paramètres. Cette relation permet de passer de la densité *a priori*, qui reflète l'état de connaissance initial, à la densité *a posteriori*, qui réexprime cette connaissance après avoir testé un certain nombre de prédictions. La somme au niveau du dénominateur est réalisée sur toutes les configurations $\mathbf{p}$ de $\mathcal{M}$.

Il existe un certain nombre de règles appelées priors ($P(\mathbf{p})$) qui définissent la distribution *a priori*, c'est-à-dire l'ensemble des modèles possibles. L'utilisation de priors permet de réduire l'espace des configurations et définit notre état de connaissance sur les paramètres physiques. Par exemple, le profil de température dans la zone de transition n'est pas connu mais des bornes supérieure et inférieure peuvent être définies. Seuls les modèles qui appartiennent à la distribution *a priori* sont testés et comparés aux données.

La vraisemblance s'exprime à partir de la fonction coût (ou misfit) $\mathcal{S}(\mathbf{d}, A(\mathbf{p}))$ qui mesure l'écart entre les données observées $\mathbf{d}$ et les données synthétiques $A(\mathbf{p})$:

$$L(\mathbf{d}|\mathbf{p}) \propto \exp[-\mathcal{S}(\mathbf{d}, A(\mathbf{p}))]. \quad (3.6)$$

Plus l'écart entre les données observées $\mathbf{d}$ et les données calculées $A(\mathbf{p})$ est faible, plus la vraisemblance est importante. Ainsi, la résolution du problème inverse par la méthode bayésienne ne nécessite que le calcul du problème direct et de la fonction coût. Le calcul du misfit prend en compte l'incertitude assimilée aux données. L'incertitude sur les données détermine directement la forme de la distribution *a posteriori* et par là même celle des échantillons générés à partir de cette distribution. Pour estimer l'écart entre les valeurs observées et calculées, les normes les plus couramment utilisées sont les normes L_1 et L_2 dont l'expression est :

$$\mathcal{S}(\mathbf{d}, A(\mathbf{p})) = \sum_{i=1}^N \frac{|d_i - A(\mathbf{d})_i|^l}{\sigma_i^l}, \quad (3.7)$$

où d_i et $A_i(\mathbf{p})$ correspondent respectivement au $i^{\text{ème}}$ composant du vecteur des données observées $\mathbf{d}$ et calculées $A(\mathbf{d})_i$. σ_i est l'incertitude associée à la $i^{\text{ème}}$ observation. L'exposant $l = 1$ et $l = 2$ correspondent aux normes L_1 et L_2 , respectivement.

3.2.3 Distribution *a posteriori* des paramètres

Nous avons vu que la formule de Bayes procure une estimation de la probabilité *a posteriori* (ou marginale) d'une configuration donnée de paramètres $P(p_i = x|\mathbf{d})$ en tenant compte des données. Cette dernière est obtenue en faisant la somme de toutes les probabilités que les configurations $\mathbf{p}$ pour le $i^{\text{ème}}$ paramètre prennent la valeur x :

$$P(p_i = x|\mathbf{d}) = \frac{\sum_{\mathbf{p} \in \mathcal{M}_x^i} L(\mathbf{d}|\mathbf{p})P(\mathbf{p})}{\sum_{\mathbf{p} \in \mathcal{M}} L(\mathbf{d}|\mathbf{p})P(\mathbf{p})} \quad (3.8)$$

où $\mathcal{M}_x^i$ représente toutes les configurations dont le $i^{\text{ème}}$ composant est x . La distribution de probabilité correspondant à un paramètre donné peut ainsi être construite à partir de la probabilité marginale calculée pour chacune des valeurs qu'il peut prendre.

L'évaluation de la probabilité marginale est cependant difficile à calculer en pratique car le nombre de configurations est égal à $\prod_{i=1}^m |\mathcal{M}_i|$, où $|\mathcal{M}_i|$ est le nombre de valeurs que le paramètre i peut prendre. Le calcul explicite des probabilités associées à chaque paramètre à partir des équations (3.5) et (3.8) est impossible à réaliser dès qu'augmente le nombre de paramètres. Cependant, il est possible d'en estimer une approximation, à partir de l'étude des modèles échantillonnés par l'algorithme. En effet, une approximation de la probabilité $P(p_i = x)$ que le $i^{\text{ème}}$ paramètre p_i prenne la valeur x est donnée par le rapport du nombre de réalisations de l'événement $p_i = x$ sur le nombre total de configurations visitées au cours de la marche aléatoire. Si ce rapport est calculé pour toutes les valeurs qu'est susceptible de prendre un paramètre donné, il est possible de

construire la distribution de probabilité marginale, qui correspond à une approximation de la distribution *a posteriori* de ce paramètre. La figure 3.1 permet de visualiser une densité de probabilité jointe pour deux paramètres. Les deux probabilités marginales associées sont également représentées.

Dans la pratique, une méthode d'échantillonnage stochastique particulière de l'espace des paramètres est utilisée, appelée méthode de Monte Carlo par chaînes de Markov. Celle-ci est l'objet de la section suivante.

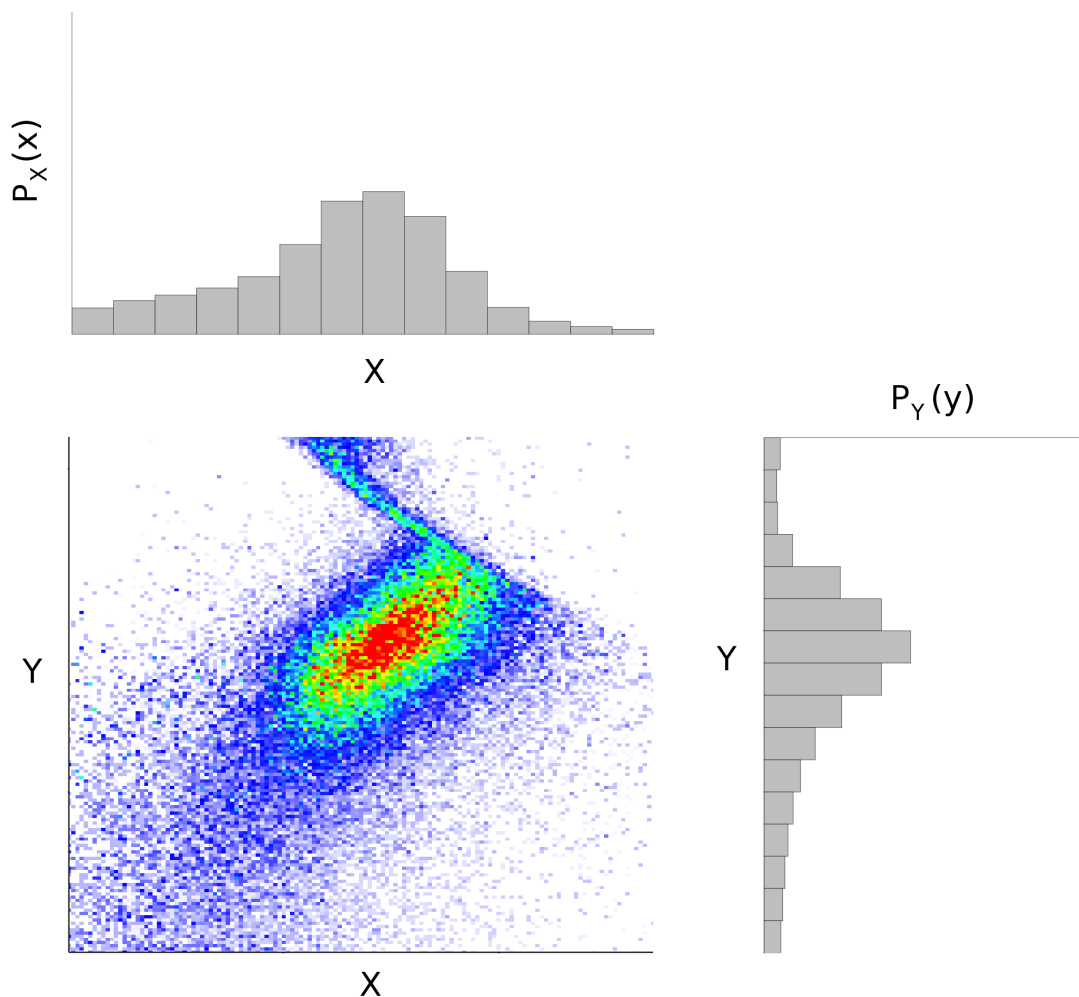


FIG. 3.1 : Graphiques représentant la densité de probabilité jointe sur les variables aléatoires X et Y , ainsi que les deux probabilités marginales associées $P_X(x)$ et $P_Y(y)$.

3.3 Méthode de Monte Carlo par chaînes de Markov (McMC)

3.3.1 Chaînes de Markov

3.3.1.1 Principe

Parmi les nombreuses méthodes de Monte Carlo qui échantillonnent aléatoirement l'espace des paramètres pour estimer la valeur des probabilités marginales, une méthode de Monte Carlo par chaînes de Markov (l'abréviation McMC pour *Markov chain Monte Carlo* est couramment employée) a été choisie.

Cette méthode utilise une marche aléatoire pour échantillonner l'espace des paramètres (l'espace des modèles $\mathcal{M}$) selon la probabilité $P(\mathbf{p}|\mathbf{d})$. Ainsi, en particulier, la valeur moyenne des paramètres est approximée par la valeur moyenne des échantillons.

La méthode McMC produit une suite de nombres aléatoires dépendants, appelée chaîne de Markov. Si la chaîne satisfait un certains nombre de critères, elle converge vers une distribution, appelée distribution d'équilibre.

3.3.1.2 Définitions et propriétés

Définition d'une chaîne de Markov

On suppose une suite de variables aléatoires indépendantes $X^{(1)}, X^{(2)}, \dots, X^{(\infty)}$. L'espace des états X associé est noté $\Omega = \{x_1, \dots, x_k\}$. Ω est considéré comme étant fini et discret. Un processus aléatoire au cours duquel sont successivement visités les états $(X^{(1)}, X^{(2)}, \dots, X^{(t)}, X^{(t+1)}, \dots)$ constitue une chaîne de Markov s'il satisfait la condition :

$$P(X^{(t+1)} = x_j | X^{(t)} = x_i, X^{(t-1)} = x_{i_{t-1}}, \dots, X^{(1)} = x_{i_1}) = P(X^{(t+1)} = x_j | X^{(t)} = x_i). \quad (3.9)$$

Cette condition implique que l'état visité au pas $t+1$ ne doit dépendre que de l'état au pas t , et d'aucune autre étape antérieure.

Une simulation particulière de la chaîne de Markov est appelée *réalisation*. Si la chaîne est simulée plusieurs fois, elle engendre des réalisations différentes étant donné que la transition d'un état à un autre est donné en terme de probabilité.

Homogénéité

Une chaîne de Markov est déterminée par une distribution de probabilité initiale $P(X^{(1)} = x_{i_1})$ et par une probabilité de transition $P(X^{(t+1)} = x_j | X^{(t)} = x_i)$. Une chaîne de Markov est dite *homogène* si la probabilité $P(X^{(t+1)} = x_j | X^{(t)} = x_i)$ est identique à chaque étape (chaque pas) et est indépendante de l'indice t . La probabilité de transition est alors associée à la matrice $\mathbf{P}$ dont l'élément i, j représente la probabilité de passer de l'état x_i à l'état x_j :

$$P_{i,j} = P(X^{(t+1)} = x_j | X^{(t)} = x_i). \quad (3.10)$$

Et les probabilités de transition satisfont la condition de normalisation :

$$\sum_{j \in \Omega} P_{i,j} = 1. \quad (3.11)$$

Distribution d'équilibre

Si la chaîne de Markov visite un grand nombre d'états, c'est-à-dire si $t + 1 \rightarrow \infty$, les valeurs de $X^{(t+1)}$ vont continuer à fluctuer indéfiniment et il n'est pas possible de trouver une valeur limite fixée pour $X^{(t+1)}$. Soient les vecteurs $\pi^{(1)}, \dots, \pi^{(t+1)}$ décrivant la distribution de probabilité associée aux pas 1, ..., $t + 1$ de la chaîne. Les éléments du vecteur $\pi^{(t+1)}$ sont

$$\pi^{(t+1)} = (\pi_1^{(t+1)}, \dots, \pi_k^{(t+1)}) = (P(X^{(t+1)} = x_1), \dots, P(X^{(t+1)} = x_k)). \quad (3.12)$$

D'après la définition de la matrice de transition :

$$\pi_j^{(t+1)} = \pi_i^{(t)} P_{i,j}. \quad (3.13)$$

Le vecteur π est appelé *distribution d'équilibre* (ou bien distribution stationnaire) s'il satisfait la condition :

$$\pi = \pi \mathbf{P}. \quad (3.14)$$

Par définition, si $\pi^{(t)} \rightarrow \pi$ quand $t \rightarrow \infty$ quelle que soit la distribution de probabilité initiale $\pi^{(1)}$, alors π est la distribution d'équilibre de la chaîne. Autrement dit, partant d'une distribution initiale quelconque, la chaîne de Markov doit converger vers une distribution d'équilibre si on la fait tourner suffisamment longtemps. La distribution sera d'autant plus proche de la distribution stationnaire recherchée que le nombre d'itérations est grand. Si à un moment donné, correspondant à l'état t , la distribution de la chaîne de Markov est égale à la distribution d'équilibre, alors toutes les distributions suivantes seront égales à π . Il s'agit d'un résultat fondamental de la théorie des chaînes de Markov. Mais ce comportement de convergence n'est valable que pour les chaînes de Markov qui

satisfont les propriétés spécifiques d'irréductibilité et d'apériodicité. On peut démontrer que, pour toute chaîne de Markov irréductible et apériodique, il existe une unique distribution d'équilibre (théorème d'existence et d'unicité).

Irréductibilité

Une chaîne de Markov est dite *irréductible* si, à partir de n'importe quel état x_i de l'espace Ω , il existe un chemin de probabilité non nul de la matrice de transition qui permet d'atteindre tout autre état x_j de cet espace, et réciproquement. En d'autres termes, une chaîne de Markov est irréductible si tous les états de l'espace peuvent être atteints à partir de tous les autres états. Théoriquement, tous les états de l'espace Ω peuvent être visités, moyennant un nombre de pas (d'itérations) suffisant. Mais le caractère irréductible de la chaîne n'est que théorique. En pratique, surtout lorsque la chaîne a convergé, certains états sont si peu accessibles que la probabilité qu'ils soient acceptés est quasiment nulle.

Apériodicité

L'*apériodicité* est une caractéristique de la chaîne de Markov qui atteste l'existence d'un état donné tel que pour tous les états postérieurs, la probabilité de rester dans le même état que le précédent est strictement positive. Une condition suffisante pour qu'une chaîne soit apériodique est que les éléments diagonaux de la matrice de transition soient non nuls :

$$P_{i,i} > 0, \forall i. \quad (3.15)$$

En revanche, une chaîne est dite *périodique* s'il existe des régions de l'espace des états qu'elle visite périodiquement.

Réversibilité

Pour qu'une chaîne soit considérée comme *réversible*, il ne suffit pas que celle-ci puisse être parcourue dans un sens ou dans l'autre. Il faut que la matrice de transition soit la même quel que soit le sens de parcours de la chaîne, de sorte que la condition suivante soit vérifiée :

$$P(X^{(t+1)} = x_j | X^{(t)} = x_i) = P(X^{(t)} = x_j | X^{(t+1)} = x_i). \quad (3.16)$$

Une condition suffisante de réversibilité de la chaîne est la condition de balance détaillée (ou condition de balance microscopique) ([Sambridge & Mosegaard, 2002](#)) :

$$\pi_i P_{i,j} = \pi_j P_{j,i}, \forall i, j \in \Omega. \quad (3.17)$$

Théorème d'ergodicité

Le *théorème d'ergodicité* stipule qu'une chaîne de Markov homogène, apériodique et irréductible, est indépendante de sa configuration et converge vers une distribution d'équilibre unique si et seulement si la condition de balance détaillée (équation 3.17) est établie. Ainsi, pour échantillonner l'espace des paramètres, il suffit de construire une chaîne de Markov dont l'état initial est tiré au sort et dont la distribution de probabilité d'équilibre est $P(\mathbf{p}|\mathbf{d})$.

3.3.2 L'algorithme de Metropolis-Hastings

Dans le contexte de l'inversion probabiliste, on cherche à déterminer de quelle façon, à partir d'une certaine densité de probabilité donnée (la densité *a posteriori*), construire une chaîne de Markov dont elle serait la densité d'équilibre. L'algorithme de Metropolis-Hastings (*e. g.* Metropolis et al., 1953; Hastings, 1970) permet de générer une chaîne de Markov qui converge vers une distribution donnée. Il s'agit de l'algorithme de base qui sera utilisé dans la suite de ce travail pour déduire la distribution de probabilité des paramètres du modèle à partir de courbes de dispersion des ondes de surface. Chaque nouveau modèle est choisi de façon à être dans le voisinage du modèle précédent. Pour déterminer la configuration des paramètres au pas suivant, l'algorithme de Metropolis-Hastings comprend trois étapes : un tirage au sort, le calcul du problème direct et de la fonction coût associée au modèle, et enfin une étape d'acceptation/rejet du modèle proposé selon une règle établie qui dépend de la valeur de la fonction coût. Après une exploration initiale, la valeur de l'écart aux observations (misfit) guide la recherche pour choisir un nouveau jeu de paramètres, et ainsi de suite.

Si le prior est uniforme à l'intérieur d'un intervalle, la probabilité marginale $P(p_i = x|\mathbf{d})$ est directement égale au rapport des vraisemblances :

$$P(p_i = x|\mathbf{d}) = \frac{\sum_{\mathbf{p} \in \mathcal{M}_x} L(\mathbf{d}|\mathbf{p})}{\sum_{\mathbf{p} \in \mathcal{M}} L(\mathbf{d}|\mathbf{p})}. \quad (3.18)$$

Dans ce cas, considérons qu'au pas t la chaîne de Markov vaille $X^{(t)} = x$. La valeur $X^{(t+1)} = x'$ de la configuration suivante est déterminée par la procédure suivante. À partir de la valeur actuelle $X^{(t)}$, un nouveau modèle $X^{(t+1)}$ est proposé à partir d'une densité de probabilité *a priori* $g(x'|x)$. Une règle est ensuite définie pour savoir si le modèle proposé doit être accepté ou rejeté. Si $\mathcal{S}(\mathbf{d}, P(X^{(t)})) > \mathcal{S}(\mathbf{d}, P(X^{(t+1)}))$, le modèle proposé est accepté et ajouté à l'ensemble des échantillons de la distribution *a posteriori*. En effet, la valeur du nouveau misfit est plus petite que la précédente, ce qui

signifie que le modèle proposé est plus susceptible de rendre compte des données. Si $\mathcal{S}(\mathbf{d}, P(X^{(t)})) < \mathcal{S}(\mathbf{d}, P(X^{(t+1)}))$, ce nouveau modèle n'est pas forcément rejeté bien qu'il dégrade la fonction coût. Le modèle est accepté si les rapports des vraisemblances, la probabilité $L(\mathbf{d}|X^{(t+1)})/L(\mathbf{d}|X^{(t)})$, est plus grande qu'un nombre aléatoire tiré dans une distribution uniforme entre 0 et 1. Si tel n'est pas le cas, le modèle est rejeté.

À partir de l'équation (3.6) la fonction de vraisemblance à l'étape $t + 1$ peut être écrite de la manière suivante :

$$L_{t+1}(x) = L(x) = k \cdot \exp[S_{t+1}(x)], \quad (3.19)$$

où k est une constante de normalisation et L la fonction de vraisemblance. Le candidat x' est accepté avec une probabilité :

$$\alpha(x'|x) = \min \left\{ 1, \frac{L(x')g(x|x')}{L(x)g(x'|x)} \right\}. \quad (3.20)$$

Notons que la densité de probabilité $L(x)$ dont on souhaite obtenir des échantillons n'apparaît que sous la forme de rapports $L(x')/L(x)$. Il n'est donc pas nécessaire de calculer l'expression exacte de la constante de normalisation de $L(x)$, indépendante de x , étant donné qu'elle se simplifie lors du calcul du rapport.

Plus le misfit est grand, plus il y a de chance que le modèle proposé soit rejeté par l'algorithme. Accepter une configuration avec une probabilité $\alpha < 1$ est avantageux car ce procédé permet d'éviter que la marche aléatoire reste prise au piège dans des minima locaux. Ce risque est très élevé lorsque l'on travaille avec des problèmes non-linéaires, comme en sismologie. Un autre avantage est que seuls des rapports de probabilité sont manipulés. En effet, le dénominateur de l'équation (3.5) n'a pas besoin d'être calculé.

La méthode McMC convient particulièrement aux problèmes dont la dépendance des paramètres par rapport aux données est fortement non-linéaire, puisque l'algorithme d'échantillonnage ne nécessite jamais de linéariser le problème. Si l'algorithme réalise un nombre suffisant d'itérations, les échantillons vont fournir une bonne approximation de la distribution *a posteriori* des paramètres du modèle, c'est-à-dire $P(\mathbf{p}|\mathbf{d})$. D'implémentation relativement simple, l'algorithme de Metropolis-Hastings constitue la base de nombreuses méthodes d'inversion non-linéaires.

3.3.3 Échantillonnage de l'espace des paramètres

Le nouveau modèle est proposé en perturbant aléatoirement le dernier modèle qui a été accepté. Par exemple, si la $t^{\text{ème}}$ valeur de la chaîne est $X^{(t)} = x$, alors la valeur du

pas suivant peut être défini comme étant égal à $x' = x + u$, où u est un nombre aléatoire obtenu provenant d'une densité de proposition $g(x'|x)$. En pratique, une distribution de probabilité gaussienne centrée sur la valeur actuelle du paramètre x est généralement utilisée :

$$g(x'|x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(x' - x)^2}{2\sigma^2} \right], \quad (3.21)$$

où σ est l'écart-type de la fonction gaussienne.

L'efficacité de la chaîne dépend de façon critique d'un bon choix de la valeur de l'écart-type σ . La variance de la distribution a un impact important sur la capacité de l'algorithme à générer une chaîne d'échantillons représentatifs en un nombre raisonnable d'itérations. Comme l'a montré [MacKay \(2003\)](#), la taille des perturbations n'affecte pas la solution (la distribution *a posteriori* obtenue), mais plutôt l'efficacité d'échantillonnage de l'algorithme. Si la largeur de la gaussienne est trop grande, alors la chaîne rejettera tous les modèles proposés pendant un grand nombre d'itérations. En revanche, si la largeur de la gaussienne est trop petite, la chaîne explorera des régions de grande probabilité très lentement, avec le risque d'être piégée dans un minimum local. Pour fixer des tailles de gaussienne adéquates, une approche est de réaliser des tests préliminaires. L'algorithme est testé avec différentes tailles de gaussienne. Seules celles qui produisent une acceptation (c'est-à-dire le nombre de modèles acceptés par l'algorithme divisé par le nombre total de modèles proposés) entre 25 et 50% peuvent être employées ([Gelman et al., 1996](#)). La valeur du pas avec lequel la chaîne de Markov se déplace est complètement dépendante du problème particulier étudié.

Cependant, cette procédure peut ne pas très bien fonctionner dans des situations où différents pas sont appropriés dans différents endroits de l'espace des paramètres. Des procédures selon lesquelles la taille du pas évolue au cours des itérations ont été proposées, mais elles ne semblent converger vers la solution que sous certaines conditions. Une approche est de changer la taille du pas en se basant sur les taux d'acceptation des itérations précédentes. Mais cette méthode est coûteuse en temps de calcul car l'algorithme passe trop de temps dans les régions de l'espace des paramètres où un petit pas est requis pour obtenir un taux d'acceptation suffisamment élevé. Une adaptation vers de plus petits pas au fur et à mesure des itérations a été testée par [Andrieu & Moulines \(2005\)](#). La méthode de Metropolis par *delayed rejection* ([Green & Mira, 2001](#)) implique différentes tailles de pas pour différentes régions de l'espace des paramètres. [Neal \(2005\)](#) a décrit une méthode selon laquelle la taille du pas est choisie de façon adaptative selon un petit nombre d'alternatives. Mais l'ensemble de ces techniques n'est pas facile à manipuler lorsqu'un grand nombre de paramètres est utilisé.

3.4 Ce qu'il faut retenir

Dans ce chapitre, la notion d'inversion probabiliste a été introduite. Celle-ci permet, à partir d'observations d'un phénomène physique, d'estimer une densité de probabilité des paramètres du modèle théorique qui décrit ce phénomène. La densité de probabilité est obtenue en se basant sur le théorème de Bayes, qui relie les densités de probabilité *a priori* et *a posteriori*, en prenant en compte les données. Dans la plupart des cas pratiques en géophysique, le calcul de la densité de probabilité *a posteriori* nécessite le calcul d'intégrales qui ne peuvent pas être résolues de façon analytique, ni par des méthodes d'intégration numérique classiques. C'est pourquoi l'utilisation de la méthode de Monte Carlo par chaînes de Markov traitée dans ce chapitre s'avère extrêmement adaptée aux problèmes non-linéaires car elle produit des échantillons de la densité de probabilité *a posteriori*. Parmi les méthodes de MCMC, l'algorithme de Metropolis-Hastings constitue une procédure qui engendre des chaînes de Markov dont la distribution d'équilibre est égale à la distribution de probabilité *a posteriori*. Cette méthode d'inversion nous paraît bien adaptée au problème posé, fortement non-linéaire, car elle ne nécessite pas de linéarisation, comme dans le cas de méthodes plus classiques telles que les moindres carrés.

Chapitre 4

Implémentation et application pratique à la zone de transition

4.1 Introduction

L'objectif de ce chapitre est de présenter en détail l'algorithme développé au cours de cette thèse pour l'inversion des vitesses de phase des ondes de surface. Nous avons choisi d'employer une méthode probabiliste, dont les fondements sont expliqués dans le chapitre 3. Cette méthode repose sur l'échantillonnage de nombreux modèles dont les paramètres sont les points de Bézier des profils radiaux représentant la température et l'anisotropie. Le calcul des données à partir des paramètres du modèle est exposé dans le chapitre 2.

Ce chapitre est organisé de la façon suivante. Dans un premier temps, les conditions aux limites du problème sont précisées. Notre objet d'étude est la zone de transition, mais le calcul des vitesses de phase nécessite que l'on introduise des valeurs pour les paramètres sismiques au-dessus et au-dessous de cette région. Les conditions de raccord des profils entre la zone de transition et les manteaux supérieur et inférieur sont décrites. Ensuite, les différents aspects du nouvel algorithme qui a été développé sont présentés. Les principales étapes de l'algorithme ainsi que les détails techniques sur la création de nouveaux modèles le long d'une chaîne de Markov et la manière d'atteindre le régime stationnaire sont précisés. Cet algorithme présente un fort potentiel pour la parallélisation, puisque chaque chaîne de Markov peut être placée sur un processeur particulier d'une machine parallèle, et ainsi explorer de façon indépendante l'espace des modèles.

L'algorithme de Metropolis-Hastings étant une méthode très générale, les choix particuliers que nous avons réalisés pour l'étude de la zone de transition sont exposés. Ces choix concernent les conditions *a priori* sur les paramètres des modèles tels que l'espace des paramètres autorisé, le nombre de paramètres et la manière de construire les

profils 1D. Les densités de probabilité *a priori* sont présentées et discutées. Ce chapitre s'achève en donnant des exemples de détails pratiques sur certains facteurs qui influent sur l'efficacité de l'algorithme, tels que la valeur de l'écart-type (σ) de la fonction gaussienne utilisée pour tirer au sort chaque nouveau paramètre (équation 3.21).

4.2 Conditions aux limites

4.2.1 La croûte et le manteau supérieur

Pour étudier le manteau terrestre, il n'est pas possible de faire abstraction des couches superficielles de la Terre puisque les ondes de surface de Love et de Rayleigh sont principalement sensibles à la croûte (*e. g.* Montagner & Jobert, 1988; Bozdog & Trampert, 2007). Cette sensibilité des fonctions propres des différents modes aux 250 premiers kilomètres de la Terre est illustrée sur les figures 2.6 et 2.7. Les modèles globaux 1D sont des modèles de référence qui sont représentatifs de la Terre dans sa globalité, mais leurs paramètres sismiques ne sont pas capables de décrire des hétérogénéités locales. C'est particulièrement vrai pour la croûte et le manteau supérieur, où les différences des modèles tomographiques par rapport aux modèles globaux sont de plusieurs pourcents (*e. g.* Nataf & Ricard, 1996). Des modèles 3D des couches superficielles sont maintenant disponibles, comme par exemple le modèle CRUST2.0 (Bassin et al., 2000). Cependant, des corrections inappropriées des structures crustales peuvent engendrer de fausses signatures dans le manteau sous-jacent (Boschi & Ekström, 2002; Ferreira et al., 2010; Panning et al., 2010).

Certaines études essayent de contourner cette difficulté en permettant de petites variations de vitesses et de la profondeur du Moho autour d'un modèle crustal choisi *a priori* (Li & Romanowicz, 1996; Shapiro & Ritzwoller, 2002; Visser et al., 2008b; Khan et al., 2009). Mais réaliser des perturbations autour d'un modèle préexistant obtenu avec des données et des théories différentes pourrait introduire des biais dans le résultat de l'inversion. La particularité de ce travail n'est pas de retrouver le profil V_S des couches superficielles en détail, mais de ne pas introduire des biais en profondeur liés à des incohérences entre les différents jeux de paramètres.

Notre région d'intérêt est la zone de transition *s. l.* (350- 1000 km de profondeur). Dans cette étude, une stratégie originale est choisie pour prendre en compte les paramètres sismiques de la croûte et du manteau supérieur. En raison de la grande variété minéralogique des premières centaines de kilomètres de la Terre, et de l'impossibilité d'appliquer les lois thermodynamiques globales aux conditions de pressions et de températures de la lithosphère, seuls V_S et ξ sont inversés au-dessus de la zone de transition, comme le font de

nombreuses études tomographiques. Cela signifie que l'étape qui consiste à utiliser la température (déterminée de façon aléatoire), ainsi que l'équation d'état de Birch-Murnaghan, pour calculer les vitesses sismiques et la masse volumique dans un second temps, n'est pas implémentée dans la partie supérieure du modèle (< 350 km). De cette manière, le modèle de référence de croûte et de manteau supérieur sert seulement à définir des bornes de recherche de la marche aléatoire et son choix n'est pas aussi crucial que dans des inversions linéarisées basées sur des perturbations d'un modèle de référence. Concrètement, un certain nombre de points de Bézier n_{V_S} , n_T et n_ξ sont tirés aléatoirement dans les domaines Ω_1 et Ω_2 (figure 4.1). Ω_1 peut être assimilé à la croûte et à la partie supérieure du manteau, Ω_2 à la zone de transition.

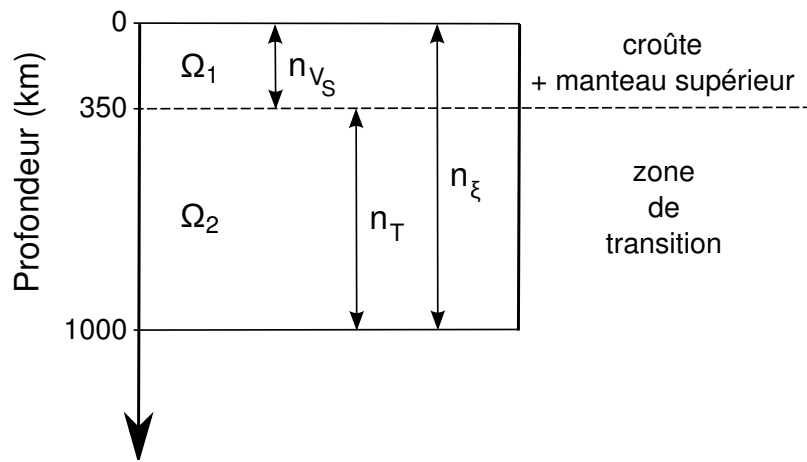


FIG. 4.1 : Les points de Bézier correspondant aux paramètres du modèle en V_S , T et ξ sont tirés aléatoirement à l'intérieur des domaines Ω_1 et Ω_2 , dont les gammes de profondeurs sont différentes.

4.2.2 Conditions de raccord

Au point de raccord entre les domaines Ω_1 et Ω_2 , la valeur de V_S à la base du manteau supérieur n'est pas choisie aléatoirement, mais est calculée afin d'obtenir une continuité avec le profil de vitesse de cisaillement de la zone de transition issu directement de la température. Elle est directement calculée grâce aux lois thermodynamiques à partir de la valeur de la température tirée au sort à cette profondeur.

De 1500 km de profondeur au centre de la Terre, les valeurs des vitesses sismiques et de la masse volumique sont celles du PREM. Pour assurer une continuité entre les paramètres sismiques dans la zone de transition et le PREM, un point de Bézier est fixé à 1500 km de profondeur et sa température a été fixée de telle manière qu'elle reproduise des valeurs sismiques équivalentes à celle du PREM à cette profondeur. Des tests numériques

montrent que la valeur de la température à cette profondeur n'influence pas notre région d'intérêt (350- 1000 km de profondeur).

De la même manière, un raccord est réalisé à 1500 km de profondeur pour l'anisotropie. Les études qui utilisent les ondes de surface suggèrent que l'anisotropie est extrêmement faible, voire nulle à cette profondeur (*e. g.* [Panning & Romanowicz, 2006](#); [Kustowski et al., 2008](#)), c'est pourquoi ξ est fixée à la valeur de 1 à 1500 km de profondeur.

4.2.3 Priors sur les paramètres de l'inversion

La formulation bayésienne permet de prendre en compte l'information *a priori* sous la forme d'une distribution de probabilité $P(\mathbf{p})$. Elle représente les informations disponibles sur le modèle. Dans ce travail, l'information *a priori* correspond aux valeurs de température, de vitesses et d'anisotropie que nous pensons être raisonnables au regard des études précédentes. Une des critiques récurrente faite à l'inversion bayésienne est que les utilisateurs adaptent le prior de façon à obtenir la solution qu'ils souhaitent et espèrent. En effet, le résultat de l'inversion est influencé par la forme de la distribution *a priori*, dont le choix est subjectif et dépend entièrement de l'utilisateur. Cependant, cette critique est également vraie pour la méthode des gradients (voir section 1.2.3), dans laquelle il faut faire des choix sur les facteurs qui contrôlent le lissage et l'amortissement des paramètres, de même que sur les opérateurs de covariance. En fait, il s'agit d'un problème récurrent pour toutes les procédures d'inversion.

Dans cette thèse, nous avons essayé d'employer des priors les moins restrictifs possibles afin de laisser le maximum de poids aux données et de peu affecter les résultats par des informations externes au problème direct et aux données. Considérons le vecteur $\mathbf{x} = [\mathbf{V}_S, \mathbf{T}, \xi]$ constitué de N points de Bézier ($N = n_{V_S} + n_T + n_\xi$, avec n_{V_S} , n_T et n_ξ la taille des vecteurs $\mathbf{V}_S$, $\mathbf{T}$ et ξ). Si l'on considère que le prior possède une valeur constante sur un intervalle donné $\Delta x = x_{\max} - x_{\min}$, la probabilité $P(x_i)$ est

$$P(x_i) = \begin{cases} 1/\Delta x & \text{si } x_{\max} > x_i > x_{\min} \\ 0 & \text{sinon.} \end{cases} \quad (4.1)$$

Les bornes $x_{\max}$ et $x_{\min}$ sont choisies de façon arbitraire mais raisonnable par rapport à la connaissance que l'on a de ces paramètres. Et puisque la valeur de x pour chaque point de Bézier est indépendante,

$$P(\mathbf{x}) = \prod_{i=1}^N P(x_i). \quad (4.2)$$

Pour la position des points de Bézier en profondeur, c'est-à-dire le vecteur $\mathbf{z} = [\mathbf{z}_{V_S}, \mathbf{z}_T, \mathbf{z}_\xi]$, une distribution uniforme conduit de la même manière à

$$P(\mathbf{z}) = \prod_{i=1}^N P(z_i). \quad (4.3)$$

En estimant que les distributions de probabilité *a priori* pour chacun des N points de Bézier sont totalement indépendantes les unes des autres, alors cette probabilité s'écrit

$$P(\mathbf{p}) = P(\mathbf{x})P(\mathbf{z}), \quad (4.4)$$

où $\mathbf{p} = [\mathbf{V}_S, \mathbf{T}, \xi, \mathbf{z}_{V_S}, \mathbf{z}_T, \mathbf{z}_\xi]$. Ces expressions sont valables uniquement dans le cas où les distributions de probabilité *a priori* sur les paramètres des modèles sont uniformes sur des intervalles. Nous avons essayé de respecter au maximum cette liberté de positionnement des points de Bézier dans l'espace des paramètres en utilisant des priors larges. Cependant, nous avons jugé nécessaire de rajouter quelques contraintes supplémentaires pour obtenir des modèles physiquement réalistes. Le détail de ces contraintes ainsi que leurs effets sur les densités de probabilité sont décrits dans la section 4.4.1.4, où nous verrons qu'il est relativement difficile d'obtenir un prior parfaitement neutre.

4.3 L'algorithme

4.3.1 Les principales étapes

Après avoir déterminé les valeurs initiales des paramètres de façon aléatoire (instant $t = 0$), l'algorithme procède de façon itérative. Les différentes étapes de l'algorithme sont représentées sur la figure 4.2 et peuvent être résumées de la manière suivante :

- À chaque itération de la chaîne, le modèle à l'itération $t - 1$ est modifié selon une distribution gaussienne et un nouveau modèle est proposé à l'itération t de façon à satisfaire les conditions *a priori*.
- Les points de Bézier discrets sont reliés par des courbes de Bézier cubiques.
- Le problème direct est résolu :
 - V_P et ρ dans Ω_1 sont déterminés par des relations d'échelle (voir section 4.4.1.3).
 - V_S , V_P et ρ dans Ω_2 sont calculés grâce aux diagrammes de phases et à l'équation d'état de Birch-Murnaghan (voir section 2.2).
 - V_{SV} et V_{SH} sont estimés dans Ω_1 et Ω_2 à partir de ξ et des vitesses isotropes (voir section 2.3).
 - Les courbes de dispersion de Love et de Rayleigh sont calculées (voir section 2.4).

- Algorithme de Metropolis-Hastings (e. g. [Metropolis et al., 1953](#); [Hastings, 1970](#)) : le modèle est accepté ou rejeté selon la probabilité d'acceptation $\alpha(x'|x)$ (équation 3.20).
 - Si le modèle proposé à t est rejeté, le modèle à l'itération $t - 1$ est retenu pour la prochaine itération et ajouté à l'ensemble des modèles échantillonnés.
 - Si le modèle à t est accepté, ce modèle devient le modèle de départ de l'itération suivante ($t + 1$), à partir duquel un nouveau modèle est tiré au sort, et rejoint la collection de modèles qui sera utilisée pour calculer la densité de probabilité marginale (que l'on appellera *pdf* pour *probability density function*).

Selon la nature du prior et de la fonction employée pour échantillonner l'espace des paramètres, l'algorithme va généralement accepter les modèles qui présentent une valeur de misfit plus faible (ou une *likelihood* plus grande). Cependant, il peut également accepter ce que nous pouvons considérer comme de « moins bons » modèles, dont le misfit est plus important que celui du modèle précédent dans la chaîne (voir section 3.3.2). Ce com-

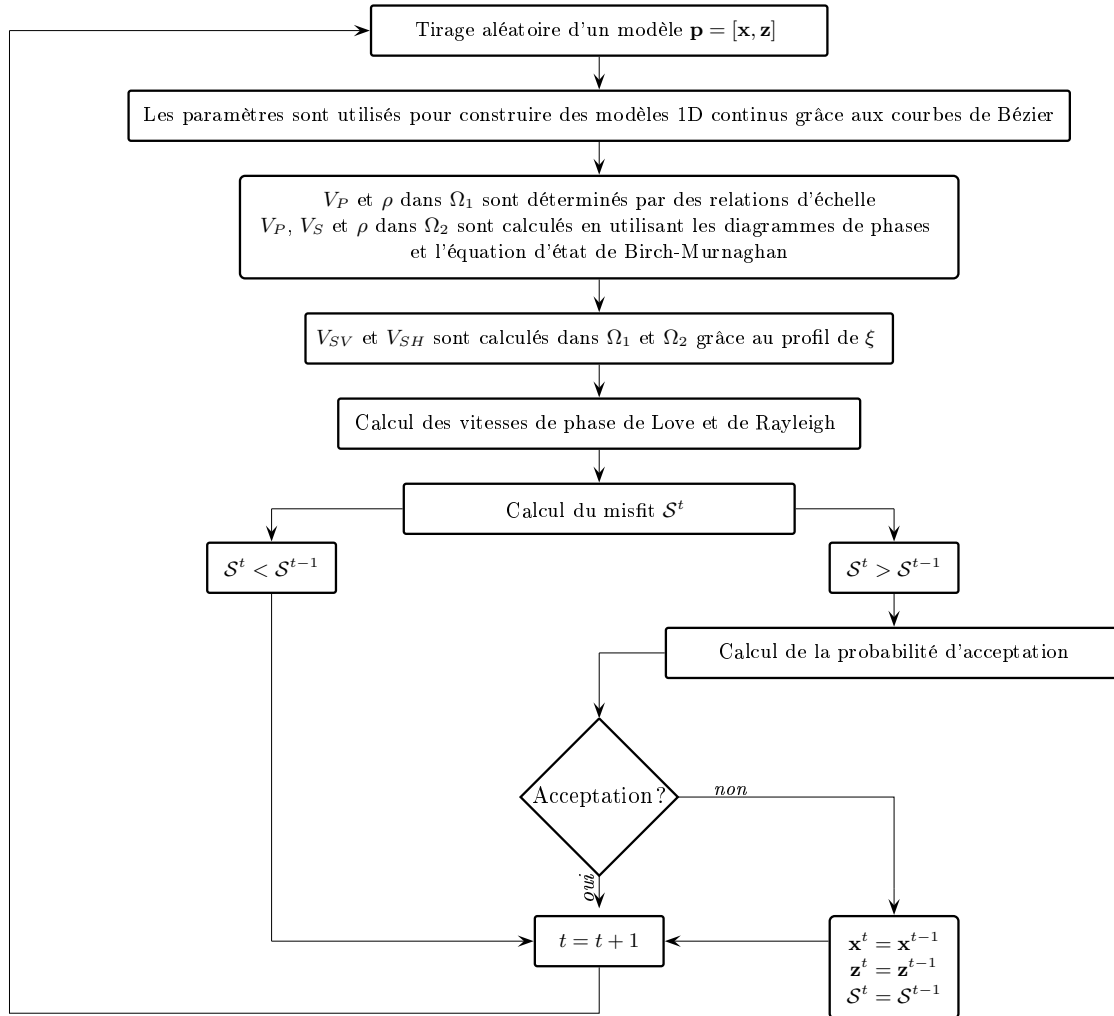


FIG. 4.2 : Les différentes étapes de l'algorithme élaboré au cours de cette thèse.

portement est très important car il permet à la chaîne de sortir de minima locaux. En effet, la forme de la fonction coût peut être telle qu'elle présente de nombreux minima locaux, en plus d'un minimum global (voir figure 1.12). De plus, c'est grâce à cet échantillonnage qu'une distribution de probabilité représentative des paramètres du modèle peut être obtenue, compte-tenu des incertitudes sur les données.

4.3.2 Générer de nouveaux modèles tout au long d'une chaîne de Markov

Nous avons vu dans la section 3.3.3 qu'une distribution gaussienne $g(x'|x)$ centrée sur le paramètre actuel à l'instant t , est utilisée pour le tirage au sort du nouveau paramètre à l'instant $t + 1$ (équation 3.21). Cette densité de proposition gaussienne est également appliquée pour les profondeurs des points de Bézier z . En pratique, les paramètres se situent sur des intervalles finis (entre $x_{\min}$ et $x_{\max}$), définis par les conditions *a priori*. Si à partir de l'état actuel x_i^t la nouvelle valeur proposée x_i^{t+1} est en-dehors de l'intervalle, alors une réflexion est réalisée sur le bord de l'intervalle pour respecter la symétrie de la distribution.

Pour choisir une nouvelle configuration des paramètres V_S à partir de la configuration précédente dans le domaine Ω_1 , l'argument d'invariance de Tarantola (2005) est appliqué. L'invariance implique l'obtention de mesures égales, qu'il s'agisse de vitesses ou de lenteurs. Avec une distribution en $\log(V_S/V_{S_0})$, une distribution similaire est obtenue qu'il s'agisse de la vitesse ou de la lenteur que nous recherchons. Tout au long de cette étude, la paramétrisation en $\log(V_S/V_{S_0})$ est adoptée dans le domaine Ω_1 . Dans la suite du manuscrit, je continuerai de nommer par V_S les paramètres des vitesses de cisaillement dans la croûte et le manteau supérieur, mais il est implicitement considéré qu'il s'agit des vitesses en log normal. Dans cette étude, $V_{S_0} = 1$ km/s est fixée. Cette paramétrisation a déjà été employée par Khan & Mosegaard (2002) avec les données sismiques lunaires.

4.3.3 Construction des courbes de Bézier de classe C^1

Les courbes de Bézier sont calculées en considérant que la direction du vecteur tangent local au niveau d'un point de Bézier donné dépend de la position des deux points localisés juste au-dessus et au-dessous de ce point. Deux cas particuliers peuvent se poser et sont illustrés sur la figure 4.3, qu'il s'agisse d'un profil en T , V_S ou ξ . Si l'on considère le point de Bézier 2 de la figure 4.3a, le gradient avec le point 1 est de signe opposé au gradient formé avec le point 3. Si les gradients de part et d'autre d'un point de Bézier sont de signes opposés, le vecteur tangent local est fixé de façon verticale. En revanche, si les gradients entre les points 1 et 2 d'une part et les points 2 et 3 d'autre part, sont

de même signe, comme illustré sur la figure 4.3b, la direction du vecteur tangent local est déterminée en calculant la moyenne des deux gradients. Au niveau des extrémités du profil (points 1 et 4), le vecteur tangent est systématiquement orienté vers le point suivant ou précédent, respectivement.

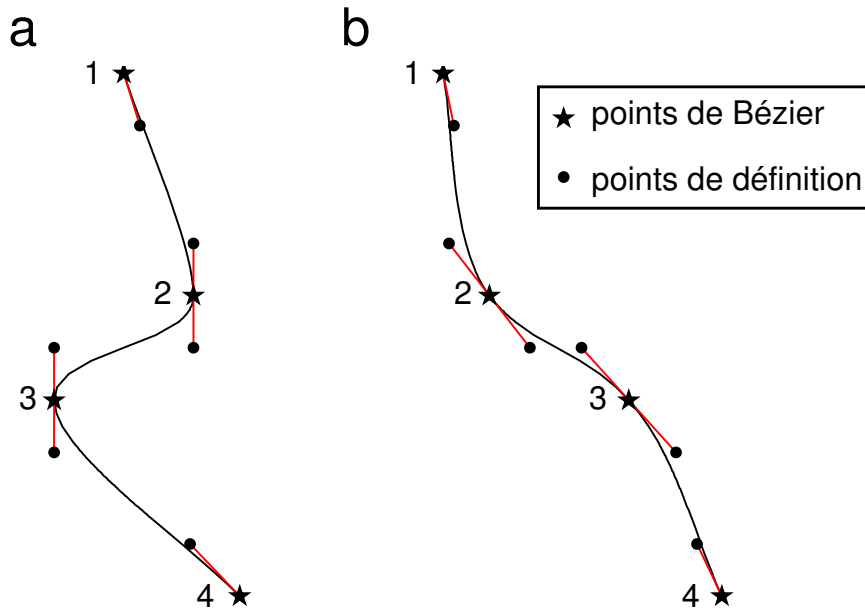


FIG. 4.3 : Représentation schématique des points de Bézier, des points de définition et des vecteurs tangents locaux (en rouge) pour la construction d'un profil 1D de température, d'anisotropie ou de vitesse de cisaillement. Trois courbes de Bézier cubiques se succèdent en profondeur. (a) Cas où les gradients de part et d'autre des points 2 et 3 sont de signes différents. (b) Cas où les gradients de part et d'autre des points 2 et 3 sont de même signe.

4.3.4 Optimisation de l'algorithme

La convergence statistique de l'algorithme vers de faibles valeurs de misfit est surveillée grâce à un certain nombre d'indicateurs tels que le taux d'acceptance des modèles, et l'efficacité de l'échantillonnage est optimisée en ajustant l'écart-type des fonction gaussiennes. Une faiblesse des méthodes MCMC est qu'elles deviennent rapidement coûteuses en temps de calcul. En effet, lorsque le modèle est défini par un nombre important de paramètres, le nombre de modèles échantillonnés nécessaires pour explorer l'espace des paramètres devient particulièrement grand (Tarantola, 2005). Dans notre cas, les courbes de dispersion ont besoin d'être calculées à chaque fois qu'un modèle est proposé, et lorsque trop de modèles doivent être testés, le temps d'exécution de l'algorithme devient prohibitif. Il est clair que la méthode la plus générale pour résoudre un problème non-linéaire nécessite une exploration intensive de l'espace des paramètres, puisque la densité de pro-

tabilité dans l'espace des modèles contient toute l'information sur le système étudié. C'est une raison pour laquelle les méthodes de Monte Carlo sont peu utilisées en sismologie.

Un second écueil est que les échantillons produits par une chaîne de Markov sont corrélés, c'est-à-dire qu'ils suivent un chemin dans l'espace des paramètres. Lorsque le nombre d'itérations est trop petit, la chaîne de Markov va seulement décrire la distribution *a posteriori* dans une région donnée de l'espace des modèles. Ce phénomène peut conduire à une sous-estimation de la taille de l'incertitude sur les paramètres du modèle, à cause d'une couverture inadéquate de l'espace des paramètres.

Cependant, un avantage de l'algorithme mis en œuvre dans cette thèse est sa relative facilité à paralléliser, dans le sens où une chaîne peut échantillonner l'espace des modèles indépendamment des autres. Le même algorithme peut être exécuté plusieurs fois sur des processeurs indépendants d'une machine parallèle (Rosenthal, 2000). Des racines de nombres aléatoires différentes sont utilisées, ce qui signifie que la configuration initiale des paramètres est différente selon le processeur considéré, et que des chemins différents sont suivis dans l'espace des modèles. Un autre avantage de la parallélisation est qu'elle produit naturellement des échantillons indépendants. Les distributions *a posteriori* issues de chaque processeur peuvent être ensuite rassemblées pour déterminer la distribution *a posteriori* finale des paramètres. Ce type de parallélisation est relativement pratique à mettre en œuvre car il nécessite peu de communications entre les processeurs.

Pour illustrer les bénéfices de la parallélisation, des exemples simples d'expériences numériques menées avec les algorithmes séquentiel et parallèle sont montrées dans les deux sections suivantes.

4.3.4.1 Avantage en temps de calcul

Supposons que l'on dispose de 2 heures de temps de calcul disponible pour déterminer la pdf de la température pour un jeu de vitesses de phase donné. La figure 4.4 montre l'ensemble des points de Bézier testés par l'algorithme (y compris ceux qui n'ont pas été acceptés) pour un seul processeur et pour 4 processeurs. Deux points de Bézier sont fixés à 350 et 1000 km de profondeur. Les autres points du même profil peuvent varier librement mais doivent être éloignés d'au moins 80 km les uns des autres. Pour un seul processeur, seulement 6000 échantillons ont été produits en 2 heures, alors que $6000 \times 4 = 24\,000$ modèles ont été générés par l'algorithme utilisé en parallèle. Notons que si nous avions voulu estimer les vitesses de phase du même nombre d'échantillons avec l'algorithme séquentiel, 8 heures de calcul auraient été nécessaires.

Dans le cas où un seul processeur est employé, l'espace des paramètres a été visité de manière moins complète. Dans cet exemple, l'échantillonnage a surtout été focalisé vers la limite supérieure du prior en température. En revanche, l'utilisation de plusieurs processeurs permet de couvrir quasiment l'ensemble du domaine possible.

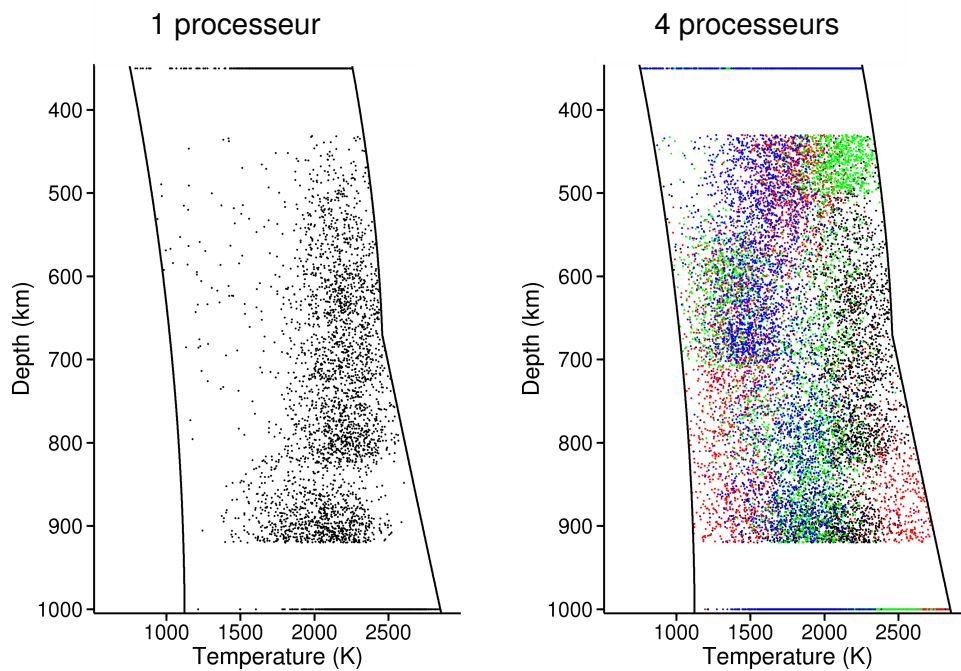


FIG. 4.4 : Ensemble des configurations de points de Bézier testées durant 6000 itérations avec un seul processeur (à gauche) et 4 processeurs (à droite). Chaque couleur correspond aux points échantillonnés par un processeur particulier. Les courbes noires représentent les bornes du domaine en température.

4.3.4.2 Avantage en performance

Un autre avantage de la parallélisation de l'algorithme réside dans le fait que des échantillons indépendants sont naturellement produits. La figure 4.5 montre les distributions des échantillons acceptés dans l'espace des paramètres obtenues avec un algorithme séquentiel et un algorithme parallélisé pour le même nombre d'itérations (soit 1500). Pour le cas séquentiel, les 1500 modèles sont montrés. Pour les 4 processeurs, 375 modèles de chaque processeur ont été sélectionnés aléatoirement, ce qui donne également au total 1500 modèles. Les étoiles noires représentent les points de Bézier du modèle synthétique à retrouver. Les échantillons d'une chaîne de Markov sont corrélés, c'est-à-dire qu'ils suivent un chemin dans l'espace des modèles. Ainsi, en utilisant l'algorithme en séquentiel, on observe que certaines régions de l'espace des paramètres sont occultées. Par exemple, vers 500 km de profondeur la distribution des points de Bézier est très resserrée si un unique processeur est employé. La distribution est beaucoup plus étalée dans la recherche avec 4 processeurs, ce qui signifie que l'incertitude sur la valeur des paramètres est sous-estimée avec un seul processeur. C'est pourquoi l'estimation de la distribution de probabilité *a posteriori* est améliorée si un grand nombre de chaînes indépendantes explorent l'espace des paramètres.

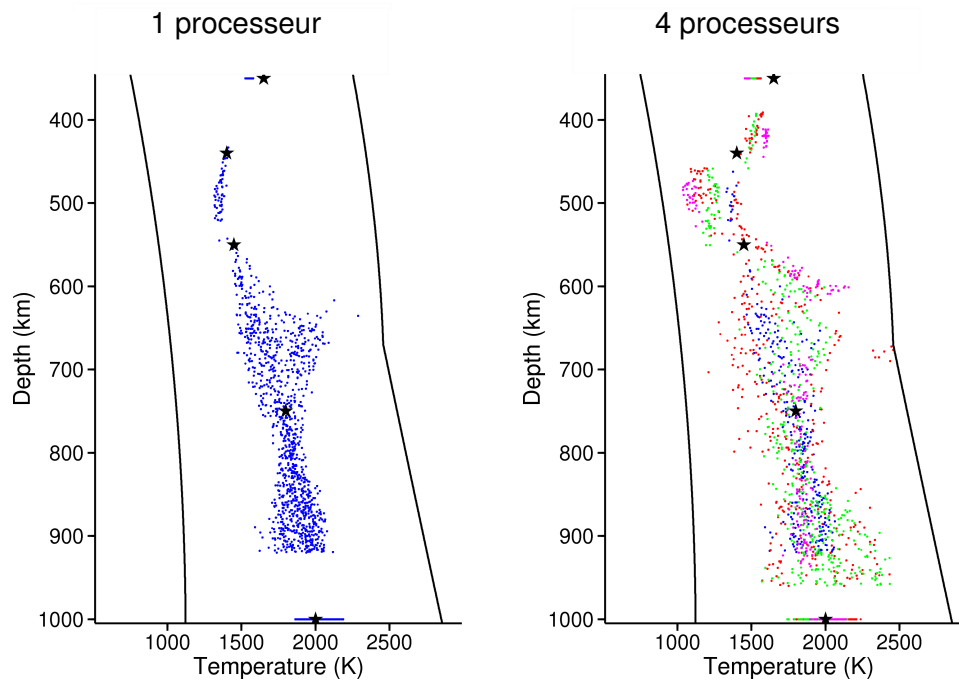


FIG. 4.5 : Ensemble des configurations des points de Bézier acceptées avec un seul processeur (à gauche) et 4 processeurs (à droite). 1500 modèles sont montrés. Chaque couleur correspond aux points échantillonnés par un processeur particulier. Les courbes noires représentent les bornes du domaine en température. Les étoiles noires correspondent aux points de Bézier du modèle synthétique recherché.

4.3.5 Atteindre la période stationnaire

En pratique, une procédure classique est de conserver uniquement les échantillons obtenus après un certain nombre d'itérations, lorsque le misfit est devenu suffisamment faible. Cette période, appelée *période de chauffe* (ou *burn-in*, en anglais), est d'une durée variable selon le problème physique posé, les conditions *a priori*, la configuration initiale des paramètres et le nombre de paramètres. La période pendant laquelle les échantillons sont ensuite collectés est appelée *période stationnaire*.

Deux des grandes questions que l'utilisateur se pose lorsqu'il emploie une méthode MCMC sont :

- À quel moment la période stationnaire est-elle atteinte ?
- Combien de modèles doivent être collectés pour que la pdf soit représentative ?

Ces deux questions sont l'objet de nombreuses recherches en statistique bayésienne. Pour une revue sur les différents diagnostics de convergence et le nombre d'itérations nécessaires, le lecteur pourra se référer à [Cowles & Carlin \(1995\)](#) ou encore [Brooks & Roberts \(1998\)](#).

Les diagnostics traditionnels de la convergence d'une chaîne de Markov reposent sur le fait que pour chaque paramètre individuel, l'évolution de sa valeur en fonction du

nombre d'itérations ne doit pas présenter de tendance particulière une fois la période de chauffe achevée. En d'autres termes, la trace de chaque paramètre de l'inversion doit approximer un bruit blanc. Cependant, nous verrons que ce diagnostic traditionnel présente ses limites pour notre problème particulier, lié au fait que la position des points de Bézier n'est pas fixée *a priori*, mais est bien une inconnue de l'inversion. Par exemple, admettons que le profil de température recherché corresponde à un profil adiabatique, c'est-à-dire à une droite. Les points de Bézier peuvent être placés à n'importe quelle profondeur, du moment que leurs températures coïncident avec celles de l'adiabat, le profil adiabatique formé par les courbes de Bézier sera identique puisqu'il s'agira d'une droite. Dans ce cas, les valeurs des paramètres T et z_T peuvent prendre des valeurs très différentes et leurs traces ne correspondront pas forcément à un bruit blanc. Pourtant, la valeur du misfit sera systématiquement faible. Ainsi, dans ce travail, l'atteinte (ou non) de la période stationnaire n'est pas estimée selon la trace des paramètres. Nous verrons dans la suite que le choix de la taille du saut entre le paramètre à l'instant t et le paramètre à l'instant $t + 1$ via la fonction gaussienne est également crucial pour l'efficacité de la recherche aléatoire.

Dans le cas de l'algorithme parallélisé employé dans cette thèse, l'utilisation de plusieurs chaînes indépendantes permet de favoriser l'atteinte de l'état stationnaire ainsi que d'obtenir des modèles indépendants. La figure 4.6 montre l'évolution de la valeur du misfit en fonction du nombre d'itérations pour 20 processeurs, et donc 20 configurations initiales différentes. Au bout de 60 000 itérations, on observe que certains processeurs ont toujours des valeurs de misfits relativement grandes, ce qui signifie que la rapidité de l'algorithme à rencontrer une configuration favorable dépend fortement de la configuration initiale et du chemin pris par la chaîne. Dans certains cas, il faudrait effectuer un très grand nombre d'itérations avant d'atteindre des modèles favorables.

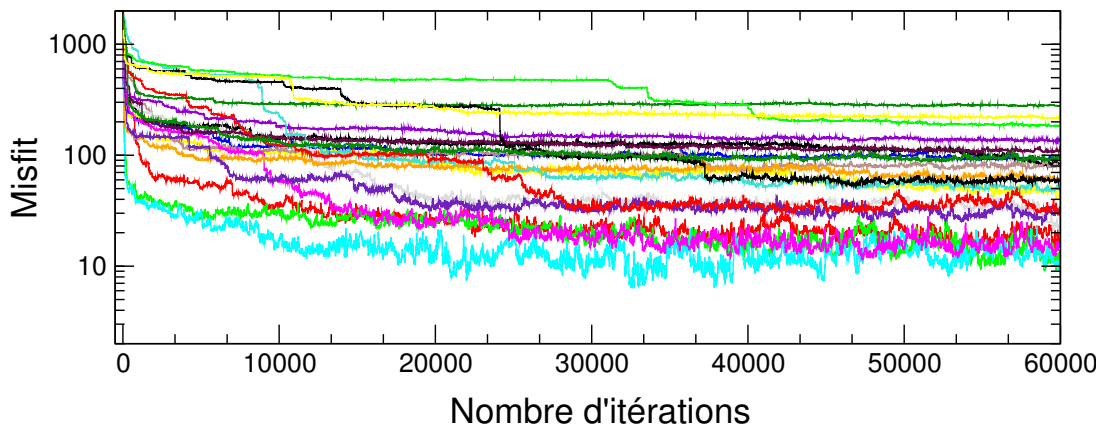


FIG. 4.6 : Exemples de l'évolution de la valeur du misfit pour 20 processeurs. Le misfit est représenté avec une échelle logarithmique.

En fait, l'idéal serait de :

1. faire tourner l'algorithme le plus longtemps possible,
2. obtenir le plus grand nombre de modèles indépendants en utilisant un grand nombre de processeurs.

Nous allons voir dans la section suivante comment l'algorithme peut être employé en parallèle pour atteindre efficacement la période stationnaire et pour échantillonner des modèles indépendants.

4.3.6 Une procédure en trois étapes

J'ai choisi d'exploiter la prédisposition de l'algorithme à la parallélisation en divisant le schéma d'inversion en trois étapes. Les deux premières étapes sont dédiées à la recherche d'une configuration favorable, à partir de laquelle la troisième étape est engagée. Cette dernière étape est la période stationnaire.

- Première étape : L'algorithme est exécuté sur un très grand nombre de processeurs $nproc_1$ avec des générateurs de nombres aléatoires différents, mais sur un nombre d'itérations $niter_1$ relativement faible. L'objectif est d'explorer le plus largement possible l'espace des paramètres, car il existe une forte sensibilité de la chaîne par rapport au modèle initial. Les écart-types des fonctions gaussiennes sont suffisamment grands pour effectuer de grands sauts d'une itération à l'autre. Tous les paramètres sont modifiés à chaque itération. À la fin de cette première étape, la configuration correspondant au plus petit misfit obtenu au cours des $niter_1$ itérations est sélectionnée pour chaque processeur. Seuls 20 % des configurations correspondant aux meilleurs misfits sont sélectionnés pour la deuxième étape.
- Deuxième étape : Puisque seules les configurations les plus favorables ont été conservées, on considère que certaines parties des modèles correspondant sont assez proches du modèle recherché. Par conséquent, les $nproc_2$ processeurs sont utilisés pour ne changer qu'un seul couple de paramètres à chaque itération, c'est-à-dire la valeur et la position d'un couple (x_i, z_i) . Pour chaque itération, il est d'abord décidé quel couple de paramètres sera ensuite perturbé. Tous les couples sont équiprobables, mêmes ceux situés aux limites du domaine en profondeur, dont la profondeur est fixée mais pas la valeur en V_S , T ou ξ . La stratégie de modifier un seul couple à la fois permet de préserver la plupart des caractéristiques du modèle à l'instant t , qui a montré une bonne adéquation avec les données. Les écart-types des fonctions gaussiennes sont réduits pour atteindre un taux d'acceptance des modèles d'environ 20-25 %. Durant cette étape, l'objectif est

de laisser l'algorithme réaliser suffisamment d'itérations pour atteindre de faibles valeurs de misfit, c'est pourquoi le nombre d'itérations est 10 fois supérieur à celui de l'étape 1 ($niter_2 = 10 \times niter_1$). Seules 25 % des meilleures configurations passeront à l'étape 3.

- Troisième étape : Les $nproc_3$ meilleurs modèles constituent le point de départ de la dernière étape, que nous appelons *période stationnaire*. Le même nombre d'itérations que pour l'étape 2 est appliqué ($niter_3 = niter_2$). Seuls les modèles échantillonnés durant la période stationnaire sont utilisés pour déterminer les pdfs et pour étudier les résultats. Parmi ceux-ci, tous ne sont pas employés pour calculer les pdfs *a posteriori*. Les modèles sont collectés toutes les t itérations pour obtenir un ensemble d'échantillons indépendants sur une même chaîne. C'est ce que l'on appelle en anglais le *thinning* de la chaîne de Markov.

La figure 4.7 montre l'évolution de la valeur du misfit au cours des 3 étapes pour un processeur donné. On voit que l'étape 1 permet de diminuer considérablement la valeur du misfit. Le pourcentage d'acceptation minimum des modèles n'est pas respecté, ce qui implique que l'algorithme peut rester longtemps bloqué sur une même configuration des paramètres. C'est pourquoi la trace du misfit forme par endroits des sortes de plateaux. Mais l'avantage est qu'il est possible de passer très rapidement d'une configuration à une autre. Les 10 000 premières itérations de l'étape 2 permettent de diminuer rapidement

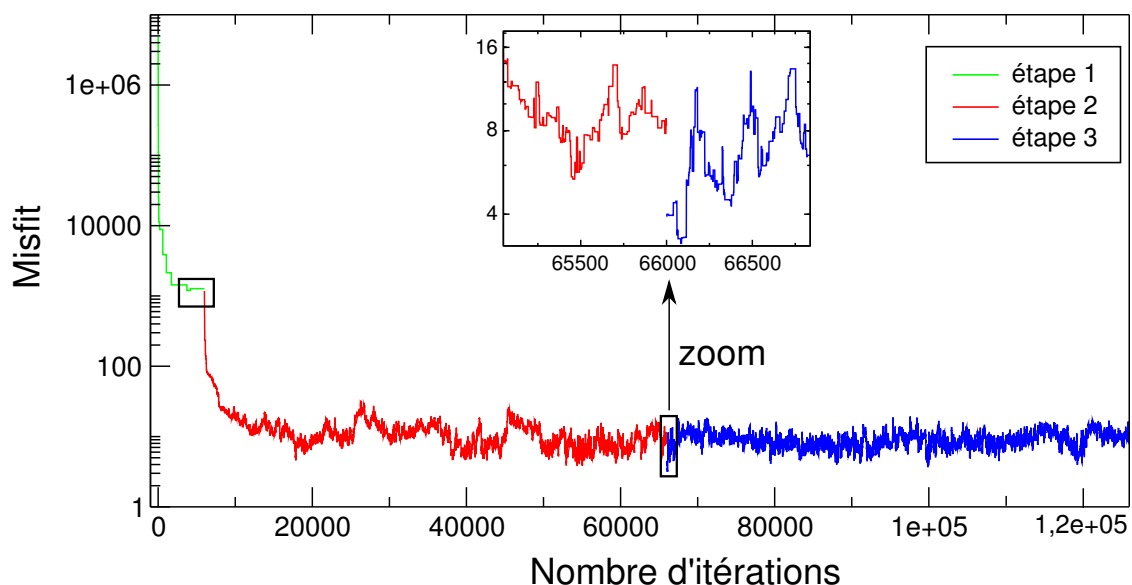


FIG. 4.7 : Exemple de l'évolution de la valeur du misfit pour un processeur donné au cours des 3 étapes de la procédure. Le misfit est représenté avec une échelle logarithmique. Les rectangles noirs indiquent les passages entre les différentes étapes.

le misfit. L'étape 2 pourrait être assimilée au *burn-in* traditionnel, si ce n'était pas le meilleur misfit rencontré au cours des 60 000 itérations mais le dernier misfit obtenu qui était retenu pour le début de la période stationnaire. Pendant cette dernière période le misfit est suffisamment faible pour commencer à collecter des échantillons. Les rectangles noirs sur la figure 4.7 montrent qu'il n'existe pas nécessairement de continuité entre la fin de l'étape 1 et le début de l'étape 2, ainsi qu'entre la fin de l'étape 2 et le début de la période stationnaire. Ce comportement est lié au fait que c'est la meilleure configuration rencontrée pendant une étape qui est sélectionnée comme étant le point de départ de la suivante, et non la configuration qui correspond à la dernière itération.

Pendant l'étape 1, l'objectif est de balayer l'espace des modèles avec de grands pas. Au bout d'un certain nombre d'itérations, la valeur du misfit ne diminue plus et atteint un plateau. Ce comportement de l'algorithme signifie qu'étant donné l'écart-type important imposé pour la fonction gaussienne, la probabilité de trouver une combinaison de paramètres qui puisse satisfaire encore mieux les données est faible, c'est pourquoi il est nécessaire d'affiner la recherche des modèles avec l'étape 2. En pratique, $niter_1$ est défini comme étant le nombre d'itérations au bout duquel les chaînes atteignent ce plateau. Le nombre d'itérations nécessaires lors de l'étape 2 est fixé en testant plusieurs inversions sur des données synthétiques et réelles, et en vérifiant que les configurations correspondant aux meilleurs misfits obtenus produisent bien des vitesses de phase qui se situent à l'intérieur des barres d'erreur estimées sur les données. Ce nombre d'itérations constitue la limite minimale et est fortement dépendant de la taille des barres d'erreur. $niter_3$ est calibré en observant l'évolution de la pdf en fonction du nombre d'itérations. On considère que l'exécution peut être arrêtée lorsque la pdf est quasiment similaire même en augmentant le nombre d'itérations de façon significative. Ce nombre d'itérations est évidemment accordé par rapport au nombre de processeurs employés (1600 au départ de la première étape). Un certain nombre de tests préliminaires ont été réalisés afin de trouver le meilleur équilibre entre le nombre de processeurs et le nombre d'itérations.

4.3.7 Temps de calcul

Il est difficile de donner un temps de calcul précis pour cet algorithme. Le temps de calcul dépend principalement de trois facteurs :

- le temps de calcul du problème direct à chaque fois qu'un modèle est testé,
- le temps mis par la chaîne de Markov pour atteindre la période stationnaire,
- le temps écoulé pour que suffisamment de modèles soient échantillonnés.

Ces trois facteurs dépendent eux-mêmes de plusieurs éléments qui sont spécifiques à chaque problème : le nombre de données, leurs incertitudes, la complexité de la distribution à retrouver, l'étendue du prior et encore la forme de la densité de proposition d'un

nouveau modèle. La distribution *a priori* sur le modèle définit la taille de l'espace des modèles et ainsi le temps nécessaire à la chaîne pour arriver à la période stationnaire. Les temps de calcul sont détaillés pour chaque étape dans la section 5.2. Dans cette étude, le problème direct est d'une durée d'environ 2 s si les courbes de dispersion de Love et de Rayleigh sont calculées jusqu'au mode 3. Cette durée n'est pour le moment pas réductible, d'où l'intérêt de paralléliser les chaînes.

4.4 Choix particuliers pour l'étude de la zone de transition

4.4.1 Contraintes *a priori*

Une représentation schématique des contraintes *a priori* sur les paramètres de l'inversion est montrée sur la figure 4.8. Les profils décrivant la vitesse de cisaillement dans les 350 premiers kilomètres (V_S), la température dans la zone de transition (T) et l'anisotropie radiale (ξ), sont paramétrés avec des courbes de Bézier. Les paramètres de l'inversion sont les coordonnées des trois jeux de points de Bézier, c'est-à-dire les vecteurs $\mathbf{V}_S$, $\mathbf{T}$, et ξ , ainsi que les vecteurs correspondant à la localisation de ces points en profondeur $\mathbf{z}_{V_S}$, $\mathbf{z}_T$ et $\mathbf{z}_\xi$. Au total, entre 27 et 35 paramètres sont considérés.

4.4.1.1 Prior sur la température

Prior 1

Le profil de température dans la zone de transition est uniquement défini entre 350 et 1000 km de profondeur. Ainsi, deux points de Bézier sont nécessairement fixés à 350 et 1000 km de profondeur. Il existe un prior joint entre la profondeur et la température car la température dépend de la profondeur des points de Bézier. Les bornes minimales en température ont été estimées à partir de la modélisation de l'état thermique à l'intérieur des plaques plongeantes (*e. g.* Tagawa et al., 2007; Ganguly et al., 2009), et l'équation polynomiale de Hirshmann (2000) et Hirshmann et al. (2009) qui décrit la position du solidus d'une péridotite sèche est utilisée comme limite supérieure. Les points de Bézier sont libres de varier entre 350 et 1000 km de profondeur. Les températures de ces points sont échantillonnées entre les courbes noires représentées sur la figure 4.8, qui correspondent aux valeurs extrêmes permises.

Prior 2

Pour ne pas échantillonner deux points proches en profondeur qui présentent une grande différence de température, un second prior a été défini pour tirer au sort les points de Bézier entre 350 et 1000 km de profondeur. Il s'ajoute au prior 1. L'application de ce

second prior pour l'échantillonnage d'un nouveau point est illustré sur la figure 4.9. La température de ce nouveau point dépend de la position dans le diagramme (température-profondeur) du point situé juste au-dessus de celui-ci et du point à 1000 km de profondeur. Dans l'exemple de la figure 4.9a, le point du dessus se situe à 350 km de profondeur. Le gradient de température entre le nouveau point et le point situé au-dessus/au-dessous, doit être dans cet exemple au maximum de ± 15 K/km. Les intersections entre ce second prior et les bornes minimales et maximales en température du prior 1 (en rouge sur la figure

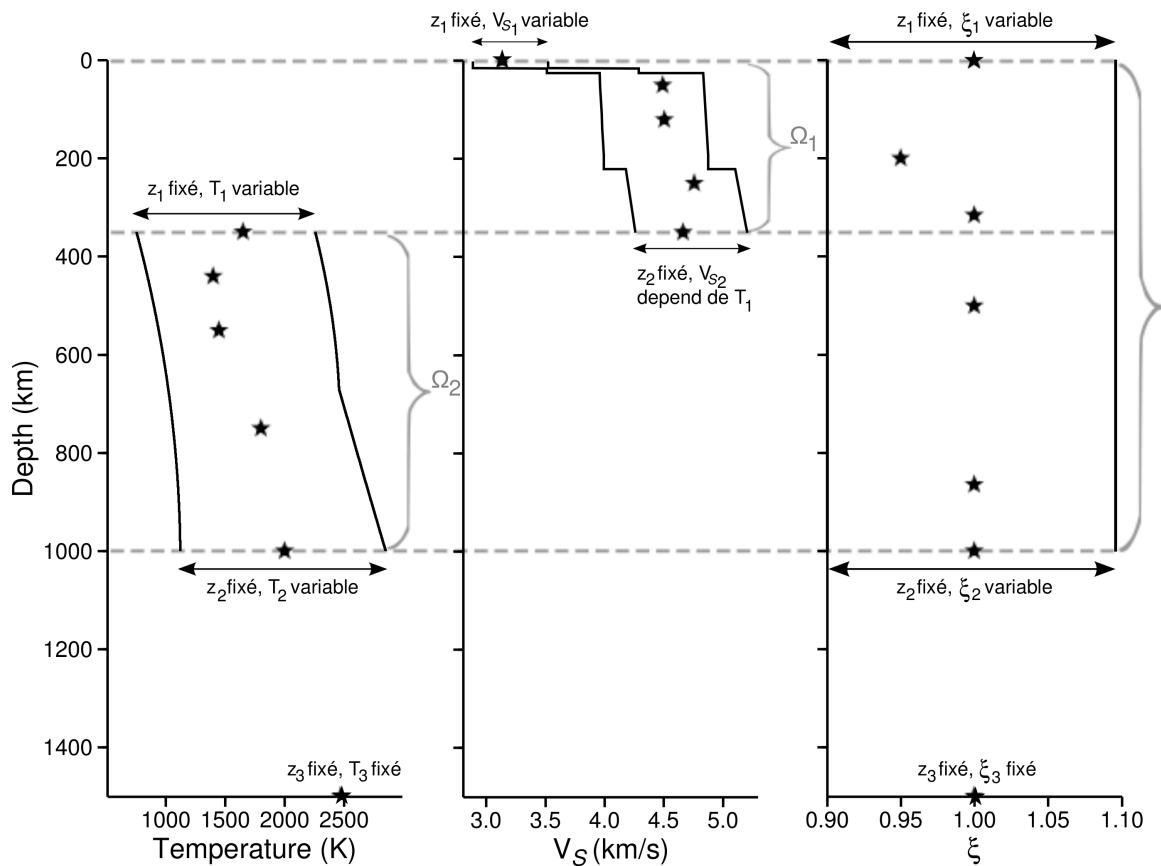


FIG. 4.8 : Représentation schématique des contraintes *a priori* sur les points de Bézier (c'est-à-dire les paramètres du modèle) pour T , V_S et ξ . Les étoiles noires correspondent aux points de Bézier. Les lignes en pointillés définissent les profondeurs à 0, 350 et 1000 km. Les lignes noires représentent les bornes de l'espace des modèles. Les accolades indiquent les intervalles de profondeurs où les points sont échantillonnés de façon aléatoire. Les bornes en profondeur pour l'échantillonnage des points en T , V_S et ξ sont 350- 1000, 0- 350 et 0- 1000 km, respectivement. Pour assurer une continuité entre les valeurs du PREM à 1500 km de profondeur, les valeurs de température (T_3) et d'anisotropie (ξ_3) sont fixées selon les valeurs du PREM. La vitesse V_{S2} à 350 km est directement calculée selon la valeur de la température T_1 à cette même profondeur. Pour l'exemple, seuls 6 points sont représentés en température, mais leur nombre peut être de 6 à 10.

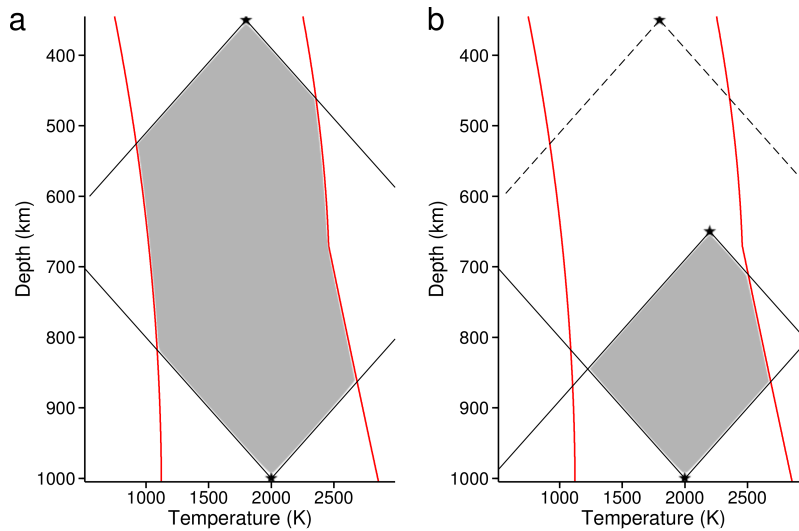


FIG. 4.9 : Représentation schématique des bornes minimales et maximales en température lors de l'échantillonnage d'un nouveau point de Bézier. Les limites du prior 1 sont représentées par les lignes rouges. Les limites du second prior correspondent aux lignes noires. Les surfaces grisées définissent l'ensemble des localisations possibles pour un nouveau point de Bézier. (a) Ensemble des possibilités pour le placement d'un nouveau point entre 350 et 1000 km de profondeur. (b) Ensemble des possibilités pour le placement d'un nouveau point entre 650 et 1000 km de profondeur.

4.9), définissent l'espace des paramètres à l'intérieur duquel le nouveau point peut être choisi de façon aléatoire. L'espace des paramètres possibles est représenté par les surfaces grisées. Supposons que le nouveau point ait été placé à une profondeur de 650 km et à une température de 2200 K (figure 4.9b). C'est à partir de ce point et du point à 1000 km de profondeur qu'un nouveau point va être choisi, et ainsi de suite.

Nombre de points de Bézier

L'exemple de la figure 4.8 montre un profil de température qui comprend 6 points de Bézier, mais l'algorithme fonctionne en parallèle avec des nombres de points différents. Le nombre de points n_T pour l'ensemble du profil peut être égal à 6, 7, 8, 9 ou 10. n_T ne varie pas durant l'exécution de l'algorithme. Les inversions transdimensionnelles (*e. g.* Green, 1995; Bodin et al., 2012), où le nombre de paramètres est également une inconnue, constituent une grande avancée dans les méthodes d'inversion probabilistes. Dans le futur, ce type d'algorithme pourrait être adapté à notre problème pour ne pas avoir à choisir le nombre de points à utiliser en V_S , T et ξ . L'utilisation de différents nombres de paramètres laisse la possibilité d'obtenir différents degrés de détails. Des valeurs raisonnables de n_T ont été choisies afin de prévenir le risque de surparamétrer les modèles et d'obtenir un degré de détails qui ne peut être atteint par les données. Les ondes de surface étant peu sensibles dans la zone de transition à de rapides changements de la structure physique,

une condition *a priori* est posée sur la distance minimale entre deux points de Bézier. La distance entre deux points ne peut pas être plus petite qu'une valeur d_z égale à 80, 60, 40, 20 et 10 km pour 6, 7, 8, 9 et 10 points, respectivement. Nous avons vu dans la section 2.5.2 que selon la valeur de la norme du vecteur tangent local, les polynômes de Bézier pouvaient produire des profils très lisses ou beaucoup plus escarpés (voir figure 2.11b). La taille de la composante de ces vecteurs en profondeur, appelée Δ , est fixée de façon à ce que $\Delta = 0.5 d_z$. Cela signifie que plus la valeur de Δ est grande, plus le profil est lisse et a la possibilité d'être courbé. Cette information *a priori* permet de détecter à la fois des variations à grande et petite longueurs d'onde. Un récapitulatif de ces différents paramètres est présenté dans la table 4.1.

	T	ξ	V_S
nombre de points de Bézier	$n_T = 6 \text{ à } 10$	$n_\xi = 7$	$n_{V_S} = 5$
gamme de profondeurs (km)	350 à 1000	0 à 1000	0 à 350
Δ (km)	40 à 5	20	10
d_z (km)	80 à 10	40	20

TAB. 4.1 : Synthèse des différents paramètres qui contraignent le placement des points de Bézier.

Le fait que les positions des points de Bézier fassent partie des paramètres de l'inversion permet une grande souplesse. Supposons qu'un profil de température soit composé de 6 points de Bézier. Trois points sont obligatoirement fixés à 350, 1000 et 1500 km de profondeur, il reste donc 3 points à placer entre 430 et 920 km (et non entre 350 et 1000 km puisqu'il faut prendre en compte le paramètre Δ). Sur la figure 4.10a sont représentées de façon indépendante les positions de ces 3 points de Bézier après un tirage aléatoire de 50 000 modèles.

Tous les modèles qui satisfont le prior 1 ainsi que la contrainte imposée par le paramètre Δ sont acceptés. Si l'intervalle des profondeurs 430- 920 km est divisé en 3 parties égales, représentées par des lignes en pointillées sur la figure 4.10a, et que le pourcentage de points dans chacun de ces intervalles est calculé, on observe qu'il n'est pas équivalent à 100 %. Lorsque la profondeur des points est tirée au sort, les points de Bézier sont ensuite systématiquement classés par ordre de profondeur croissante, sinon la probabilité dans chaque intervalle serait de 33 %. Les mêmes figures sont montrées pour 5 points de Bézier (figure 4.10b). Pour le 3^{ème} point, qui est globalement situé au centre de la gamme de profondeurs autorisées, seulement la moitié des points échantillonnés (50.9 %) sont restreints dans l'intervalle 618- 732 km. Ce résultat est très intéressant car même si globalement la localisation des points est de plus en plus profonde en fonction de son numéro, il montre que les points de Bézier ne restent pas complètement confinés à des profondeurs spécifiques. Par exemple, le 3^{ème} point de la figure 4.10 peut se placer aussi bien à 500

qu'à 850 km. C'est au niveau du premier et du dernier point que le pourcentage est le plus important ($\geq 80\%$), quel que soit le nombre de points, car leurs déplacements sont limités par les bornes supérieure et inférieure en profondeur du domaine en température.

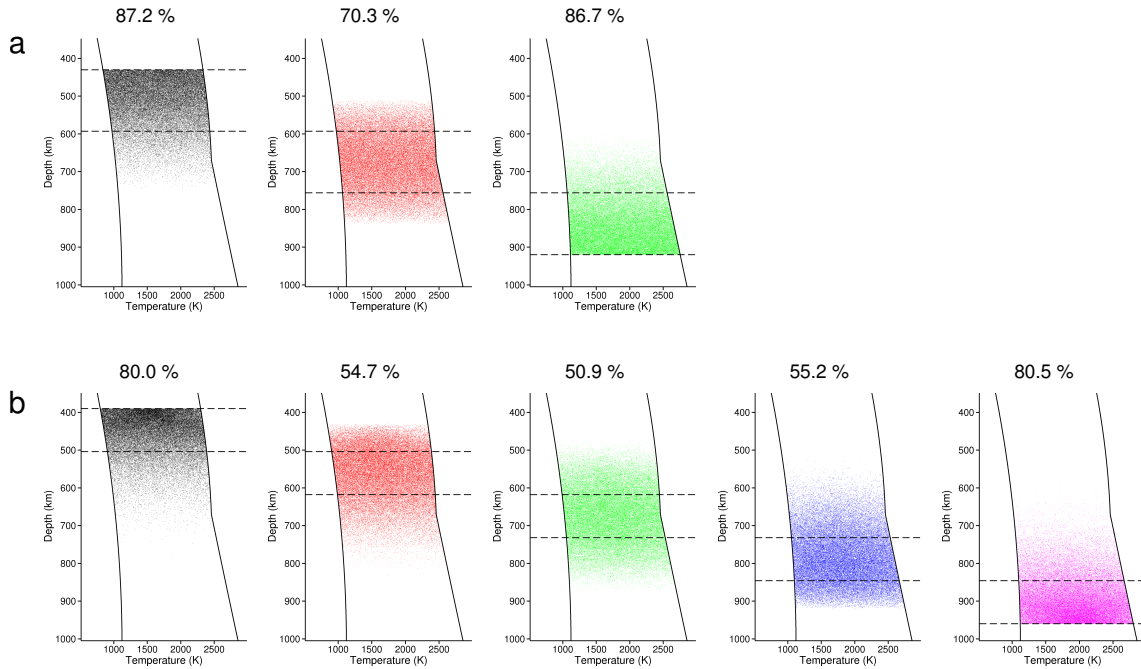


FIG. 4.10 : Position de chaque point de Bézier pour un tirage aléatoire de 50 000 modèles constitués de (a) 3 et (b) 5 points de Bézier. Les lignes en pointillés délimitent les gammes de profondeurs à l'intérieur desquelles le pourcentage de points est calculé par rapport à l'ensemble des points échantillonnés.

4.4.1.2 Prior sur l'anisotropie

La valeur de ξ est distribuée de façon uniforme entre 0.9 et 1.1. Les profils sont formés de 7 points de Bézier. Les valeurs limites ont été choisies pour encadrer la plupart des gammes de valeurs obtenues par des études récentes utilisant les ondes de surface (*e. g.* [Panning & Romanowicz, 2006](#); [Kustowski et al., 2008](#); [Visser et al., 2008b](#); [Khan et al., 2009, 2011](#)). Une anisotropie de cisaillement négative entre 200 et 600 km de profondeur a été trouvée par [Panning & Romanowicz \(2006\)](#) et [Visser et al. \(2008b\)](#). Selon les études récentes, même si l'anisotropie radiale semble décroître au fur et à mesure que l'on s'approche de la zone de transition, le prior n'est pas réduit en fonction de la profondeur car il peut provoquer un lissage artificiel du profil d'anisotropie ([Khan et al., 2009](#)). Deux points de Bézier sont fixés à la surface et à 1000 km de profondeur. d_z vaut 40 km.

4.4.1.3 Prior sur les paramètres sismiques

La zone de transition

Aucune information *a priori* sur les vitesses sismiques et la masse volumique n'est appliquée dans la zone de transition car il s'agit de paramètres de seconde génération du modèle qui dépendent uniquement des paramètres primaires (température, composition) et des lois thermodynamiques. Seule la zone de transition est définie en terme de température et le calcul des vitesses de phase par le programme MINEOS nécessite un modèle de croûte et de manteau supérieur.

V_S dans la croûte et le manteau supérieur

V_S est directement inversé entre la surface et 350 km de profondeur. Les points de Bézier peuvent varier entre $\pm 10\%$ des valeurs du PREM (il s'agit des lignes noires sur la figure 4.8). Ici nous insistons sur le fait que le modèle de référence sert seulement à définir des bornes de recherche.

Le nombre de points de Bézier est $n_{V_S} = 5$. La valeur de V_S à 350 km de profondeur n'est pas choisie aléatoirement. Elle est directement calculée à partir de la valeur de la température à cette profondeur grâce aux lois thermodynamiques. Pour éviter de créer une discontinuité artificielle du profil de vitesse de cisaillement entre le manteau supérieur et la zone de transition, le point de Bézier localisé juste au-dessus de 350 km de profondeur est échantillonné de telle manière que le gradient de vitesse entre ce point et le point à 350 km de profondeur soit entre $\pm 2.5 \times 10^{-3} \text{ s}^{-1}$.

V_P et ρ dans la croûte et le manteau supérieur

Les relations d'échelle sont souvent utilisées dans les inversions en profondeur pour réduire le nombre de paramètres à ceux les mieux résolus, notamment V_{SV} et V_{SH} (e. g. Ekström & Dziewonski, 1998; Shapiro & Ritzwoller, 2002; Gung et al., 2003; Paning & Romanowicz, 2006). La perturbation de la vitesse des ondes P est reliée à celle des ondes S par la relation $R_{S/P} = \partial \ln V_S / \partial \ln V_P$. De la même manière, la perturbation de la masse volumique ρ dépend de la perturbation de la vitesse des ondes S selon $R_{\rho/S} = \partial \ln \rho / \partial \ln V_S$. Cependant, les facteurs employés dans ces relations présentent des valeurs et des origines diverses. Une synthèse de plusieurs valeurs de $R_{S/P}$ et $R_{\rho/S}$ employées dans différentes études est présentée dans le tableau 4.2.

Les valeurs des rapports peuvent être issues d'expériences de laboratoire avec prise en compte de l'anélasticité (Karato, 1993), ou non (Isaak et al., 1989; Isaak, 1992). Vacher et al. (1996) ont calculé les valeurs de $R_{\rho/S}$ en utilisant un modèle minéralogique simplifié du manteau supérieur pour des températures prédites par une convection isovisqueuse et ont obtenu des valeurs allant de 0.42 en surface et diminuant jusqu'à 0.37 à 660 km

de profondeur. Le paramètre $R_{S/P} = 0.5$ employé par [Gung et al. \(2003\)](#) et [Panning & Romanowicz \(2006\)](#) est basé sur des valeurs publiées par [Masters et al. \(2000\)](#), tandis que le paramètre $R_{\rho/S} = 0.3$ est fixé selon l'approximation que les effets thermiques dominant à la fois les anomalies de vitesses et de masse volumique. L'inversion jointe de données de gravité et d'un modèle tomographique par [Deschamps et al. \(2001\)](#) fournit une valeur de $R_{\rho/S}$ très faible par rapport à celles utilisées dans les autres études.

Malgré le fait que les relations d'échelle soient différentes d'une étude à une autre, et qu'elles soient basées sur des hypothèses différentes, selon [Ekström & Dziewonski \(1998\)](#), [Gung et al. \(2003\)](#) et [Beghein \(2010\)](#), les relations d'échelles utilisées n'influencent que très peu le résultat de l'inversion de V_{SV} et V_{SH} . Bien que le modèle minéralogique utilisé dans cette étude ne soit pas pertinent pour la partie superficielle de la Terre, nous avons calculé les deux rapports pour une pyrolite et une olivine dans les 350 premiers kilomètres de la Terre. Ces rapports valent $R_{S/P} = 0.8$, $R_{\rho/S} = 0.45$ et $R_{S/P} = 0.75$, $R_{\rho/S} = 0.37$ pour une pyrolite et une olivine, respectivement. Ces valeurs sont relativement proches des résultats de [Isaak et al. \(1989\)](#) et [Isaak \(1992\)](#), issus des extrapolations de données expérimentales sur l'olivine à haute température et pression ambiante. Dans ce travail, les valeurs de [Isaak et al. \(1989\)](#) et [Isaak \(1992\)](#) ont été sélectionnées pour calculer les profils V_P et ρ entre la surface et 350 km de profondeur à partir du profil V_S déterminé aléatoirement.

Référence	$R_{\rho/S}$	$R_{S/P}$
Isaak et al. (1989); Isaak (1992)	0.35-0.45	0.82-0.83
Karato (1993)	0.2-0.3	
Vacher et al. (1996)	0.42	
Ekström & Dziewonski (1998)		0.55
Deschamps et al. (2001)	0.05	
Shapiro & Ritzwoller (2002)	0.25	0.5
Gung et al. (2003); Panning & Romanowicz (2006)	0.3	0.5

TAB. 4.2 : Valeurs des rapports $R_{S/P}$ et $R_{\rho/S}$ selon différentes études.

4.4.1.4 Distributions de probabilité *a priori*

Bien que des bornes relativement larges soient employées dans cette étude, il est impossible d'obtenir un prior qui n'engendre aucun impact sur le résultat de l'inversion. L'objectif est qu'il soit aussi faible que possible. La figure 4.11 présente la densité de probabilité marginale *a priori* pour T , V_S et ξ . Dans ce cas, tous les modèles en bon accord avec les conditions *a priori* décrites dans les sections précédentes ont été acceptés. Au total, 250 000 modèles ont été échantillonnés. Les résultats sont montrés à la fois pour les points de Bézier (en haut) et les courbes de Bézier (en bas). Une composition pyrolitique est considérée.

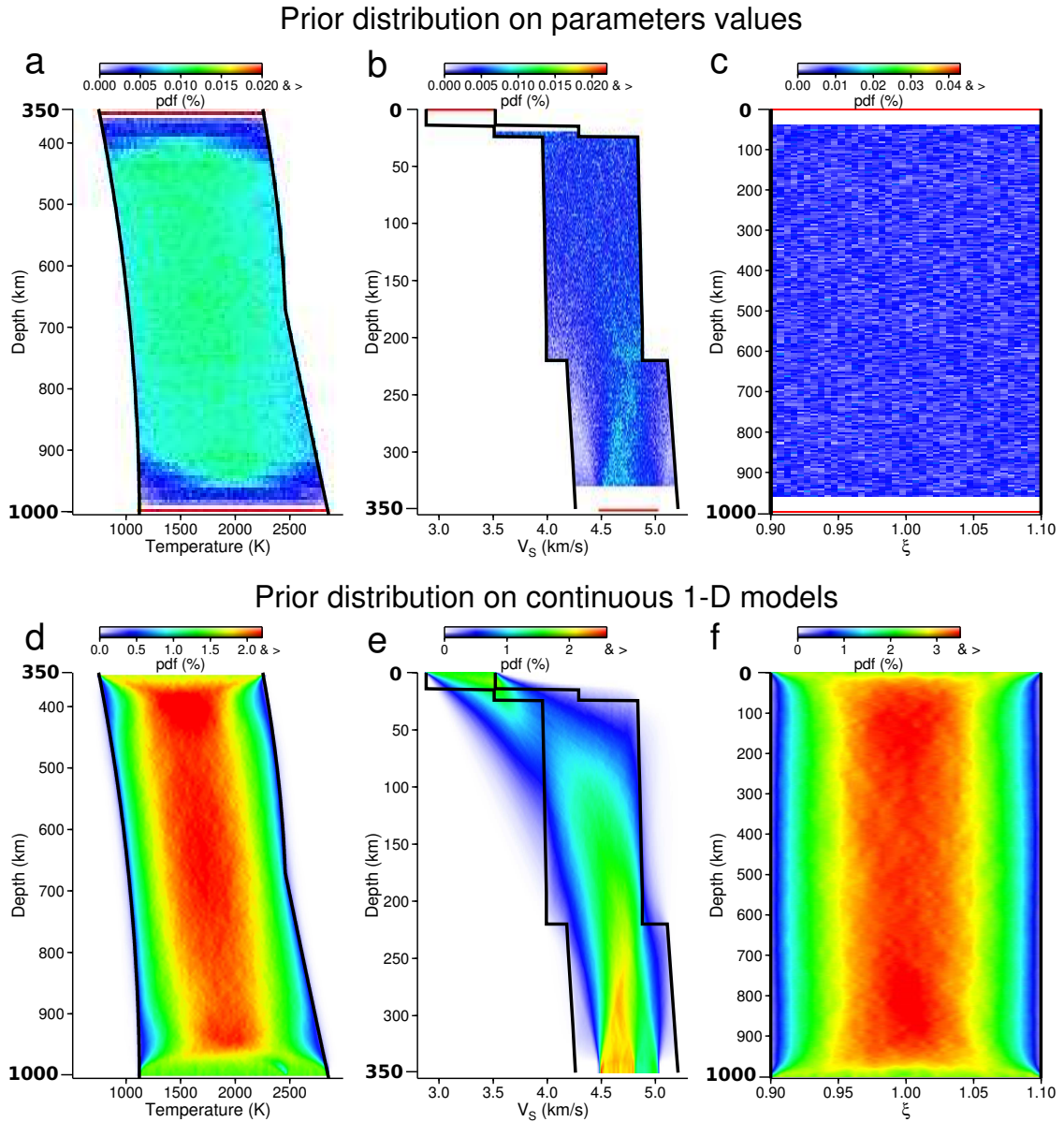


FIG. 4.11 : Densités de probabilité marginales pour T (à gauche), V_S (au milieu) et ξ (à droite), en considérant que tous les modèles échantillonnés en accord avec les conditions *a priori* ont été acceptés. Les résultats sont montrés pour les valeurs discrètes des paramètres (en haut) et pour les modèles 1D (en bas). Du bleu au rouge, la densité de probabilité est de plus en plus importante. Les lignes noires représentent les valeurs minimales et maximales autorisées. Les valeurs des pdfs sur les paramètres discrets sont calculées en divisant le nombre de points par case par le nombre total de points. Les tailles des cases pour T , V_S et ξ sont $25 \text{ K} \times 5 \text{ km}$, $0.01 \text{ km/s} \times 2 \text{ km}$ et $0.005 \times 5 \text{ km}$, respectivement. Les valeurs des pdfs pour les modèles 1D sont calculées en comptant le nombre de profils à l'intérieur de chaque intervalle de température, à chaque profondeur. Pour une profondeur donnée, la somme de toutes les pdfs sur tous les intervalles de température vaut 100 %.

Les distributions marginales de la solution *a priori* sont représentées par des cartes couleur sur la figure 4.11. La distribution marginale de T , V_S ou ξ à une profondeur donnée correspond à un histogramme construit à partir de l'ensemble des modèles acceptés. En pratique, le résultat correspond à la juxtaposition des histogrammes à chaque profondeur, ce qui permet de construire une carte de densité. Bien que toutes les configurations autorisées par le prior aient la même probabilité d'être acceptées, nous allons voir que les contraintes requises durant l'échantillonnage produisent des densités de probabilité non uniformes. Seule la pdf sur les points de Bézier représentant l'anisotropie (figure 4.11c) montre une distribution uniforme, puisque toutes les valeurs entre 0.9 et 1.1 à chaque profondeur ont la même probabilité d'être acceptées.

La pdf est quasiment uniforme pour les points de Bézier en V_S (figure 4.11b), excepté entre 190 et 330 km de profondeur. Ce comportement est lié aux contraintes sur les deux points les plus profonds. Le point à 350 km de profondeur est contraint par la valeur de la température à cette profondeur et la loi thermodynamique. Ainsi, la gamme des valeurs de V_S possibles est plus restreinte que les bornes du prior, en considérant une composition réaliste de type pyrolitique. Cette gamme se situe entre 4.5 et 5.0 km/s. Par conséquent, la gamme de valeurs de V_S possibles pour le point situé juste au-dessus de 350 km de profondeur est plus restreinte et les bornes du prior sont moins bien échantillonnées au-dessous d'environ 190 km de profondeur.

La pdf pour les points en température (figure 4.11a) a été calculée pour des profils contenant 6, 7, 8, 9 et 10 points de Bézier. 50 000 modèles ont été tirés au sort dans chaque cas. Le fait d'utiliser des nombres de points différents avec des valeurs d_{zT} spécifiques, conduit à une restriction de la pdf entre 350- 430 et 920- 1000 km de profondeur. Dans ces deux gammes de profondeur la pdf est de plus en plus faible au fur et à mesure que l'on s'approche des extrémités. Le centre du domaine est légèrement mieux échantillonné à cause de l'information *a priori* qui empêche la sélection de deux points voisins présentant une grande différence de température (prior 2).

La distribution des courbes de Bézier produit clairement des pdfs non uniformes, même pour ξ (figure 4.11f). Les valeurs situées entre 0.95 et 1.05 sont mieux échantillonnées que les bords. Pour une profondeur donnée on observe ensuite une diminution de la pdf jusqu'aux limites du domaine. La distribution des courbes en température présente le même comportement (figure 4.11d). Ce type de distribution est appropriée car l'objectif est de retrouver un modèle moyen le long d'un trajet source-récepteur et on estime que les profils localisés au niveau des extrémités ont une faible probabilité d'exister. Les deux distributions sont caractérisées par une forme en « I » à cause de l'estimation de la direction du vecteur tangent local à 350 et 1000 km de profondeur. Comme mentionné dans la section 4.3.3, la direction de ces vecteurs dépend de la position du point le plus proche. La distribution en V_S (figure 4.11e) sort des bornes du prior. En effet, ces bornes sont

uniquement valables pour le tirage au sort des points de Bézier. Lorsque les points sont reliés entre eux par les courbes de Bézier, il est possible que les profils de vitesses de cisaillement puissent sortir des bornes. Ce type de construction permet d'obtenir un vaste panel de profils afin de prendre en compte les variations importantes de vitesse dans les couches superficielles.

Dans l'exemple de la figure 4.11, les modèles tirés au sort n'ont pas été comparés aux données. Seul l'impact des informations *a priori* est visible. Dans les chapitres 5 et 6, l'algorithme est testé avec des données synthétiques et réelles. Seules les densités de probabilités marginales dont les valeurs sont différentes de celles des densités de probabilités marginales *a priori* décrites dans cette section, pourront être considérées comme étant un apport d'informations lié aux données.

4.4.2 Modalités de tirage au sort de nouveaux modèles

Un nouveau modèle est déterminé en perturbant aléatoirement le dernier modèle qui a été accepté. Bien que la manière dont ce nouveau modèle est choisi soit essentiellement arbitraire, des choix inappropriés peuvent conduire à une exploration très lente de l'espace des modèles, de telle manière que le temps de convergence de la chaîne soit très long. Dans cette section nous allons étudier l'influence de différents facteurs sur le résultat de l'inversion.

4.4.2.1 Écart-type des fonctions gaussiennes

Le choix de la valeur de l'écart-type (σ) de la fonction gaussienne $g(x'|x)$ (équation 3.21) est crucial car il va définir l'efficacité d'échantillonnage de la chaîne. En pratique, il est recommandé de choisir l'écart-type le plus grand possible, tout en maintenant un taux d'acceptance suffisamment élevé (Mosegaard, 1998). Pour des distributions normales, un taux d'acceptance de 23 % pour des problèmes multidimensionnels signifie que l'algorithme est performant (Gelman et al., 2003). Une approche classique est de réaliser des tests préliminaires avec différentes tailles de gaussiennes jusqu'à obtenir un taux d'acceptance raisonnable. Suite à ces tests, des valeurs de σ différentes ont été choisies pour T , ξ , V_S , z_T , z_ξ et z_{V_S} . Nous allons illustrer l'impact des tailles de gaussienne sur la recherche des points de Bézier pour la température.

On cherche à retrouver le profil de température de la figure 4.12a, formé par 5 points de Bézier. Pour se déplacer dans l'espace des paramètres, un écart-type entre 50 et 100 K semble être une gamme de valeurs raisonnable, sachant que les incertitudes expérimentales sur la température dans la zone de transition sont de l'ordre de ± 100 -150 K (Ito & Takahashi, 1989). La valeur $\sigma = 70$ (K ou km) est choisie pour les deux fonctions

gaussiennes sur la température et la profondeur. Une recherche aléatoire a été réalisée sur 95 000 itérations. Pour évaluer l'impact de la valeur de σ sur l'efficacité de la recherche, le même algorithme a été employé mais en imposant un écart-type égal à 5 (K ou km). Sur la figure 4.13a sont représentées les traces des paramètres, en profondeur et en température, correspondant aux points de Bézier numérotés de 1 à 3 sur la figure 4.12a. Les traces noires, rouges et vertes correspondent aux points 1, 2 et 3, respectivement. Dans le cas où $\sigma = 70$, la profondeur et la température d'un point de Bézier donné peuvent varier de façon très importante en seulement quelques itérations, ce qui signifie que l'algorithme explore de façon efficace l'espace des paramètres. En revanche, dans le cas où $\sigma = 5$, les oscillations des paramètres d'une itération à une autre ont une amplitude beaucoup plus faible et l'algorithme met beaucoup de temps pour bouger d'une valeur à une autre.

Les figures 4.12b et 4.12c présentent la distribution des points de Bézier échantillonnés dans le diagramme (température, profondeur). Les isocontours de la densité de probabilité *a posteriori* des points sont également indiqués. Dans les deux cas, les points de Bézier acceptés par l'algorithme se situent sur la courbe qui représente le profil de température. Les points de Bézier recherchés numérotés 1 et 2 sur la figure 4.12a correspondent à des valeurs de pdfs importantes. À première vue, l'algorithme est capable de retrouver la solution en utilisant les deux valeurs d'écart-type. Cependant, la distribution des points est beaucoup plus étalée au-dessous de 600 km de profondeur pour le cas

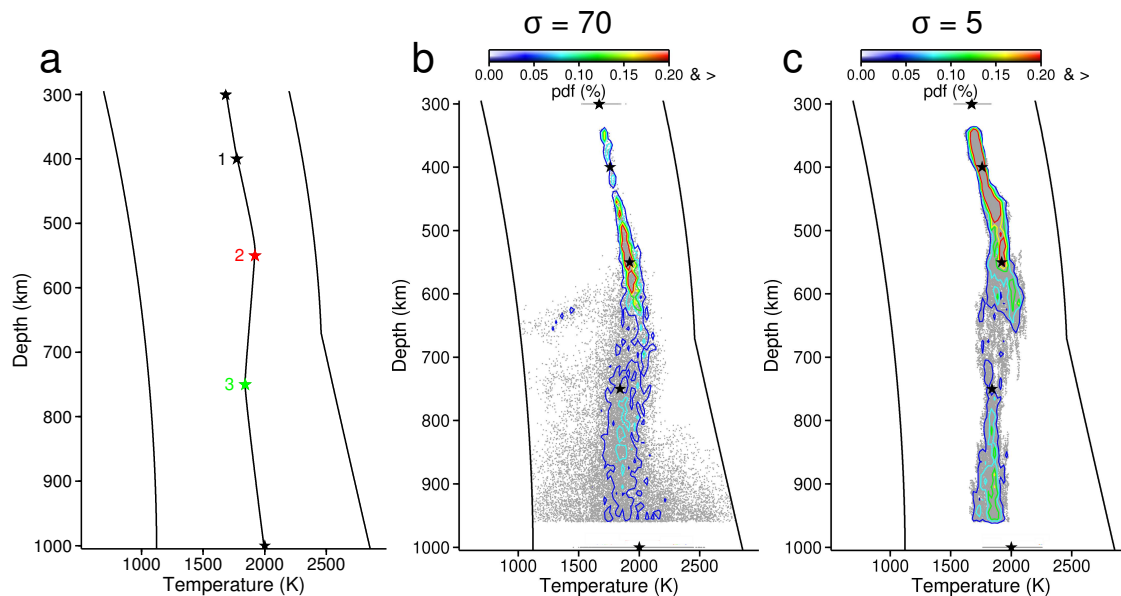


FIG. 4.12 : (a) Profil de température synthétique et ses points de Bézier associés. Les couleurs des points numérotés 1, 2 et 3 correspondent au code couleur des figures 4.13a et 4.13b. (b) et (c) représentent la distribution des points de Bézier après l'inversion du modèle (a) ainsi que les pdfs. Des écart-types de $\sigma = 70$ et $\sigma = 5$ ont été employés pour (b) et (c), respectivement.

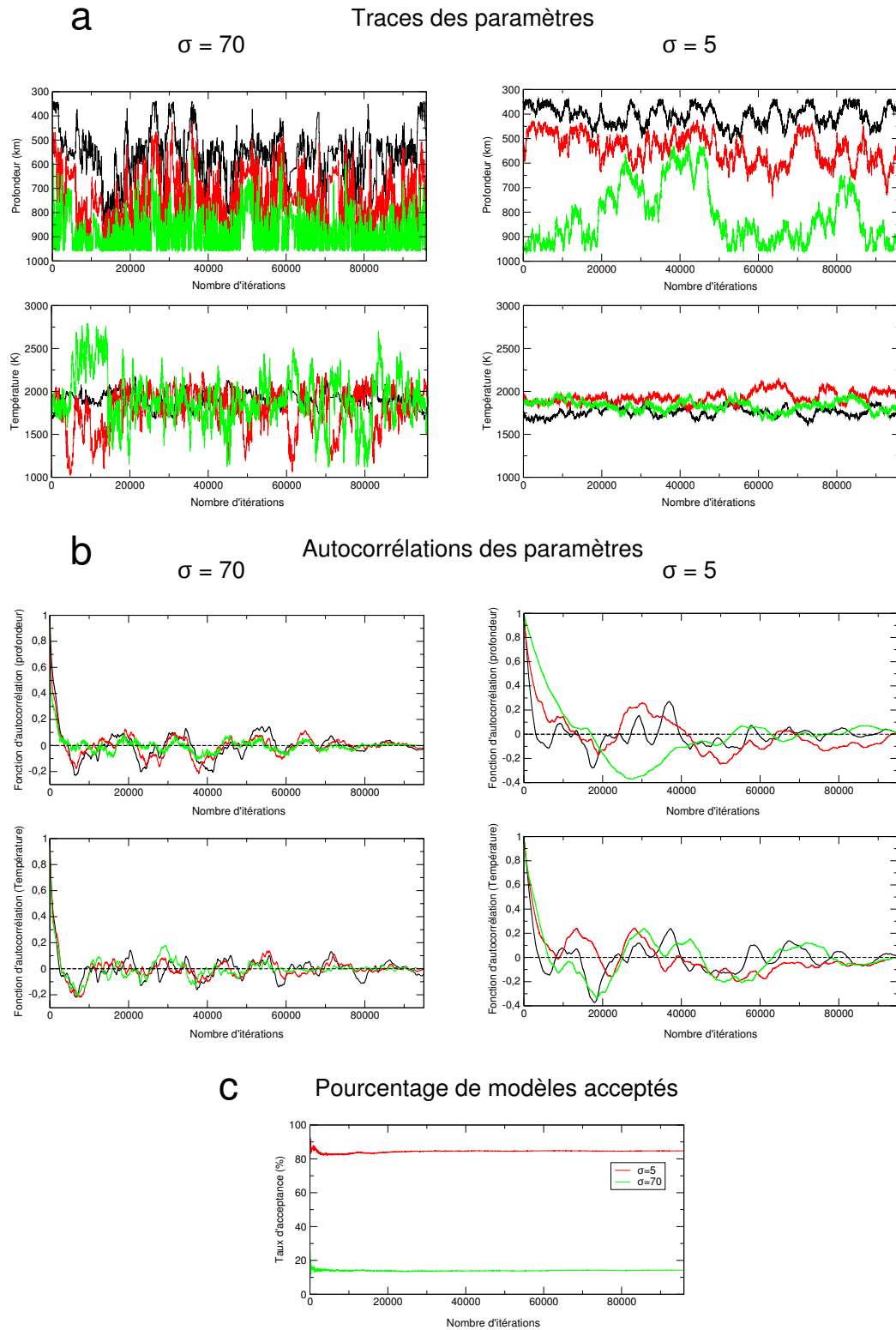


FIG. 4.13 : (a) Traces des paramètres numérotés 1, 2 et 3 sur la figure 4.12a en température et en profondeur, pour $\sigma = 70$ et $\sigma = 5$. (b) Fonctions d'autocorrélation de ces paramètres. (c) Pourcentage de modèles acceptés pour $\sigma = 70$ et $\sigma = 5$.

$\sigma = 70$ que pour le cas $\sigma = 5$, ce qui montre que, dans ce dernier cas, tout une partie de l'espace des paramètres a été occultée, et qu'il manque une partie de la solution. D'après ces résultats, on observe que les chaînes de Markov tendent à générer des modèles d'autant plus corrélés que la marge de mouvement des paramètres est faible. Plus σ est petit, plus il faudra d'itérations avant que la chaîne puisse s'éloigner et échantillonner efficacement l'espace des paramètres.

L'effet de mémoire de la chaîne peut être quantifié par l'évaluation de la fonction d'autocorrélation pour chaque séquence de paramètres. La fonction d'autocovariance nécessite d'abord d'être calculée. Le calcul de cette fonction pour une séquence d'un paramètre correspond au calcul de la moyenne des produits scalaires de cette séquence avec elle-même, décalée de $k = (1, 2, \dots, \text{niter})$ échantillons, où niter correspond au décalage maximal pour lequel est déterminée la fonction d'autocovariance. Pour un paramètre p_i donné, échantillonné pendant niter itérations, cette fonction s'écrit :

$$C_k = \frac{1}{\text{niter} - k} \sum_{t=1}^{\text{niter}-k} (p_i(t) - \bar{p}_i)(p_i(t+k) - \bar{p}_i), \quad (4.5)$$

où p_i est le $i^{\text{ème}}$ paramètre du vecteur $\mathbf{p}$ et $\bar{p}_i$ est la moyenne de toutes les valeurs prises par p_i . La fonction d'autocorrélation r_k est :

$$r_k = \frac{C_k}{C_0}. \quad (4.6)$$

L'évolution de r_k rend compte de l'état de corrélation de la séquence. Pour $k = 0$, r_k est égal à 1 car la séquence est corrélée avec elle-même. r_k tend vers 0 d'autant plus rapidement que ses éléments sont décorrélés.

La représentation des fonctions d'autocorrélation des paramètres permet de donner des indications sur le nombre d'itérations requis pour que les valeurs des paramètres échantillonnées par l'algorithme soient décorrélées, ainsi que sur la valeur de l'écart-type de la gaussienne à employer. Les fonctions d'autocorrélation pour les trois couples (profondeur, température) sont représentées sur la figure 4.13b. Un décalage maximal de $k = 5000$ pour la profondeur et de $k = 3500$ pour la température est nécessaire pour obtenir un jeu de paramètres indépendants, si un écart-type de 70 est employé. Le décalage est beaucoup plus important dans le cas où $\sigma = 5$. En effet, il faut attendre respectivement 18 300 et 8300 itérations pour la profondeur et la température, avant que les valeurs des paramètres soient décorrélées.

Un indicateur supplémentaire quant au choix de la valeur des écart-types est le taux d'acceptance des modèles. La figure 4.13c montre que le taux d'acceptance est d'environ 18 % pour $\sigma = 70$ et 83 % pour $\sigma = 5$. Évidemment, cette valeur n'est pas exactement la même selon la racine de générateur de nombre aléatoire employée, car le chemin parcouru

dans l'espace des paramètres est différent, mais en général elle varie très peu. L'ensemble de ces informations indique qu'un écart-type de 70 (K et km) paraît effectivement adapté à notre problème. Des tests similaires ont été réalisés pour ξ , V_S , z_ξ et z_{V_S} . Les valeurs des écart-types de chaque fonction gaussienne sont récapitulées dans le tableau 4.3.

	T (K)	z_T (km)	ξ	z_ξ (km)	V_S	z_{V_S} (km)
σ étape 1	150	100	0.03	150	0.05	100
σ étapes 2 et 3	70	70	0.01	100	0.02	70

TAB. 4.3 : Valeurs des écart-types des fonctions gaussiennes utilisées pour les 3 étapes de la procédure.

4.4.2.2 Échantillonnage continu/discontinu

Jusqu'à présent nous avons considéré que l'espace des états possibles pour un paramètre donné était continu dans un intervalle. Qu'en est-il si l'espace des modèles est supposé discret ? En effet, diminuer le nombre de valeurs possibles pourrait être un moyen de diminuer le nombre d'itérations nécessaires pour obtenir une distribution de probabilité *a posteriori* représentative. Un test a été réalisé en estimant que ξ pouvait uniquement prendre des valeurs discrètes tous les 0.02, de 0.9 à 1.1. Ces 11 valeurs possibles sont considérées comme étant équiprobables. Pour 16 racines de nombres aléatoires différentes (une racine par processeur), le même modèle synthétique a été inversé en T , ξ et V_S dans le cas où ξ est échantillonnée de façon continue, et dans le cas où ξ ne peut prendre que des valeurs discrètes. L'algorithme a fonctionné pendant 60 000 itérations et les taux d'acceptance sont dans les deux cas autour des 20 %, bien que celui-ci soit globalement plus faible dans le cas de l'échantillonnage discret. Pour chaque processeur, le plus petit misfit obtenu au cours de ces 60 000 itérations a été sélectionné. Les résultats sont synthétisés sur la figure 4.14. Excepté pour le processeur numéro 4, tous les misfits obtenus avec un tirage en continu sont inférieurs à ceux calculés avec un tirage discontinu, ce qui indique une plus grande efficacité de l'échantillonnage continu.

Dans cet exemple, seules les valeurs de ξ ont été sélectionnées de façon discrète, mais cet effet est encore plus important lorsque tous les paramètres de l'inversion sont forcés de prendre uniquement des valeurs discrètes, et les misfits obtenus sont encore plus grands. Les taux d'acceptance restent extrêmement faibles, ce qui signifie que les paramètres ne peuvent pas bouger facilement dans l'espace des modèles et la marche aléatoire est en quelque sorte bridée. Pour atteindre une configuration favorable, il se peut que la chaîne doive transiter par des configurations qui produisent des misfits plus élevés, mais dont les valeurs des paramètres sont relativement proches des valeurs vraies. Si seules des valeurs discrètes sont autorisées, le passage par ces états transitoires est

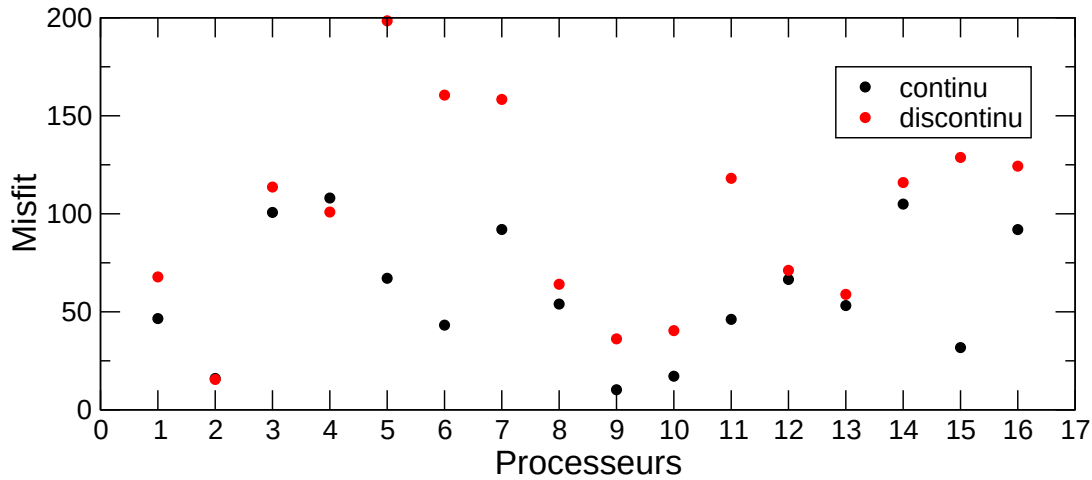


FIG. 4.14 : Représentation des plus faibles misfits obtenus lors d’une inversion de T , ξ et V_S réalisée pour 16 racines de nombres aléatoires différentes (16 processeurs). 60 000 itérations ont été effectuées. Les résultats pour un tirage aléatoire continu (points noirs) et discontinu (points rouges) des points de Bézier représentant ξ sont montrés.

difficile. La probabilité de sélectionner une combinaison de nouvelles valeurs qui soient acceptées par l’algorithme de Metropolis-Hastings est donc très faible.

Dans ce travail, l’algorithme est construit de façon à ce que les paramètres soient échantillonnés avec la plus haute résolution possible, ainsi, l’échantillonnage est systématiquement effectué avec une fonction gaussienne continue.

4.4.2.3 Tirage au sort des paramètres

Plusieurs stratégies de tirage au sort des paramètres sont possibles. Tous les couples de paramètres (x, z) peuvent être modifiés à chaque itération, ou bien un seul couple (x_i, z_i) à chaque fois. Modifier l’ensemble des paramètres permet de passer rapidement à une configuration complètement indépendante. En revanche, comme l’ont souligné [Khan & Mosegaard \(2002\)](#), la stratégie de changer un seul couple (x_i, z_i) à chaque pas permet, lorsqu’une configuration satisfaisante a été atteinte (le misfit correspondant est faible), de préserver une partie de ses caractéristiques à l’itération suivante.

Sur la figure 4.15 sont représentées les fonctions d’autocorrélation de la température pour 3 points de Bézier qui constituent les paramètres d’un profil de température. Nous insistons qu’il s’agit là uniquement des configurations retenues par l’algorithme, et non de toutes celles testées. L’inversion a été réalisée seulement pour la température et un écart-type de la fonction gaussienne de 70 (K et km) est fixé. La fonction d’autocorrélation tend vers zéro d’autant plus rapidement que ses éléments sont décorrélés. Lorsque tous les paramètres sont modifiés à chaque itération, il faut attendre environ 2300 itérations avant que toutes les chaînes « perdent la mémoire » de la configuration initiale. En

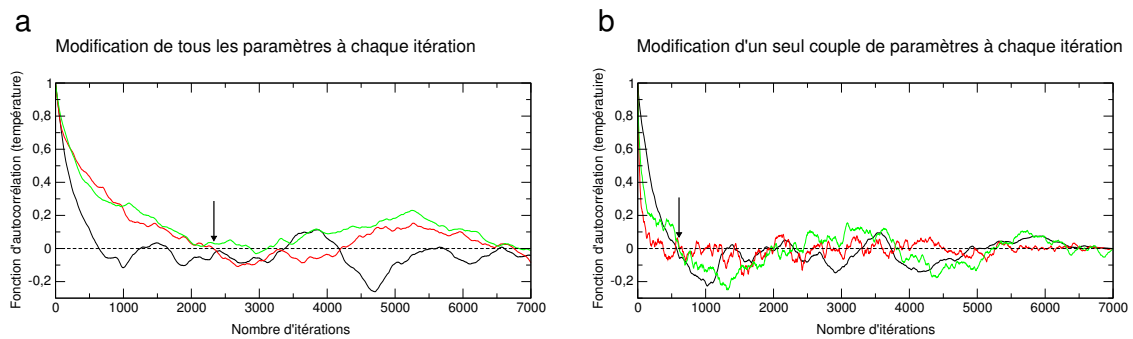


FIG. 4.15 : Autocorrélation de la température pour 3 points de Bézier. (a) L'ensemble des paramètres (T et z_T) est modifié à chaque itération. (b) Un seul couple de paramètres (T_i, z_{T_i}) est changé par itération. Tous les couples sont équiprobables et l'écart-type de la fonction gaussienne vaut 70 (K et km). Les flèches indiquent le nombre d'itérations au bout desquelles toutes les fonctions d'autocorrélation ont atteint au moins une fois la valeur de zéro.

d'autres termes, c'est seulement après 2300 itérations qu'un jeu de paramètres indépendants est obtenu. En revanche, il ne faut que 600 itérations si un seul couple de paramètres est changé à chaque pas. Le taux d'acceptance des modèles est d'environ 20 % dans ce cas, mais n'est que de 2 % si tous les paramètres sont modifiés. Pour atteindre un taux d'acceptance proche des 20 %, il faudrait diminuer l'écart-type à une valeur de 10, ce qui ne permet pas aux paramètres de réaliser de grands pas dans l'espace des modèles d'une itération à une autre.

D'après ce test, changer tous les paramètres à la fois implique un faible taux d'acceptance, excepté dans le cas d'écart-types très faibles. Ce dernier est un critère important pour s'assurer que suffisamment de modèles sont acceptés et ainsi obtenir une pdf représentative. Cependant, même si ce taux paraît correct (proche de 23%), ce n'est pas pour autant que la convergence vers le régime stationnaire va être plus rapide. Effectivement, le nombre de modèles acceptés peut être satisfaisant, mais l'écart-type tellement faible qu'il faudrait réaliser un nombre d'itérations considérable pour atteindre une configuration des paramètres favorable et parvenir à la période stationnaire. Nous avons choisi de tirer parti des deux types de tirage au sort décrits dans cette section. Dans l'étape 1, nous avons choisi de conserver de grands pas, bien que le taux d'acceptance soit très faible, afin de tester un maximum de configurations différentes. L'algorithme conserve la même configuration de paramètres pendant un grand nombre d'itérations (voir figure 4.7), mais l'avantage est que toutes les configurations testées sont extrêmement différentes les unes des autres, ce qui permet une large exploration de l'espace des modèles. Puis le meilleur jeu de paramètres trouvé à la fin de l'étape 1 est affiné avec l'étape 2, où cette fois-ci un seul couple de paramètre est modifié à chaque itération avec des écart-types qui produisent des taux d'acceptance satisfaisants.

4.4.3 La fonction coût (misfit)

4.4.3.1 Normes L_1 et L_2

Le choix de l'expression de la fonction coût est crucial, car nous avons vu dans le paragraphe 3.3.2 que le fait qu'un modèle soit accepté ou rejeté par l'algorithme de Metropolis-Hastings repose entièrement sur la valeur du misfit. En effet, il s'agit du seul critère quantitatif que nous possédons qui permet de juger la crédibilité d'un modèle testé par rapport aux données, compte-tenu des incertitudes associées. Les expressions les plus couramment employées sont les normes L_1 et L_2 (équation 3.7). Tarantola (2005) encourage l'utilisation de la norme L_1 , qui est plus robuste si des valeurs aberrantes sont contenues dans les données. L'utilisation de la norme L_1 revient à considérer que l'incertitude estimée suit une distribution de Laplace. L'emploi de la norme L_2 est basé sur l'hypothèse que le bruit est gaussien.

J'ai étudié la sensibilité de ces deux expressions à de petites variations des points de Bézier, en température et en profondeur, pour un profil adiabatique réaliste (une température potentielle de 1600 K et un gradient de 0.4 K/km). Le profil adiabatique considéré est représenté sur la figure 4.16e. Il est composé de 5 points de Bézier localisés tous les 100 km, de 400 à 800 km. Ces points sont modifiés un à un en température et en profondeur. Les résultats sont présentés sur les figures 4.16 et 4.17, à chaque fois pour les normes L_1 et L_2 . La valeur en abscisse des graphiques δm est un estimateur de la « distance » du point modifié par rapport au point de Bézier du profil adiabatique de référence et vaut

$$\delta m = \sqrt{(T_{\text{ref}} - T)^2 + (z_{\text{ref}} - z)^2}. \quad (4.7)$$

Chaque point de Bézier du profil adiabatique est représenté par une couleur différente. On considère que $\Delta z = z_{\text{ref}} - z$ et $\Delta T = T_{\text{ref}} - T$.

Les figures 4.16a et c présentent la valeur du misfit pour un déplacement des points en température de ± 40 K et pour $\Delta z = 0$ km, tandis que les figures 4.16b et d correspondent à un mouvement en profondeur de ± 20 km pour $\Delta T = 0$ K. De manière générale, lorsque $\Delta z = 0$ km et $\Delta T = 0$ K, c'est-à-dire si l'on modifie soit la température, soit la profondeur pour un point de Bézier donné, la fonction $\delta m = f(\text{misfit})$ pour la norme L_2 montre une forme parabolique, tandis que celle de la norme L_1 présente un comportement linéaire. Avec l'utilisation d'une norme L_2 , il va être facile pour une chaîne de Markov de converger et de rester aux alentours de la valeur vraie. En revanche, pour la norme L_1 , la chaîne va avoir plus de mal à se dégager de la valeur recherchée et à échantillonner autour pour rendre compte de l'incertitude sur les paramètres. De plus, l'algorithme rejettera plus de modèles avec une norme L_1 car pour un modèle proche de celui de référence, les valeurs de misfit sont plus importantes que pour une norme L_2 . C'est pour ces raisons que la norme L_2 est préférée dans la suite de ce travail.

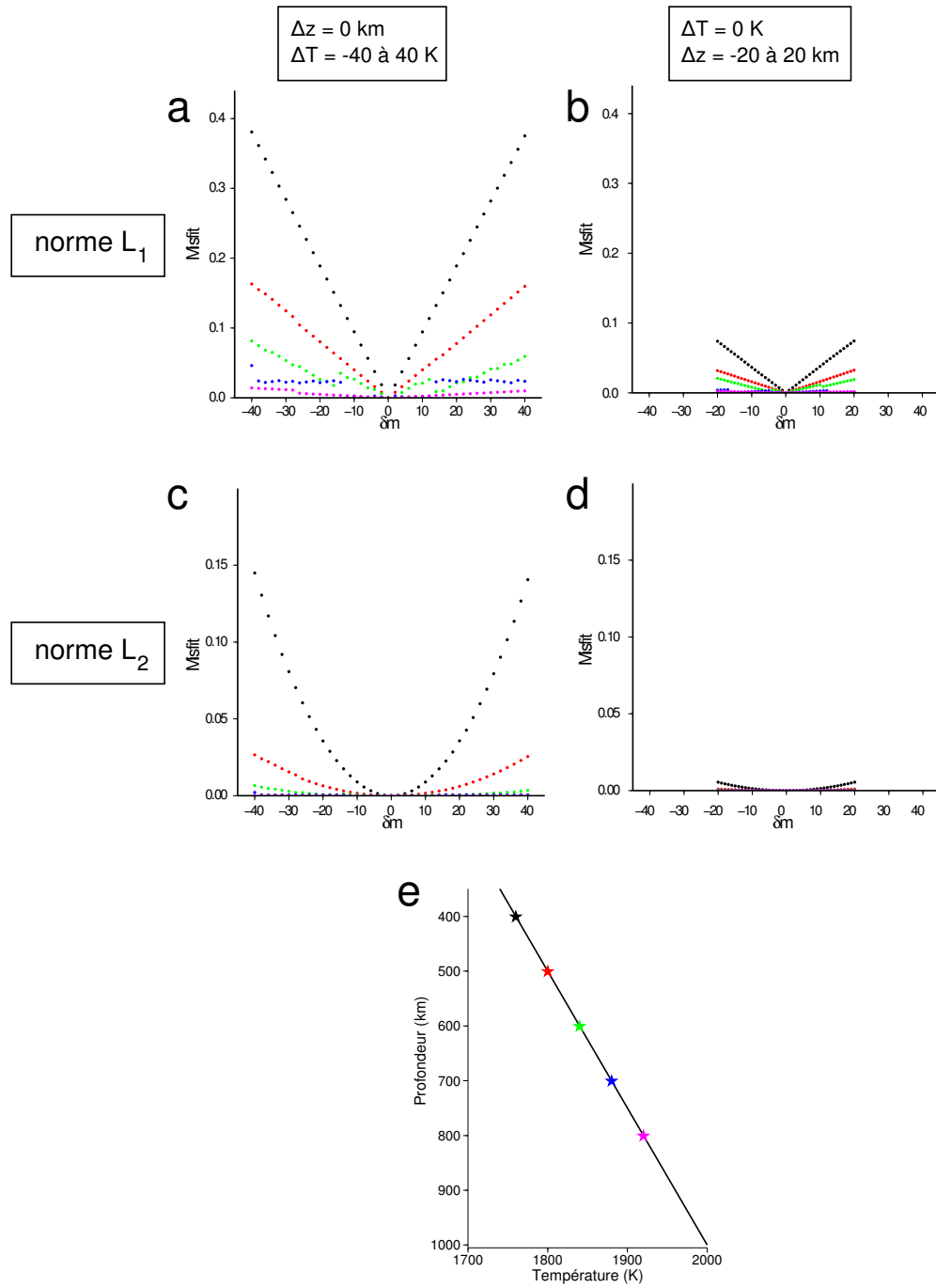


FIG. 4.16 : Influence de la modification des points de Bézier d'un profil adiabatique dans le diagramme (température-profondeur) sur la valeur du misfit pour les normes L_1 et L_2 . (a) et (c) sont les résultats pour la norme L_1 , et (b) et (d) pour la norme L_2 . (e) représente le profil adiabatique de référence et ses points de Bézier. À chaque point correspond une couleur. (a) et (c) montrent l'évolution du misfit pour Δz fixée aux profondeurs des points représentés en (e) mais pour ΔT variable. (b) et (d) décrivent les valeurs du misfit pour ΔT fixée aux températures des points représentés en (e) mais pour Δz variable.

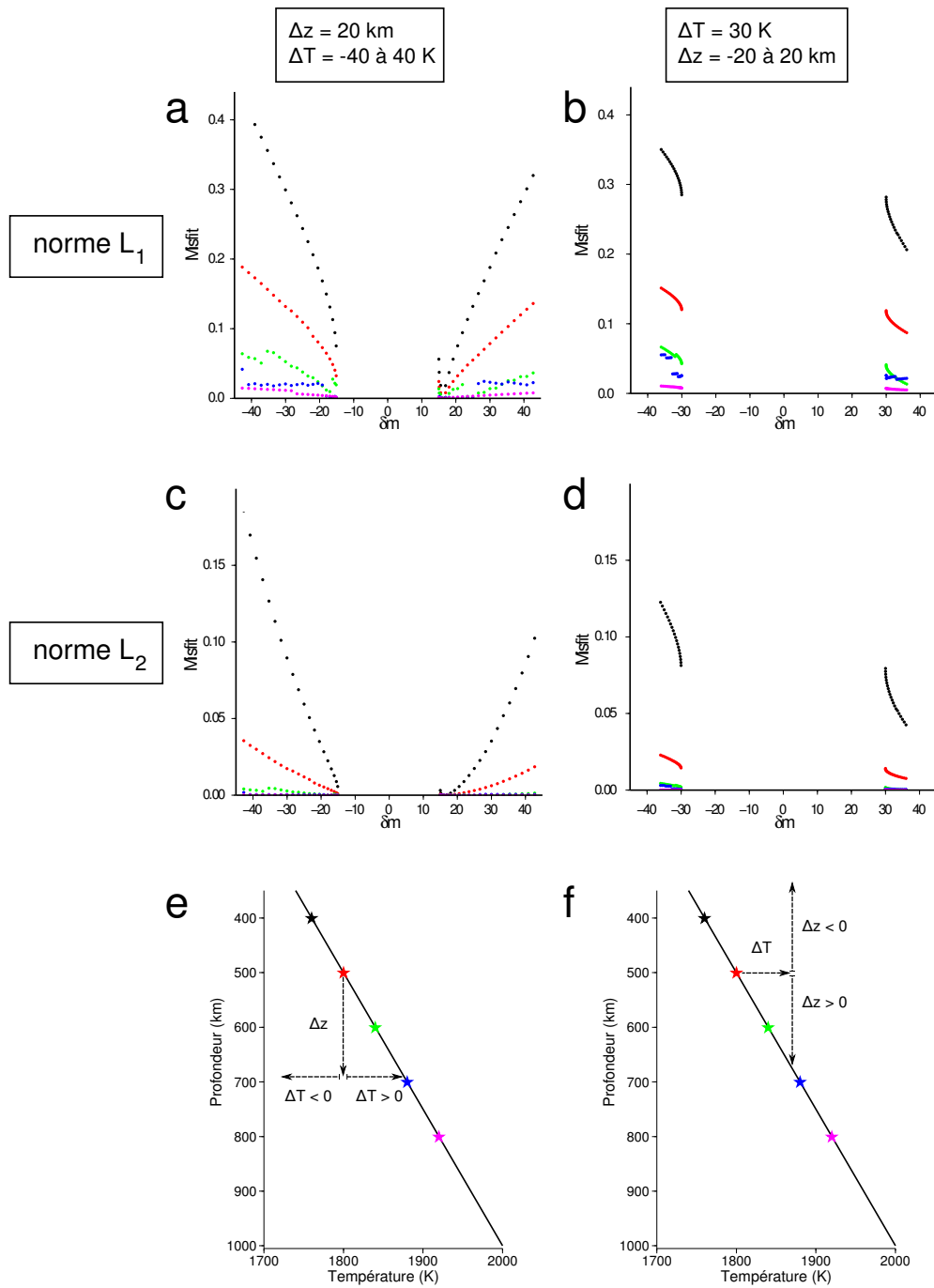


FIG. 4.17 : Influence de la modification des points de Bézier d'un profil adiabatique dans le diagramme (température-profondeur) sur la valeur du misfit pour les normes L_1 et L_2 . (a) et (c) sont les résultats pour la norme L_1 , et (b) et (d) pour la norme L_2 . (e) et (f) représentent le profil adiabatique de référence et ses points de Bézier. À chaque point correspond une couleur. (a) et (c) montrent l'évolution du misfit pour Δz fixée mais pour ΔT variable. (b) et (d) décrivent les valeurs du misfit pour ΔT fixée mais pour Δz variable.

Si les points de Bézier sont maintenant fixés à $\Delta z = 20$ km (figures 4.17a et c) et $\Delta T = 30$ K (figures 4.17b et d), et que l'on change respectivement la température et la profondeur des points de Bézier, les fonctions $\delta m = f(\text{misfit})$ ne sont plus symétriques. Dans le cas où $\Delta z = 20$ km, une augmentation de la température ($\Delta T > 0$) produit des valeurs de misfit plus faibles, car les points de Bézier se rapprochent du profil adiabatique (figure 4.17e). De la même manière, pour $\Delta T = 30$ K, c'est en s'enfonçant en profondeur ($\Delta z > 0$) que les points de Bézier retombent sur le profil adiabatique (figure 4.17f). Ainsi, il apparaît que dans le cas d'un profil de température relativement lisse, des positions différentes des points de Bézier peuvent conduire au même modèle 1D. Comme mentionné dans la section 4.3.5, cet exemple simple renforce la conviction pour cette étude particulière de ne pas se fier uniquement à la trace des paramètres pour détecter la convergence de l'algorithme.

Un autre aspect important est que pour des valeurs de ΔT et de Δz données, le fait de bouger les points de Bézier les plus proches de la surface produit un impact plus important sur les valeurs du misfit que le mouvement des points les plus profonds, ce qui est cohérent avec la décroissance de la sensibilité des ondes de surface avec la profondeur. Notons que les points situés sur le profil adiabatique à 600 km (en vert) et 700 km (en bleu) de profondeur n'ont pas des valeurs de misfit qui se comportent de façon linéaire en fonction de la « distance » δm . Ce comportement est lié aux transitions de phase dans les systèmes olivine et pyroxènes-grenat. Dans la section suivante, nous allons illustrer l'importante influence de la discrétisation de ces transitions en termes de discontinuités sismiques sur la valeur du misfit.

4.4.3.2 Influence de la discrétisation des profils de vitesses

Les profils de température sont paramétrés en courbes de Bézier, ce qui implique que la valeur de la température est connue pour n'importe quelle profondeur. Cependant, il n'est pas possible d'obtenir des profils de V_S , V_P et ρ continus dans la zone de transition, puisque leurs valeurs résultent de calculs non-linéaires (voir section 2.2). Ils doivent donc être discrétisés en fonction de la profondeur. J'ai étudié l'impact du choix de la discrétisation des profils de vitesses sismiques sur le calcul des vitesses de phase et du misfit.

Les transitions de phase les plus complexes ont lieu entre 600 et 700 km de profondeur. Sur un profil de température adiabatique, un des points de Bézier a été placé à 610 km de profondeur pour une température de 1810 K. Il constitue le profil de référence. 6 nouveaux modèles de température ont été construits en diminuant la température à 610 km de profondeur tous les 50 K. Les profils de vitesse de cisaillement correspondant ainsi que les valeurs du misfit ont été calculés. Les calculs sont réalisés uniquement

pour le mode fondamental, pour 34 périodes couvrant la gamme 50- 273 s. La figure 4.18a présente les résultats pour une discrétisation de V_S tous les 1 km et tous les 10 km. Dans le cas où la discrétisation est fine, plus la température est éloignée de sa température de référence à 1810 K, plus la valeur du misfit augmente. Le gradient de vitesse vers 660 km de profondeur est également cohérent. La transformation minéralogique de la ringwoodite en pérovskite et magnésiowüstite étant endothermique, la discontinuité est de plus en plus profonde au fur et à mesure que la température diminue.

Cependant, dans le cas où la discrétisation est plus grossière, l'évolution du misfit en fonction de la température n'est pas linéaire. En effet, une température de 1660 K fournit un meilleur misfit que pour une température de 1710 K. Le graphique représentant V_S montre que les discontinuités ont été trop grossièrement échantillonnées, ce qui crée des discontinuités artificielles qui ne rendent pas compte de la complexité des modèles. Le même effet est produit lorsque des profils discrétisés tous les 10 km sont comparés à un profil de référence échantillonné tous les kilomètres (figure 4.18b). Le modèle correspondant à la température « vraie » n'est même pas celui qui présente le plus faible misfit. En revanche, si les profils discrétisés tous les 1 km sont comparés au profil de référence échantillonné tous les 10 km, une tendance linéaire est retrouvée.

Ces tests montrent qu'il est indispensable d'échantillonner finement les gammes de profondeurs où ont lieu les transitions de phase, sous peine d'obtenir des effets non-linéaires liés à l'échantillonnage et de biaiser les résultats de l'inversion. En effet, une discrétisation fine autour de la discontinuité à 660 km de profondeur permet ici d'obtenir un comportement linéaire du misfit en fonction de la modification du profil en température (figure 4.18c). Cependant, ces conclusions restent valides pour chaque mode pris séparément. Un résultat supplémentaire est que le mode fondamental, théoriquement peu sensible aux profondeurs supérieures à 400 km, est capable de détecter des variations de profondeur de la discontinuité à 660 km liées à des variations de température.

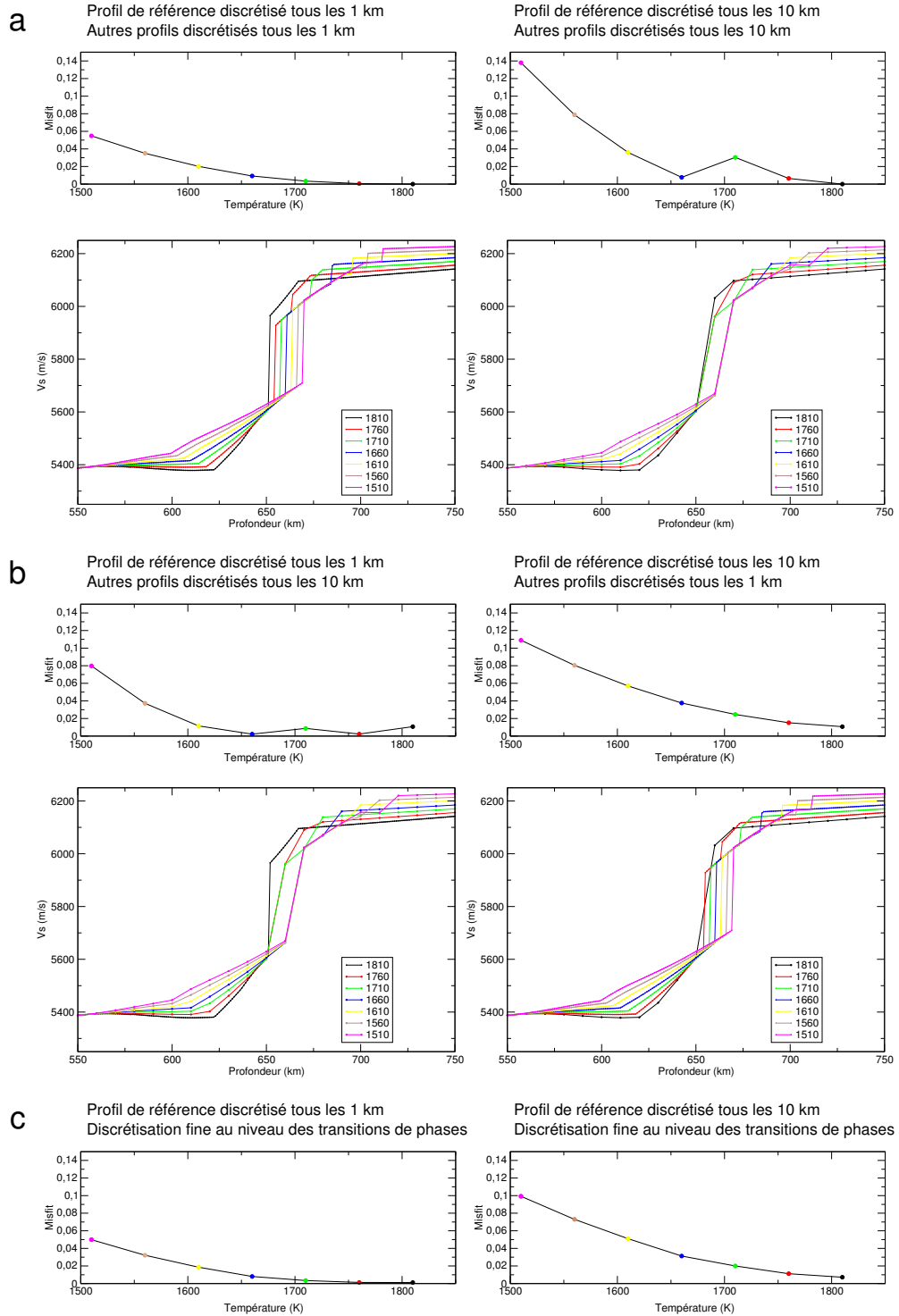


FIG. 4.18 : Influence de la discrétisation de V_S sur les valeurs du misfit. La température d'un point de Bézier sur un profil adiabatique à 610 km de profondeur est modifiée tous les 50 K (le profil de référence est en noir et les autres sont en couleurs). Comparaisons pour (a) une même discrétisation entre le profil de référence et les autres profils, (b) une discrétisation différente, (c) une discrétisation fine au niveau des transitions de phase.

4.5 Ce qu'il faut retenir

La méthodologie présentée dans ce chapitre est une stratégie générale d'inversion des courbes de dispersion des ondes de surface. Le problème inverse est traité avec une méthode de recherche des paramètres entièrement non-linéaire. Bien que certaines conditions *a priori* soient nécessaires pour conserver des profils de T , V_S et ξ réalistes, relativement peu de contraintes sont imposées. L'utilisation d'un nombre de paramètres variable pour la température permet de ne pas imposer un degré de lissage des modèles. Nous avons vu que les pdfs sur les points de Bézier sont relativement uniformes. La parallélisation est indispensable si l'on souhaite conserver un large panel de modèles possibles et ainsi échantillonner adéquatement l'espace des paramètres. L'utilisation de plusieurs processeurs permet de mieux déterminer la variabilité de la gamme de solutions possibles. Réaliser une procédure en trois étapes présente un avantage en temps de calcul, mais aussi en performance. Lors de la première étape, l'espace des paramètres est exploré avec de grands pas, ce qui permet d'explorer rapidement l'espace des modèles. L'utilisation de plusieurs racines de nombres aléatoires permet également de ne pas se focaliser sur une région particulière de l'espace des paramètres. Pour les étapes 1 et 2, seules les meilleures configurations sont conservées, c'est-à-dire celles qui présentent les misfits les plus faibles. Cette stratégie favorise l'atteinte de la période stationnaire pour les processeurs restants. Pour que l'efficacité de la procédure soit optimale, nous avons vu que des tests préliminaires étaient nécessaires, notamment sur les modalités de proposition de nouveaux modèles. Dans le chapitre suivant, l'intégralité de la procédure est testée sur des données synthétiques.

Chapitre 5

Inversion de données synthétiques

5.1 Introduction

La procédure d'inversion est ici validée sur des données synthétiques. Dans cette thèse, ce ne sont pas les données brutes qui sont utilisées, c'est-à-dire le sismogramme, mais les données géophysiques dérivées de ces données brutes : les vitesses de phase. Les tests synthétiques sont indispensables pour évaluer la justesse et la robustesse de la méthode, ainsi que pour mettre en valeur des caractéristiques inhérentes à la théorie qui pourraient être retrouvées sur les données réelles, en considérant des incertitudes réalistes sur les courbes de dispersion. Seuls quelques résultats, obtenus dans des modèles atypiques sont présentés dans ce chapitre. Les modèles synthétiques à retrouver ont été choisis en raison de la difficulté qu'ils présentaient à l'algorithme d'inversion.

Les résultats d'inversion de deux modèles synthétiques sont décrits en détail. Le résultat de la méthode MCMC employée étant un ensemble de modèles, les différentes manières de représenter et d'analyser les distributions *a posteriori* sont également discutées.

Il n'existe pas une manière unique de paramétrer le système étudié. Ainsi, les résultats présentés dans ce travail reflètent la paramétrisation particulière employée.

L'ensemble des calculs, y compris sur les données réelles, a été réalisé au CCIPL¹ (Centre de Calcul Intensif des Pays de Loire) et au CINES² (Centre Informatique National de l'Enseignement Supérieur).

¹<http://www.ccipl.univ-nantes.fr>

²<http://www.cines.fr/>

5.2 Les modèles utilisés

L'algorithme est testé avec des données synthétiques calculées pour des profils T , ξ et V_S connus. Une composition pyrolitique est considérée car il s'agit d'une composition généralement compatible avec une large gamme d'observations pétrologique, géochimique et sismologique (Ringwood, 1975). La teneur en fer est considérée comme étant constante à travers l'ensemble de la zone de transition et vaut $\text{Fe}\# = 0.11$ (voir section 2.2.3).

Pour illustrer le potentiel de l'algorithme à retrouver la structure de la zone de transition, deux modèles synthétiques en T , ξ et V_S ont été choisis. Ils sont présentés sur la figure 5.1. Les points de Bézier sont représentés avec des étoiles noires et les profils 1D avec des courbes noires. Cette convention sera conservée dans toute la suite du manuscrit.

En température, les modèles ont été volontairement éloignés d'un profil adiabatique (en pointillés sur la figure 5.1), pour démontrer la capacité de l'algorithme à détecter des variations de température importantes, lorsque le trajet des ondes entre la source et le récepteur croisent une plaque plongeante ou un point chaud par exemple.

Actuellement, les différentes études s'accordent sur le fait que la Terre est radialement anisotrope dans les 220 premiers kilomètres, comme dans le PREM. Cependant, il n'existe pas de réel consensus sur le fait que de l'anisotropie radiale soit présente au-dessous de cette zone anisotrope. C'est pourquoi nous avons testé l'habileté de l'algorithme à détecter de l'anisotropie à la fois dans le manteau supérieur (modèle 1) et dans la zone de transition (modèle 2). Le modèle 1 a été choisi de telle manière que la perturbation sur les vitesses de phase produite par une variation de température comparée à un profil adiabatique soit deux fois supérieure à une variation provoquée par l'anisotropie. Avec le modèle 2, nous avons voulu tester la capacité des ondes de surface à détecter de l'anisotropie radiale dans la zone de transition en plaçant une anomalie à 500 km de profondeur. Son amplitude est similaire à celle du modèle 1 et est égale à $\xi = 0.95$ à son maximum.

Enfin, en ce qui concerne les profils de V_S entre la surface et 350 km de profondeur, une rupture de pente conséquente a été placée dans la première centaine de kilomètres pour les 2 modèles. Elle symbolise dans ces tests la transition lithosphère-asthénosphère. Entre 80 et 250 km de profondeur, le profil du modèle 2 est placé relativement proche de la limite inférieure du prior.

L'objectif est de déterminer si l'algorithme est capable de détecter l'ensemble de ces caractéristiques et de décorréler les effets des différents paramètres. Le nombre de paramètres utilisés ainsi que les contraintes employées pour calculer les courbes de Bézier sont resumés dans le tableau 5.1. À travers le manteau, un gradient adiabatique de 0.4 K/km est supposé (Katsura et al., 2010). Une liberté importante est laissée pour le

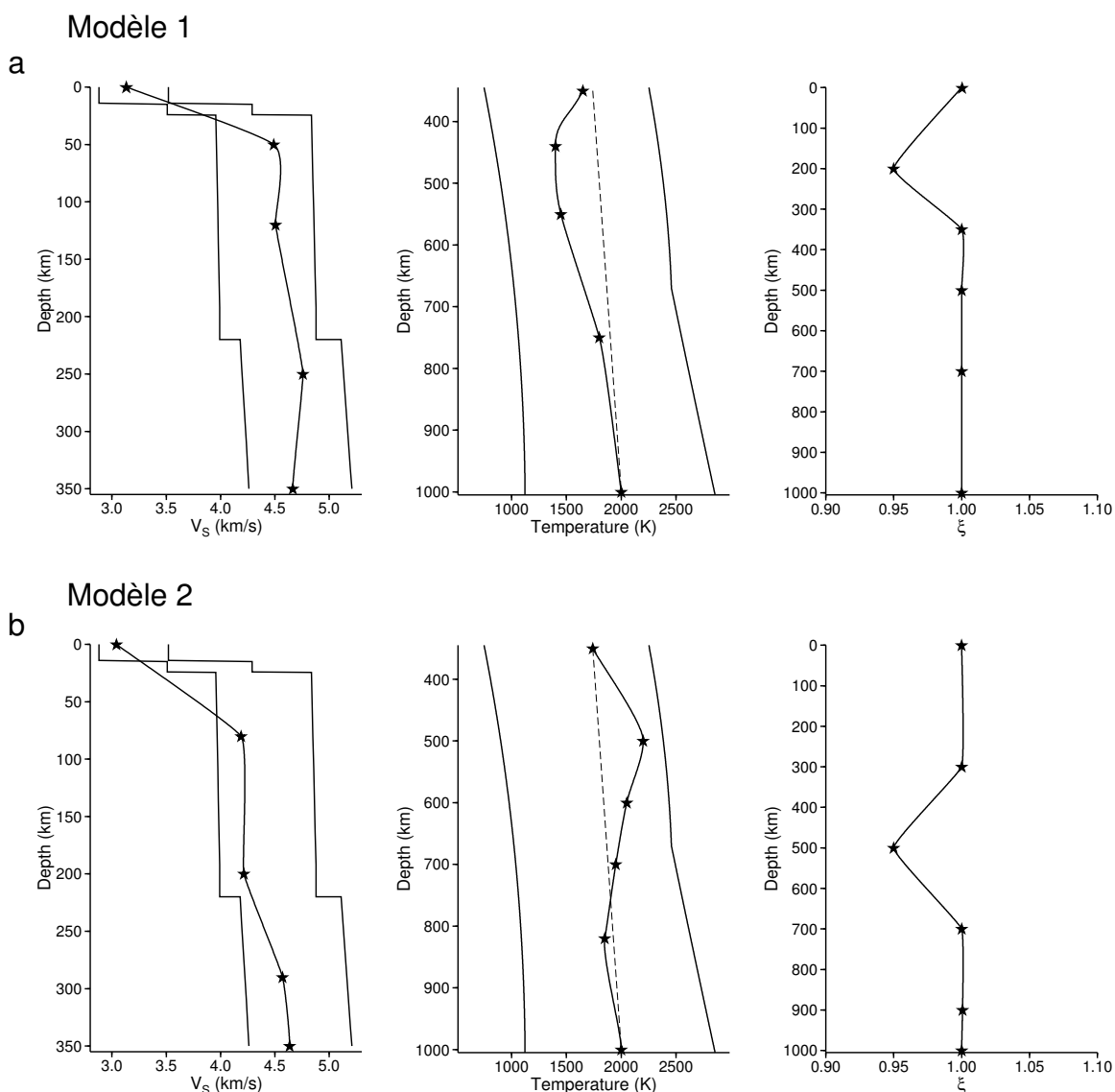


FIG. 5.1 : Représentation des modèles synthétiques 1 (a) et 2 (b) en V_s , T et ξ . Les profils correspondent aux courbes noires. Les étoiles noires sont les points de Bézier. La ligne en pointillé est un adiabat à 1600 K avec un gradient de température de 0.4 K/km.

placement des points de Bézier en température, puisqu'un gradient maximal entre deux points consécutifs de 15 K/km est autorisé. Cette valeur, bien qu'irréaliste, a été choisie dans le but d'évaluer si les modèles d'entrée pouvaient être retrouvés par la procédure d'inversion en considérant des contraintes très faibles sur les paramètres.

Le nombre de processeurs et d'itérations réalisées, ainsi que le temps de calcul pour les trois étapes de la procédure d'inversion sont indiqués dans le tableau 5.2. Pour l'étape 1, les 1600 processeurs sont divisés en 5 groupes égaux, qui fonctionnent chacun respectivement avec 6, 7, 8, 9 et 10 points de Bézier pour la température. Au total, 4.8

millions de modèles ont été échantillonnés durant la période stationnaire. Pour éliminer des modèles corrélés dans l'ensemble de la solution, puisqu'un seul couple de paramètres est modifié à chaque itération, seul un modèle sur 40 a été sélectionné pour analyser les résultats, ce qui donne un ensemble de 120 000 modèles.

	T	ξ	V_S
nombre de points de Bézier	6 à 10	7	5
gamme de profondeurs (km)	350- 1000	0- 1000	0- 350
distance minimale d_z entre 2 points consécutifs (km)	80 à 10	40	20
taille du vecteur tangent local Δ (km)	40 à 5	20	10
gradient maximal entre 2 points consécutifs	15 K/km		

TAB. 5.1 : Synthèse des contraintes employées pour le positionnement des points de Bézier et des points de définition.

Pour obtenir de l'information sur la zone de transition, l'utilisation des modes harmoniques est indispensable car ils échantillonnent l'intervalle entre 250 et 1000 km de profondeur, ce qui complète la forte sensibilité du mode fondamental aux couches superficielles (voir section 2.4.3). Les données synthétiques employées sont les courbes de dispersion de Love et de Rayleigh du mode fondamental et des trois premiers harmoniques. Elles ont été calculées par le problème direct détaillé dans le chapitre 2 à partir des 2 modèles définis au préalable (figure 5.1). Toutes les courbes de dispersion sont inversées de façon simultanée. En moyenne, 35 vitesses de phase par mode sont inversées. Elles couvrent la gamme de périodes 45- 273 s. Des incertitudes réalistes sur les vitesses de phase sont prises en compte. Elles sont fixées d'après l'étude de [Beucler et al. \(2003\)](#), où des mesures des vitesses de phase des modes harmoniques jusqu'à l'ordre 6 ont été réalisées. Ces incertitudes varient entre 0.04 et 2.4 % par rapport aux vitesses de phase synthétiques.

	étape 1	étape 2	étape 3
nproc	1600	320	80
niter	6000	60 000	60 000
temps de calcul	2 h	20 h	20 h

TAB. 5.2 : Nombre de processeurs **nproc** et nombre d'itérations **niter** employés au cours des trois étapes de la procédure d'inversion. Le temps de calcul est également indiqué.

5.3 Les différents types de représentation

La densité de probabilité *a posteriori* est la solution du problème inverse considéré. Une fois les modèles échantillonnés, la question suivante se pose : comment analyser la distribution de ces modèles ? Si l'on souhaite regarder l'information sur un paramètre à une profondeur fixée, les probabilités marginales 1D sont appropriées, mais elles ne permettent pas de visualiser les informations sur les autres paramètres. C'est pourquoi les représentations les plus employées sont les cartes de densité de probabilité marginales (2D), comme celles présentées sur la figure 4.11 montrant les densités de probabilité *a priori*.

La solution à la formulation bayésienne des problèmes inverses résulte en une densité de probabilité et non en un unique estimateur, bien que ceux-ci puissent être obtenus par intégration numérique. Il existe une assez grande liberté pour estimer des profils 1D uniques à partir de l'ensemble de la solution, par exemple la moyenne, la médiane, le mode, etc. Ce type d'estimateur a déjà été utilisé par [Shapiro & Ritzwoller \(2002\)](#), [Verhoeven et al. \(2009\)](#), ou encore [Bodin et al. \(2012\)](#). Dans ce cas, il faut s'assurer que les marginales 1D possèdent une distribution gaussienne. À partir de la densité de probabilité des courbes de Bézier, qui tient compte de l'ensemble des modèles 1D, il semblerait assez naturel que le modèle appartenant à la région de plus grande probabilité corresponde au modèle le plus représentatif. Il s'agit du mode de la distribution marginale, qui suit le maximum de la distribution en fonction de la profondeur. En réalité, il ne correspond à aucun modèle individuel pris dans l'ensemble de la solution. Cependant, ce modèle n'est pas en général celui le plus probable, puisque nous devons garder à l'esprit que la distribution marginale formée de modèles 1D a été construite en utilisant une juxtaposition d'histogrammes (c'est-à-dire des marginales 1D). Par exemple, cela signifie que si l'on souhaite obtenir la température la plus probable à une profondeur donnée, elle correspond à la température la plus probable à cette profondeur particulière uniquement, sans regarder l'information sur les autres paramètres du modèle.

Des représentations supplémentaires en plus des cartes de densité de probabilité apportent un réel intérêt quant à l'analyse des modèles échantillonnés. Les résultats sur la répartition des points de Bézier doivent également être visualisés, en plus des résultats sur les profils 1D. En effet, nous avons vu sur la figure 4.11 que les pdfs *a priori* sur les profils 1D ne sont pas uniformes. Bien que la construction des courbes de Bézier soient indispensables pour calculer les courbes de dispersion des ondes de Love et de Rayleigh, elles ne correspondent pas strictement aux paramètres du modèle. Ce sont bien les points de Bézier qui constituent les seuls paramètres sur lesquels l'algorithme réalise des tirages aléatoires.

Avant de discuter des résultats de l'inversion des 2 modèles synthétiques, nous allons d'abord détailler les intérêts des 4 différents types de représentation sur lesquels nous allons nous baser pour analyser l'ensemble des résultats. Les mêmes représentations sont utilisées pour les données synthétiques et réelles, et sont employées pour T , ξ , V_S , ainsi que pour les vitesses sismiques calculées à partir de la température par les lois thermodynamiques. Les 4 représentations sont détaillées sur les figures 5.2 et 5.3 pour les résultats en température des modèles 1 et 2, respectivement :

- La répartition des points de Bézier et sa densité de probabilité (figure 5.2a) :

La répartition des points de Bézier acceptés par l'algorithme est montrée en gris. Cette figure permet de voir facilement les zones où les points sont resserrés, ce qui signifie que les paramètres sont plutôt bien contraints, des zones où ils sont étalés, qui révèlent des régions où de nombreuses valeurs des paramètres peuvent rendre compte des données. La profondeur des points n'étant pas fixée, cette représentation est également utile pour voir si les points se concentrent à des profondeurs particulières ou non. Les iso-contours de la densité de probabilité des points de Bézier apportent une information supplémentaire quant à l'espace des paramètres qui a été le plus souvent échantillonné. L'information manquante sur cette figure est qu'il n'est pas possible de savoir quels points sont liés ensemble pour former un profil 1D.

- La répartition des points de Bézier et les valeurs du misfit (figure 5.2b) :

De la même manière que pour la figure 5.2a, tous les points de Bézier sont représentés. L'échelle de couleur correspond à la valeur du misfit. Celle-ci s'étend de 0 à 100 %, car les valeurs du misfit sont mises à l'échelle selon la plus faible et la plus grande valeur rencontrées, ce qui permet de voir comment les points de faible misfit se répartissent entre eux. L'information apportée par le misfit est souvent négligée dans les études bayésiennes, puisque la théorie est basée sur des probabilités, et donc sur la répétition (ou non) de certaines valeurs des paramètres. Cependant, comme mentionné dans la section 4.4.3, le misfit est la seule expression mathématique qui indique si les vitesses de phase du modèle testé sont proches de celles recherchées. Par exemple, si l'algorithme reste longtemps pris au piège dans un minimum local, la probabilité que le paramètre prenne cette valeur sera grande, mais le misfit sera plus élevé que pour le minimum global. Grâce à la parallélisation de l'algorithme, ce risque est minimisé, mais il est intéressant d'observer si les régions de forte probabilité coïncident effectivement avec les régions de faible misfit.

Modèle 1

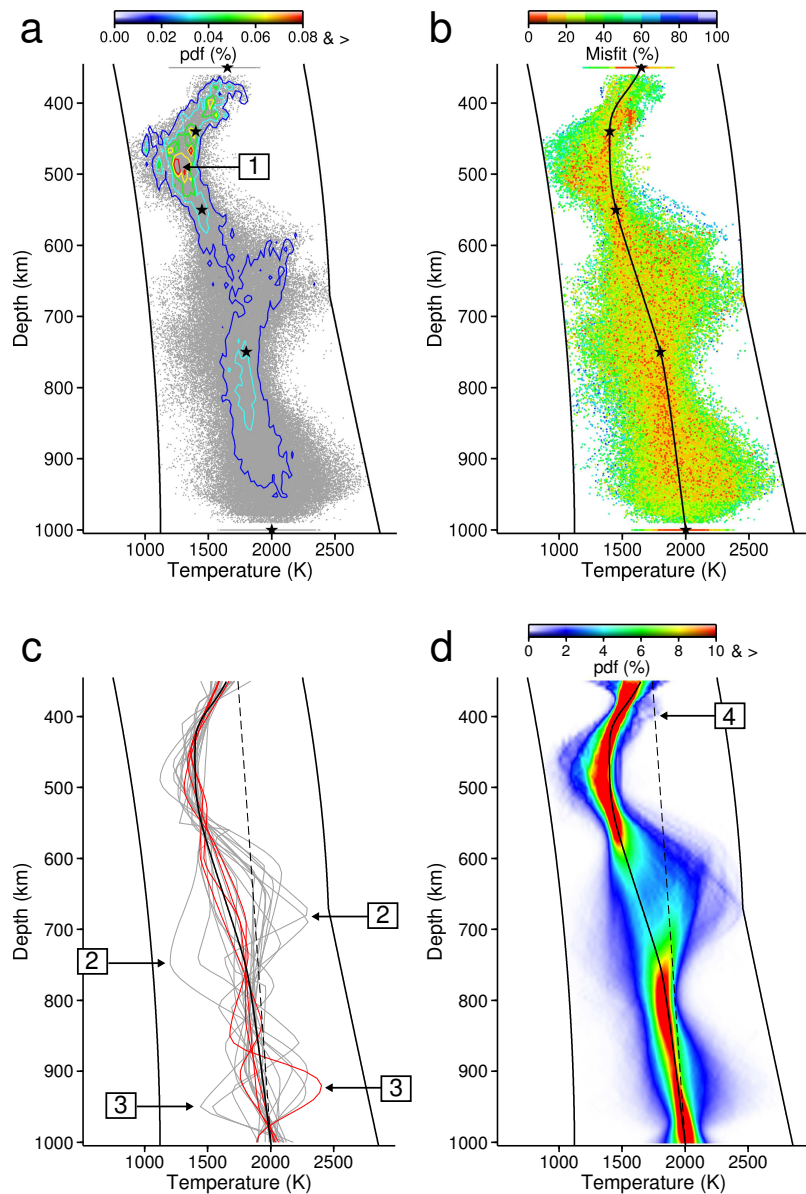


FIG. 5.2 : Résultats en température de l'inversion probabiliste du modèle 1. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier. La ligne en pointillé sur les figures (c) et (d) est un adiabat à 1600 K avec un gradient de température de 0.4 K/km. Les étoiles sur les figures (a) et (b) sont les points de Bézier décrivant les modèles à retrouver.

Modèle 2

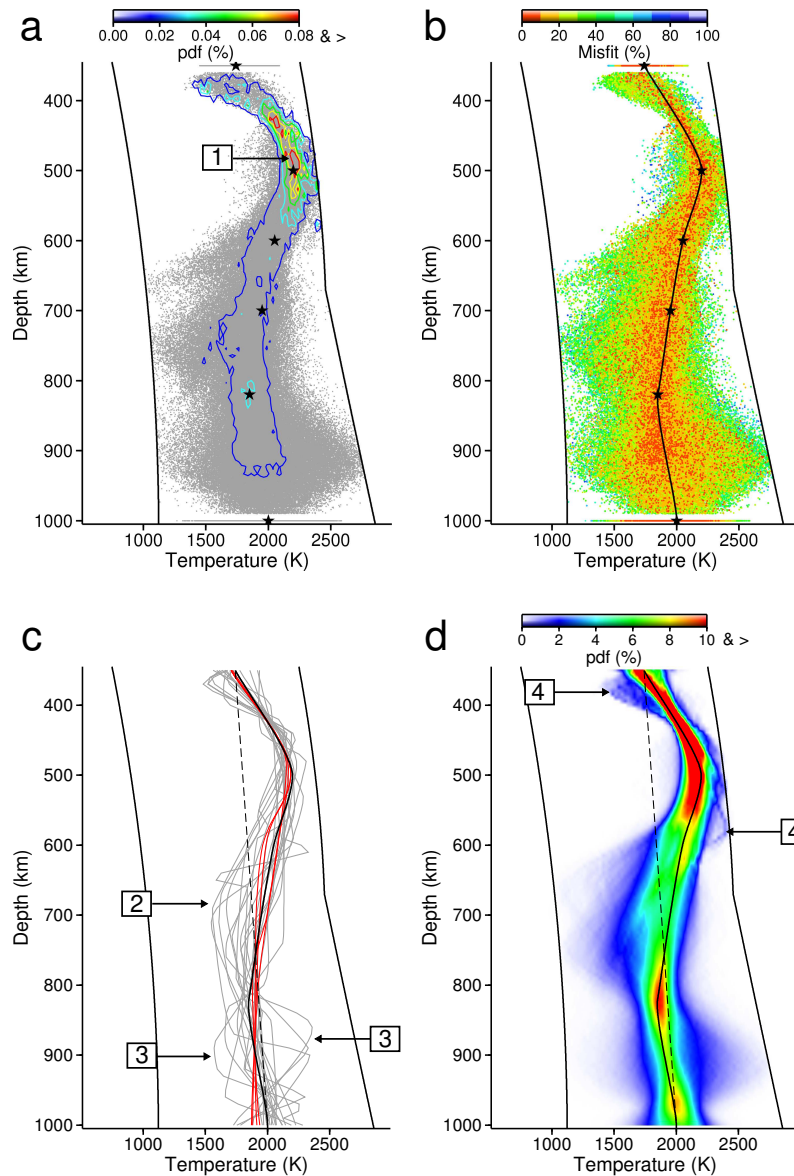


FIG. 5.3 : Résultats en température de l'inversion probabiliste du modèle 2. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier. La ligne en pointillé sur les figures (c) et (d) est un adiabat à 1600 K avec un gradient de température de 0.4 K/km. Les étoiles sur les figures (a) et (b) sont les points de Bézier décrivant les modèles à retrouver.

- Les modèles 1D choisis aléatoirement et les meilleurs modèles (figure 5.2c) :
 Cette figure présente en gris 20 modèles choisis de façon aléatoire parmi l'ensemble des modèles acceptés. Ce type de représentation est recommandée par Mosegaard & Tarantola (1995) et Tarantola (2005). Il s'agit évidemment d'un petit nombre de modèles, et ils ne peuvent pas être utilisés pour réaliser une étude statistique, mais c'est un moyen pratique d'appréhender la diversité des modèles sélectionnés. Ces modèles peuvent montrer des caractéristiques 1D différentes, mais qui produisent toutes une vraisemblance élevée, c'est-à-dire une bonne adéquation aux données. La visualisation de ces modèles est très utile pour se rendre compte des parties des modèles qui sont les mieux résolues et/ou qui correspondent à des caractéristiques récurrentes. En rouge sont montrés les 4 profils correspondant aux 4 meilleurs misfits rencontrés durant les marches aléatoires. Pour les modèles synthétiques, nous verrons qu'ils sont en général proches du modèle recherché.
- La densité de probabilité des courbes de Bézier (figure 5.2d) :
 Cette dernière représentation est la plus couramment employée. Puisqu'à partir d'une courbe de Bézier, les valeurs de T , ξ et V_S sont définies pour chaque profondeur, un histogramme correspondant à la distribution de T , ξ ou V_S à chaque profondeur peut être construit. La juxtaposition de ces histogrammes fournit une carte 2D de densité de probabilité. Cette représentation donne une vue d'ensemble des chemins les plus souvent empruntés par les profils, ainsi que les régions de l'espace des paramètres où ils se recoupent fréquemment. Nous verrons que cette représentation présente une tendance à lisser les résultats et à centrer les pdfs par rapport au prior.

5.4 Résultats

5.4.1 Température

Les résultats en température des modèles 1 et 2 sont représentés sur les figures 5.2 et 5.3. Les deux distributions de probabilité sur les courbes de Bézier contiennent bien les modèles recherchés. Les points de Bézier sont répartis autour du profil à retrouver et leur pdf montre que les plus fortes valeurs sont réparties le long du modèle synthétique. Les misfits sont faibles pour les points situés sur la courbe de Bézier recherchée. Les valeurs de densité de points les plus importantes sont localisées au niveau des profondeurs où les profils montrent des variations (étiquettes 1) : une diminution de 376 K et une augmentation de 400 K comparé à un profil adiabatique pour les modèles 1 et 2, respectivement.

Les pdfs sur les profils 1D ont des valeurs plus importantes entre 350 et 600 km de profondeur que dans le reste de la zone de transition. Les probabilités marginales du modèle 2 à 410 et 660 km de profondeur sont comparées sur la figure 5.4. Seules les températures *a posteriori* dont la fréquence est supérieure à la distribution *a priori* sont considérées comme étant de l'information contenues dans les données. La température recherchée à 410 km de profondeur est sans ambiguïté la plus fréquemment échantillonnée. Ce n'est pas aussi évident pour la température à 660 km de profondeur, où toutes les températures entre 1700 et 2050 K possèdent des fréquences d'échantillonnage comparables. Cette large distribution au-dessous de 600 km de profondeur est liée aux transformations minéralogiques. Différentes transitions de phase et transformations minéralogiques ont lieu pour des variations réalistes de la température du manteau dans les deux sous-systèmes olivine et pyroxènes-grenat (voir section 1.1.1). Sur les figures 5.2c et 5.3c, on observe que les profils correspondant aux meilleurs misfits sont très proches des modèles recherchés. Les profils choisis aléatoirement pourraient être regroupés en plusieurs familles, ce qui est révélateur de la non-linéarité entre les paramètres du modèle et les données. En général, ils s'éloignent du modèle synthétique entre 600 et 800 km de profondeur (étiquettes [2]) et vers 900 km (étiquettes [3]). Il s'agit des profondeurs où les points de Bézier sont les plus étalés dans l'espace des paramètres. La distribution des points est plus resserrée juste au-dessus de 1000 km de profondeur à cause de la condition *a priori* sur le point de Bézier à 1500 km.

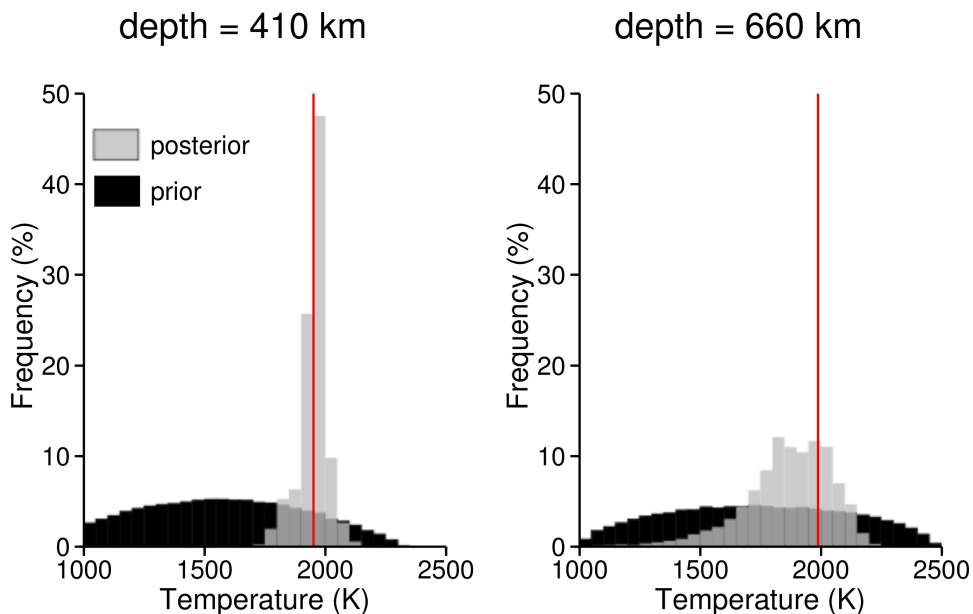


FIG. 5.4 : Probabilités marginales *a priori* (en noir) et *a posteriori* (en gris) sur la température à 410 et 660 km de profondeur pour le modèle 2. Les valeurs vraies sont montrées en rouge (1951 K à 410 km et 1987 K à 660 km). La taille des intervalles est de 50 K.

J'ai réalisé une inversion similaire pour le modèle 1 mais en considérant que la Terre n'était composée que d'olivine pure, sans transition de phase (figure 5.5). Comparée à la figure 5.2, la pdf sur les courbes de Bézier est beaucoup plus resserrée. Il en est de même pour la distribution des points de Bézier. La pdf maximale sur les points épouse

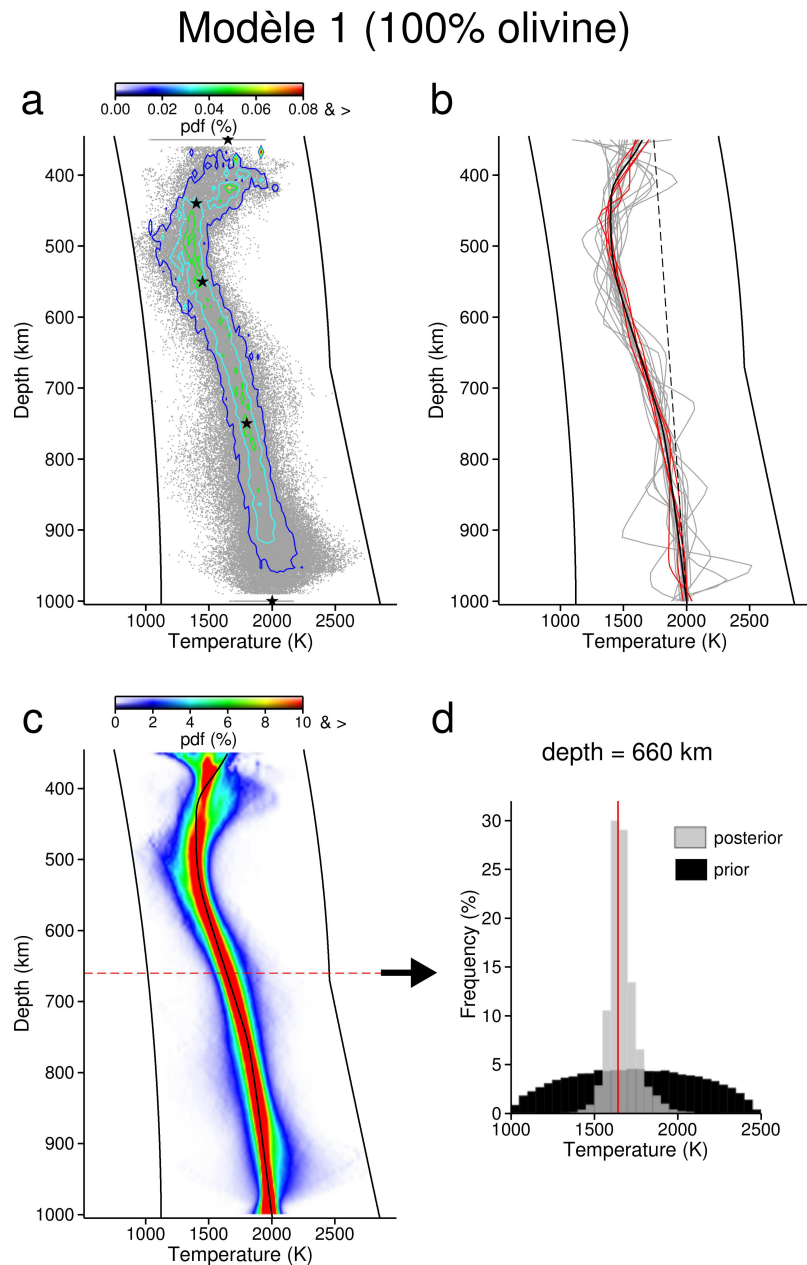


FIG. 5.5 : Résultats en température de l'inversion probabiliste du modèle 1, en considérant une zone de transition composée à 100 % d'olivine pure. (a) Répartition des points de Bézier et sa pdf. (b) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). La ligne en pointillé est un adiabat à 1600 K avec un gradient de température de 0.4 K/km. (c) Pdf des courbes de Bézier. (d) Probabilité marginale à 660 km de profondeur. La valeur vraie, montrée en rouge, vaut 1641 K.

la forme du modèle synthétique. Les profils représentés sur la figure 5.5b ne montrent pas de grandes oscillations entre 600 et 800 km de profondeur. En effet, la probabilité marginale à 660 km révèle que la vraie valeur est retrouvée à 50 K près (figure 5.5d). La forme de l'histogramme est très différente de celui de la figure 5.4 pour un modèle pyrolitique à cette même profondeur, ce qui prouve que la large distribution des modèles entre 600 et 800 km de profondeur sur les figures 5.2 et 5.3 est bien liée aux différentes transformations minéralogiques. En revanche, des oscillations persistent au-dessous de 800 km de profondeur, de la même manière que pour les résultats du modèle avec une composition pyrolitique. Celles-ci sont attribuées à la diminution de la sensibilité des ondes de surface à ces profondeurs.

La figure 5.6 est révélatrice du fait que plusieurs modèles avec des caractéristiques différentes dans la zone de transition sont capables de décrire les mêmes données, compte-tenu des incertitudes sur ces données. Cette figure représente la pdf des vitesses de phase de tous les modèles acceptés par l'algorithme lors de l'inversion du modèle 2. Les distributions se situent entièrement à l'intérieur des barres d'erreur, ce qui montre clairement que tous les modèles échantillonnés sont capables d'expliquer les données. Ici seuls le

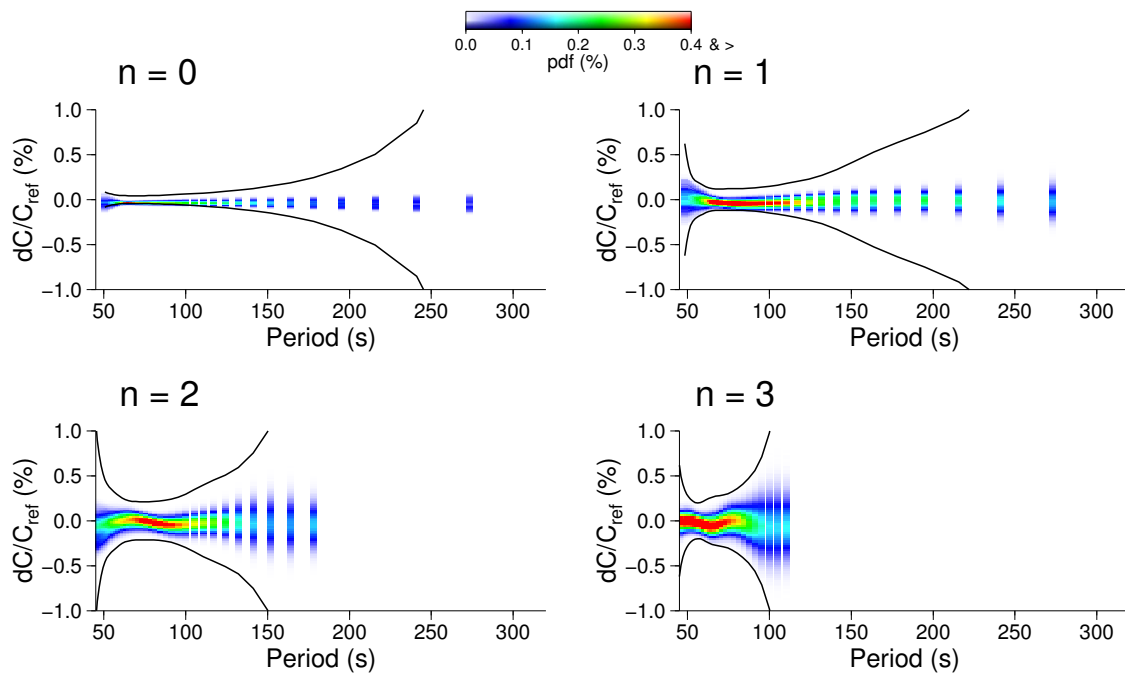


FIG. 5.6 : Distributions *a posteriori* des vitesses de phase du mode fondamental et des modes harmoniques des ondes de Love pour le modèle 2. Les résultats sont montrés en pourcentage de la vitesse de phase testée (C) comparée à la vitesse de phase du modèle 2 (C_{ref}). Les courbes noires correspondent aux incertitudes sur les données de vitesses de phase. Pour les longues périodes, ces incertitudes ne sont pas représentées au-delà de $\pm 1.0\%$.

mode fondamental et les modes harmoniques des ondes de Love sont présentés, mais les pdfs sur les courbes de dispersion des ondes de Rayleigh se situent également dans les barres d'erreur. Cette bonne adéquation entre les vitesses de phase recherchées et les vitesses de phase testées est un argument supplémentaire quant à l'investigation de la zone de transition à partir d'un grand nombre de modèles, plutôt qu'en sélectionnant simplement le modèle qui présente la vraisemblance la plus grande, et d'en déduire la structure physique de la zone de transition directement à partir de ce modèle unique.

Les nombres de points de Bézier sélectionnés pour chacun des 80 processeurs engagés dans la période stationnaire sont synthétisés dans le tableau 5.3. Les modèles 1 et 2 sont composés de 6 et 7 points de Bézier, respectivement. Le nombre correct de points n'est pas celui qui est le plus fréquemment utilisé. Pour le modèle 2, il n'existe pas un nombre de points significativement plus échantillonné par rapport aux autres. Le modèle 1 est plus lisse que le modèle 2, ce qui peut expliquer pourquoi les profils composés de 10 points sont les moins échantillonnés. Lorsque le nombre de points est plus élevé que celui du modèle synthétique, les points de Bézier s'ajustent le long du profil recherché et la forme globale du profil 1D est préservée. Cela explique pourquoi les points de Bézier sont échantillonnés à toutes les profondeurs. En fait, le nombre de points employé n'est pas décisif car la profondeur des paramètres n'est pas fixée. La surparamétrisation des modèles peut engendrer des oscillations supplémentaires des profils (étiquettes [4]), mais les valeurs de misfit correspondant sont élevées. Ces modèles sont effectivement assez éloignés des modèles recherchés et reflètent des régions de faible probabilité.

nombre de points de Bézier	6	7	8	9	10
modèle 1 (6 points)	18.75	17.50	23.75	27.15	12.50
modèle 2 (7 points)	22.50	20.00	18.75	18.75	20.00

TAB. 5.3 : Nombre de points de Bézier (en %) sélectionnés pour les processeurs engagés dans la période stationnaire.

5.4.2 Anisotropie radiale

Les figures 5.7 et 5.8 présentent les résultats sur l'anisotropie radiale. Le résultat important qui ressort de ces figures est que les deux perturbations à 200 et 500 km de profondeur pour les modèles 1 et 2, respectivement, sont retrouvées aux bonnes profondeurs (étiquettes [1]). Des amas de points de Bézier avec des valeurs de misfit faibles sont présents à ces profondeurs, ce qui signifie que les courbes de dispersion sont sensibles aux singularités des profils d'anisotropie aussi bien dans le manteau supérieur que dans la zone de transition.

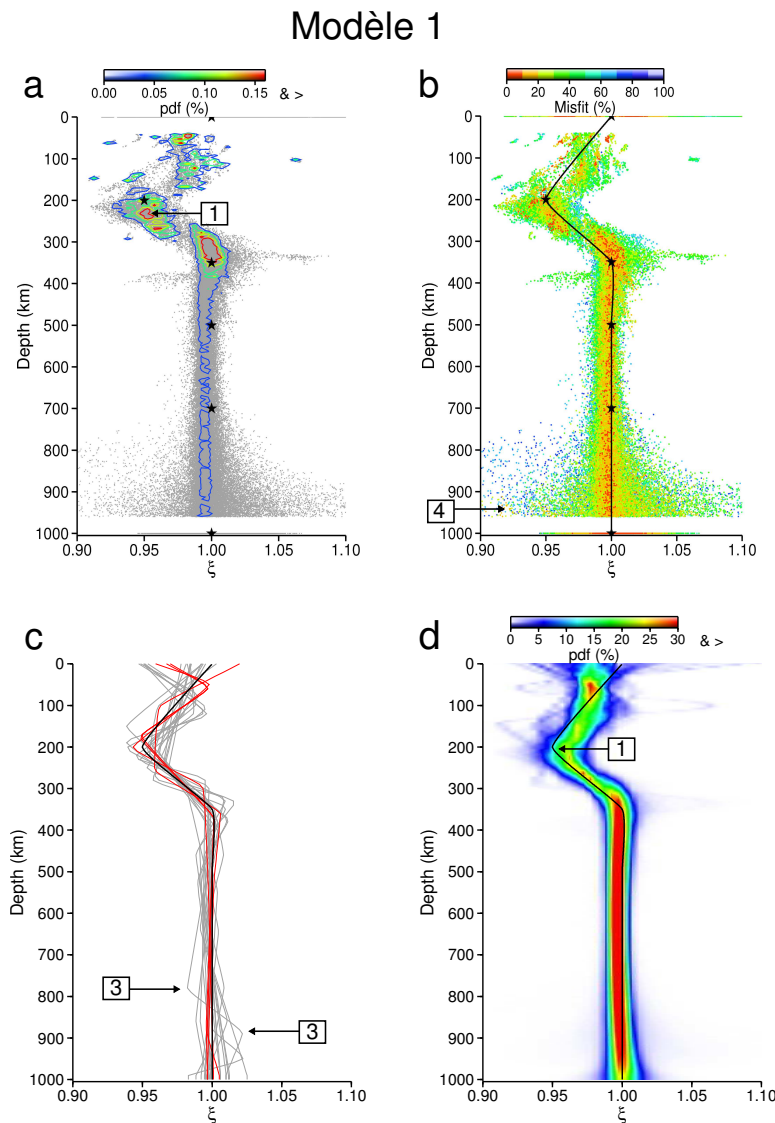


FIG. 5.7 : Résultats en anisotropie de l'inversion probabiliste du modèle 1. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier.

La représentation de l'information en terme de densité de probabilité des profils 1D est trompeuse car elle sous-estime l'amplitude des perturbations. La figure 5.9 montre les probabilités marginales à la fois sur les points de Bézier et sur les courbes de Bézier, à 500 km de profondeur pour le modèle 2. Les deux distributions sont unimodales. Les valeurs les plus échantillonnées sont $\xi = 0.95$ pour les points et $\xi = 0.96$ pour les courbes, la valeur vraie étant $\xi = 0.95$. Bien que non présentées, les mêmes caractéristiques sont retrouvées pour le modèle 1. Au niveau du maximum d'amplitude des perturbations, les pdfs des courbes de Bézier sont clairement décalées vers de plus grandes valeurs de ξ

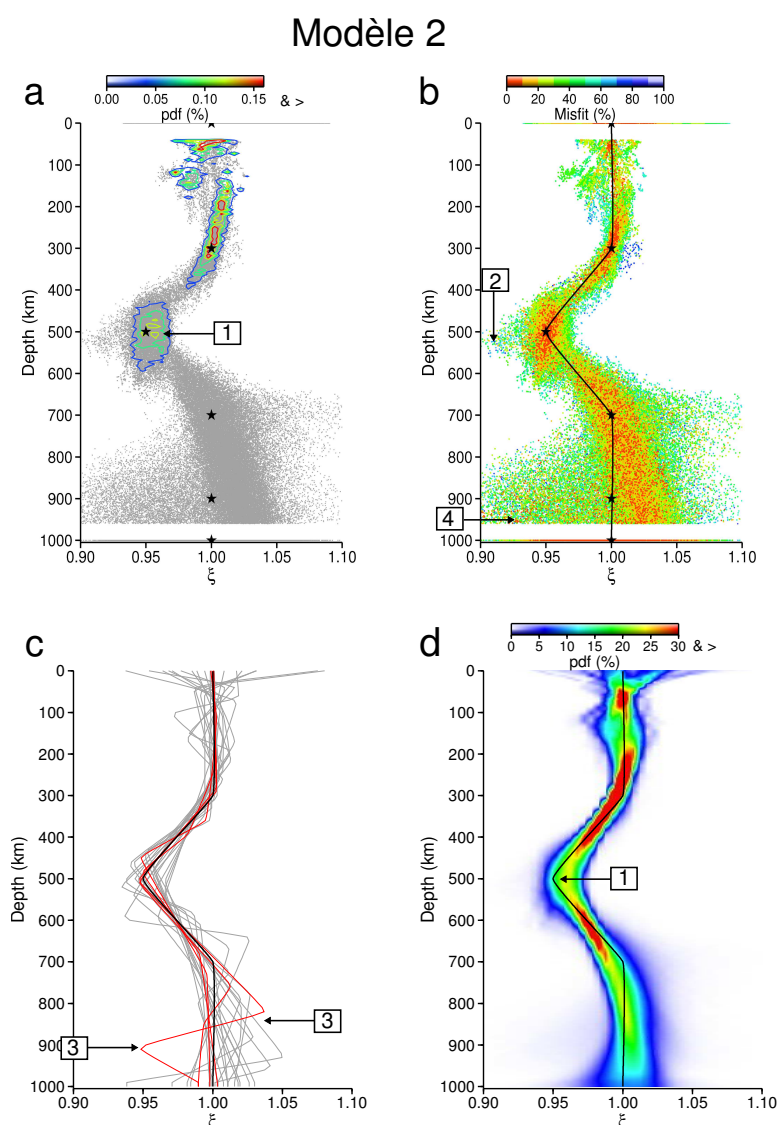


FIG. 5.8 : Résultats en anisotropie de l'inversion probabiliste du modèle 2. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier.

(étiquettes 1). En revanche, les répartitions des points de Bézier montrent que la valeur $\xi = 0.95$ et ses alentours sont bien échantillonnés, et que les valeurs de misfit correspondantes sont faibles. De plus, les profils 1D représentant les meilleurs misfits retrouvent l'amplitude attendue. L'espace des paramètres est même échantillonné jusqu'à la borne inférieure du prior égale à 0.9, même si les misfits de ces points sont élevés (étiquette 2). Cette information n'est pas du tout visible sur la pdf des profils 1D. En fait, la pdf des courbes de Bézier est très utile pour se donner une idée de la collection de profils 1D acceptés, mais cette représentation a tendance à lisser les résultats, notamment lorsque les

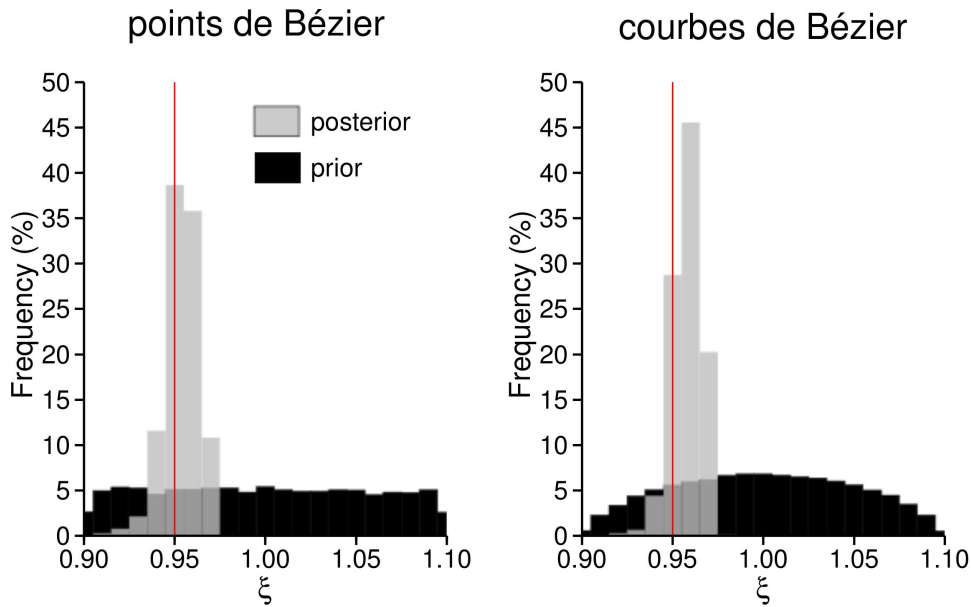


FIG. 5.9 : Probabilités marginales *a priori* (en noir) et *a posteriori* (en gris) sur l'anisotropie à 500 km de profondeur pour le modèle 2. Les résultats sont présentés pour les points de Bézier et pour les courbes de Bézier. La valeur vraie est montrée en rouge et vaut $\xi = 0.95$. La taille des intervalles est de 0.01.

variations des profils synthétiques en fonction de la profondeur sont brutales. Par conséquent, pour l'étude des données réelles, nous donnerons plus de crédit à l'analyse des points de Bézier.

La vingtaine de modèles 1D représentée est assez similaire aux modèles synthétiques, excepté au-dessous de 800 km de profondeur. En effet, certains modèles oscillent autour de la valeur $\xi = 1$ (étiquettes [3]). Passée cette profondeur, les points de Bézier présentent une dispersion plus forte sur l'ensemble des valeurs possibles. La répartition des points est assez symétrique pour le modèle 1. Bien que certains points soient assez éloignés du profil à retrouver, les misfits correspondant sont relativement faibles, c'est-à-dire inférieurs à 20 % (étiquettes [4]). Notons encore une fois que cet évasement des points en profondeur n'est que très peu visible sur les pdf des profils 1D.

Théoriquement, dans le PREM, les noyaux de sensibilité du mode fondamental et du premier harmonique des ondes de Love et de Rayleigh par rapport à l'anisotropie radiale sont très sensibles aux 200 premiers kilomètres de la Terre. Les modes d'ordre supérieur permettent d'obtenir une résolution jusqu'à 1000 km de profondeur. Pour illustrer la sensibilité des différents modes à l'anisotropie radiale, une inversion a été réalisée en prenant en compte uniquement le mode fondamental, puis une autre en ajoutant le premier harmonique, puis le deuxième, et ceci jusqu'au troisième. Les répartitions des points de Bézier issues de ces 4 inversions ainsi que les valeurs des misfits associés sont montrées sur la figure 5.10. Clairement, peu d'information sur l'anisotropie radiale peut être

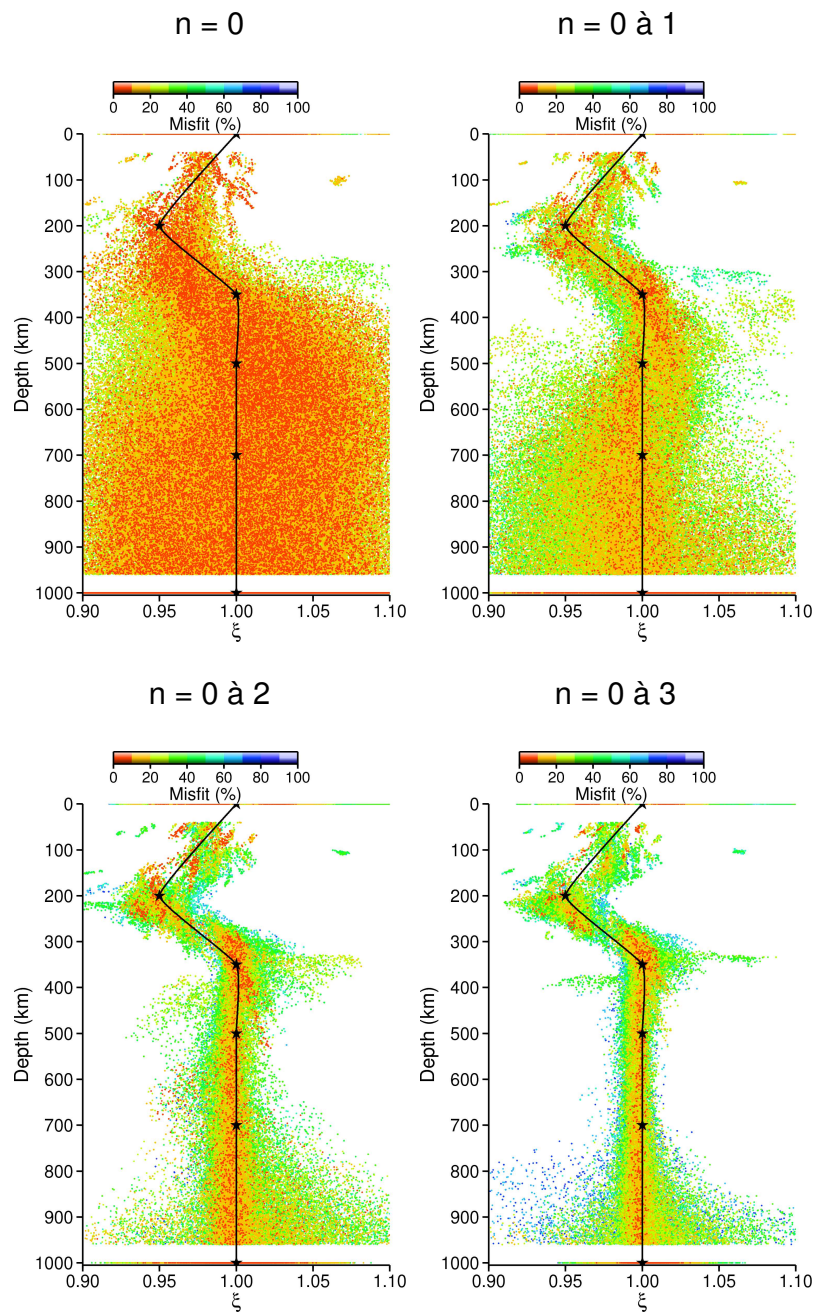


FIG. 5.10 : Répartition des points de Bézier représentant l'anisotropie pour le modèle 1. Les inversions ont été d'abord réalisées avec uniquement le mode fondamental, puis en rajoutant mode après mode jusqu'à l'ordre 3.

obtenue en utilisant uniquement le mode fondamental. On observe que ξ est globalement inférieure à 1 dans les 300 premiers kilomètres. Plus en profondeur, les points de Bézier sont étalés sur l'ensemble du prior. Les misfits sont toujours très faibles, ce qui signifie que la position des points de Bézier n'influence que très peu la valeur du misfit. Si le premier harmonique est incorporé, une bonne résolution est obtenue jusqu'à environ 400 km

de profondeur. Si les modes 2 et 3 sont progressivement rajoutés, la répartition des points de Bézier dans l'espace est de plus en plus resserrée en profondeur et les meilleurs misfits sont centrés sur le profil à retrouver. La répartition des points est cohérente puisque ce sont les points situés sur le pourtour de la distribution qui possèdent les misfits les plus élevés.

En résumé, grâce à la prise en compte des modes harmoniques, la méthode probabiliste employée dans ce travail montre un clair potentiel pour détecter de l'anisotropie radiale dans la zone de transition et dans le manteau supérieur, si celle-ci a une importance, *via* le problème direct, dans l'espace des données.

5.4.3 Vitesse de cisaillement dans la croûte et le manteau supérieur

Les figures 5.11 et 5.12 montrent que les profils de vitesse de cisaillement sont bien retrouvés dans les deux cas, même lorsque le profil est proche des limites du prior, comme vers 210 km de profondeur pour le modèle 2 (étiquette [1]). Les changements du gradient de vitesse à 50 et 250 km de profondeur pour le modèle 1, et à 80, 200 et 290 km de profondeur pour le modèle 2, sont caractérisés par des densités élevées de points de Bézier, ce qui signifie que les données sont très sensibles aux changements de pente dans les couches superficielles.

À 50 km (modèle 1) et 80 km de profondeur (modèle 2), où le changement de pente est le plus important (étiquettes [2]), une corrélation est observée entre la profondeur des points de Bézier et les valeurs de V_S . Cette corrélation est logique puisqu'elle signifie que les données peuvent être expliquées de manière identique si la discontinuité est profonde et V_S est élevée, ou bien si la discontinuité est située plus en surface avec une valeur de V_S plus faible. Cette caractéristique a déjà été mise en évidence avec une méthode probabiliste par Bodin et al. (2012). En utilisant des fonctions récepteurs en tant que données synthétiques, les auteurs ont montré qu'il existait une corrélation entre la profondeur de la discontinuité qu'ils avaient définie comme étant le Moho et la valeur de V_S .

Les autres points de Bézier sont localisés le long du modèle synthétique. Pour le modèle 1, certains points sont éloignés du profil recherché et correspondent à des oscillations des modèles 1D (étiquettes [3]). Leurs valeurs de pdf étant faibles, ils sont donc peu probables.

Les points de Bézier sont plus concentrés dans l'espace des paramètres que pour les résultats sur T et ξ , et ils ne sont pas répartis à toutes les profondeurs. Ils sont également moins nombreux. Cela signifie que les points de Bézier en V_S étaient déjà dans une configuration favorable à la fin de l'étape 2 de l'algorithme, et que celui-ci n'a pas eu besoin de

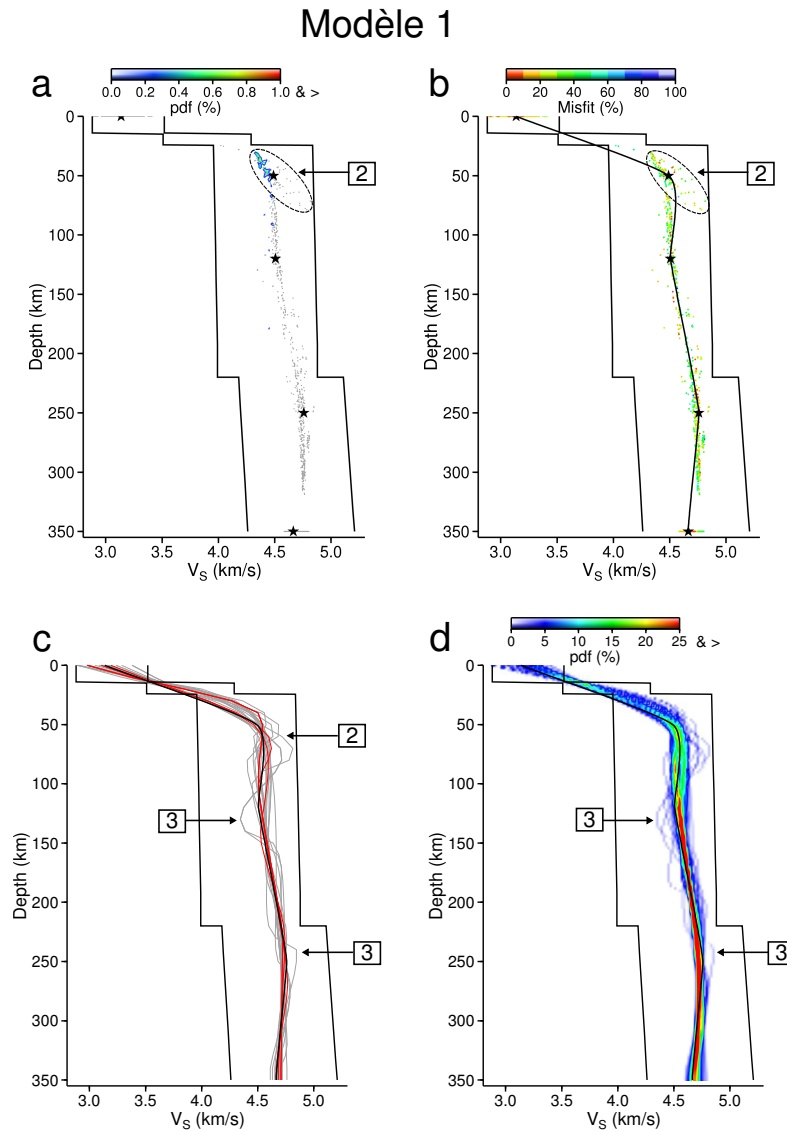


FIG. 5.11 : Résultats en vitesse de cisaillement de l'inversion probabiliste du modèle 1. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier.

les modifier durant la période stationnaire. Les distributions des modèles sont resserrées autour des modèles synthétiques, ce qui est en bon accord avec la sensibilité importante du mode fondamental aux couches superficielles de la Terre. La figure 5.6 montre que les incertitudes sur le mode fondamental sont très faibles (entre 0.04 et 0.065 %) à l'intérieur de la gamme de période 50-100 s. Le mode fondamental pèse très lourd dans le calcul de la fonction coût, c'est pourquoi le misfit est amélioré de façon significative si le profil de V_S est proche de celui recherché.

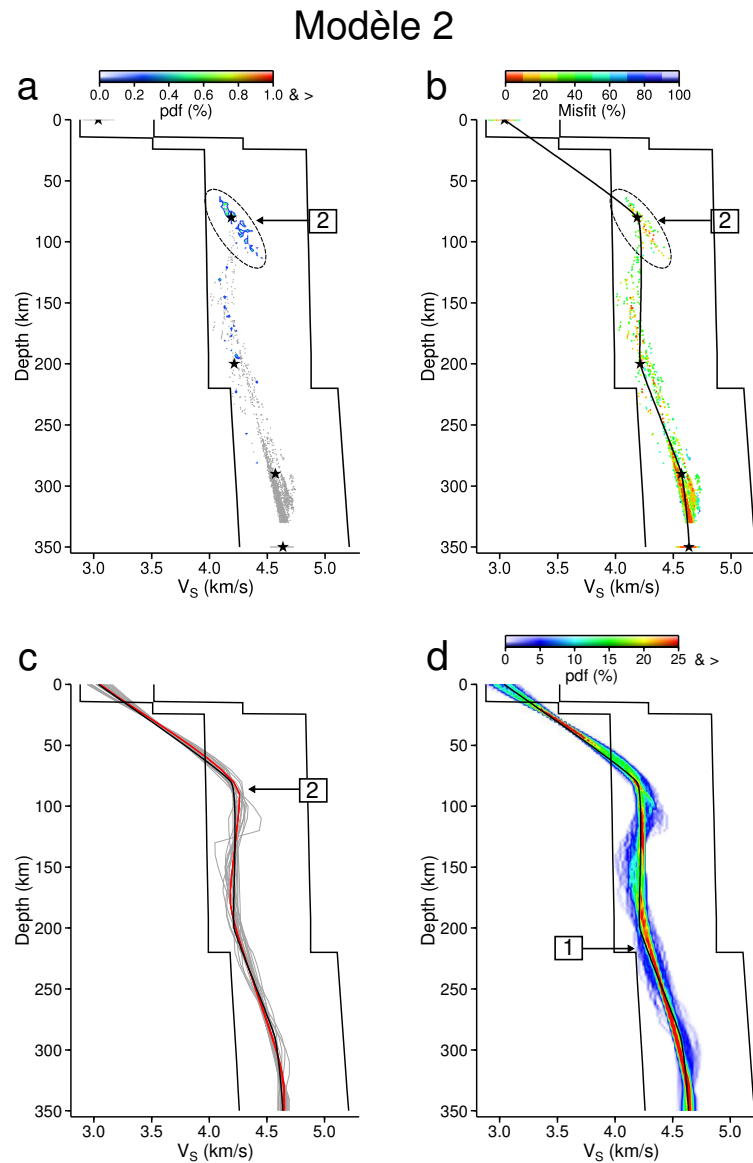


FIG. 5.12 : Résultats en vitesse de cisaillement de l'inversion probabiliste du modèle 2. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier.

5.4.4 Corrélations entre les paramètres

Lorsque plusieurs types de paramètres sont inversés, la question de savoir s'il existe des corrélations entre ces différents types de paramètres est soulevée (Khan et al., 2006, et références qui suivent). Des tests de corrélations ont été réalisés sur les paramètres primaires et secondaires de l'inversion. Des exemples de ces tests sur les résultats du modèle 1 sont présentés sur la figure 5.13 pour les paramètres primaires et sur la figure 5.14 pour les paramètres secondaires. Les résultats sont montrés sous la forme de densité de probabilité marginale 2D entre deux paramètres, à des profondeurs particulières. Le terme de « paramètres » est ici légèrement abusif car il ne s'agit pas strictement des paramètres de l'inversion, qui sont les points de Bézier, mais les valeurs des profils 1D à ces profondeurs. Les échelles utilisées en abscisse et en ordonnée correspondent aux intervalles de valeurs autorisées pendant l'inversion à la profondeur particulière étudiée.

5.4.4.1 Paramètres primaires

Sur la figure 5.13a sont présentées les corrélations entre T et ξ à 500 et 660 km de profondeur. La valeur vraie $\xi = 1$ est bien retrouvée dans les deux cas et correspond au maximum de la pdf. Comme nous l'avons déjà mentionné, la distribution est beaucoup plus étalée en température à 660 qu'à 500 km, mais la valeur recherchée est située dans la gamme de température où la pdf est la plus élevée. Les paramètres T et ξ ne montrent aucune sorte de corrélation.

C'est à 200 km de profondeur que se situe le maximum d'amplitude en anisotropie. La figure 5.13b présente la corrélation entre V_S et ξ à cette profondeur. La pdf maximale est très localisée. Comme expliqué dans la section 5.4.2, elle indique $\xi = 0.96$, ce qui signifie que la vraie valeur est retrouvée à 1 % près. La valeur de V_S la plus probable est très précise et correspond à 0.2 % près à la valeur du profil synthétique. Aucune corrélation n'est remarquée entre V_S et ξ dans la croûte et le manteau supérieur.

Ces résultats impliquent que T et ξ dans la zone de transition d'une part, et que V_S et ξ dans la partie superficielle de la Terre d'autre part, peuvent être retrouvées de façon indépendante.

5.4.4.2 Paramètres secondaires

Qu'en est-il au niveau des corrélations de T et ξ dans la zone de transition avec les paramètres secondaires, c'est-à-dire les valeurs de V_S isotrope issues de l'équation d'état de Birch-Murnaghan ? Les figures 5.14a et 5.14b présentent les corrélations entre V_S et ξ , et entre V_S et T , respectivement, à 500 et 660 km de profondeur.

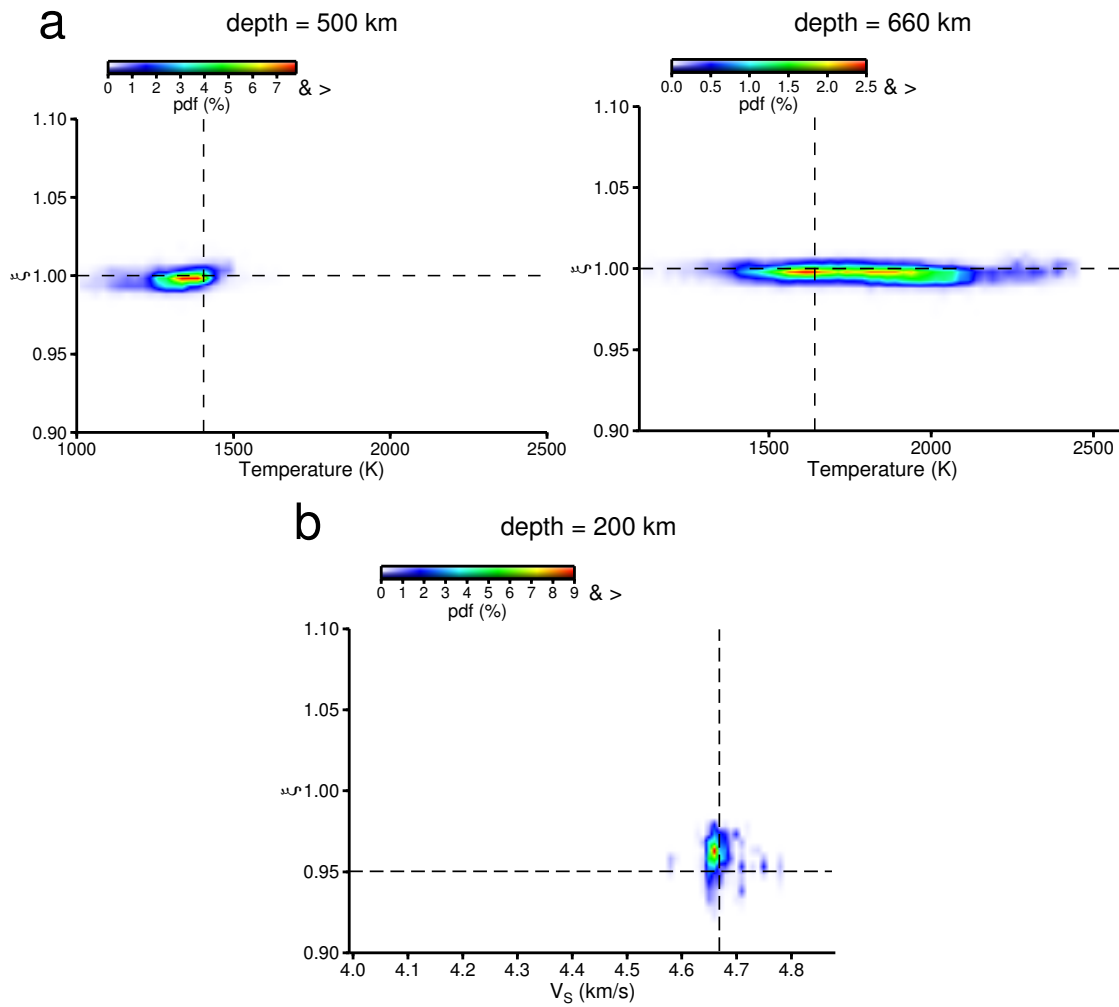


FIG. 5.13 : Densités de probabilité marginales montrant les corrélations entre les valeurs des paramètres primaires issues de l'inversion du modèle 1 à des profondeurs fixées. Les valeurs vraies sont montrées en pointillés. Les échelles en abscisse et en ordonnée correspondent à la taille du prior. (a) Corrélations entre T et ξ à 500 et 660 km de profondeur. (b) Corrélation entre V_S et ξ à 200 km de profondeur.

À 500 km de profondeur, il n'existe pas de corrélation entre V_S et ξ . Les vraies valeurs sont retrouvées de façon précise : à environ 0.2 % près pour V_S et ξ . En revanche, il n'est pas surprenant d'observer que la valeur de V_S dépend fortement de la température. Plus la température diminue, plus V_S augmente, et *vice-versa*.

De la même manière, aucune corrélation n'est à noter entre V_S et ξ à 660 km de profondeur. La pdf maximale est située au niveau de la vraie valeur de V_S . En revanche, une caractéristique surprenante par rapport aux densités de probabilité marginales étudiées jusqu'alors est que la distribution est clairement bimodale. En effet, il existe un deuxième minima vers 6.0 km/s, en plus du minimum global correspondant à la valeur à retrouver. Ces deux extrema sont également bien visibles au niveau de la corrélation entre V_S et T .

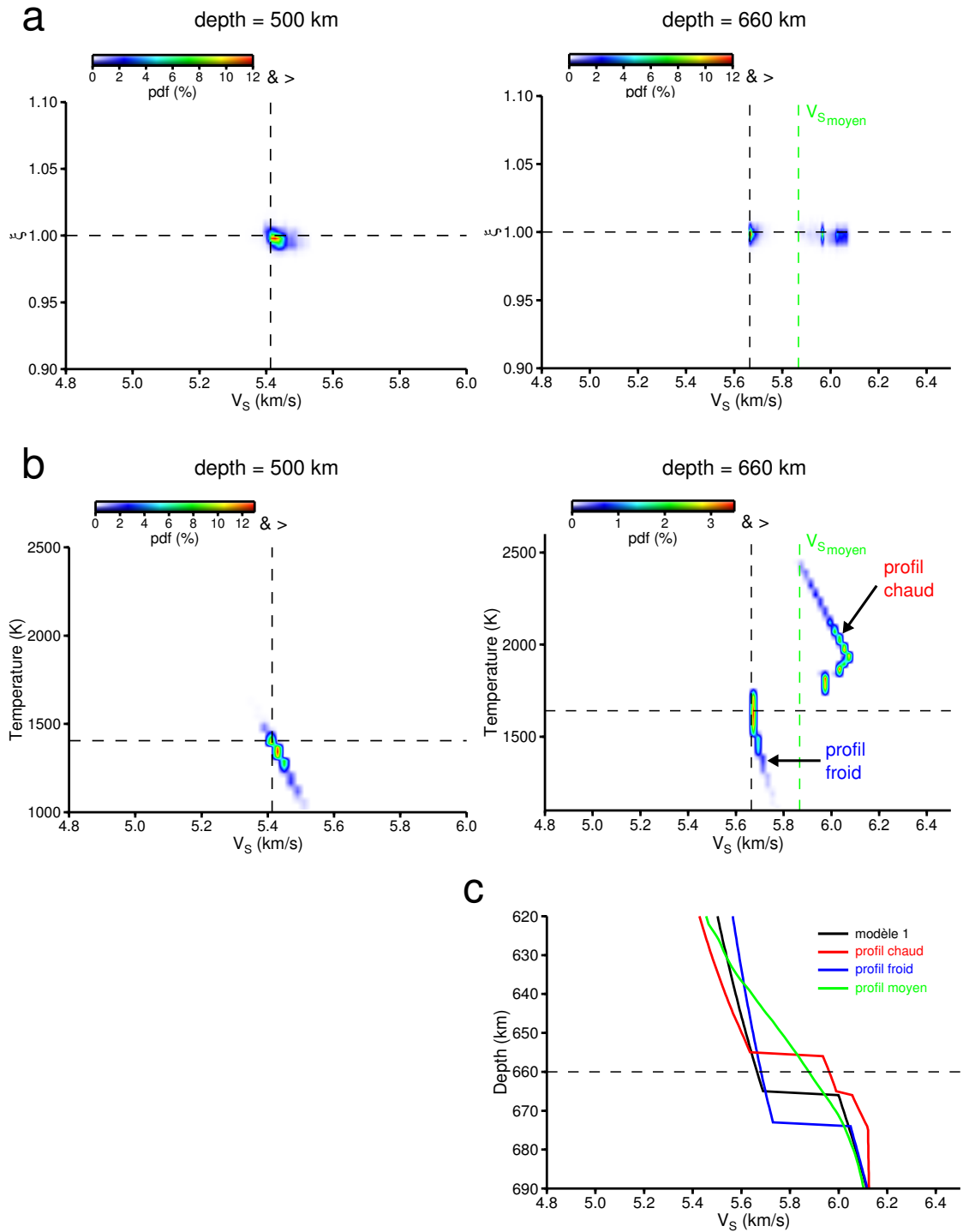


FIG. 5.14 : Densités de probabilité marginales montrant les corrélations entre les paramètres primaires et secondaires issus de l'inversion du modèle 1 à 500 et 660 km de profondeur. Les valeurs vraies sont montrées en pointillés. (a) Corrélations entre V_S et ξ . (b) Corrélations entre V_S et T . (c) Évolution de la profondeur de la discontinuité sismique engendrée par la transformation de ringwoodite en pérovskite et magnésioiwüstite selon la température.

Même si l'on retrouve la même corrélation entre V_S et T que pour la pdf à 500 km de profondeur, la valeur vraie correspond plutôt à un profil « froid » et le second minima à 6.0 km/s à un profil « chaud ». La figure 5.14c montre que ce comportement est relié aux transformations minéralogiques. À 660 km de profondeur, la discontinuité sismique engendrée par la transformation de la ringwoodite en pérovskite et magnésio-wüstite pour le modèle 1 n'a pas encore été franchie. Étant donné que cette transformation est endothermique, un profil « froid » présente une discontinuité encore plus profonde. En revanche, la discontinuité a lieu à des profondeurs plus superficielles pour un profil « chaud », ce qui a pour conséquence de remonter la discontinuité. Ainsi, à 660 km de profondeur, toute la ringwoodite a été transformée. La valeur de V_S pour un profil « chaud » est donc nettement plus élevée que la valeur du modèle 1.

Comme le font remarquer Bodin et al. (2012), bien que la pdf maximale, c'est-à-dire le mode de la distribution, corresponde effectivement à la valeur de V_S recherchée, le fait que la distribution soit bimodale et dissymétrique renforce la conviction qu'il est difficile de rendre compte de l'ensemble de la solution du profil de vitesse en présentant uniquement la moyenne des valeurs de V_S avec l'écart-type associé. En effet, la moyenne est située entre les 2 valeurs extrêmes, à un endroit où en fait aucun modèle en V_S n'a été échantillonné par l'algorithme (figures 5.14a et b), ce qui donne l'impression que la discontinuité est lissée (figure 5.14c).

5.5 Tests supplémentaires

Dans cette section nous avons voulu évaluer l'impact de différents facteurs sur les résultats de l'inversion. Seuls les résultats sur le modèle 2 sont présentés pour ne pas surcharger cette partie.

5.5.1 Influence de la taille des incertitudes sur les données

La taille des incertitudes sur les données produit un impact important sur l'étendue de la distribution à l'intérieur du prior. Pour illustrer cet effet, le modèle 2 a été inversé pour différentes valeurs d'incertitudes sur les vitesses de phase. La figure 5.15 synthétise la valeur de l'écart-type à 1σ d'intervalle de confiance pour la température en fonction de l'incertitude sur les données, pour différentes profondeurs. Pour une profondeur donnée, l'écart-type augmente linéairement en fonction de l'incertitude attribuée aux vitesses de phase. La valeur de 1σ est de plus en plus importante en fonction de la profondeur, ce qui est cohérent avec l'accumulation des transformations minéralogiques et de la diminution de la sensibilité des ondes de surface.

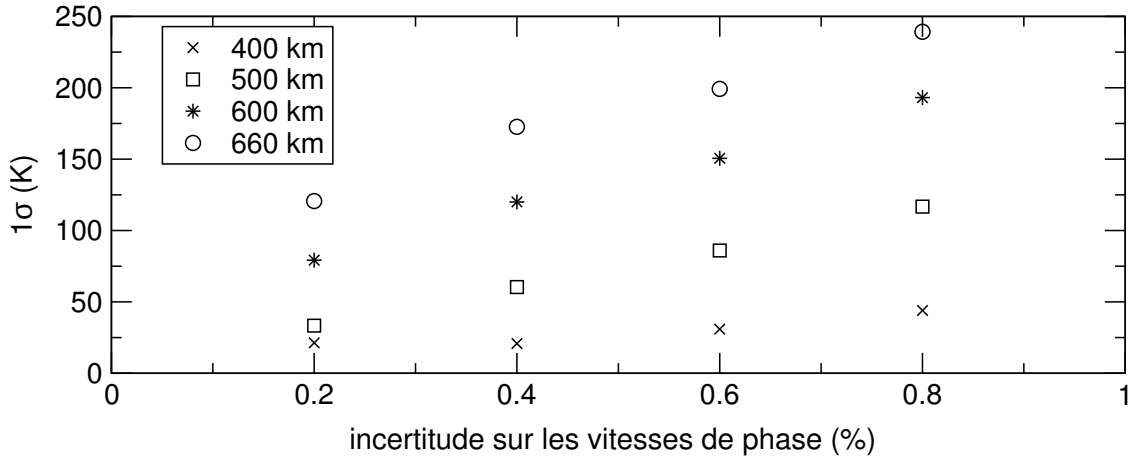


FIG. 5.15 : Évolution de l'écart-type à 1σ sur T en fonction des incertitudes sur les données pour différentes profondeurs.

La figure 5.16 montre les pdfs sur les profils 1D en V_S , T et ξ pour une faible incertitude sur les vitesses de phase (0.15 %). Dans ce cas extrême, le profil en V_S est très bien contraint. Pour la température, la distribution est centrée sur le profil recherché, ce qui n'est pas le cas lorsque des incertitudes réalistes sont considérées (voir figure 5.3d). Le profil d'anisotropie radiale est également bien situé au centre de la distribution. Cependant, les variations abruptes à 500 et 700 km de profondeur ont tendance à être lissées. Comme nous l'avons expliqué dans la section 5.4.2, cet effet est intrinsèquement lié à ce type de représentation.

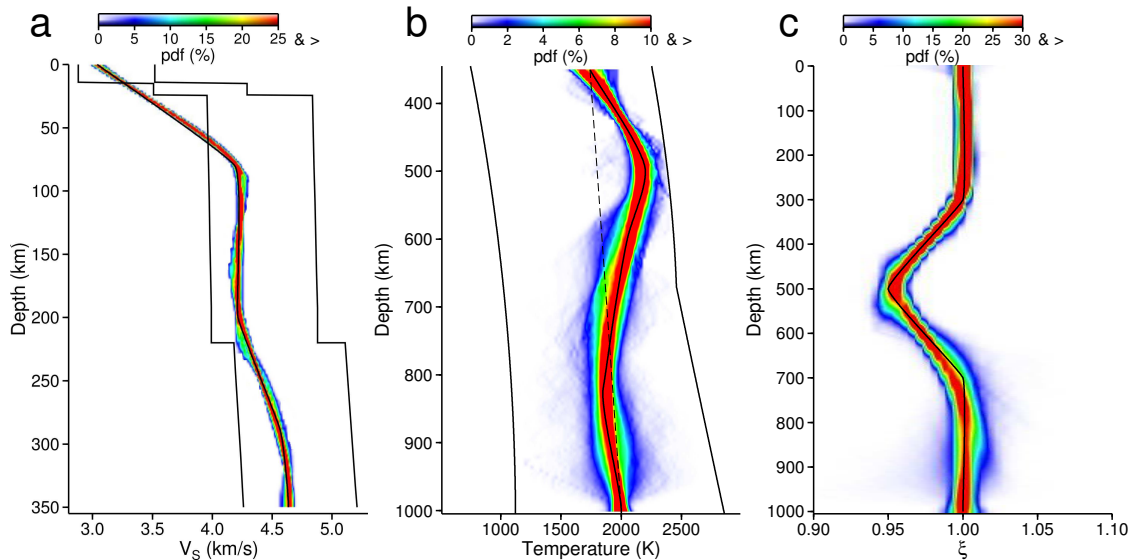


FIG. 5.16 : Résultats de l'inversion du modèle 2 en utilisant une incertitude sur les vitesses de phase de 0.15 %. La ligne en pointillé sur la figure (b) est un adiabat à 1600 K avec un gradient de température de 0.4 K/km.

5.5.2 Résultats avec une norme L_1

La figure 5.17 montre les pdfs sur les profils 1D obtenus pour le modèle 2 en utilisant un misfit défini par une norme L_2 et une norme L_1 . Pour ces deux tests, la même méthodologie est employée. Que les incertitudes sur les données soient considérées comme étant gaussiennes ou laplaciennes, les résultats sont relativement similaires.

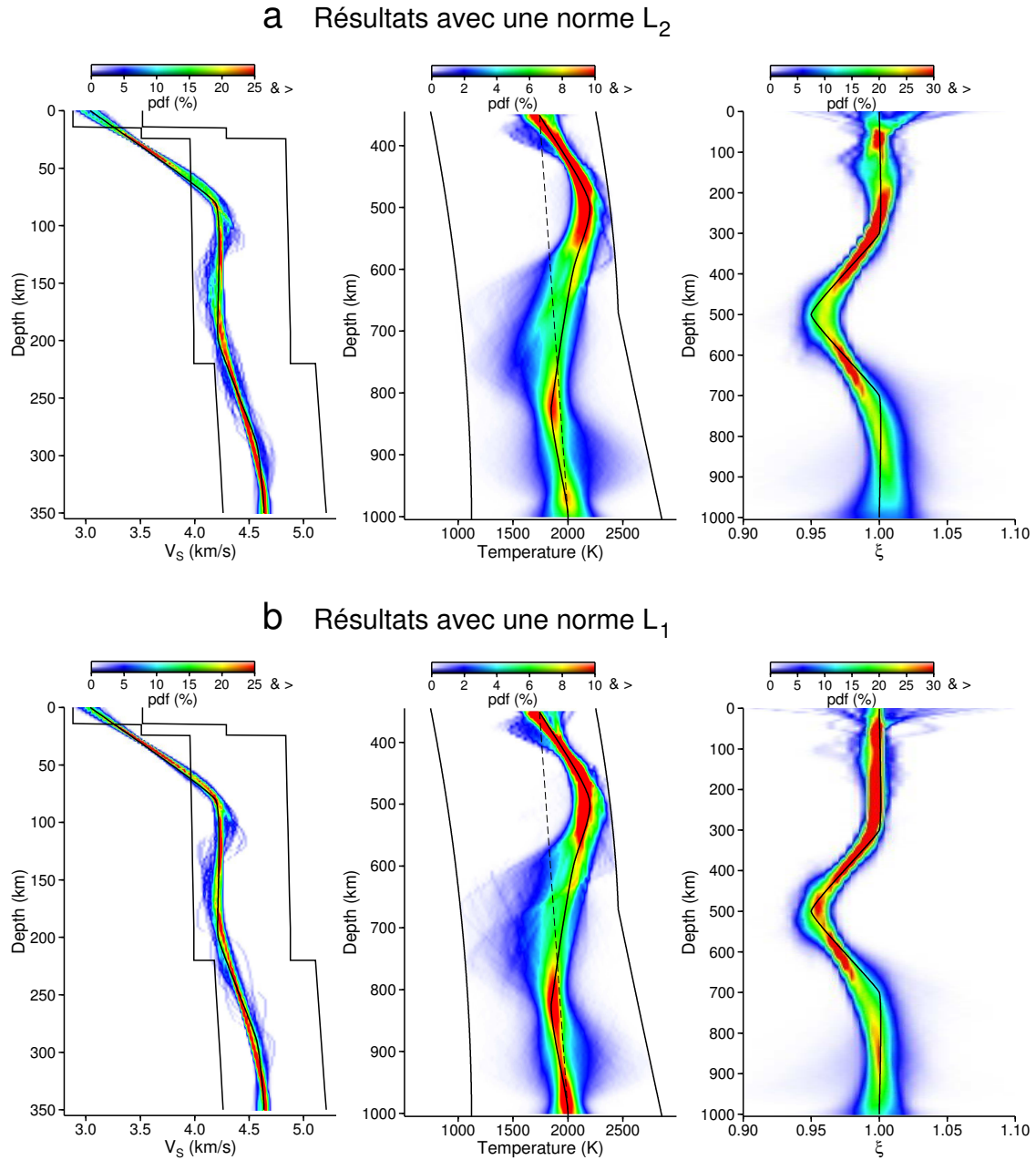


FIG. 5.17 : Résultats de l'inversion du modèle 2 en utilisant (a) une norme L_2 et (b) une norme L_1 , selon la même procédure. Les lignes en pointillé indiquent un adiabat à 1600 K avec un gradient de température de 0.4 K/km.

Les vitesses de phase correspondent à des observables secondaires, qui sont extraites à partir du sismogramme enregistré à l'emplacement d'une station sismique. Il existe de nombreuses méthodes de mesure de la dispersion des ondes de surface, telles que les analyses par inversion de formes d'ondes, les analyses par paires de stations ou encore les analyses en nombre d'ondes (*p. e.* [Beucler et al., 2003](#); [Visser et al., 2007](#)). Ces méthodes reposent sur la comparaison entre le sismogramme observé et le sismogramme synthétique. Pour un mode donné, les vitesses de phase forment une même branche et ne constituent pas des données complètement indépendantes les unes des autres (voir figure 2.8). Pour des vitesses de phase estimées le long d'un trajet source-station, il est peu probable que pour une même courbe de dispersion, il existe une ou plusieurs vitesses de phase dont les valeurs soient très éloignées de la tendance globale de la courbe de dispersion. Bien que ces tests aient été réalisés sur des données synthétiques, dépourvues de valeurs aberrantes, l'utilisation de la norme L_1 dans le calcul du misfit ne paraît pas indispensable dans cette étude.

5.5.3 Profondeurs des paramètres fixés

Nous avons vu dans les sections 5.4.1 et 5.4.2 que pour V_S les points sont plutôt regroupés en amas, alors que pour T et ξ , les points de Bézier sont répartis tout le long du profil recherché. Dans ce cas, on peut se demander si faire varier la profondeur des points de Bézier en T et ξ est indispensable, puisqu'ils ne restent pas systématiquement confinés aux profondeurs où il existe des variations importantes en température et en anisotropie. Nous avons voulu tester l'influence d'une paramétrisation avec des points de Bézier à des profondeurs fixées sur le résultat de l'inversion du modèle 2. Les résultats sont présentés sur la figure 5.18 pour des profondeurs fixées aux profondeurs avec lesquelles le modèle 2 a été construit, et sur la figure 5.19 pour des profondeurs fixées à des valeurs quelconques.

Si les points de Bézier sont directement placés aux bonnes profondeurs, les plus faibles valeurs de misfit encadrent bien les valeurs de T et ξ recherchées (figures 5.18b et d). Pour ξ , la pdf est quasiment symétrique de part et d'autre du profil synthétique (figure 5.18c). La valeur de ξ à 500 km de profondeur est parfaitement retrouvée sur la pdf. Comme expliqué dans la section 5.4.2, il existe un évasement de la distribution à partir de 800 km. Pour la température, la pdf est également maximale le long du modèle recherché (figure 5.18a). Comme nous l'avons décrit dans la section 5.4.1, un étalement de la distribution est observé au centre de la zone de transition et au-dessous de 800 km de profondeur.

Dans le cas où les profondeurs des paramètres sont fixées à des profondeurs différentes de celles du modèle d'entrée, les points de Bézier pour l'anisotropie sont fixés tous les 200 km, de 0 à 1000 km de profondeur (figure 5.19b). De la surface à 500 km

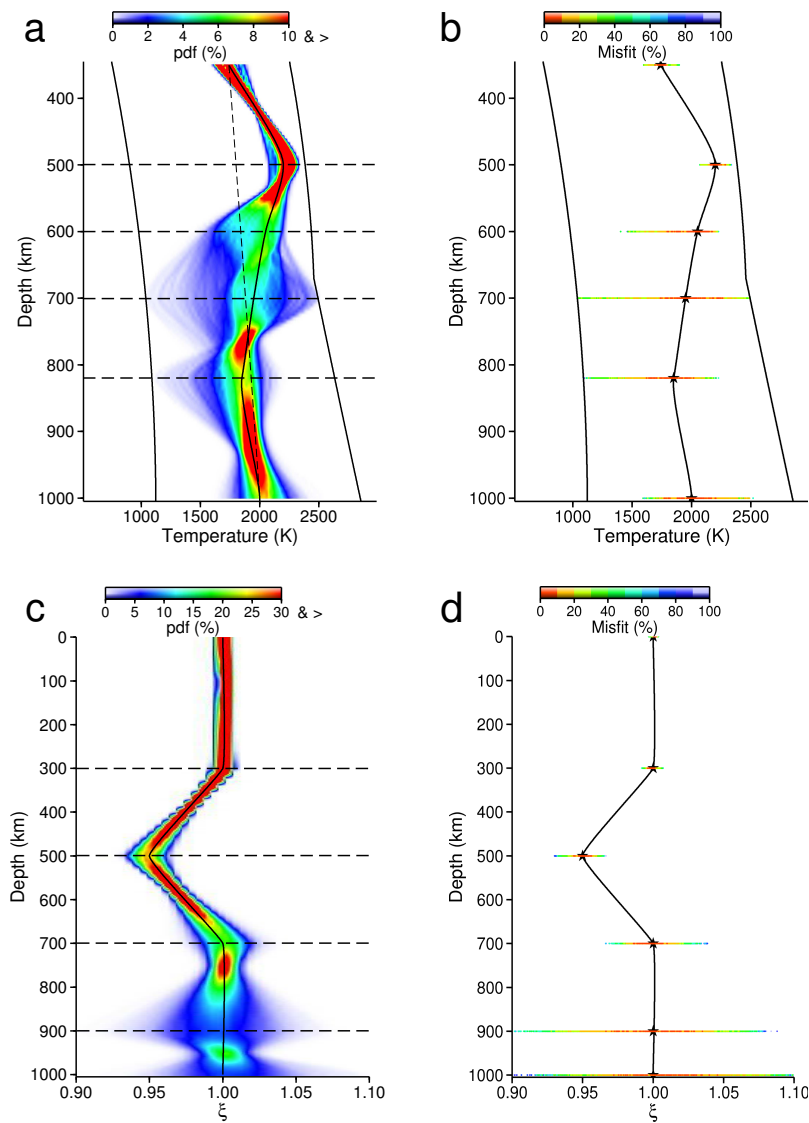


FIG. 5.18 : Résultats (a, b) en température et (c, d) en anisotropie de l'inversion du modèle 2, pour des profondeurs fixées similaires à celles qui ont été utilisées pour construire le modèle 2. Les pointillés horizontaux sur les figures (a) et (c) correspondent aux profondeurs attribuées aux points de Bézier. La ligne en pointillé sur la figure (a) indique un adiabat à 1600 K avec un gradient de température de 0.4 K/km.

de profondeur, les courbes de Bézier s'ajustent sur le profil du modèle 2. Cependant, les courbes ne sont pas capables de reproduire le « coude » qui se situe à 300 km de profondeur (étiquette [1]). L'amplitude maximale de l'anisotropie correspond à celle recherchée ($\xi = 0.95$), mais elle est placée 100 km plus profond (étiquette [2]). De 550 à 700 km, la valeur de ξ est sous-estimée (étiquette [3]). Entre 700 et 1000 km se produit un phénomène de compensation (étiquette [4]). Le seul point autorisé dans cette gamme de profondeur se situe à 800 km. Pour compenser le $\xi < \xi_{\text{vrai}}$ au-dessus de 700 km, la distribution de ξ va jusqu'à un maximum de 1.05 à 800 km. Puis ξ diminue et atteint une valeur de 1

vers 900 km, pour finalement repasser plus en profondeur à des valeurs inférieures à 1. Ce comportement en « zigzag » correspond bien à un phénomène de compensation car si l'on calcule la moyenne de chaque profil entre 700 et 1000 km de profondeur, celle-ci se situe autour de 1.01. Cette valeur est légèrement supérieure à la valeur recherchée de 1 dans cette région, mais elle montre bien que ce comportement singulier de la distribution a pour objectif de contrebalancer le $\xi < \xi_{\text{vrai}}$ entre 550 et 700 km. En fait, fixer la profondeur des points de Bézier comprend un risque important de ne pas retrouver les anomalies d'anisotropie à la bonne profondeur, et surtout de voir apparaître des anomalies qui n'existent pas en réalité.

En ce qui concerne la température, le domaine entre 350 et 1000 km de profondeur a été discrétisé de façon régulière en fonction de la profondeur. Plus le nombre de points de Bézier constitutifs d'une courbe est important, plus la discrétisation est fine. Malgré le fait que la distribution de ξ soit relativement éloignée du modèle recherché, la densité de probabilité en température (figure 5.19a) contient le modèle espéré. Cependant, les régions de faible probabilité, représentées en bleu, sont caractérisées par de nombreuses oscillations autour du profil recherché. Les extrémités de ces oscillations sont très « pointues » et ne sont pas physiquement très réalistes. D'autre part, elles rappellent le phénomène de compensation à 800 km de profondeur remarqué sur la distribution des profils d'anisotropie.

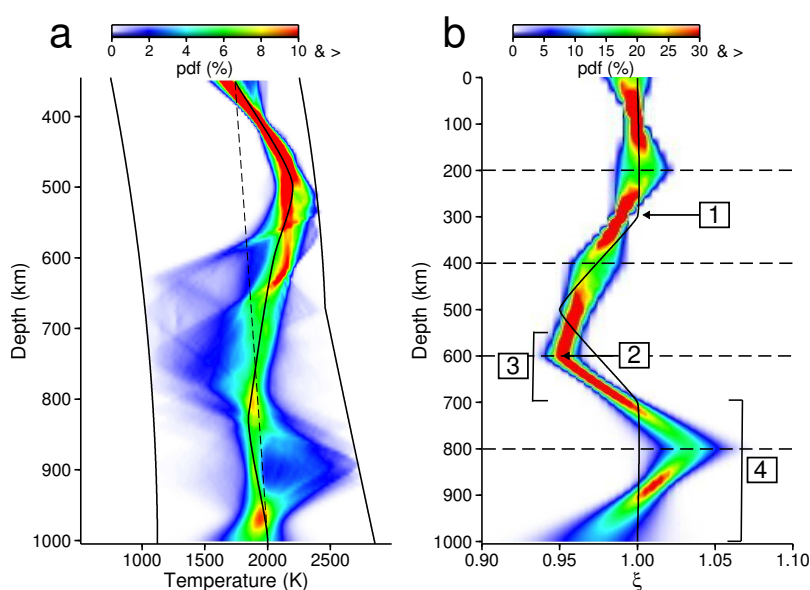


FIG. 5.19 : Résultats (a) en température et (b) en anisotropie de l'inversion du modèle 2, pour des profondeurs fixées des paramètres de façon arbitraire. La ligne en pointillé sur la figure (a) indique un adiabat à 1600 K avec un gradient de température de 0.4 K/km. Les pointillés horizontaux sur la figure (b) correspondent aux profondeurs attribuées aux points de Bézier.

La figure 5.20 permet d'expliquer l'origine de ces oscillations. Les modèles échantillonnés durant la période stationnaire ont été triés selon leur nombre de points de Bézier. Des densités de probabilité indépendantes ont été estimées pour chaque nombre de point. La figure 5.20a montre les pdfs lorsque la profondeur des paramètres est variable, tandis que la figure 5.20b présente les résultats de l'inversion quand les profondeurs sont fixées. Dans le cas où les profondeurs sont libres de varier aléatoirement, les distributions obtenues avec 6, 7, 8, 9 et 10 points sont lisses et relativement semblables. La forme globale du modèle 2 est conservée. Comme expliqué dans la section 5.4.1, le nombre de points de Bézier n'a que peu d'influence sur la distribution. En revanche, ce comportement n'est du pas du tout retrouvé lorsque les profondeurs sont fixées. Plus le nombre de points est important, plus la distribution réalise des « zigzags » autour du modèle recherché. Les oscillations s'intensifient à partir de 600 km de profondeur, à cause de la non-linéarité liée aux transitions de phase et aux transformations minéralogiques dans cette région. Clairement, il s'agit du même phénomène de compensation que pour l'anisotropie. Celui-ci est bien visible sur la distribution avec 7 points. À 740 km de profondeur, une forte diminution de la température (étiquette [1]) peut être compensée par une forte augmentation de la température à 870 km (étiquette [2]). Fixer les profondeurs donne lieu à des modèles physiquement peu réalistes. Par exemple, pour la distribution avec 9 points, on observe une solution bimodale aux alentours de 600 km de profondeur. Une première famille de modèles suit bien le profil recherché (étiquette [3]), tandis qu'une seconde famille montre une diminution importante de la température (étiquette [4]), puis revient s'ajuster sur le profil recherché à 720 km, et continue en profondeur avec des oscillations. Il est peu probable de trouver un tel profil de température dans la zone de transition.

Bien que faire varier la profondeur des points de Bézier rajoute des paramètres dans le processus d'inversion, ce procédé paraît indispensable pour obtenir des modèles réalistes, d'autant plus que les vitesses de phase des distributions en T et ξ de la figure 5.19 sont toutes comprises dans les barres d'erreur des données synthétiques du modèle 2. Cependant, le maximum de la pdf en température contient le modèle recherché (figure 5.19a). Une solution pour conserver des profondeurs fixes serait d'introduire des conditions de lissage sur la température. Par exemple, un prior gaussien quadratique pourrait être implémenté pour pénaliser les variations de température trop importantes entre deux ou trois points consécutifs (Besag et al., 1995). Un tel prior dépend d'un coefficient, appelé *hyperparamètre*, qui contrôle l'amplitude du lissage sur le profil. Plus cette valeur est grande, plus les fortes variations sont pénalisées, et plus grande est l'influence du prior comparée à la vraisemblance. En d'autres termes, cet hyperparamètre contrôle les importances relatives du prior et de la fonction de vraisemblance dans la densité de probabilité *a posteriori*. Dans l'algorithme élaboré au cours de cette thèse, le paramètre de lissage peut être assimilé au gradient maximum en température que l'on attribue entre deux points consécutifs (voir section 4.4.1.1).

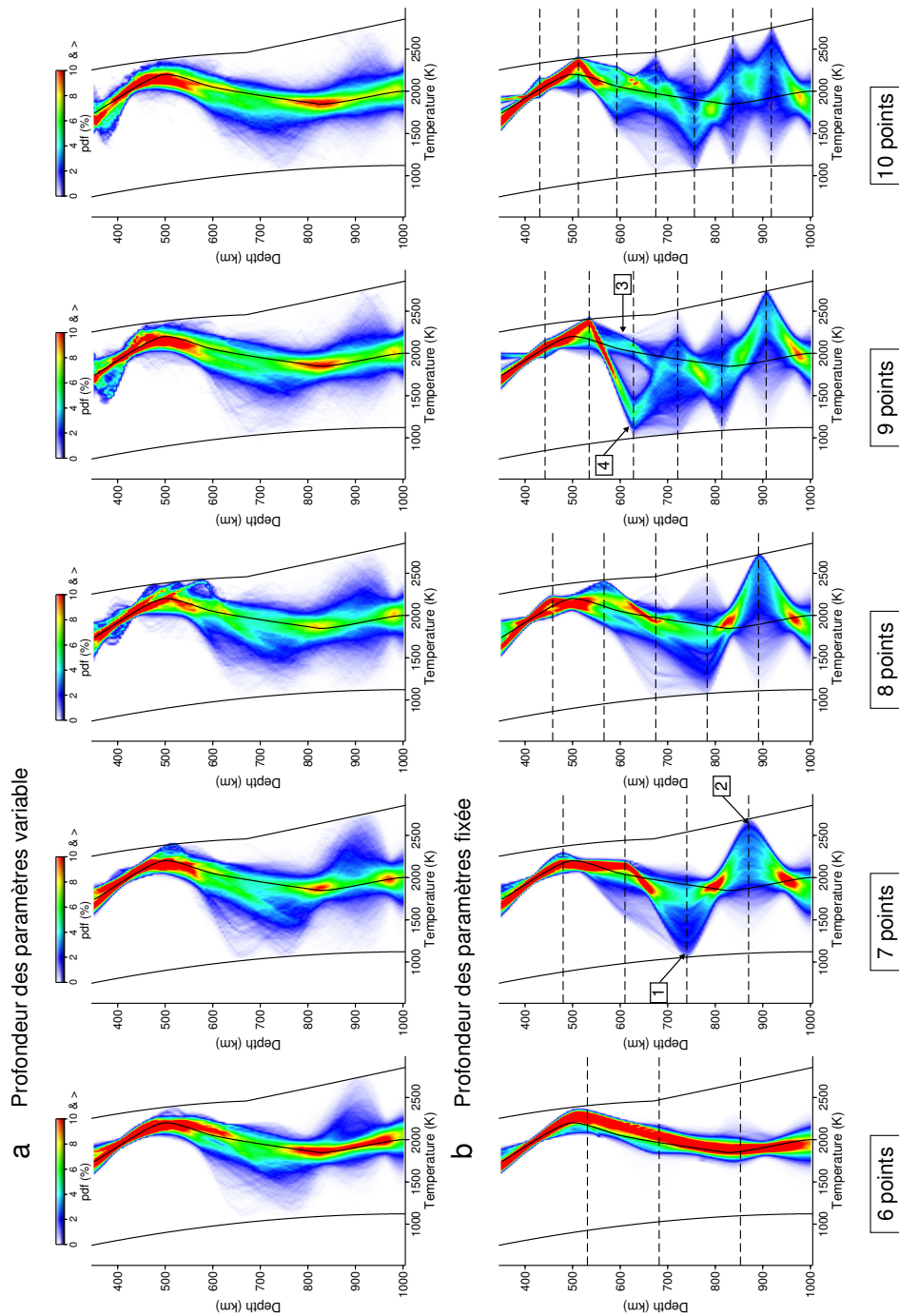


FIG. 5.20 : Densités de probabilité en température pour l'inversion du modèle 2 en considérant que les profondeurs des paramètres sont (a) variables et (b) fixées. Les profils sélectionnés lors de l'étape stationnaire ont été triés selon leur nombre de points de Bézier et sont représentés de façon indépendante. Les pointillés horizontaux sur la figure (b) correspondent aux profondeurs attribuées aux points de Bézier. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte.

5.5.4 Influence du lissage des profils en température

Contrairement à l'introduction d'un hyperparamètre, imposer un gradient maximum entre deux points de Bézier consécutifs en profondeur permet de donner un sens physique au lissage des profils en température. Jusqu'à présent, un gradient extrême de ± 15 K/km a été employé pour observer le comportement de l'algorithme lorsque très peu d'informations *a priori* sont considérées. Deux gradients plus faibles, de 5 et 1.5 K/km, sont testés. Le modèle 2 est composé de 7 points de Bézier. Les gradients de température entre 2 points consécutifs pour ce modèle sont tous inférieurs ou égaux à 1.5 K/km, excepté entre les 2 points les plus superficiels (à 350 et 500 km de profondeur), où le gradient est égal à 3.1 K/km (voir figure 5.22b). Quel est le comportement de l'algorithme lorsqu'un gradient inférieur à celui recherché est utilisé ?

Les résultats pour les gradients de 5 et 1.5 K/km sont présentés respectivement sur les figures 5.21 et 5.22. Avec un gradient de 5 K/km, on remarque les mêmes particularités qu'avec l'utilisation d'un gradient à 15 K/km, à savoir :

- les points de Bézier, ainsi que les pdfs des points et des courbes épousent la forme du modèle 2,
- les misfits les plus faibles sont localisés autour du profil recherché,
- les pdfs sont maximales jusqu'à environ 550 km de profondeur,
- il existe un étalement de la distribution entre 600 et 800 km.

La seule différence avec les résultats obtenus avec un gradient de 15 K/km est que les points de Bézier sont moins étalés dans l'espace des paramètres (figure 5.21b). La distribution de température avec le gradient à 1.5 K/km (figure 5.22) ne montre pas les mêmes caractéristiques. De 550 à 1000 km de profondeur, la distribution est très resserrée autour du modèle 2. Ce n'est pas le cas entre la surface et 550 km. Effectivement, étant donné que le gradient de température entre le point 1 et le point 2 (figure 5.22b) est supérieur au gradient autorisé pour la recherche des modèles, les profils échantillonnés n'ont pas la possibilité de reproduire cette portion de courbe, et ils choisissent un chemin qui minimise au mieux l'écart entre ces 2 points. Le gradient de température entre 350 et 550 km de profondeur pour les meilleurs modèles est d'environ 1.4 K/km (figure 5.22c), soit quasiment la valeur maximale autorisée. Le tableau 5.4 montre le pourcentage de points retenus pendant la période stationnaire pour les 2 gradients testés. Quelle que soit la valeur du gradient, le nombre de points de Bézier employé n'est pas un facteur limitant dans le résultat de l'inversion. Bien qu'elles ne soient pas présentées, les pdfs sur les vitesses de phase obtenues pour les gradients à 5 et 1.5 K/km ne dépassent pas les incertitudes imposées sur les courbes de dispersion synthétiques.

Quoique pour les tests synthétiques, des modèles de température extrêmes aient été choisis, les résultats des tests présentés dans cette section indiquent qu'un choix sévère du gradient maximal de température autorisé entre 2 points consécutifs peut constituer une

condition *a priori* très forte en comparaison de l'information contenue dans les données. Ce choix peut entraîner l'incapacité de visualiser l'entière complexité de la distribution de température.

nombre de points de Bézier	6	7	8	9	10
5 K/km	20.00	17.50	23.75	20.00	18.75
1.5 K/km	26.25	17.50	18.75	16.25	21.25

TAB. 5.4 : Nombre de points de Bézier (en %) sélectionnés pour les processeurs engagés dans la période stationnaire.

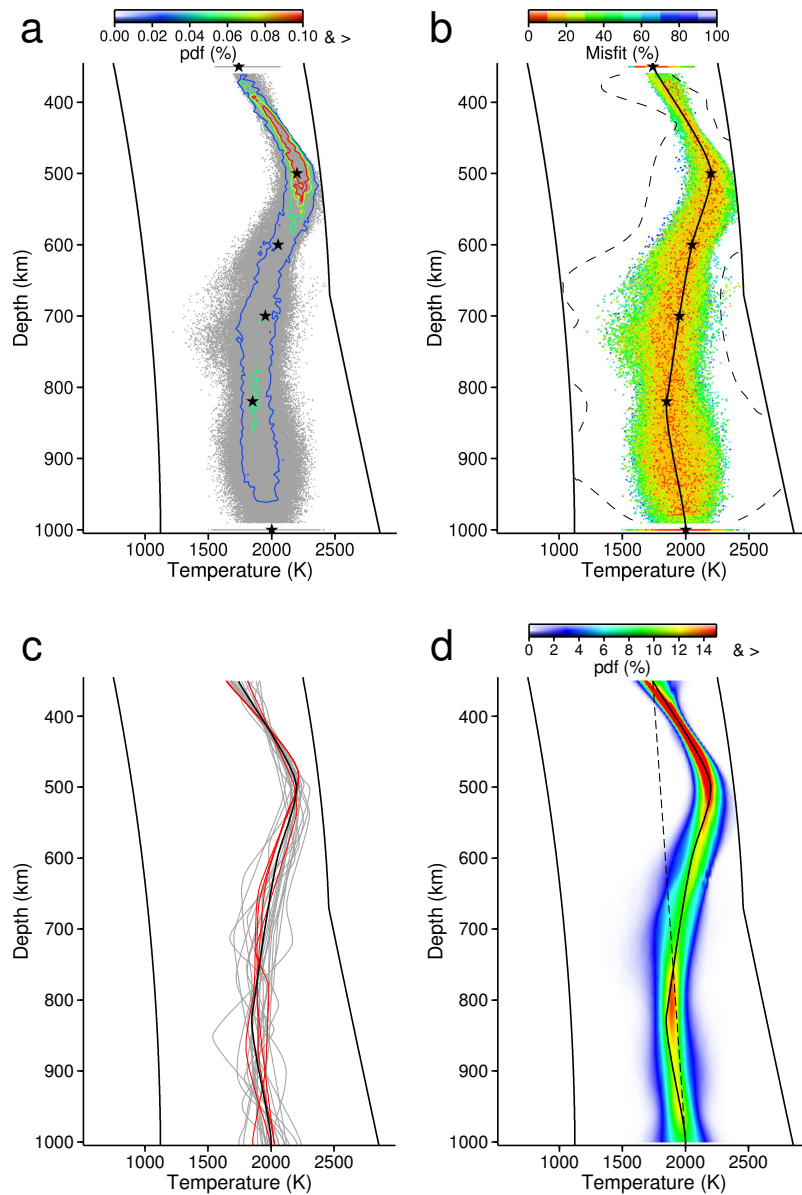


FIG. 5.21 : Résultats en température de l'inversion probabiliste du modèle 2 avec un gradient maximal de ± 5 K/km autorisé entre 2 points consécutifs. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. Les contours en pointillés correspondent aux contours de la distribution des points dans le cas où le gradient vaut ± 15 K/km (voir figure 5.3). (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier. La ligne en pointillé indique un adiabat à 1600 K avec un gradient de température de 0.4 K/km.

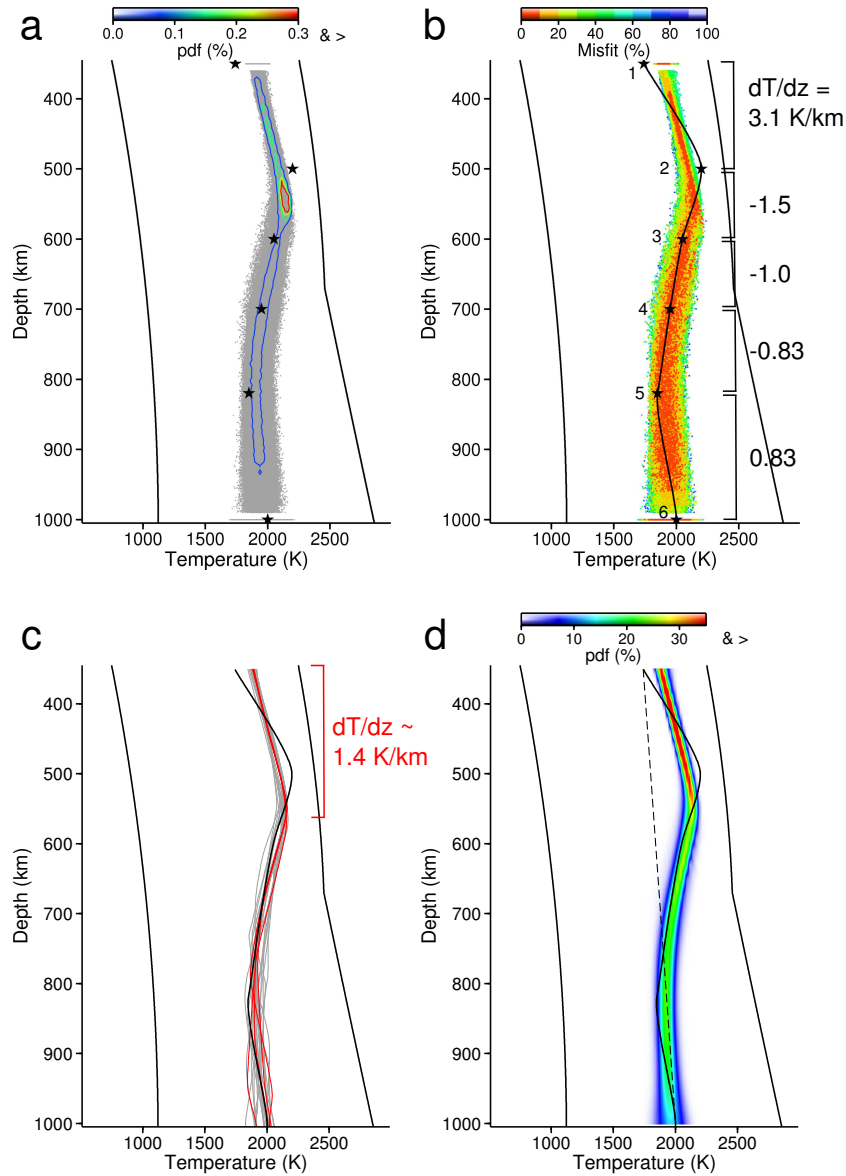


FIG. 5.22 : Résultats en température de l'inversion probabiliste du modèle 2 avec un gradient maximal de $\pm 1.5 \text{ K/km}$ autorisé entre 2 points consécutifs. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier. La ligne en pointillé indique un adiabat à 1600 K avec un gradient de température de 0.4 K/km .

5.6 Ce qu'il faut retenir

Une des principales critiques adressées aux méthodes McMC est que les résultats sont influencés par des choix subjectifs, comme la forme de la distribution *a priori* ou la densité de proposition des paramètres. Dans les exemples montrés dans ce chapitre, les distributions des paramètres sont clairement influencées par les données plutôt que par l'information *a priori*, sauf dans le cas d'un gradient de température trop restrictif pour la recherche des modèles. Les résultats sur les tests synthétiques montrent la capacité de la méthode d'inversion de détecter des variations de température dans la zone de transition, d'anisotropie radiale entre la surface et 1000 km de profondeur, et de vitesse de cisaillement dans le manteau supérieur. Les paramètres primaires de l'inversion ne présentent aucun signe de corrélation. À cause des différentes transitions de phase et transformations minéralogiques qui ont lieu au cœur de la zone de transition, la distribution des modèles en température est plus étalée à partir de 600 km de profondeur. Les incertitudes sur les données ont également une importante influence sur l'étendue de la distribution dans l'espace des paramètres. Puisque la profondeur des paramètres n'est pas fixée, les points de Bézier se placent naturellement le long des courbes recherchées pour reproduire la forme globale de la courbe de Bézier espérée, ce qui apporte une très nette amélioration par rapport à une paramétrisation fixe en profondeur.

Chapitre 6

Inversion de données réelles

6.1 Introduction

La méthode développée dans cette thèse est testée sur des données réelles, à savoir des vitesses de phase intégrées le long d'un trajet source-station. Dans un premier temps, les résultats de l'inversion de courbes de dispersion réelles pour un trajet particulier sont analysés en détail. L'ensemble de la discussion de ce chapitre s'articule autour de ce que la méthode est capable ou non de discriminer. Comme dans le chapitre précédent, une attention spéciale est donnée à la comparaison entre les différents types de représentation des résultats.

Nous montrons que la distribution des modèles en température obtenue par l'intermédiaire des ondes de surface est extrêmement dépendante de la valeur de la pente de Clapeyron de la transformation minéralogique ringwoodite $\rightarrow$ pérovskite + magnésio-wüstite, qui constitue un prior important lié à la minéralogie. Puis une comparaison est réalisée entre la méthode développée dans cette thèse, pour laquelle la distribution en température est directement obtenue à partir des données selon le schéma $\{C\} \rightarrow \{V_{S_{\text{isotrope}}}, \xi\} \rightarrow \{T\}$, et une autre méthode généralement employée pour obtenir de l'information sur la température dans le manteau, qui est réalisée en deux étapes $\{C\} \rightarrow \{V_{S_{\text{isotrope}}}, \xi\}$ et $\{V_{S_{\text{isotrope}}}\} \rightarrow \{T\}$ (voir section 1.2.5). Nous verrons que les distributions obtenues sont significativement différentes selon la méthode utilisée, ce qui est important pour l'interprétation géodynamique des modèles sismologiques.

6.2 Les données

L'algorithme est utilisé pour déterminer les distributions de température et d'anisotropie radiale le long du trajet Vanuatu-Californie (figure 6.1), qui a déjà été l'objet de plusieurs études (*e. g.* Cara, 1979; van Heijst & Woodhouse, 1997; Beucler et al., 2003). Ce trajet est particulièrement intéressant car une forte anisotropie radiale positive a été remarquée sous la plaque Pacifique (Ekström & Dziewonski, 1998; Kustowski et al., 2008; Nettles & Dziewonski, 2008; Khan et al., 2009). La figure 6.2 montre une coupe du modèle tomographique 3D S20RTS (Ritsema et al., 1999) le long du trajet Vanuatu-Californie. Ce modèle présente à la fois des anomalies de vitesses positives et négatives vers 660 km de profondeur, liées aux plaques plongeantes au niveau des îles Vanuatu et aux points chauds placés le long du trajet. D'après les nombreuses anomalies de vitesses localisées au cœur de la zone de transition, il est difficile d'estimer quelle pourrait être la distribution de température moyenne le long de ce trajet. Effectivement, on peut s'imaginer que les anomalies « froides » et « chaudes » traversées par les ondes sismiques peuvent se compenser pour former une collection de modèles répartis autour d'un profil adiabatique. Au contraire, les ondes sismiques peuvent être particulièrement sensibles aux anomalies « froides » situées en début de trajet, ou à certains points chauds, et révéler une distribution de température relativement éloignée d'un profil adiabatique.

Le jeu de données employé est composé des vitesses de phase intégrées des ondes de Love et de Rayleigh jusqu'à l'ordre 3, calculées par Visser et al. (2008a). Les mesures des vitesses de phase ont été obtenues selon une approche par l'espace des modèles (Visser et al., 2007). L'espace de recherche des modèles a été centré autour d'un profil de vitesse des ondes de cisaillement 1D de référence. Les dérivées de Fréchet de ce modèle de référence ont été utilisées pour calculer les sismogrammes synthétiques correspondant

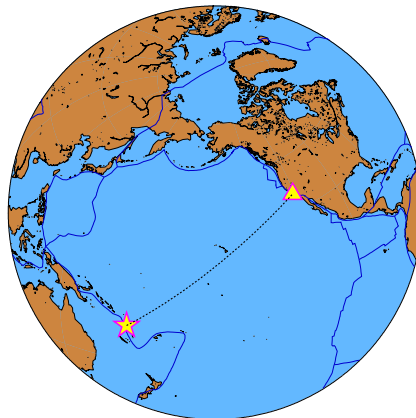


FIG. 6.1 : Illustration du trajet emprunté par les ondes pour le séisme considéré, ayant lieu dans les Vanuatu (latitude : -13.3° , longitude : 166.5°) et enregistré en Californie à la station SCZ (latitude : 36.6° , longitude : -121.4°) du réseau GEOSCOPE.

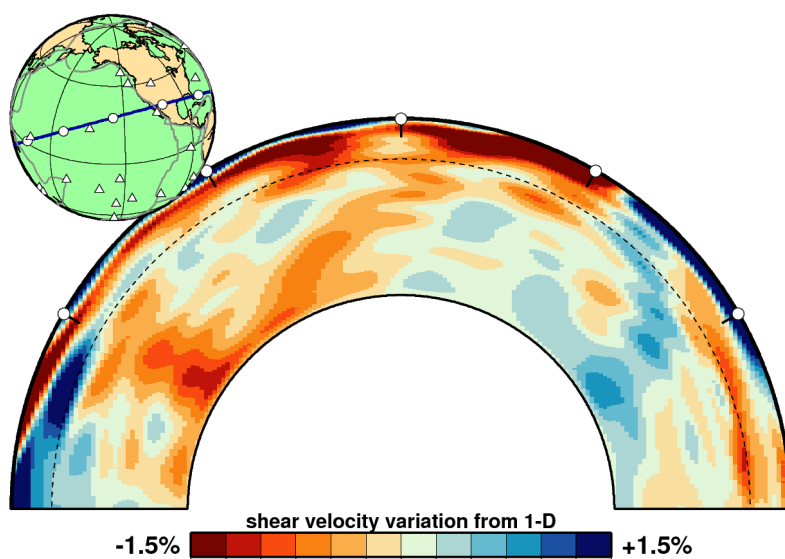


FIG. 6.2 : Coupe réalisée le long du trajet Vanuatu-Californie à travers le modèle 3D S20RTS (d'après Ritsema et al., 1999). Les régions colorées en rouge (bleu) indiquent que V_S est plus faible (plus élevée) que la vitesse de cisaillement moyenne à cette profondeur. La ligne en pointillé indique la profondeur de 670 km. La ligne bleue sur le globe terrestre correspond à la projection en surface de la coupe. Les triangles montrent la position des points chauds. Cette image a été réalisée grâce au script fourni par J. Ritsema (<http://www.earth.lsa.umich.edu/jritsema/research.html>).

aux perturbations de ce modèle de référence lors de l'échantillonnage de l'espace des modèles. La recherche des modèles a été réalisée par l'algorithme de voisinage (*neighbourhood algorithm*, en anglais), qui appartient à la famille des méthodes de Monte Carlo (Sambridge, 1999a,b). Cet algorithme minimise une fonction coût entre le sismogramme synthétique et le sismogramme réel dans différentes fenêtres de temps et de fréquences. Le choix de ces fenêtres est basé sur les valeurs des vitesses de groupe¹. Pour chaque mode et chaque fréquence, les probabilités marginales des vitesses de phase de l'ensemble des modèles peuvent être déterminées. La distribution de chaque vitesse de phase étant gaussienne, la mesure de la vitesse de phase et de son incertitude sont définies comme étant la moyenne et l'écart-type de la distribution. Cette méthode présente l'avantage de mesurer le mode fondamental et les modes harmoniques de façon automatique, sans nécessité de regrouper les séismes ou les stations.

Au total, 107 valeurs de vitesses de phase pour plusieurs fréquences et plusieurs modes, sont inversées simultanément pour retrouver les distributions de température et d'anisotropie radiale le long du grand cercle (tableau 6.1). La gamme de période em-

¹Contrairement à la vitesse de phase, qui correspond à la vitesse de propagation d'une seule phase associée à une fréquence, la vitesse de groupe peut être définie comme étant la vitesse de propagation du paquet d'ondes associé à cette fréquence.

mode	Rayleigh		Love	
	nombre de données	gamme de périodes (s)	nombre de données	gamme de périodes (s)
$n = 0$	16	35.1- 175.4	13	43.1- 174.0
$n = 1$	16	35.1- 172.5	13	43.4- 176.7
$n = 2$	15	35.0- 149.3	13	35.0- 115.4
$n = 3$	11	35.0- 87.7	10	35.0- 78.7

TAB. 6.1 : Nombre de vitesses de phase des ondes de Rayleigh et de Love employées et gamme de périodes concernées pour chaque mode.

ployée se situe entre 35 et 175 s et les incertitudes sont comprises entre 0.29 et 0.77 % des vitesses de phase considérées. De la même manière que pour les tests synthétiques, une composition pyrolitique est fixée. Les contraintes sur les profils radiaux en T , ξ et V_S sont détaillées dans le tableau 6.2. Afin d’obtenir des profils de température physiquement réalistes, le gradient de température maximal entre deux points de Bézier consécutifs est fixé à ± 1.5 K/km. Cette valeur est largement supérieure à la valeur du gradient adiabatique du manteau de l’ordre de 0.4 K/km (Katsura et al., 2010), et permet de laisser un maximum de liberté au placement des points de Bézier dans l’espace des paramètres. Les bornes du prior en V_S et ξ sont légèrement différentes de celles employées dans le chapitre précédent. Puisqu’une anisotropie radiale importante est susceptible d’être trouvée, la borne supérieure de ξ est étendue à 1.3 au lieu de 1.1. Étant donné que nous étudions un trajet océanique, des vitesses sismiques supérieures à celles du PREM peuvent être envisagées dans les couches superficielles, c’est pourquoi un intervalle de valeurs de V_S dans les 25 premiers kilomètres plus important est choisi. Imposer des bornes de recherche relativement larges permet d’anticiper des phénomènes de compensation entre les différents paramètres, qui pourraient se produire si certains paramètres ne peuvent pas atteindre les valeurs adéquates.

	T	ξ	V_S
nombre de points de Bézier	6 à 10	7	5
gamme de profondeurs (km)	350- 1000	0- 1000	0- 350
distance minimale d_z entre 2 points consécutifs (km)	80 à 10	40	20
taille du vecteur tangent local Δ (km)	40 à 5	20	10
gradient maximal entre 2 points consécutifs	1.5 K/km		

TAB. 6.2 : Synthèse des contraintes employées pour le positionnement des points de Bézier et des points de définition.

6.3 Résultats

6.3.1 Vitesse de cisaillement dans la croûte et le manteau supérieur

Les résultats en V_S sont présentés sur la figure 6.3. Entre 50 et 350 km de profondeur, la distribution se situe au centre du prior et est assez proche des valeurs du PREM. En surface, les valeurs de V_S sont élevées par rapport au PREM, ce qui est cohérent avec un trajet océanique. Contrairement aux tests synthétiques, toutes les profondeurs sont échantillonnées par les points de Bézier. Ce phénomène est lié aux plus grandes incertitudes estimées sur les données réelles pour les courtes périodes, comparées à celles utilisées pour les tests synthétiques.

Entre 50 et 150 km de profondeur, V_S décroît progressivement pour venir finalement se raccorder au PREM (étiquettes [1]). Il s'agit d'une zone à faible vitesse qui a été décrite sous les océans dans plusieurs études (*e. g.* Gaboret et al., 2003; Gung et al., 2003; Nettles & Dziewonski, 2008).

Entre 150 et 200 km de profondeur, les valeurs sont équivalentes à celles du PREM, puis un changement de pente se produit et les profils viennent se raccorder au PREM juste au-dessus de 300 km (étiquettes [2]). Le changement de pente est clairement visible sur les points de Bézier, où ils se divisent en 2 branches (étiquettes [3]). Cette modification du gradient est assimilée à la discontinuité de Lehmann², qui se produit à 220 km de profondeur dans le PREM. Cette discontinuité est ici totalement lissée car les ondes de surface ne sont pas aussi sensibles à la localisation exacte des discontinuités, comme le sont les phases réfléchies ou converties, par exemple.

6.3.2 Température

La figure 6.4 montre que la distribution en température correspond globalement à un adiabat et qu'elle possède des valeurs inférieures au profil adiabatique à 0.4 K/km avec une température potentielle de 1600 K. Comme pour les tests synthétiques, les points de Bézier sont répartis à toutes les profondeurs et les profils ne sont pas dépendants du nombre de points utilisé (voir tableau 6.3). La température est particulièrement bien contrainte entre 350 et 550 km de profondeur, où les pdfs des distributions en points et en courbes de Bézier sont maximales. Les distributions sont plus étalées au cœur de la zone de transition, entre 550 et 750 km de profondeur. Il s'agit des mêmes caractéristiques que celles observées sur les tests synthétiques (voir section 5.4.1). Sur la figure 6.4c, 3

²La discontinuité de Lehmann est associée à un changement de l'anisotropie sismique. Elle pourrait refléter une modification du mécanisme responsable de l'orientation des cristaux dans le manteau, bien qu'il existe d'autres possibilités (voir Deuss & Woodhouse, 2004). Cette discontinuité n'est pas prise en compte dans le AK135 (Kennett et al., 1995).

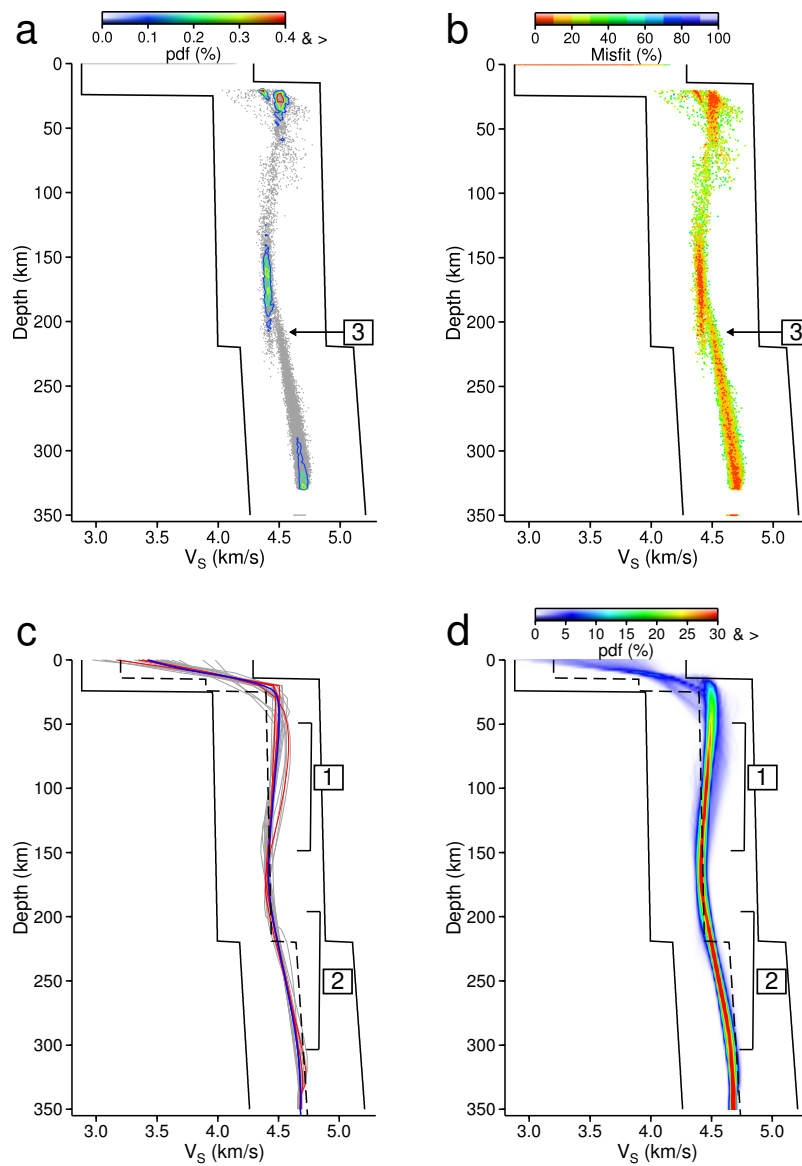


FIG. 6.3 : Résultats en vitesse de cisaillement. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). Le modèle représenté en bleu est le profil médian. (d) Pdf des courbes de Bézier. Les lignes en pointillés sur les figures (c) et (d) correspondent aux valeurs du PREM.

familles de profils peuvent être distinguées. La première famille (en rouge) correspond quasiment à un adiabat et se rapproche de la médiane de la distribution (en bleu). Le gradient du profil médian est d'environ 0.22 K/km, ce qui est légèrement plus faible que le gradient supposé dans le manteau (Katsura et al., 2010). Si les valeurs du profil médian dans la zone de transition sont extrapolées jusqu'en surface, une température potentielle d'environ 1600 K est obtenue. La seconde famille de modèles (en vert) et la troisième (en

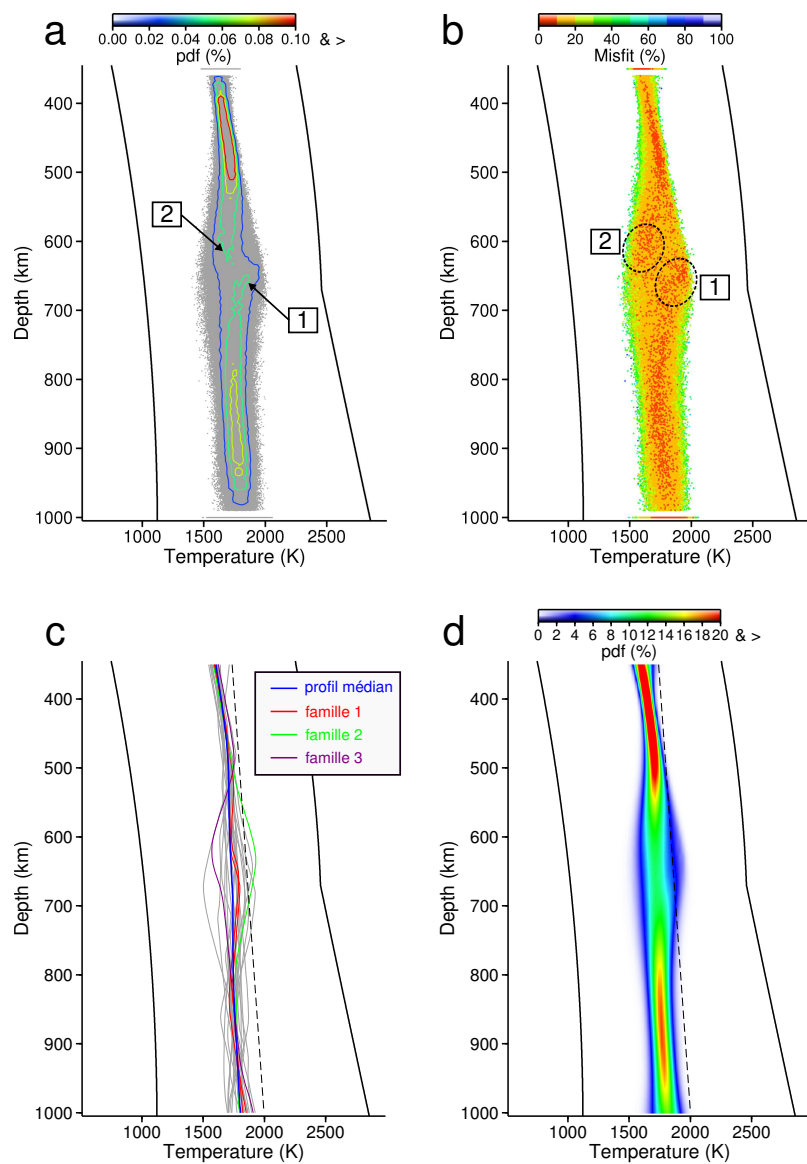


FIG. 6.4 : Résultats en température. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge, vert et violet). Le modèle représenté en bleu est le profil médian. (d) Pdf des courbes de Bézier. Les lignes en pointillés sur les figures (c) et (d) correspondent à l'adiabat à 1600 K avec un gradient de 0.4 K/km.

violet), sont caractérisées respectivement par une température supérieure et inférieure à celle du profil médian, au centre de la zone de transition. Deux des profils correspondant aux meilleurs misfits font partie des familles 2 et 3, ce qui implique qu'une augmentation ou une diminution de la température d'environ 170 K par rapport au profil médian au centre de la zone de transition peut expliquer les mêmes données qu'un profil adiabatique, compte-tenu des incertitudes sur ces données.

La pdf sur les courbes de Bézier apparaît lisse et unimodale entre 550 et 750 km de profondeur. En revanche, la pdf sur les points semble mettre en valeur 2 amas de points, l'un correspondant à une solution « froide » vers 600 km de profondeur (étiquettes [2]) et l'autre à une solution « chaude » vers 670 km (étiquettes [1]). Ces configurations sont retrouvées sur la figure 6.4b, où des amas de points avec de faibles misfits sont visibles, ce qui serait en bonne adéquation avec les familles de profils 2 et 3 de la figure 6.4c. La figure 6.5 montre les probabilités marginales à 600 et 670 km de profondeur. Les marginales sont représentées à la fois pour les points et les courbes de Bézier. Pour les courbes, les températures les plus échantillonnées à 600 et 670 km sont similaires et valent environ 1710 K. Pour les points, la température prédominante à 600 km est légèrement plus faible (1670 K). À 670 km, l'histogramme présente un pic à environ 1850 K et la distribution est

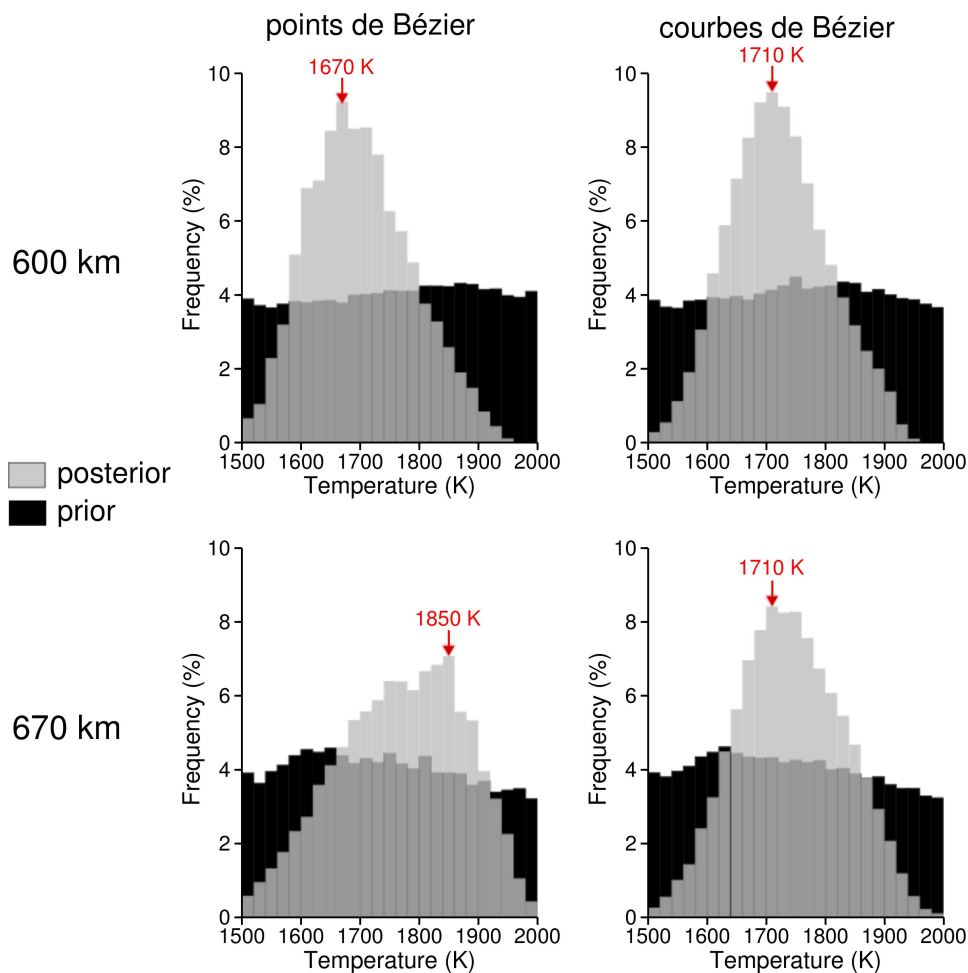


FIG. 6.5 : Probabilités marginales *a priori* (en noir) et *a posteriori* (en gris) sur la température à 600 (en haut) et 670 km de profondeur (en bas). Les résultats sont montrés pour les points de Bézier (à gauche) et les courbes de Bézier (à droite). La taille des intervalles est de 20 K. La valeur centrale de l'intervalle de température le plus échantillonné pour chaque histogramme est indiquée en rouge.

significativement différente de celle des courbes. Les résultats statistiques sur les points montrent qu’il existe bien une solution « froide » et une solution « chaude », plus profonde. Ces solutions ne sont statistiquement pas visibles en examinant la pdf des modèles 1D seule, d’où l’intérêt d’analyser différents types de représentation.

Comme dans le cas des tests synthétiques (voir section 5.4.1), la variété des profils de température au cœur de la zone de transition illustre l’incertitude induite par les transformations minéralogiques sur le résultat de l’inversion. Ce point est analysé plus en détail dans les sections 6.3.5 et 6.4.

nombre de points de Bézier	6	7	8	9	10
trajet Vanuatu-Californie	22.50	16.25	25.00	18.75	17.50

TAB. 6.3 : Nombre de points de Bézier (en %) sélectionnés pour les processeurs engagés dans la période stationnaire.

6.3.3 Anisotropie radiale

Entre la surface et 350 km de profondeur, une anisotropie radiale $\xi \sim 1.05$ est observée (figure 6.6). Cette structure apparaît robuste car les points de Bézier sont tous concentrés aux alentours de cette valeur et montrent des misfits relativement faibles.

Entre 350 et 450 km de profondeur, l’anisotropie diminue progressivement pour atteindre la valeur de 1. Ensuite la distribution reste proche de $\xi = 1$ jusqu’à 1000 km. Comme nous l’avons déjà observé avec les tests synthétiques (voir section 5.4.2), la répartition des points de Bézier devient plus éparse au-dessous de 700 km, liée à la diminution de la sensibilité des ondes de surface.

Les profils 1D représentés sur la figure 6.6c montrent que l’amplitude du signal anisotrope diminue au fur et à mesure que l’on se rapproche de la zone de transition. Le passage du manteau supérieur à la zone de transition est caractérisé sur plusieurs profils par le passage d’une région où $V_{SH} > V_{SV}$ ($\xi > 1$), à une région où $V_{SH} < V_{SV}$ ($\xi < 1$) (étiquette [1]). Deux des meilleurs profils présentent cette caractéristique. Entre 400 et 500 km de profondeur, il existe des points de Bézier avec de faibles misfits et des valeurs de ξ inférieures à 1 (étiquette [2]). Cependant, les pdfs sur les points et les courbes de Bézier sont faibles lorsque $\xi < 1$ à ces profondeurs (étiquettes [3]), ce qui signifie que ces valeurs n’ont été que peu échantillonnées. Les probabilités marginales en points et en courbes à 450 km de profondeur sont présentées sur la figure 6.7. Pour les points, l’histogramme montre clairement que $\xi = 1$ est la valeur la plus probable. Pour les courbes, des valeurs comprises entre 1 et 1.01 sont les plus échantillonnées. D’après l’ensemble de ces résultats, la zone de transition entre 450 et 1000 km de profondeur peut être considérée comme étant isotrope.

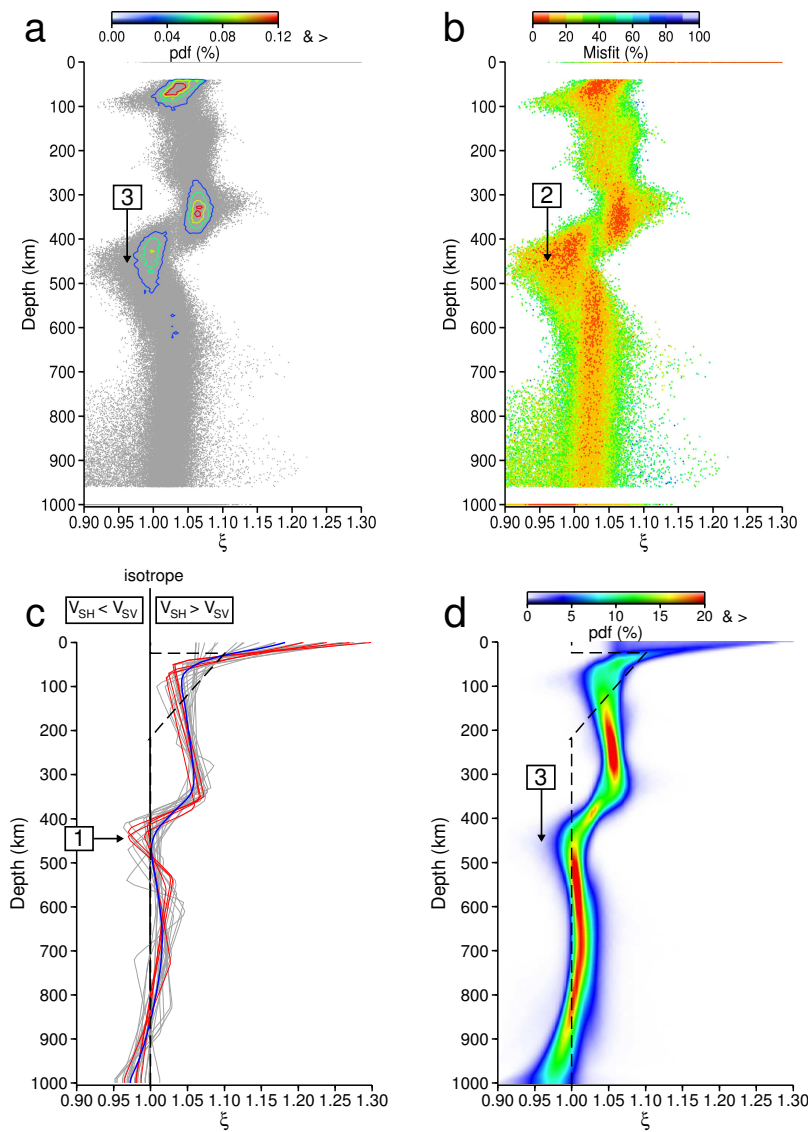


FIG. 6.6 : Résultats en anisotropie radiale. Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). Le modèle représenté en bleu est le profil médian. (d) Pdf des courbes de Bézier. Les lignes en pointillés sur les figures (c) et (d) correspondent aux valeurs du PREM.

En résumé, une forte anisotropie dans le manteau supérieur et un comportement quasi isotrope dans la zone de transition caractérisent les données de Visser et al. (2008a) pour le trajet Vanuatu-Californie.

Ce résultat dans le manteau supérieur est différent de l'anisotropie radiale du PREM, mais est en bonne adéquation avec l'importante anisotropie radiale positive dans le manteau supérieur ($V_{SH} > V_{SV}$) trouvée dans les études globales de tomographie sismique. Par exemple, l'étude de Ekström & Dziewonski (1998) a mis en évidence une forte aniso-

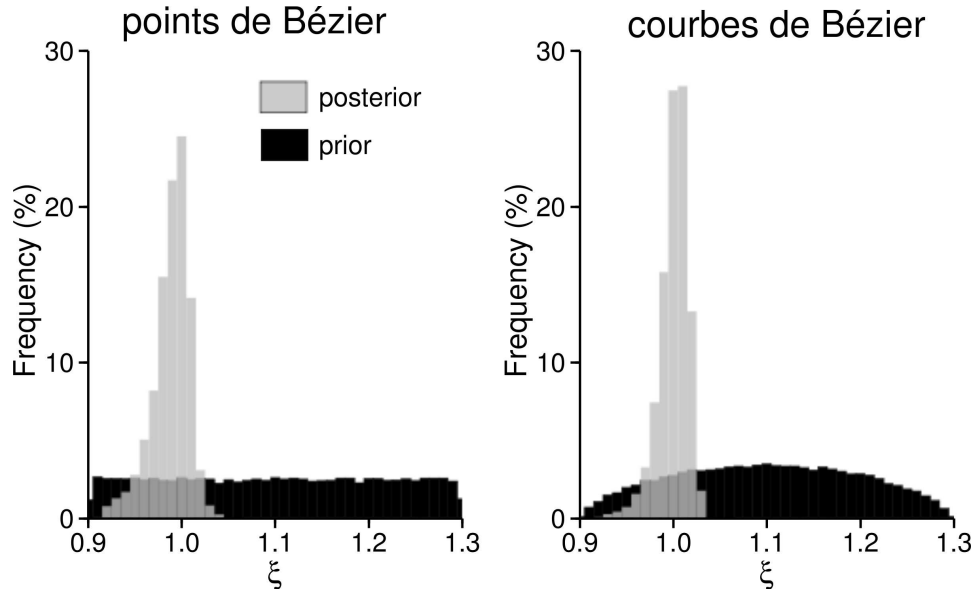


FIG. 6.7 : Probabilités marginales *a priori* (en noir) et *a posteriori* (en gris) sur l'anisotropie à 450 km de profondeur, pour les points et les courbes de Bézier. La taille des intervalles est de 0.01.

tropie radiale dans les 200 premiers kilomètres sous l'Océan Pacifique. [Panning & Romanowicz \(2006\)](#) ont également reporté une importante anomalie $V_{SH} > V_{SV}$ à 150 km de profondeur. L'origine de cette anomalie n'est pas très bien comprise, bien que plusieurs auteurs aient proposé que cette région d'anisotropie corresponde au découplage entre la lithosphère et l'asthénosphère ([Gaboret et al., 2003](#); [Gung et al., 2003](#)). Les variations de l'anisotropie radiale sont parfois utilisées pour estimer la limite entre la lithosphère et l'asthénosphère ([Yuan et al., 2011](#)). Cette limite permet ici de détecter de l'anisotropie radiale sur le trajet Vanuatu-Californie.

Dans la zone de transition, les modèles globaux 1D d'anisotropie ne s'accordent pas sur une seule et même structure. [Montagner & Kennett \(1996\)](#) trouvent une anisotropie « positive » ($V_{SH} > V_{SV}$) juste au-dessus de 670 km de profondeur, qui se change ensuite en une anisotropie « négative » ($V_{SH} < V_{SV}$) au-dessous de cette profondeur. En revanche, l'étude de [Panning & Romanowicz \(2006\)](#) montre un signal plus faible en amplitude et de signe opposé. Au travers d'une approche probabiliste, l'étude plus récente de [Visser et al. \(2008b\)](#) présente une anisotropie négative qui s'étend jusqu'à 670 km (figure 6.8). Cependant, l'amplitude de cette anomalie reste très faible et ne dépasse pas la valeur de $\xi = 0.985$ pour le modèle moyen. Compte-tenu des incertitudes sur les vitesses de phase intégrées le long d'un trajet employées dans cette thèse et la démarche probabiliste adoptée, la zone de transition sous l'Océan Pacifique paraît principalement isotrope. La figure 6.9 montre que les pdfs des courbes de dispersion des ondes de Love et de Rayleigh se situent à l'intérieur des incertitudes déterminées par [Visser et al. \(2008a\)](#).

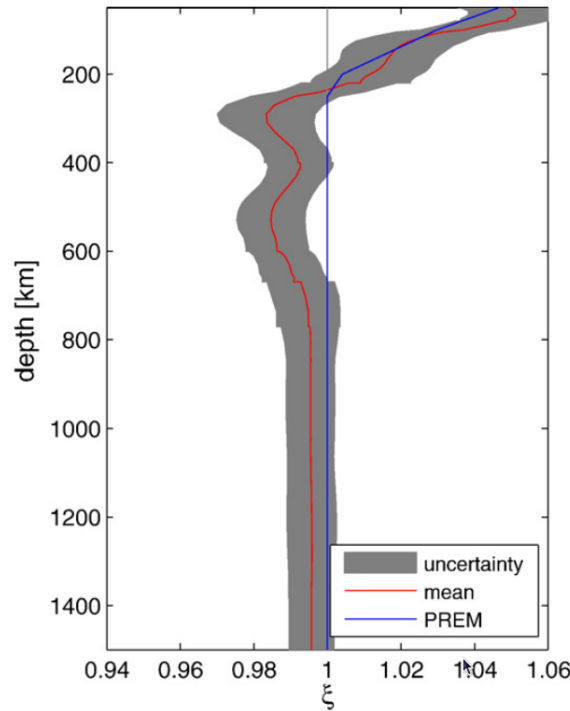


FIG. 6.8 : Modèle moyen d'anisotropie radiale de [Visser et al. \(2008b\)](#). La surface grise représente l'écart-type à 2σ .

6.3.4 Corrélation entre les paramètres primaires

Sur la figure 6.10 sont présentées les corrélations entre les paramètres primaires de l'inversion à certaines profondeurs.

La figure 6.10a montre les corrélations entre T et ξ à 450, 600 et 670 km de profondeur. À 450 km, ξ est centrée sur la valeur de 1 et T autour de 1690 K. À 600 et 670 km de profondeur, les distributions en température sont beaucoup plus étalées, ce qui est cohérent avec l'accumulation des changements de phase et des transformations minéralogiques à ces profondeurs. L'anisotropie est centrée sur $\xi = 1.01$. Aucune tendance particulière entre T et ξ n'est remarquée.

La figure 6.10b présente les corrélations entre V_S et ξ à 25 et 210 km de profondeur. À 210 km, V_S et ξ sont bien contraints. Le maximum de la pdf vaut $\xi = 1.05$, ce qui correspond au maximum d'anisotropie estimé dans le manteau supérieur. La pdf possède une forme circulaire, avec une « traînée » vers les vitesses plus faibles. Cette dernière est un résidu lié au changement de pente vers 200 km de profondeur (voir figure 6.3, étiquettes 3). En revanche, à 25 km, l'anisotropie est moins bien contrainte et s'étend de 1.05 à

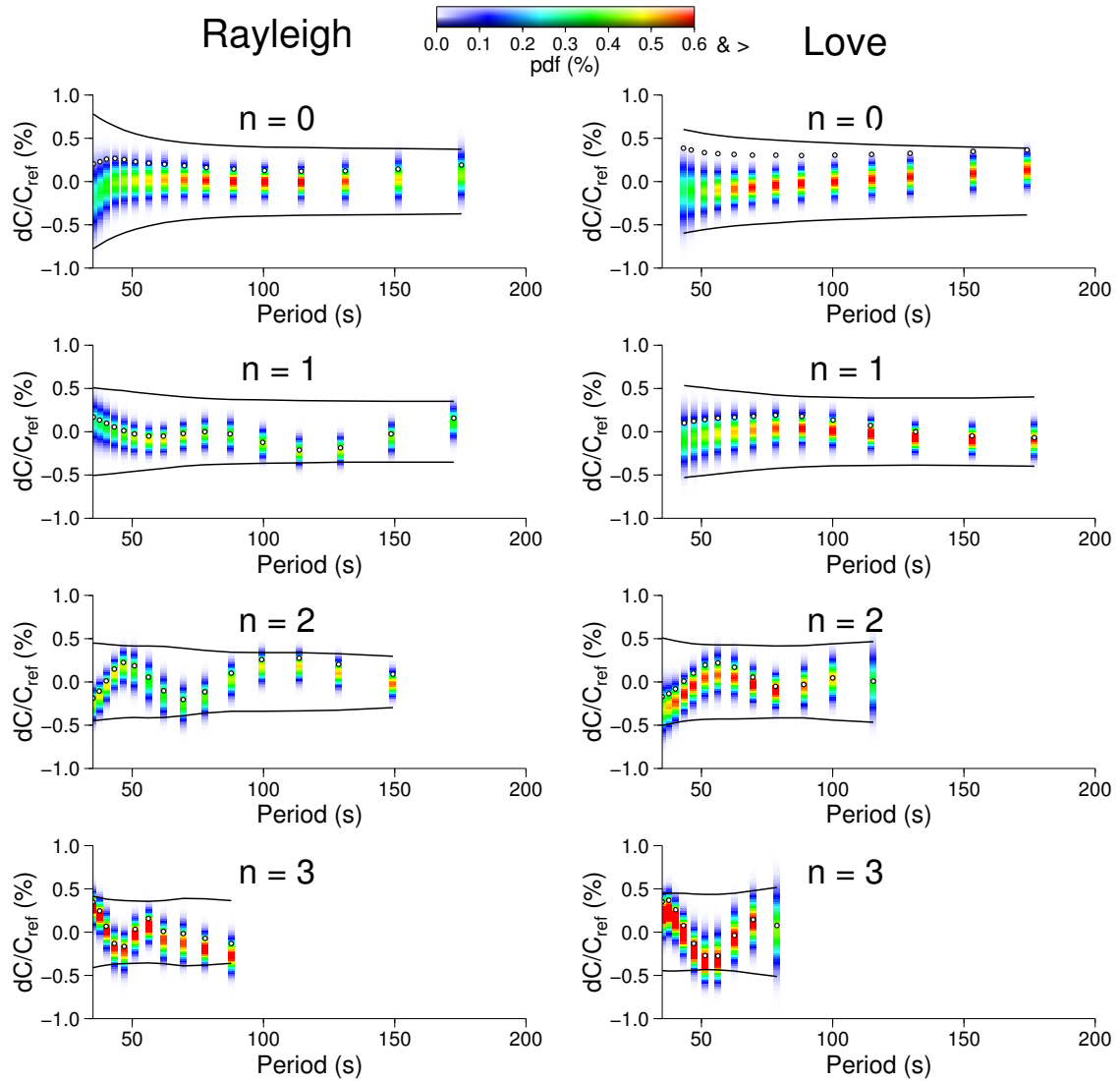


FIG. 6.9 : Distributions *a posteriori* des vitesses de phase du mode fondamental et des modes harmoniques des ondes de Rayleigh et de Love. Les résultats sont montrés en pourcentage de la vitesse de phase testée (C) comparée à la vitesse de phase réelle (C_{ref}). Les courbes noires correspondent aux incertitudes sur les vitesses de phase de [Visser et al. \(2008a\)](#). Les points blancs indiquent les vitesses de phase obtenues pour le modèle médian issu des distributions en V_S , T et ξ .

1.15. Cette large distribution peut s'expliquer par le fait que la description de la croûte n'a pas été prise en compte de manière assez fine, par exemple, en forçant l'intégration de plusieurs points de Bézier en V_S et en ξ entre la surface et 25 km de profondeur. Pour V_S , il existe plusieurs solutions localisées autour de 4.5 km/s.

D'après l'ensemble de ces résultats, nous pouvons affirmer que les paramètres primaires de l'inversion V_S , T et ξ , ne montrent aucun signe de corrélation. Les valeurs inversées sont donc bien indépendantes.

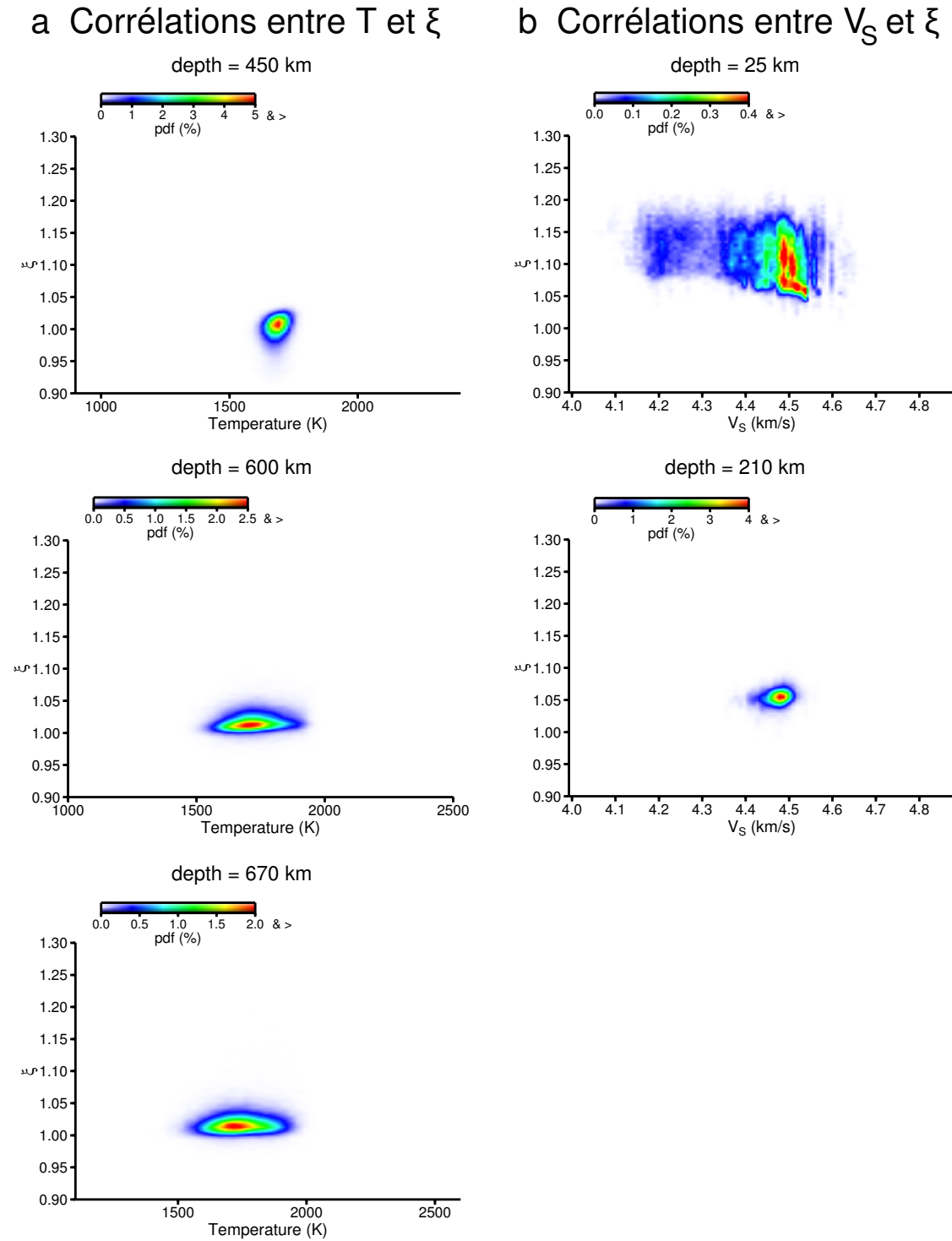


FIG. 6.10 : Densités de probabilité marginales montrant les corrélations entre les paramètres primaires de l'inversion à des profondeurs fixées. Les échelles en abscisse et en ordonnée correspondent à la taille du prior. (a) Corrélations entre T et ξ à 450, 600 et 670 km de profondeur. (b) Corrélation entre V_S et ξ à 25 et 210 km de profondeur.

6.3.5 Les distributions en V_S , V_P isotropes et ρ dans la zone de transition

Les vitesses des ondes S et P , ainsi que la masse volumique ρ entre la surface et 1000 km de profondeur sont montrées sur la figure 6.11. Ces propriétés physiques sont représentées sous la forme de pdfs. De 350 à 1000 km de profondeur, il s'agit des paramètres secondaires de l'inversion, calculés à partir des diagrammes de phases et de l'équation d'état de Birch-Murnaghan (voir section 2.2).

Dans la première moitié de la zone de transition, entre 410 et environ 550 km de profondeur, les vitesses des ondes S et P sont plus importantes que celles du PREM. Au contraire, dans la partie inférieure, juste au-dessus de la transformation minéralogique de la ringwoodite en pérovskite et magnésioiwüstite, les vitesses sismiques sont inférieures au PREM. Nos résultats montrent des gradients de vitesses plus forts dans la zone de transition que ceux du PREM.

La transition de phase olivine $\rightarrow$ wadsleyite s'étend entre 390 et 415 km de profondeur, ce qui est en bonne adéquation avec la discontinuité à 400 km du PREM. La température la plus fréquemment échantillonnée à 400 km est ~ 1650 K. Cette valeur est contenue à l'extrémité des barres d'erreurs obtenues à partir des expériences hautes pressions, où la transition de phase a lieu à $\sim 1750 \pm 100$ K (Ito & Takahashi, 1989). Une caractéristique frappante est la magnitude de cette discontinuité pour les vitesses des ondes S et P . Au contraire, le saut de vitesse de la masse volumique est du même ordre de grandeur que le PREM. Ce désaccord entre le saut de vitesse estimé à partir des valeurs des modules élastiques déterminés par les expériences de laboratoire et le saut de vitesse des modèles globaux a déjà été mis en évidence par Stixrude (1997) et Katsura et al. (2004) pour une composition pyrolitique, et plus récemment par Irifune et al. (2008), Xu et al. (2008), ou encore Khan et al. (2009) pour d'autres compositions. L'amplitude de la discontinuité pourrait être réduite en considérant un modèle minéralogique moins riche en olivine, comme proposé par Duffy et al. (1995). Dans les modèles sismiques globaux, il se pourrait que les amplitudes des sauts de vitesses aient été sous-estimées (Shearer, 2000). Cette sous-estimation pourrait être liée aux effets de fréquence finie des ondes sismiques (Jackson, 2008). Puisque les ondes de surface sont peu sensibles à la localisation et à l'épaisseur des discontinuités, nous n'irons pas plus loin dans la discussion de ce résultat.

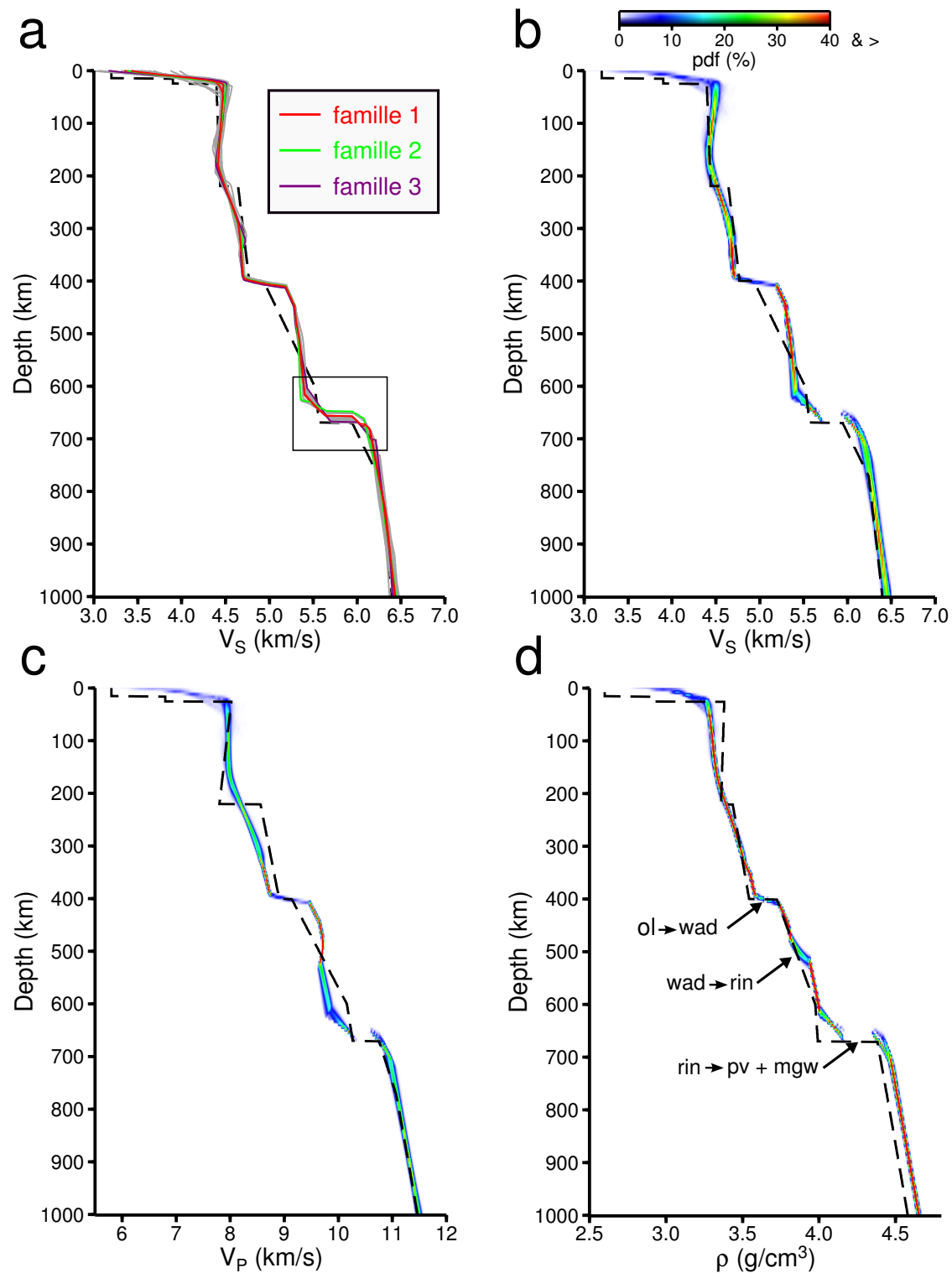


FIG. 6.11 : Pdfs en (b) V_S , (c) V_P et (d) ρ . Les pointillés noirs sont les valeurs du PREM. La figure (a) présente 20 modèles en V_S échantillonnés aléatoirement (en gris) et les 4 modèles qui correspondent aux meilleurs misfits (en rouge, vert et violet). Les structures à l'intérieur du rectangle noir sont détaillées sur la figure 6.13.

La discontinuité sismique associée à la transition de phase wadsleyite $\rightarrow$ ringwoodite (la "520") n'est pas présente dans le PREM. Elle n'est pas non plus visible sur les profils en V_S mais apparaît sur les profils de masse volumique. Cette transition a lieu de façon plus superficielle, entre 490 et 520 km de profondeur, ce qui est cohérent avec son caractère exothermique (voir tableau 1.2). En effet, puisque la distribution en température (figure 6.4) est globalement inférieure au profil adiabatique à 1600 K, il est logique que la transition ait lieu au-dessus de 520 km. Le saut de masse volumique est plus faible que dans le cas de la "410" et de la "660". Entre 490 et 520 km de profondeur, la pdf de V_P montre une légère diminution des vitesses sismiques, qui témoigne de la transition wadsleyite $\rightarrow$ ringwoodite. Ce comportement est intrinsèquement lié aux valeurs des modules élastiques et de leurs dérivées choisies pour la wadsleyite et la ringwoodite, et a déjà été mis en évidence par Cammarano et al. (2005a) ou encore Cobden et al. (2008).

Entre 600 et 700 km de profondeur, les distributions des paramètres sismiques sont plus complexes. Le problème inverse est fortement non-linéaire et les distributions *a posteriori* au cœur de la zone de transition sont loin de ressembler à une distribution gaussienne unimodale. Pour illustrer cet effet, les probabilités marginales sur V_S à 655, 660 et 665 km de profondeur sont représentées sur la figure 6.12. Ces coupes se situent aux alentours de la transformation minéralogique de la ringwoodite en pérovskite et magnésio-wüstite. L'histogramme à 660 km possède 2 maxima, ce qui signifie que la distribution marginale est influencée par les valeurs de V_S au-dessus et au-dessous de la discontinuité. Si l'on se réfère à la figure 6.11a, on observe que la transformation minéralogique a lieu sur un large intervalle de profondeurs. Globalement, les discontinuités de la distribution se situent au-dessus de celle du PREM. Sur cette figure sont représentés les profils de V_S qui correspondent aux 3 familles de profils thermiques différentes (voir section 6.3.2). Les profils de vitesse sont cohérents avec les profils de température. La transformation

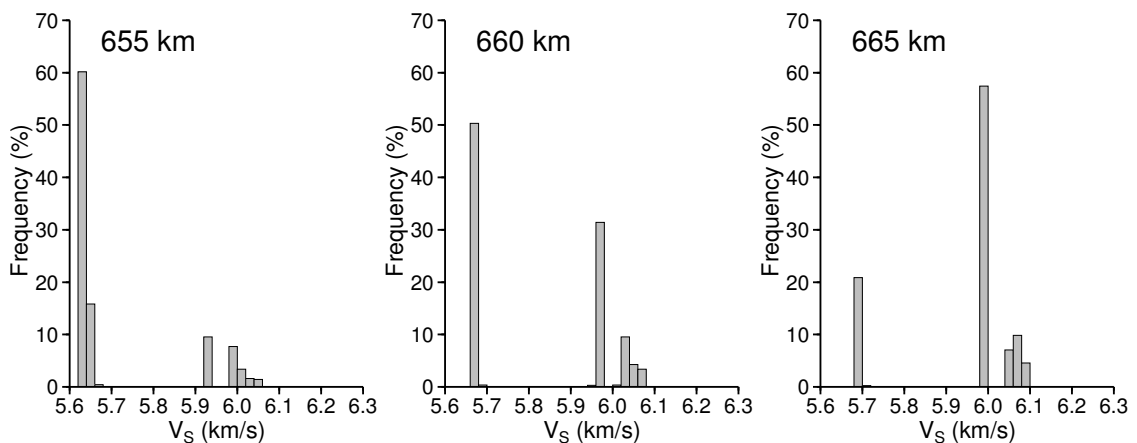


FIG. 6.12 : Probabilités marginales *a posteriori* des vitesses de cisaillement à 655, 660 et 665 km de profondeur. La taille des intervalles est de 0.02 km/s.

ringwoodite $\rightarrow$ pérovskite + magnésio-wüstite étant endothermique, c'est le profil de température le plus froid dans cette région (famille 3) qui produit la discontinuité sismique la plus profonde, vers 668 km. En revanche, le profil le plus chaud (famille 2) présente une discontinuité plus superficielle, à environ 648 km. Les discontinuités des profils de température représentatifs de la famille 1, proches du modèle médian, ont une profondeur intermédiaire (658 km).

La discontinuité vers 660 km de profondeur n'implique pas uniquement la transformation minéralogique dans le système olivine (voir section 1.1.1). Les vitesses sismiques tendent à former plusieurs gradients qui se superposent, témoignant des transformations qui ont lieu également dans le système pyroxènes-grenat. Les profils de vitesse de cisaillement pour les 3 familles de modèles sont détaillés entre 600 et 725 km de profondeur sur la figure 6.13a. Les fractions volumiques des différents minéraux en fonction de la profondeur sont représentées sur la figure 6.13b. Plus le profil de température est froid, plus les transformations dans le système pyroxènes-grenat vont s'étaler en profondeur. En effet, l'intervalle de profondeurs sur lesquelles ont lieu les transformations est de seulement 40 km pour un profil chaud (famille 2), mais atteint 100 km pour un profil froid (famille 3). La présence d'ilménite tend à ajouter des discontinuités discrètes lorsqu'elle se change en pérovskite. Les transformations du grenat ont plutôt tendance à venir insérer des gradients de vitesses. Nous insistons sur le fait qu'observer en détails ces différents gradients n'est possible que parce qu'une méthode thermodynamique impliquant des diagrammes de phases et une équation d'état est employée. En réalité, comme nous le verrons dans la section 6.5, les ondes de surface ne sont pas sensibles à chacune des contributions de ces transformations, mais ont plutôt tendance à détecter les variations de vitesses de grandes longueur d'ondes.

Au-dessous de cette zone complexe, les pdfs des paramètres sismiques entre 700 et 1000 km de profondeur (figure 6.11) montrent que les vitesses des ondes S et P sont égales ou légèrement supérieures à celles du PREM. En revanche, la masse volumique montre des valeurs supérieures d'environ 1.7 %.

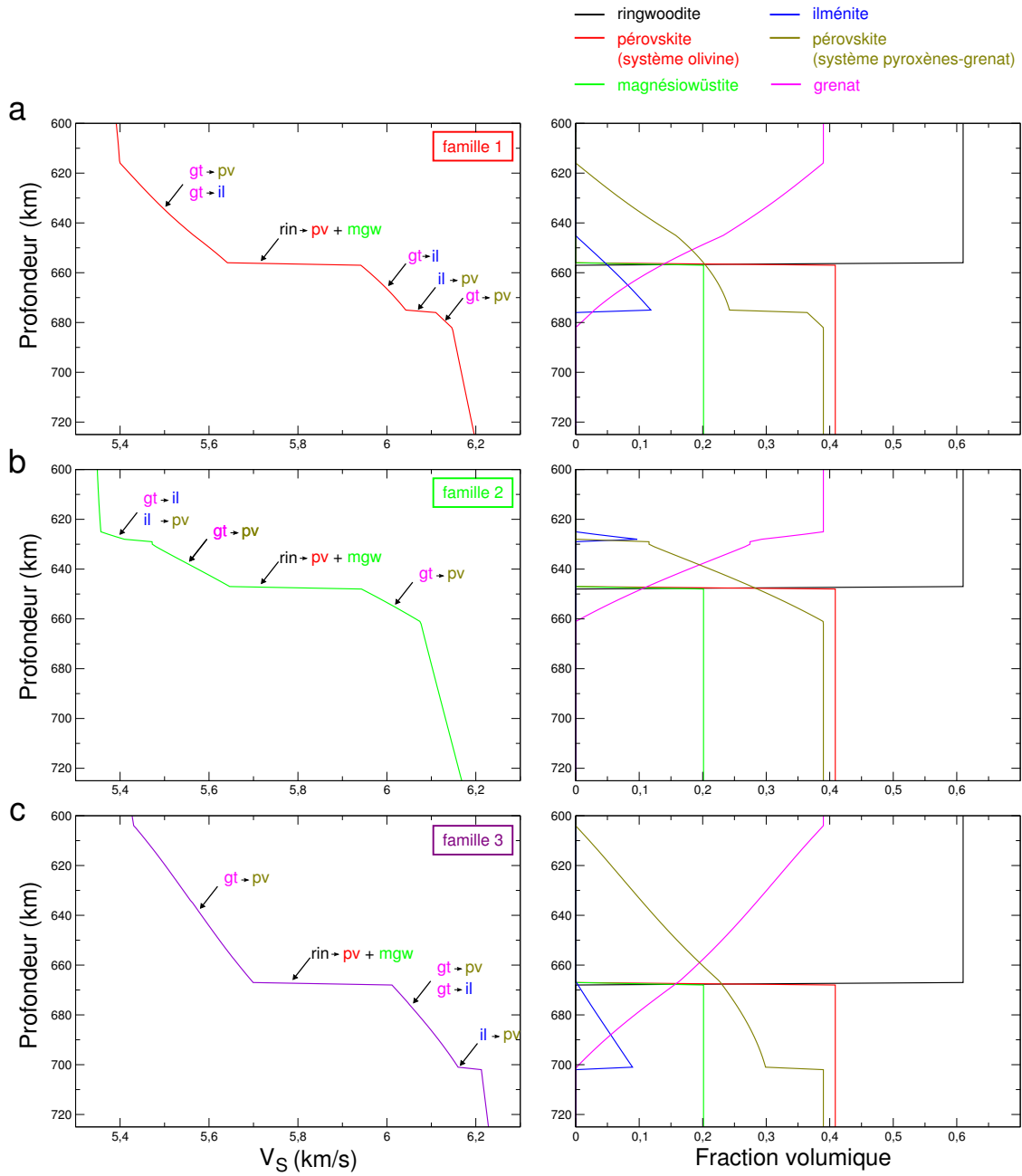


FIG. 6.13 : Détails des différentes transitions de phase et transformations minéralogiques responsables des gradients de vitesses sismiques entre 600 et 725 km de profondeur, pour les 3 familles correspondant aux meilleurs misfits. Les différentes transformations sont annotées sur le profil de vitesse de cisaillement (à gauche). Les fractions volumiques des différents minéraux en fonction de la profondeur sont également indiquées (à droite).

6.4 Influence de la valeur de la pente de Clapeyron à 660 km

Nous avons vu dans la section précédente qu'il existe globalement 3 familles de profils de température qui peuvent expliquer les vitesses de phase intégrées du trajet Vanuatu-Californie. Sur la figure 6.14a sont représentés les profils représentatifs de ces 3 familles ainsi que la distribution des points de Bézier. D'après les conditions de pression et de température, il est également indiqué pour chaque point quelle phase minérale est majoritaire dans le système olivine, c'est-à-dire soit les polymorphes de l'olivine, soit la pérovskite et la magnésioiwüstite. La pente de Clapeyron choisie pour notre étude ($\Gamma_{660} = -2.8 \text{ MPa/K}$, d'après Ito & Takahashi, 1989) est clairement mise en évidence. Nous allons voir que le fait qu'il existe 2 familles de température extrêmes, une « chaude » (famille 2) et une « froide » (famille 3) est lié à la nature endothermique de la transformation minéralogique. En effet, une pérovskite « chaude » vers 660 km de profondeur (famille 2) est capable de rendre compte des mêmes courbes de dispersion qu'une ringwoodite « froide » (famille 3).

L'influence de la valeur de la pente de Clapeyron pour la transformation minéralogique du système olivine sur la distribution en température est testée en fixant une pente de Clapeyron nulle ($\Gamma_{660} = 0 \text{ MPa/K}$, figure 6.14). La transformation a lieu de fa-

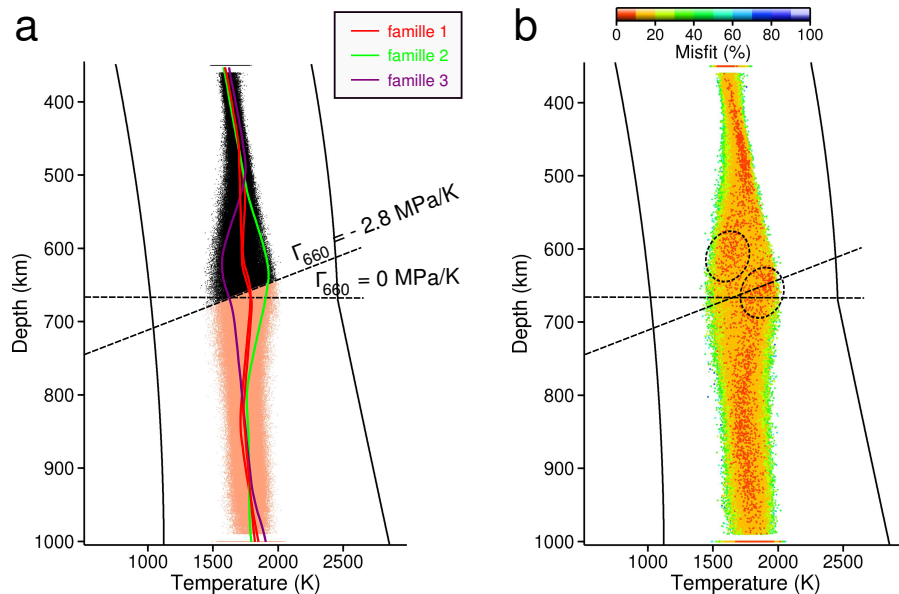


FIG. 6.14 : Distribution des points de Bézier en température. (a) Minéralogie du système olivine selon la position des points dans le diagramme (P, T). Les polymorphes de l'olivine sont représentés en noir et la pérovskite + magnésioiwüstite en saumon. (b) Valeurs du misfit.

çon instantanée lorsque le profil de température franchit le seuil de 24 GPa. Les résultats en V_S entre la surface et 350 km de profondeur et en anisotropie sont très similaires à ceux obtenus pour $\Gamma_{660} = -2.8$ MPa/K et ne sont pas montrés. Cependant, la distribution en température est modifiée de façon significative (figure 6.15). Bien que la répartition des

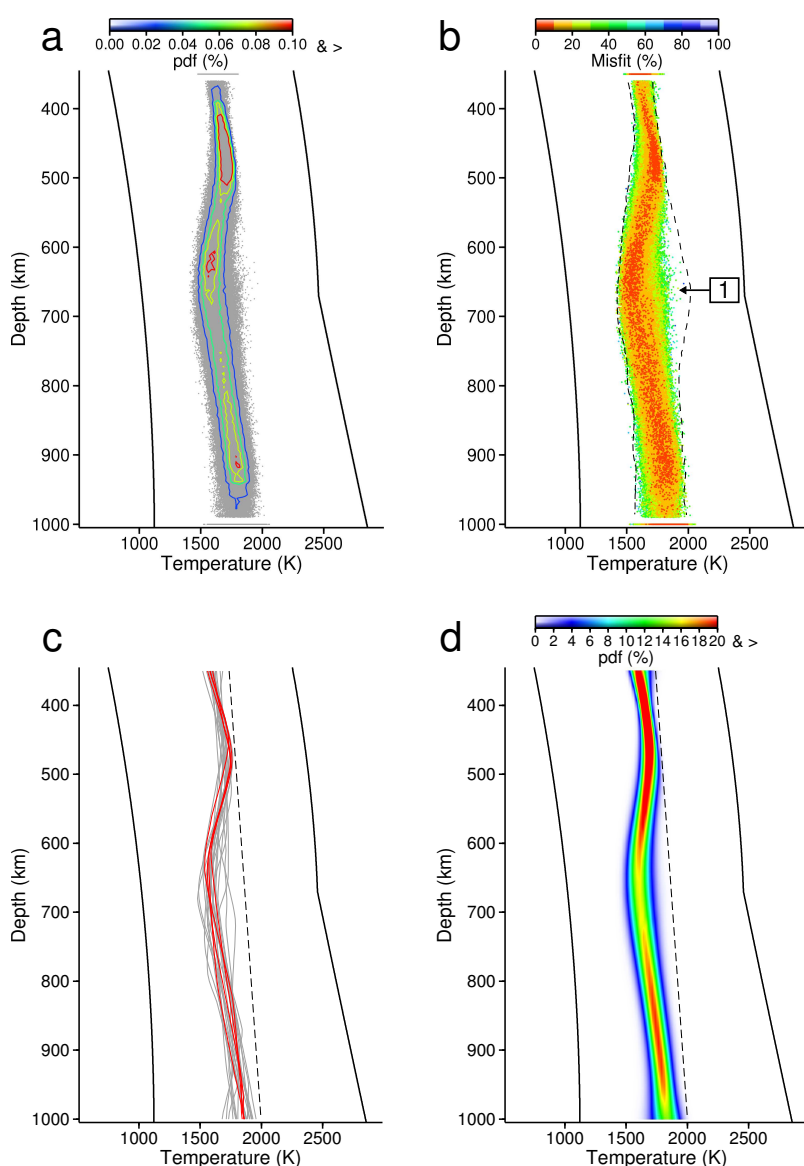


FIG. 6.15 : Résultats en température pour une transformation ringwoodite $\rightarrow$ pérovskite + magnésiowüstite fixée à 24 GPa ($\Gamma_{660} = 0$ MPa/K). Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. Les courbes en pointillés correspondent aux contours de la distribution des points de la figure 6.14b, lorsque $\Gamma_{660} = -2.8$ MPa/K. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). (d) Pdf des courbes de Bézier. Les lignes en pointillés sur les figures (c) et (d) correspondent à l'adiabat à 1600 K avec un gradient de 0.4 K/km.

points et des courbes soient similaires aux résultats obtenus avec une valeur de pente de Clapeyron issue des expériences de laboratoire jusqu'à 500 km de profondeur, l'utilisation d'une pente de Clapeyron nulle induit dans les résultats une diminution de la température au centre de la zone de transition. Cette diminution est de 270 K par rapport au profil adiabatique à 1600 K à son maximum, vers 650 km de profondeur.

Si l'on compare les meilleurs profils obtenus (figure 6.15c) aux 3 familles de profils représentées sur la figure 6.14a, on s'aperçoit que seule la famille 3 est conservée lorsque la profondeur de la transformation minéralogique est fixée. En effet, seule la transition entre une ringwoodite et une pérovskite « froides » à 660 km de profondeur peut toujours expliquer les données, et cette famille de profils est donc préservée. En revanche, un changement minéralogique se produisant de façon plus superficielle, pour lequel un profil de température « chaud » (famille 2) peut rendre compte des données avec une pente de Clapeyron de -2.8 MPa/K (figure 6.15b, étiquette [1]), produit à présent des misfits élevés. La figure 6.16 montre les vitesses de phase du mode 1 des ondes de Rayleigh et du mode 3 des ondes de Love, des profils correspondant aux familles 2 et 3, en supposant $\Gamma_{660} = 0 \text{ MPa/K}$. Les vitesses de phase de la famille 2 sortent des barres d'erreur, ce qui confirme que les données de Visser et al. (2008a) ne peuvent pas s'expliquer avec un profil chaud au centre de la zone de transition si la profondeur de la transformation minéralogique est fixée à 670 km de profondeur.

Les résultats de ce test sont révélateurs du fait que le choix du modèle minéralogique constitue une contrainte importante sur le résultat de l'inversion. En effet, la profondeur des discontinuités dépend des choix sur la minéralogie, qui conditionnent la distribution en température. Dans la section 1.1.1, nous avons constaté qu'il existe des incertitudes quant à la valeur de la pente de Clapeyron de la "660" (elle varie de -0.4 à

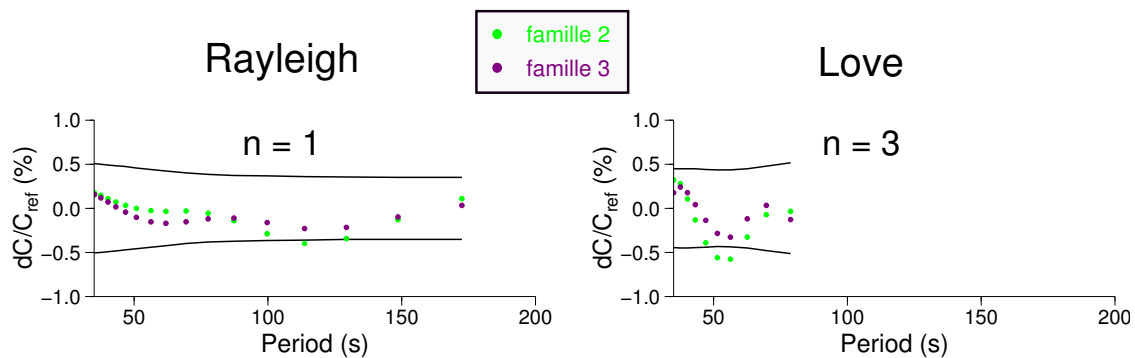


FIG. 6.16 : Valeurs des vitesses de phase des profils correspondant à la famille 2 (en vert) et à la famille 3 (en violet), en considérant $\Gamma_{660} = 0 \text{ MPa/K}$. Les résultats sont montrés en pourcentage de la vitesse de phase testée (C) comparée à la vitesse de phase réelle (C_{ref}). Les courbes noires correspondent aux incertitudes sur les vitesses de phase de Visser et al. (2008a).

–3.0 MPa/K, selon les études). Pour apporter des éléments de réponse sur la sensibilité des ondes de surface à la valeur de la pente de Clapeyron, une inversion en considérant Γ_{660} comme étant un des paramètres de l'inversion pourrait être envisagée.

6.5 Déterminer la distribution de température à partir des vitesses de cisaillement

Dans la section 1.2.5, deux approches généralement suivies pour estimer la température à partir des données sismologiques ont été présentées. Celles-ci sont représentées de façon schématique sur la figure 6.17. La méthode que nous avons choisie d'adopter est de déterminer directement la température à partir des vitesses de phase (nous l'appellerons la méthode 1). Dans ce cas, $V_{S_{\text{isotrope}}}$ n'est utilisée que comme intermédiaire pour calculer les courbes de dispersion. La deuxième méthode (méthode 2) consiste à extraire la température à partir de modèles sismiques, qui correspondent déjà à une interprétation des données. Dans cette section, nous allons décrire les 2 étapes d'inversion nécessaires pour obtenir la distribution de température avec la méthode 2, à savoir le passage des vitesses de phase aux vitesses sismiques, puis la conversion de ces vitesses sismiques en température grâce à l'équation d'état de Birch-Murnaghan et aux diagrammes de phases. Enfin, nous comparerons les résultats sur les distributions en température issues des 2 méthodes.

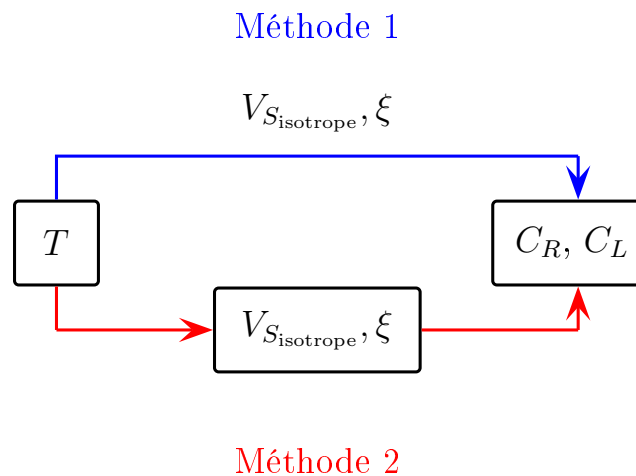


FIG. 6.17 : Illustration schématique du problème direct des deux méthodes généralement employées pour déterminer l'état thermique dans le manteau. Pour la méthode 1, une seule inversion est réalisée, tandis que la méthode 2 fait appel à deux inversions distinctes pour remonter à la distribution de température.

6.5.1 Des vitesses de phase vers la distribution en V_S

Dans un premier temps, seules V_S et ξ sont inversés. Au lieu de définir V_S uniquement entre la surface et 350 km de profondeur comme nous le faisons jusqu'alors, V_S est ici définie jusqu'à 1500 km de profondeur, comme pour ξ . Les bornes de l'espace de recherche des points de Bézier restent $\pm 10\%$ autour des valeurs du PREM. V_P et ρ sont calculés d'après les lois d'échelle (voir section 4.4.1.3). La procédure d'inversion en 3 étapes reste celle employée (voir section 4.3.6). Les contraintes utilisées pour la construction des courbes de Bézier sont détaillées dans le tableau 6.4.

	V_S	ξ
nombre de points de Bézier	11 à 15	7
gamme de profondeurs (km)	0- 1000	0- 1000
gamme de variation des paramètres	$\pm 10\%$ du PREM	0.9- 1.3
distance minimale d_z entre 2 points consécutifs (km)	50 à 10	40
taille du vecteur tangent local Δ (km)	25 à 5	20

TAB. 6.4 : Synthèse des contraintes employées pour le positionnement des points de Bézier et des points de définition.

La pdf des profils d'anisotropie est présentée sur la figure 6.18b. Elle est quasiment similaire à celle obtenue précédemment (figure 6.18a), ce qui montre encore une fois l'absence de corrélation *a posteriori* entre les paramètres de l'inversion. La distribution d'anisotropie radiale demeure très stable, quelle que soit la méthode employée.

Les résultats en V_S sont montrés sur la figure 6.19. Si on compare la pdf sur les courbes de Bézier en V_S (figure 6.19d) à celle obtenue dans la section précédente lors de l'inversion de la température (figure 6.11b), on constate que les discontinuités vers 410 et 660 km de profondeur sont lissées et sont plutôt caractérisées par des changements de gradient plutôt que par des modifications abruptes des propriétés sismiques. Cependant, ces deux distributions présentent plusieurs similitudes :

- une zone à faible vitesse entre 50 et 150 km de profondeur,
- les vitesses sismiques sont inférieures à celles du PREM au-dessous de la discontinuité de Lehmann,
- les vitesses sont supérieures au PREM entre 400 et 550 km de profondeur,
- les vitesses sont inférieures au PREM au-dessous de 550 km et juste avant la discontinuité à 660 km.

La répartition des points de Bézier (figures 6.19a et 6.19b) à l'intérieur du prior indique que V_S est moins bien résolue au-dessous de 500 km de profondeur. En effet, passée cette profondeur, les points remplissent quasiment l'ensemble du prior et des points présentant de faibles valeurs de misfit ($< 10\%$) sont placés vers les extrémités du prior. Ces

points correspondent à des oscillations des profils de vitesse au-dessous de 550 km (figure 6.19c). Plus la profondeur est importante, plus les profils se mettent à osciller. Même pour les modèles correspondant aux meilleurs misfits, les oscillations deviennent significatives à partir de 750 km de profondeur, ce qui est révélateur de la diminution de la sensibilité des ondes de surface avec la profondeur. Le nombre de points de Bézier employé pendant la période stationnaire est détaillé dans le tableau 6.5. Les profils constitués de 11 et 12 points ont été peu sélectionnés. Ce sont les modèles construits avec le nombre maximal de points autorisé, c'est-à-dire 15 points, qui ont été le plus souvent retenus. Cette tendance à sélectionner le maximum de points peut s'expliquer par la diminution de la sensibilité des données avec la profondeur, qui induit que le misfit est moins sensible au placement des points de Bézier en profondeur et crée des oscillations des profils. Ce phénomène pourrait être atténué en introduisant un paramètre de lissage des profils, comme le système de gradient maximal employé pour la température (voir section 5.5.4), mais laisser libre la position des points présente l'intérêt de détecter les régions de l'espace des modèles qui sont moins bien contraintes par les données.

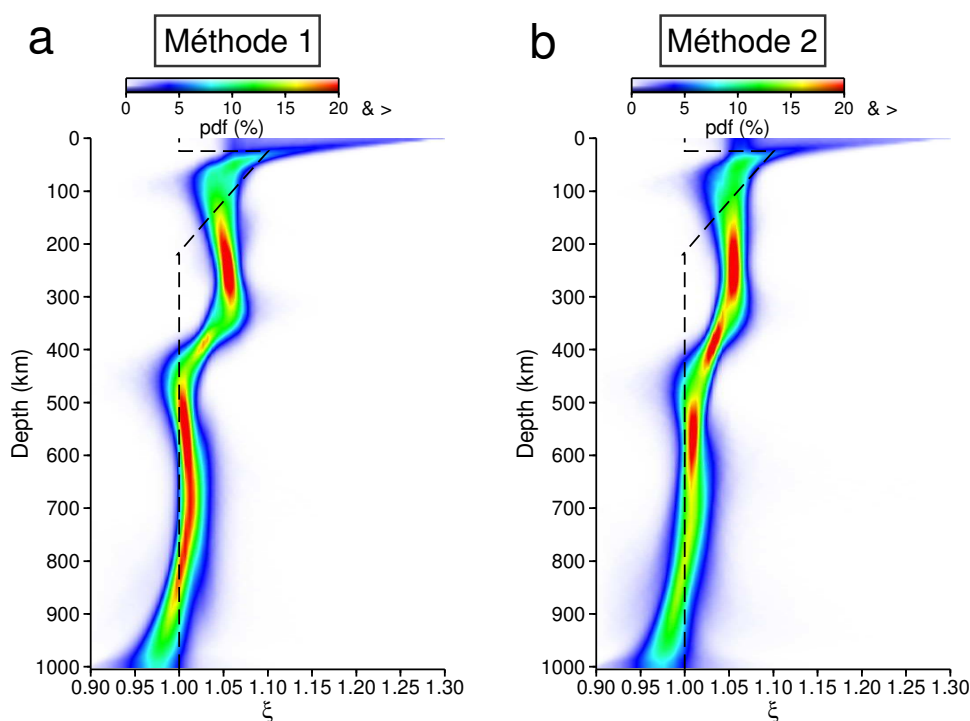


FIG. 6.18 : Résultats en anisotropie pour (a) l'inversion en V_S , T et ξ , et pour (b) l'inversion en seulement V_S et ξ . Les lignes en pointillés correspondent aux valeurs du PREM.

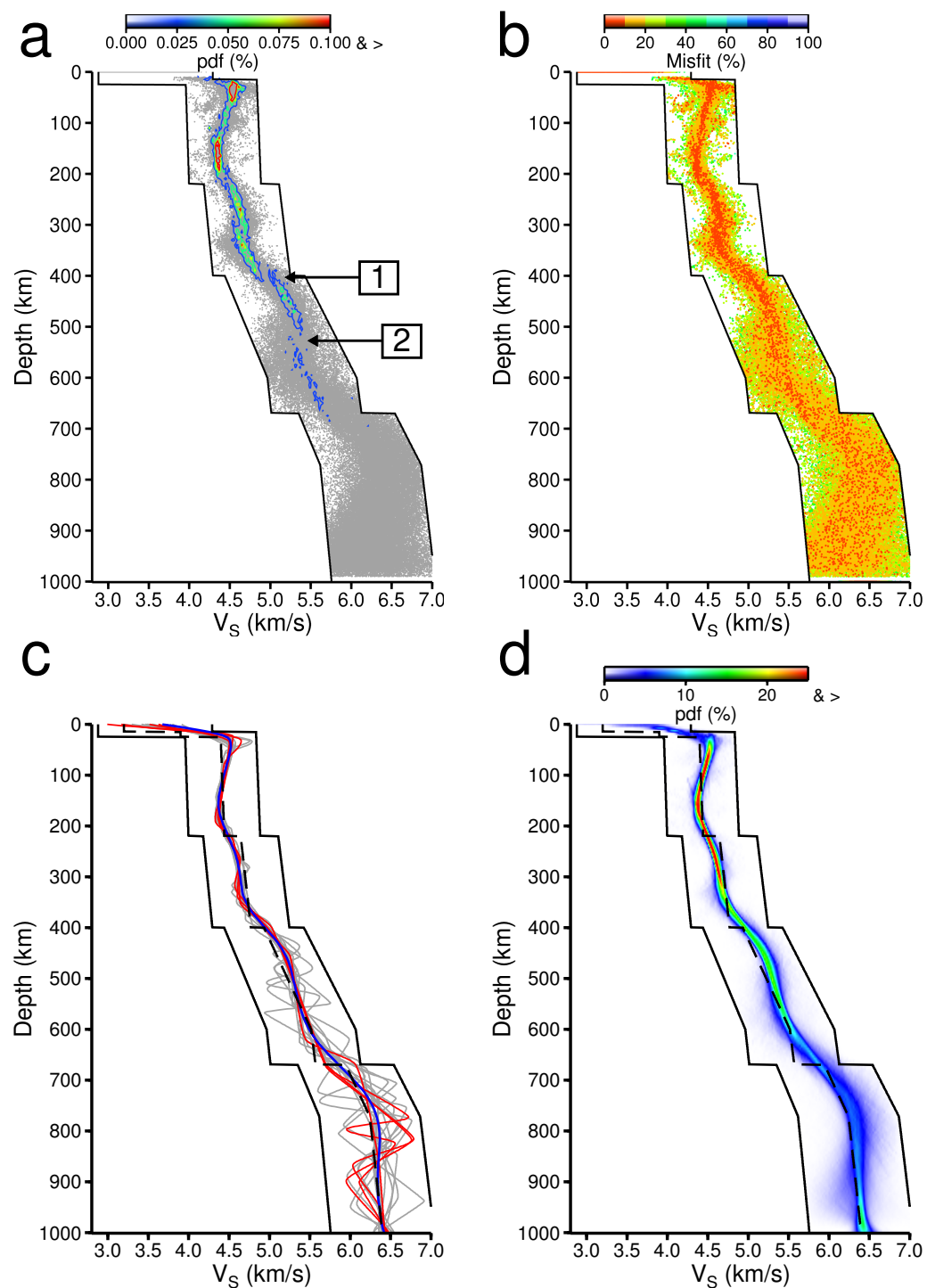


FIG. 6.19 : Résultats en V_S . Les étiquettes correspondent à des points ou des observations particulières décrites dans le texte. (a) Répartition des points de Bézier et sa pdf. (b) Répartition des points de Bézier et valeurs du misfit. (c) 20 modèles choisis aléatoirement (en gris) et les 4 meilleurs modèles (en rouge). Le modèle représenté en bleu est le profil médian. (d) Pdf des courbes de Bézier. Les lignes en pointillés sur les figures (c) et (d) correspondent aux valeurs du PREM. Les lignes noires sont les limites du prior.

nombre de points de Bézier	11	12	13	14	15
trajet Vanuatu-Californie	5.00	8.75	25.00	23.75	37.50

TAB. 6.5 : Nombre de points de Bézier (en %) sélectionnés pour les processeurs engagés dans la période stationnaire pour construire le profil V_S .

Bien que les ondes de surface soient peu sensibles aux contrastes d'impédance, la pdf sur les points de Bézier (figure 6.19a) montre deux changements de pente dans la zone de transition : l'un à 400 km (étiquette [1]) et l'autre plus étalé en profondeur, vers 550 km (étiquette [2]). Ce net changement de pente à 400 km est en bonne adéquation avec le changement de phase olivine $\rightarrow$ wadsleyite remarqué à cette même profondeur lors de l'inversion avec la méthode 1. Bien que plus diffuse, la modification du gradient de vitesse vers 550 km est à mettre en relation avec la verticalisation du profil de vitesse entre 500 et 600 km sur la figure 6.11b, et également visible sur la figure 6.19d, avant une augmentation de la valeur du gradient liée à la transformation minéralogique de ringwoodite en pérovskite et magnésio-wüstite. Les probabilités marginales à 400 km de profondeur en points et en courbes de Bézier sont représentées sur la figure 6.20. L'histogramme des points montre qu'il existe 2 maxima (flèches rouges). Au contraire, l'histogramme des courbes exhibe une distribution unimodale, dont le maximum (flèche verte) se situe entre les 2 maxima de la distribution des points de Bézier. Ce résultat signifie que les paramètres de l'inversion sont bien sensibles au saut de vitesse brutal qui a lieu à cette profondeur, mais la pdf des profils 1D lisse cette information.

La pdf sur les vitesses de phase pour les différents modes est représentée sur la figure 6.21. L'ensemble des vitesses de phase correspondant aux modèles échantillonnés se situent à l'intérieur des barres d'erreur, ce qui montre que l'inversion en V_S et ξ est tout aussi capable de rendre compte des données de Visser et al. (2008a) que l'inversion combinée de V_S , T et ξ . Si l'on compare les distributions de vitesses de phase pour les 2 méthodes (figures 6.9 et 6.21), on remarque certaines différences. Les distributions des vitesses de phases des modes fondamentaux des ondes de Rayleigh et de Love sont relativement similaires pour les résultats des 2 méthodes. En revanche, bien que la forme globale des pdfs des modes harmoniques présentent des similitudes et oscillent autour de zéro, dans le détail la position du maximum de la pdf pour une même période est différente entre les 2 méthodes.

Une observation intéressante est que les vitesses de phase issues des modèles médians en V_S et ξ (méthode 2, figure 6.21) suivent quasiment les maxima des pdfs des vitesses de phase. Ce n'est pas toujours le cas pour les vitesses de phase des modèles médians en V_S , T et ξ (méthode 1, figure 6.9), notamment pour le mode fondamental des ondes de Rayleigh ainsi que pour les modes 0 à 2 des ondes de Love. Le modèle médian (qu'il soit en V_S , T ou ξ) ne constitue pas un modèle particulier pris dans l'ensemble des

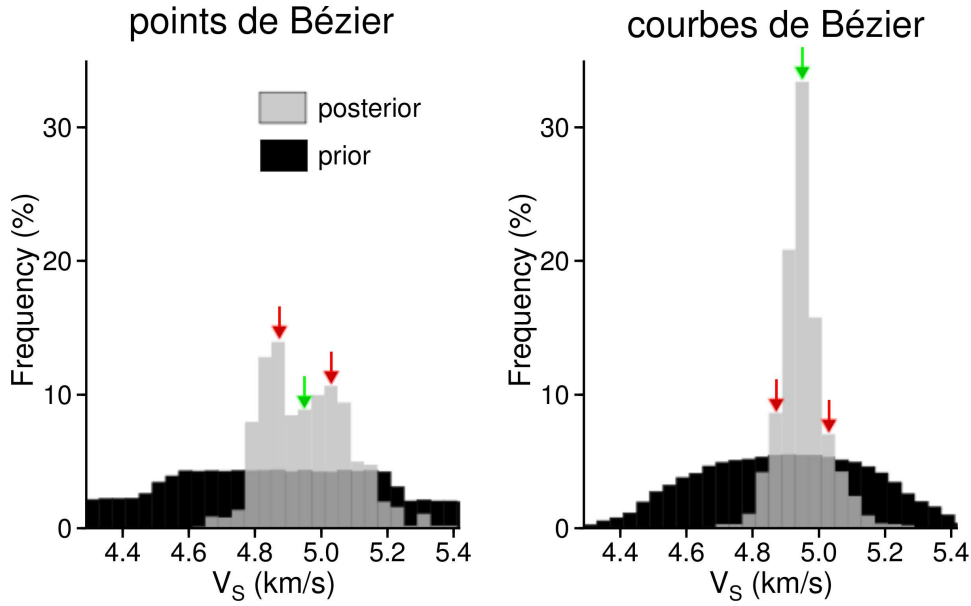


FIG. 6.20 : Probabilités marginales *a priori* (en noir) et *a posteriori* (en gris) sur V_S à 400 km de profondeur, pour les points et les courbes de Bézier. La taille des intervalles est de 0.04 km/s.

modèles. Pour la méthode 2, le profil médian en V_S est directement utilisé pour calculer les vitesses de phase (points blancs sur la figure 6.21). En revanche, pour la méthode 1, les courbes de dispersion du modèle médian sont estimées à partir du profil médian en V_S entre la surface et 350 km de profondeur, et à partir des valeurs de V_S entre 350 et 1000 km issues du modèle médian en T . Ainsi, la continuité du profil en V_S à 350 km de profondeur n'est pas forcément respectée. De plus, puisqu'il s'agit d'un problème non-linéaire, une structure sismique moyenne ne se traduit pas forcément en une structure thermique moyenne. Ces deux arguments peuvent expliquer le fait que les vitesses de phase du modèle médian de la méthode 1 ne soient pas aussi centrées sur les maxima des pdfs que pour la méthode 2.

6.5.2 De V_S vers la distribution en température

La seconde étape de la procédure consiste à déterminer quelle distribution de température est susceptible de rendre compte de la distribution en V_S obtenue précédemment, et de savoir si elle est semblable à celle retrouvée par l'inversion jointe de V_S , T et ξ . Pour ce faire, nous considérons que les données sont cette fois-ci non plus les vitesses de phase, mais le modèle médian de la distribution en V_S (figure 6.22). Les incertitudes sur le profil médian correspondent à l'écart-type à 1σ , représenté en gris foncé sur la figure 6.22.

Les figures 6.23a et b montrent les pdfs des profils de température obtenus par les méthodes 1 et 2, respectivement. Les distributions sont semblables uniquement entre 400

et 500 km de profondeur. La deuxième méthode entraîne une diminution significative de la température (~ 170 K) vers 680 km de profondeur. Le fait de réaliser une inversion en deux étapes produit un effet équivalent à celui obtenu lorsque la pente de Clapeyron de la transformation minéralogique de ringwoodite en pérovskite et magnésio-wüstite est fixée à 0 MPa/K (voir section 6.4). À 350 km, la température est d'environ 30 K inférieure à celle obtenue avec la première méthode.

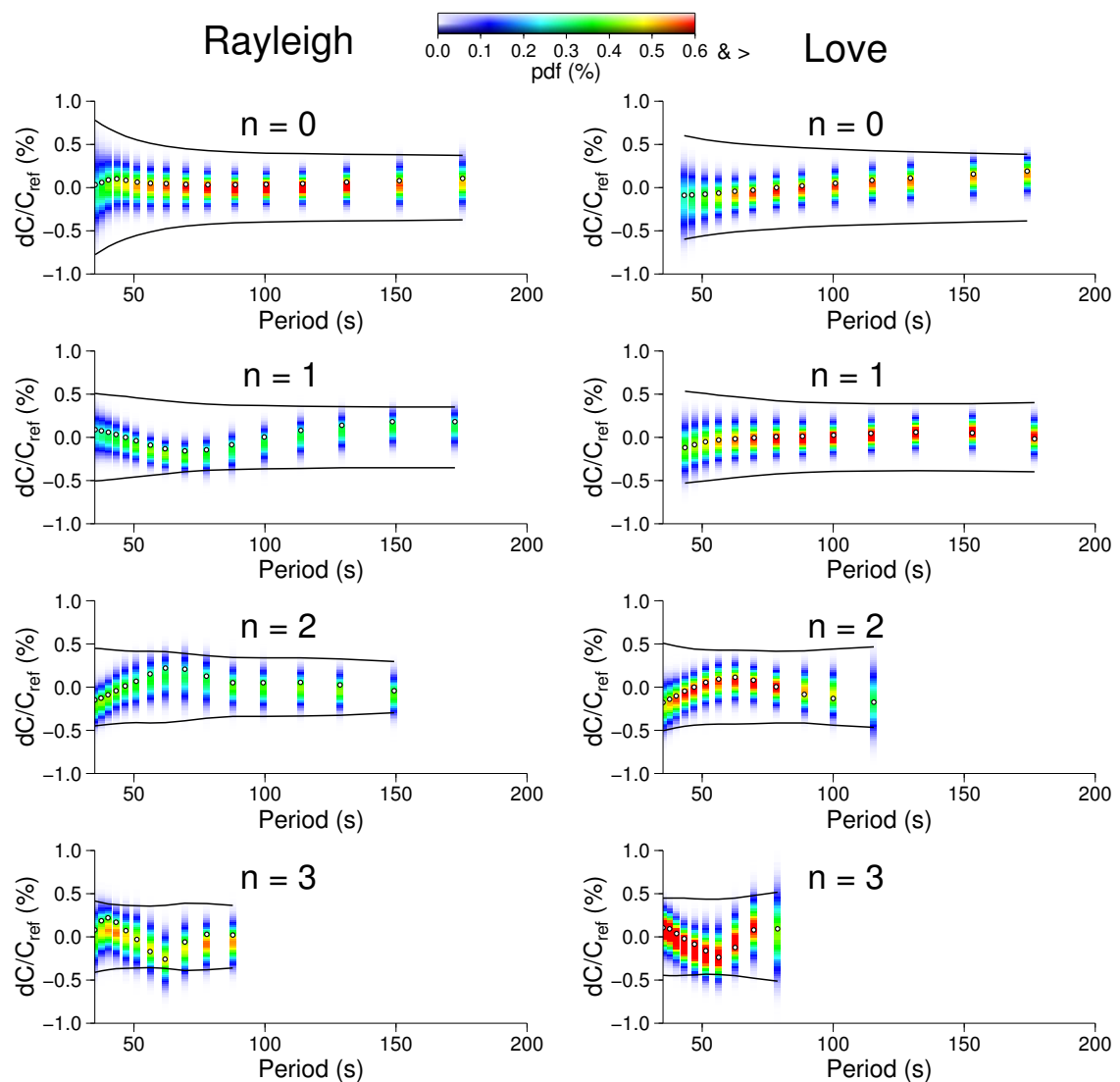


FIG. 6.21 : Distributions *a posteriori* des vitesses de phase du mode fondamental et des modes harmoniques des ondes de Rayleigh et de Love. Les résultats sont montrés en pourcentage de la vitesse de phase testée (C) comparée à la vitesse de phase réelle (C_{ref}). Les courbes noires correspondent aux incertitudes sur les vitesses de phase de [Visser et al. \(2008a\)](#). Les points blancs indiquent les vitesses de phase obtenues pour le modèle médian issu des distributions en V_S et ξ .

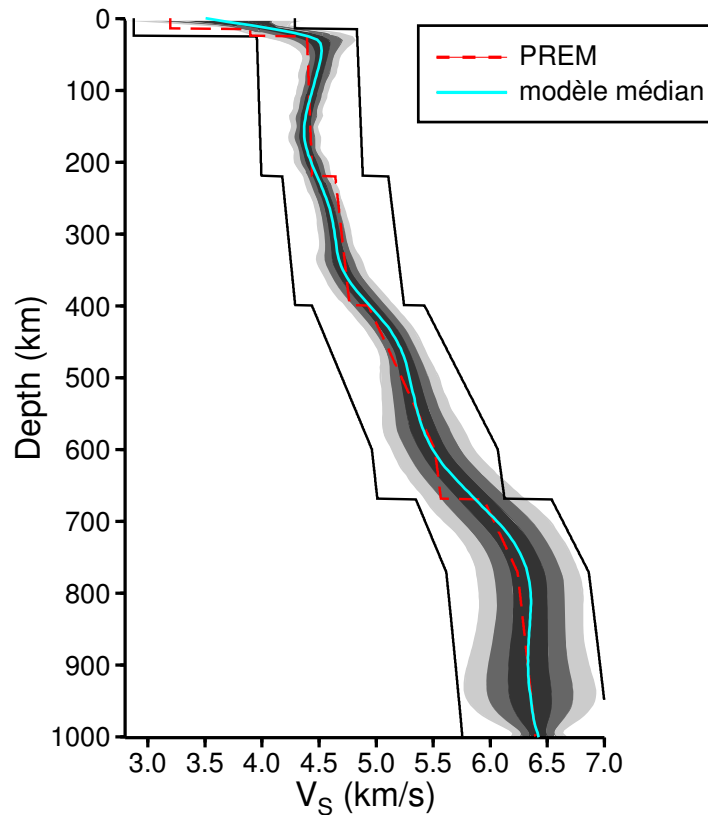


FIG. 6.22 : Profil médian de la distribution en V_S . Les surfaces grises (de la plus claire à la plus foncée) délimitent les intervalles de probabilité à 0.997, 0.95 et 0.68.

Ces différences s'expliquent en observant les figures 6.23c et d, qui représentent les pdfs des vitesses de cisaillement en fonction de la profondeur pour les deux méthodes. Les pdfs correspondent à l'écart en pourcents par rapport au profil médian présenté sur la figure 6.22. Les courbes noires indiquent les incertitudes à $\pm 1\sigma$ d'intervalle de confiance par rapport au modèle médian. Pour la distribution obtenue avec la méthode 1 (figure 6.23c), la pdf sort particulièrement des barres d'erreur au niveau des 2 principales discontinuités à 410 et 660 km de profondeur. Pour la méthode 2, l'intégralité de la distribution est confinée à l'intérieur des barres d'erreur, excepté au niveau de la transition de phase de l'olivine en wadsleyite à 410 km. Les propriétés élastiques de l'olivine et de la wadsleyite impliquent un saut de vitesse important (voir section 6.3.5) et il faudrait le diviser quasiment par 3 pour pouvoir le rendre compatible avec les barres d'erreur du modèle médian lissé. Cependant, la discontinuité liée à la formation de pérovskite peut parfaitement expliquer des données, compte-tenu des incertitudes. Si l'on regarde le résultat de la méthode 1 (figure 6.23c), seule une discontinuité située vers 680 km de profondeur peut être intégralement contenue dans l'intervalle à 1σ . Des discontinuités plus superficielles, correspondant à des températures plus élevées, sortent des barres d'erreur jusqu'à 1.5 %. C'est pourquoi la méthode 2 a retenu uniquement la discontinuité localisée à 680 km,

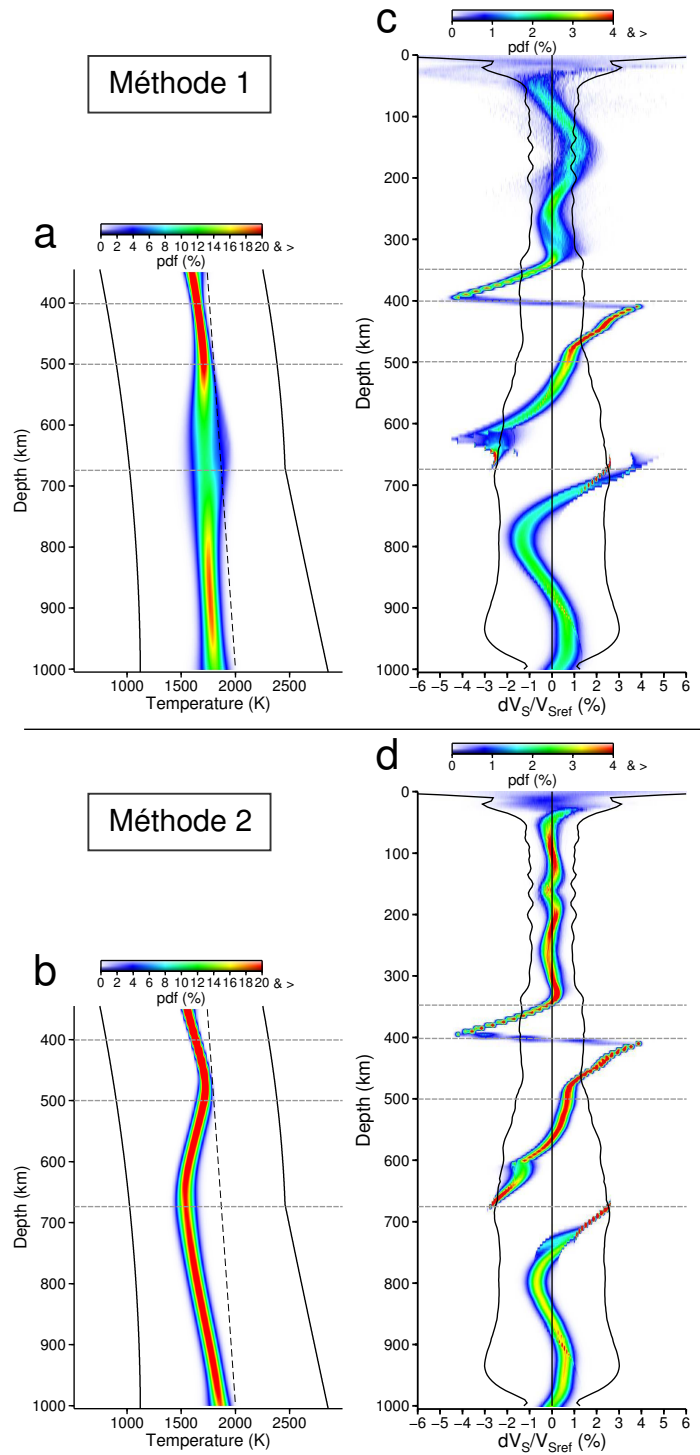


FIG. 6.23 : Pdfs en T et V_S issues des méthodes 1 et 2. Les résultats en V_S sont montrés en pourcentage de la vitesse testée (V_S) comparée à la vitesse du modèle médian (V_{Sref}) représenté sur la figure 6.22. Les lignes en pointillés sur les figures (a) et (b) correspondent à l'adiabat à 1600 K avec un gradient de 0.4 K/km. Les courbes noires sur les figures (c) et (d) représentent l'écart-type à 1σ du modèle médian.

qui est liée à des températures plus faibles que celles sélectionnées par la méthode 1. La conséquence de l'utilisation de cette méthode est que la transformation ringwoodite $\rightarrow$ pérovskite + magnésio-wüstite a lieu de façon instantanée à 680 km, ce qui équivaut à fixer une pente de Clapeyron de valeur nulle. À 350 km de profondeur, la pdf sur V_S pour la méthode 2 coïncide parfaitement avec la vitesse du profil médian ($dV_S/dV_{Sref} = 0$). En revanche, la vitesse à cette même profondeur pour la méthode 1 est légèrement plus faible que celle du modèle de référence, à cause d'une température plus élevée.

Avec cet exemple, nous montrons que les deux méthodes employées généralement pour obtenir des distributions de température dans le manteau produisent des résultats différents en terme de température, en se basant pourtant sur les mêmes données de départ, c'est-à-dire les courbes de dispersion des ondes de surface. Dans les deux cas, une signification physique au modèle sismologique a été imposée, puisque des informations issues de la physique des minéraux ont été employées. Cette disparité entre les distributions de température peut avoir d'importantes conséquences sur l'interprétation géodynamique de ces résultats.

Passer par le moins d'intermédiaires possible apparaît être une condition importante lors de l'inversion. En effet, la méthode 2 suppose une étape supplémentaire durant laquelle la température est déterminée à partir des vitesses sismiques. Pour ce faire, il faut choisir un modèle sismique 1D représentatif de la distribution en V_S . Ici le modèle médian a été choisi, et c'est à partir de ce modèle et de l'incertitude à 1σ sur la distribution en V_S que la distribution de température a été établie. Avec cette méthode, nous avons l'illusion que la profondeur de la transformation minéralogique de la ringwoodite en pérovskite et magnésio-wüstite est extrêmement bien contrainte (~ 680 km). Ce résultat est en contradiction avec le fait que les ondes de surface ne peuvent pas détecter les discontinuités avec une telle précision. En effet, nous avons vu sur la figure 6.19 qu'une inversion en V_S des vitesses de phase engendrait un lissage des discontinuités. Cette profondeur est complètement dépendante du profil V_S qui fait office de données, et qui conditionne la profondeur à laquelle a lieu la transformation de la pérovskite. De plus, le choix d'employer le modèle médian est arbitraire car le mode ou bien la moyenne de la distribution en V_S auraient pu également être utilisés, et la profondeur de la transformation aurait alors été différente. Le fait de réaliser un schéma complet d'inversion, comme celui développé dans cette thèse (méthode 1), permet de reporter directement les incertitudes sur les vitesses de phase sur la distribution en température et de visualiser l'ensemble des solutions possibles. Cette inversion fournit des résultats plus stables et plus complets.

6.6 Ce qu'il faut retenir

En considérant une composition pyrolitique, les résultats majeurs de notre méthode d'inversion appliquée aux données de [Visser et al. \(2008a\)](#) sous l'Océan Pacifique sont :

- une zone à faible vitesse au-dessus de la discontinuité de Lehmann,
- une importante anisotropie radiale positive ($V_{SH} > V_{SV}$) jusqu'à 350 km de profondeur,
- une zone de transition isotrope,
- une distribution de température quasi-linéaire dans la zone de transition, caractérisée par un gradient adiabatique de 0.22 K/km avec une température potentielle de 1600 K,
- un élargissement de la distribution de température entre 600 et 700 km de profondeur, induite par le prior sur la transformation minéralogique de la ringwoodite en pérovskite et magnésioiwüstite.

Une pente de Clapeyron nulle au niveau de la "660" favorise une diminution de la température à partir de 550 km de profondeur, ce qui illustre que les choix minéralogiques constituent des contraintes sur les résultats de l'inversion. Une procédure d'inversion en deux étapes, en déterminant d'abord une distribution en V_S et en ξ à partir des vitesses de phase, puis en estimant T à partir de V_S , montre également une diminution de la température. Ce résultat est dépendant du modèle choisi comme représentatif de la distribution en V_S (ici le modèle médian). En revanche, la méthode d'inversion développée au cours de cette thèse permet de caractériser directement les vitesses de phase en terme de température, à travers des lois thermodynamiques et des diagrammes de phases, pour chaque modèle échantillonné. Les résultats issus cette méthode d'inversion ne présentent aucun signe de corrélation. En particulier, l'anisotropie radiale peut être retrouvée indépendamment de la température, et n'est pas affectée par les contraintes *a priori* sur la minéralogie, quelle que soit la méthode employée dans ce chapitre.

Chapitre 7

Conclusions et perspectives

Dans cette thèse, j'ai présenté une méthode d'inversion des courbes de dispersion des ondes de surface en termes de température et d'anisotropie radiale dans la zone de transition. L'objectif de ce chapitre est d'aborder la méthode développée d'une manière plus globale. Après avoir rappelé les principaux résultats de la thèse, des applications potentielles du travail réalisé sont proposées, ainsi que les améliorations éventuelles qui pourraient être apportées à l'algorithme.

7.1 Synthèse

L'objectif de ce travail était double. D'une part, le but était de proposer une interprétation des données sismologiques en termes de température et d'anisotropie radiale, en introduisant les résultats théoriques et expérimentaux concernant la physique des minéraux. D'autre part, le problème inverse étant fortement non-linéaire, il fallait se donner les moyens d'obtenir une vision d'ensemble de la solution dans l'espace des paramètres, sans être dépendant d'un modèle de départ.

Le problème direct, présenté dans le chapitre 2, comprend 3 étapes. La première étape consiste à calculer les fractions volumiques des différents minéraux du manteau, étant données les conditions de pression et de température, pour un modèle minéralogique particulier. Dans la deuxième étape, par l'intermédiaire des diagrammes de phases et de l'équation d'état de Birch-Murnaghan au troisième ordre, les propriétés élastiques de chaque minéral sont déterminées, puis les vitesses sismiques et la masse volumique de l'assemblage minéralogique. La troisième étape, dans laquelle l'anisotropie radiale est introduite, permet le calcul des courbes de dispersion à l'aide du programme MINEOS.

Cette procédure a été utilisée pour inverser les courbes de dispersion du mode fondamental et des modes harmoniques (jusqu'à l'ordre 3) des ondes de Love et de Rayleigh.

La méthode d'inversion repose sur un algorithme de Monte Carlo par chaînes de Markov, détaillé dans le chapitre 3, qui réalise une marche aléatoire dans l'espace des modèles. L'information *a priori* est à la fois combinée avec les informations apportées par les données et avec les relations théoriques qui relient les données aux paramètres du modèle. Les modèles sont générés selon les informations *a priori* et la fonction de vraisemblance.

La philosophie générale de cette thèse est d'imposer le moins de contraintes possibles lors de l'échantillonnage des modèles afin de déterminer quels modèles sont les plus probables, connaissant les informations *a priori* et les données. Afin de respecter ce principe et de réaliser peu de choix arbitraires, l'algorithme a été optimisé et parallélisé (chapitre 4). Les modèles 1D en V_S , T et ξ sont paramétrés avec des courbes de Bézier, dont les points de contrôle sont les paramètres de l'inversion. L'originalité de cette approche est que les profondeurs des points sont considérées comme étant des inconnues.

L'algorithme a été testé avec des vitesses de phase synthétiques (chapitre 5) et appliqué à des données réelles (chapitre 6). Des exemples de la façon de représenter et d'analyser l'ensemble des modèles échantillonnés ont été présentés et discutés. Plusieurs représentations sont nécessaires, en plus des cartes de densité de probabilité *a posteriori*, afin d'interpréter les résultats de façon complète. Les résultats montrent que les vitesses de phase sont sensibles à la température et à l'anisotropie radiale dans la zone de transition. Les résultats sur les données réelles révèlent que la distribution de température sous l'Océan Pacifique est quasi-linéaire. Cependant, son interprétation est relativement complexe entre 550 et 700 km de profondeur, à cause de la transformation minéralogique de la ringwoodite en pérovskite et magnésio-wüstite, ce qui signifie que la structure thermique dans cette région ne peut pas être réduite à un profil 1D unique. Une importante structure anisotrope est trouvée dans le manteau supérieur sous le Pacifique, en bon accord avec les études précédentes. À plus grande profondeur, cette étude a montré qu'en considérant peu de contraintes sur l'anisotropie, et étant données les incertitudes sur les vitesses de phase considérées, la zone de transition semble être isotrope. Des incertitudes plus fines sur les vitesses de phase pourraient permettre de déterminer si un changement de signe dans l'anisotropie a réellement lieu au-dessous de 400 km de profondeur. D'autres travaux sont nécessaires pour déterminer si ces caractéristiques sont robustes. En particulier, une comparaison avec un autre jeu de données ainsi que l'étude d'un trajet perpendiculaire au trajet Vanuatu-Californie pourraient fournir des éléments de réponse.

Nous avons montré que la distribution de température dans la zone de transition est particulièrement sensible au modèle minéralogique employé. Réaliser deux inversions distinctes, c'est-à-dire d'abord obtenir la distribution en V_S à partir des vitesses de phase, puis déterminer la distribution en température à partir d'un modèle choisi arbitrairement comme étant représentatif de la distribution en V_S , donne des résultats en température significativement différents par rapport à ceux obtenus avec notre méthode. Ces différents

résultats peuvent induire d'importantes conséquences au niveau de l'interprétation géodynamique. De mon point de vue, passer par le moins d'intermédiaires possibles est préférable, car les incertitudes sur les données sont directement reportées sur les paramètres recherchés.

7.2 Avantages et critiques de la méthode

Dans cette thèse, j'ai présenté une méthode combinant une loi thermodynamique avec un schéma d'inversion non-linéaire. Celle-ci permet de déterminer les distributions de température et d'anisotropie radiale dans la zone de transition, ainsi que la vitesse de cisaillement entre la surface et 350 km de profondeur. Cette méthode est différente des méthodes tomographiques standards, car les variations des vitesses sismiques et de masse volumique dans la zone de transition sont liées aux propriétés physiques des minéraux (les modules élastiques et leurs dérivées par rapport à la pression et à la température). Dans cette approche, la taille et la localisation des discontinuités des propriétés physiques sont modélisées d'une manière physiquement réaliste, puisque leurs variations dépendent de la composition et des conditions de pression et de température du modèle particulier considéré. L'utilisation de points de Bézier comme paramètres des modèles, dont la profondeur n'est pas fixée *a priori*, constitue une paramétrisation adaptative qui permet une grande souplesse dans la recherche de la solution. Les propriétés sismiques des couches superficielles sont prises en compte à travers une large gamme de valeurs, et ne dépendent pas d'un modèle de croûte particulier. Le fait de réaliser une inversion des courbes de dispersion en termes de température, en considérant V_P , V_S et ρ uniquement comme des paramètres intermédiaires (secondaires) est important, car cette technique permet de reporter directement les incertitudes des données sur les paramètres du modèle.

Le but final des inversions en géophysique est de produire un modèle interprétable, c'est pourquoi des méthodes d'optimisation sont souvent préférées par les géophysiciens, puisque la vraie Terre est évidemment « unique ». Cependant, dans le contexte bayésien, l'ensemble de la solution ne peut pas être interprétée directement. Il existe un grand nombre de manières de produire des modèles interprétables, et ce choix est arbitraire. Différents choix peuvent conduire à différents modèles de Terre, et ainsi à différentes interprétations. Dans cette thèse, nous avons étudié la zone de transition sous la forme de modèles 1D. Si des vitesses de phase régionalisées sont inversées, une juxtaposition de distributions de profils 1D est obtenue, qui peut éventuellement par la suite être utilisée pour faire de la tomographie 2D ou 3D. Représenter les résultats devient alors encore plus complexe : doit-on prendre la médiane de chaque distribution ? le mode ? la moyenne ? En vérité, dans la formulation bayésienne, la seule véritable solution du problème inverse est la distribution *a posteriori*, et tout modèle unique dérivé de cette solution doit être considéré comme étant seulement un type de projection de la solution *a posteriori*.

Au-delà de cette question conceptuelle, la principale critique qui peut être adressée à notre méthode, et de façon plus générale aux méthodes de la famille de Monte Carlo, est le temps de calcul. Il s'agit du prix à payer si on souhaite mettre peu de contraintes sur le résultat de l'inversion. Si le modèle est défini par trop de paramètres, le nombre d'itérations nécessaire pour échantillonner la distribution *a posteriori* peut rapidement devenir gigantesque. Et puisque les données prédites doivent être calculées à chaque fois qu'un nouveau modèle est sélectionné, le temps de calcul de l'algorithme peut devenir prohibitif.

7.3 Applications potentielles de ce travail

Cette section est divisée en deux parties. Premièrement, nous allons exposer des exemples d'applications qui peuvent être réalisées en utilisant directement l'algorithme en l'état. La méthode présentée dans cette thèse n'est pas strictement restreinte à l'étude des courbes de dispersion. Puisque la procédure est basée sur une recherche directe dans l'espace des paramètres, le problème direct correspond à plusieurs sous-programmes indépendants du programme principal. Dans la deuxième partie, nous allons voir que le problème direct peut être adapté à d'autres données sismologiques, et plus généralement géophysiques.

7.3.1 En utilisant l'algorithme en l'état

Apports sur la minéralogie

Bien que la structure physique de la zone de transition soit principalement contrôlée par la température, les contributions relatives de la température et de la minéralogie ne sont pas encore très bien comprises. Notamment, nous avons vu que la valeur de la pente de Clapeyron de la transformation minéralogique de la ringwoodite en pérovskite et magnésio-wüstite a un impact important sur la distribution en température. Une première approche pourrait être d'inverser de façon probabiliste ce paramètre, entre deux valeurs extrêmes rencontrées dans la littérature, c'est-à-dire -0.4 MPa/K (Katsura et al., 2003) et -3.0 MPa/K (Akaogi & Ito, 1993). Ce test permettrait de déterminer si les ondes de surface privilégient certaines valeurs de la pente de Clapeyron, ou bien si toutes les valeurs sont capables d'expliquer les données, compte-tenu des incertitudes associées.

Nous avons testé avec des données synthétiques une inversion jointe de la température et de la minéralogie, mais les distributions *a posteriori* des fractions volumiques des minéraux du manteau sont toujours très similaires aux distributions *a priori*, ce qui signifie qu'aucune information sur la minéralogie n'a pu être extraite des données. Il est difficile de décorréler les contributions relatives des différents minéraux. La figure 7.1 montre les

profils en V_P , V_S et ρ , des principaux minéraux constitutifs de la zone de transition. On observe que les propriétés sismiques de l'olivine, du clinopyroxène et de l'orthopyroxène sont très proches, notamment pour V_S , ce qui explique pourquoi il n'est pas possible de discriminer les proportions des différents minéraux en utilisant les informations véhiculées par les ondes de surface uniquement.

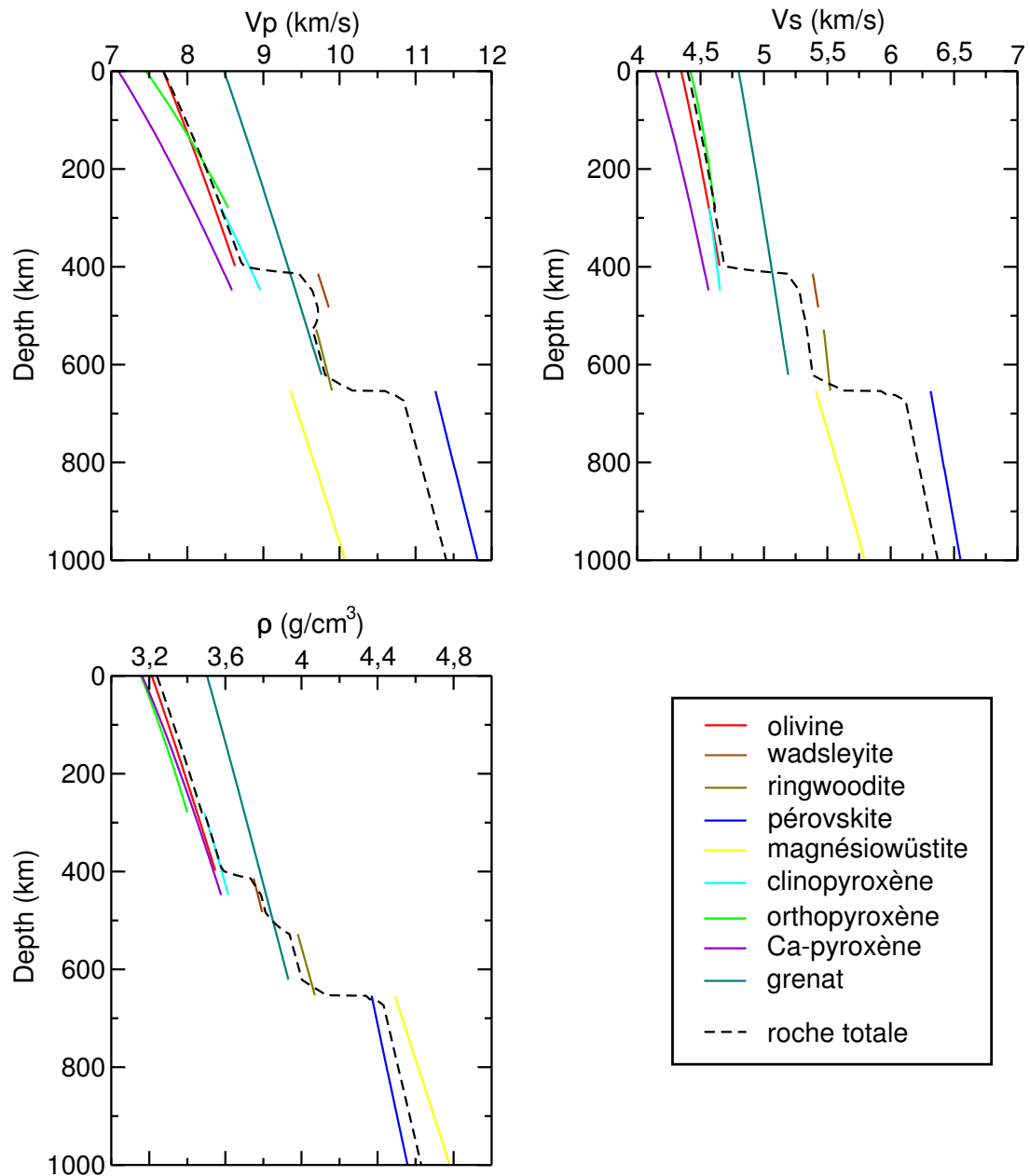


FIG. 7.1 : Vitesses sismiques et masse volumique des principaux minéraux du manteau, calculées le long d'un adiabat à 0.4 K/km avec une température potentielle de 1600 K. Les propriétés de la roche totale sont représentées en pointillés.

Pour le moment, l'algorithme ne peut être employé qu'avec une composition minéralogique fixée. Cependant, une sélection de modèles minéralogiques pourrait être testée sur les données du trajet Vanuatu-Californie, afin de quantifier l'impact de la minéralogie sur la distribution en température.

Qu'est-ce qu'un modèle moyen ?

L'algorithme développé peut également être utilisé dans le cas de vitesses de phase régionalisées. Les résultats obtenus pour des vitesses de phase intégrées pourraient être comparées aux distributions produites lors de l'inversion de vitesses de phase régionalisées, pour des points localisés le long du trajet Vanuatu-Californie. Cette étude permettrait d'identifier si les distributions pour les vitesses de phase intégrées correspondent à une « moyenne » de l'ensemble des distributions obtenues par les vitesses de phase régionalisées, ou au contraire, si elles correspondent à l'ensemble des différentes contributions des distributions estimées à partir des vitesses de phase intégrées.

Une autre application pour approfondir cette question serait de comparer un modèle moyen (moyenne le long d'un trajet), déterminé en prenant en compte les effets de propagation d'ondes 3D, avec par exemple la méthode des éléments spectraux (Capdeville et al., 2003), au résultat de l'inversion probabiliste.

Études locales

Une étude régionale plus complète pourrait être réalisée au niveau de l'Océan Pacifique, en particulier près de la ride océanique Est-Pacifique, où Gu et al. (2005), Panning & Romanowicz (2006) et Kustowski et al. (2008) ont détecté une anisotropie négative ($V_{SH} < V_{SV}$) juste au-dessus de la zone de transition.

Des sites particuliers, où une importante anisotropie radiale est suspectée, pourraient être examinés. Par exemple, Zhou et al. (2006) ont également décelé une anisotropie négative sous la ride médio-Atlantique et sous la Mer Rouge. Cette anomalie s'étendrait jusqu'à au moins 400 km de profondeur.

Des régions continentales seraient également intéressantes à étudier, comme l'Amérique du Nord, où Marone et al. (2007) et Nettles & Dziewonski (2008) ont trouvé d'importantes variations latérales de l'anisotropie. De plus, les distributions de température au niveau des cratons pourraient être comparées à celles déterminées sous les océans.

7.3.2 En modifiant l'algorithme

Applications aux formes d'ondes

La méthode de mesure des courbes de dispersion développée par Visser et al. (2007, 2008a) repose sur l'échantillonnage de profils V_S selon une approche par l'espace

des modèles, mais elle présente une très forte dépendance au modèle de référence, puisqu'à chaque itération (nouveau modèle) les mêmes dérivées de Fréchet sont employées. Il existe également une dépendance au modèle PREM, qui a été utilisé pour borner l'espace des modèles lors de la recherche effectuée par l'algorithme de voisinage.

L'emploi de sismogrammes synthétiques représente une approche plus complète pour obtenir de l'information sur la Terre profonde, puisqu'ils peuvent être directement comparés à leurs équivalents observés. La continuité directe du travail développé dans cette thèse serait d'adapter une partie du problème direct pour inverser directement la forme d'onde du sismogramme, ce qui permettrait de ne plus être tributaire de la manière dont les vitesses de phase ont été calculées. Une partie de ce travail a d'ores et déjà été effectuée. Les sismogrammes synthétiques sont calculés par le programme GEMINI (Friederich & Dalkolmo, 1995). Ce dernier est construit dans le domaine fréquentiel, ce qui présente deux avantages : il n'est pas nécessaire de définir une borne maximale en ordre radial, et les calculs pour chaque fréquence étant réalisés de façon indépendante, ils peuvent donc être effectués en parallèle.

Apport de la conductivité électrique

Un problème majeur rencontré dans l'inversion des données sismologiques est le conflit entre les effets liés à la température et ceux liés à la composition. L'interprétation des discontinuités sismiques en termes de transformations minéralogiques n'est pas unique, car les propriétés sismiques des principaux constituants du manteau (olivine et pyroxènes) sont similaires (voire figure 7.1), et parce que ces minéraux subissent des transformations de phase dans des gammes similaires de pression et de température. La seule manière de les discriminer, du moins par des moyens géophysiques, est de réaliser une inversion jointe entre les données électromagnétiques et sismologiques. En effet, ces deux types de données sont affectées de manière différente par la composition et la température. Il est alors possible de distinguer deux roches qui présentent un comportement mécanique identique (et qui ne peuvent donc pas être distinguées par les données sismologiques uniquement), mais un comportement électrique différent.

Les valeurs de conductivité électrique des minéraux constitutifs du manteau à hautes pressions et hautes températures sont disponibles (voir Xu et al., 2000 pour une synthèse). Cependant, encore peu d'études ont réalisé une inversion jointe de ces deux types de données (Verhoeven et al., 2009). De plus, l'interprétation des données électromagnétiques dans la zone de transition doit être effectuée avec prudence car la présence d'eau peut avoir un effet important sur les données.

7.4 Améliorations potentielles de l'algorithme

Les algorithmes transdimensionnels bayésiens (Green, 1995; Bodin & Sambridge, 2009; Bodin et al., 2012) constituent une classe d'algorithmes qui traitent le nombre de paramètres du modèle, comme étant une inconnue du problème, par exemple dans le cas de cette thèse le nombre de points de Bézier. Par définition, le nombre de paramètres employés affecte la solution du problème. Avec de tels algorithmes, le principe de parcimonie est appliqué : entre un modèle simple avec moins d'inconnues et un modèle plus complexe qui produisent des misfts similaires, l'échantillonnage bayésien favorisera le modèle le plus simple.

En laissant une liberté relative au nombre de paramètres constitutifs d'un profil de température et en introduisant les positions des points de Bézier en profondeur comme étant des paramètres de l'inversion, nous avons essayé de limiter les choix arbitraires. En effet, nous avons vu avec les tests synthétiques que les points venaient naturellement se placer le long du profil recherché, et que la solution ne dépendait pas du nombre de points employés. Cependant, le choix d'utiliser tel ou tel nombre de points dans le processus d'inversion est purement arbitraire, et nous n'avons pas fait varier le nombre de points pour l'anisotropie radiale et la vitesse de cisaillement dans le manteau supérieur et dans la croûte. À l'avenir, l'adaptation d'un tel algorithme permettrait de ne pas fixer *a priori* le nombre de paramètres à la fois pour V_S , T et ξ .

Annexes

Annexe A

Notations

symbole	signification	unité
A	lois physiques reliant les paramètres aux données	
A_1, A_2, A_3, A_4	les différentes théories physiques employées dans le problème direct	
a_0	coefficient d'expansion thermique	K^{-1}
B_j	courbe de Bézier cubique de classe C^1	
b_0	coefficient d'expansion thermique	K^{-2}
$b_{i,k}$	polynôme de Bernstein de degré k	
C	vitesse de phase	m/s
C_L	vitesse de phase des ondes de Love	m/s
C_R	vitesse de phase des ondes de Rayleigh	m/s
$\mathbf{C}_d$	opérateur de covariance sur les données	
C_k	fonction d'autocovariance	
$\mathbf{C}_p$	opérateur de covariance sur les paramètres	
C_p	capacité calorifique	J/kg/K
C_{ijkl} ou C_{pq}	tenseur élastique	
c	composition minéralogique	
c_0	coefficient d'expansion thermique	K
$\mathbf{d}$	vecteur des données	
d_z	distance minimale entre 2 points de Bézier consécutifs	m
d_0	coefficient d'expansion thermique	
$Fe\#$	teneur en fer (<i>iron number</i>)	
G	module de cisaillement	Pa
G_0	module de cisaillement aux conditions STP	Pa
$G_{0,Fe}$	dépendance du module de cisaillement par rapport à la teneur en fer aux conditions STP	Pa

$\mathcal{G}$	anisotropie azimutale	
g	densité de probabilité gaussienne	
K_S	module d'incompressibilité isentropique	Pa
K_{S_0}	module d'incompressibilité aux conditions STP	Pa
$K_{S_0,Fe}$	dépendance du module d'incompressibilité par rapport à la teneur en fer aux conditions STP	Pa
k	nombre d'onde	rad/m
L_1	norme L_1	
L_2	norme L_2	
$L(\mathbf{d} \mathbf{p})$	vraisemblance (<i>likelihood</i>)	
$\mathcal{L}$	chaleur latente d'une réaction	J/kg
l	ordre angulaire	
m	ordre azimutal	
n	ordre radial	
n_T	nombre de points de Bézier du vecteur $\mathbf{T}$	
n_{V_S}	nombre de points de Bézier du vecteur $\mathbf{V}_S$	
n_ξ	nombre de points de Bézier du vecteur ξ	
P	pression	Pa
P_0	pression aux conditions STP	Pa
$P(\mathbf{p})$	probabilité <i>a priori</i> sur les données	
$P(\mathbf{p} \mathbf{d})$	probabilité <i>a posteriori</i>	
$\mathcal{P}_{ji}$	point de contrôle numéro i de la courbe de Bézier j	
$\mathbf{p}$	vecteur des paramètres	
Q^{-1}	modèle d'atténuation	
r_k	fonction d'autocorrélation	
S	entropie	J/K
$\mathcal{S}$	fonction coût (<i>misfit</i>)	
T	température	K
$\mathbf{T}$	vecteur des points de Bézier en température	K
T_0	température aux conditions STP	K
U	fonction propre associée au déplacement radial des modes sphéroïdaux	
V	volume	m ³
V_P	vitesses des ondes P	m/s
V_{PH}	vitesses des ondes P polarisées horizontalement	m/s
V_{PV}	vitesses des ondes P polarisées verticalement	m/s
V_S	vitesses des ondes S	m/s
$\mathbf{V}_S$	vecteur des points de Bézier des vitesses des ondes S	m/s

V_{SH}	vitesses des ondes S polarisées horizontalement	m/s
V_{SV}	vitesses des ondes S polarisées verticalement	m/s
ϑ	température obtenue à P_0 pour un minéral donné après un chauffage isobare	K
W	fonction propre associée au déplacement horizontal des modes toroïdaux	
X	fraction volumique d'un minéral	
$\mathbf{x}$	vecteur composé de V_S , T et ξ	
$\mathbf{z}$	vecteur composé de $\mathbf{z}_{V_S}$, $\mathbf{z}_T$ et $\mathbf{z}_\xi$	
$\mathbf{z}_T$	vecteur des profondeurs correspondant aux points de Bézier T	km
$\mathbf{z}_{V_S}$	vecteur des profondeurs correspondant aux points de Bézier V_S	km
$\mathbf{z}_\xi$	vecteur des profondeurs correspondant aux points de Bézier ξ	km
α	expansivité thermique	K^{-1}
Δ	taille du vecteur tangent local	km
δ_{S_0}	paramètre d'Anderson-Grüneisen	
ϵ_{ij}	tenseur des déformations	Pa
η	paramètre anisotrope décrivant la dépendance de la propagation d'une onde selon l'angle d'incidence	
Γ	pente de Clapeyron	MPa/K
γ_{th}	paramètre de Grüneisen	
γ_{th0}	paramètre de Grüneisen aux conditions STP	
ξ	anisotropie radiale ou de polarisation des ondes S	
ξ	vecteur des points de Bézier en anisotropie radiale	
ρ	masse volumique	kg/m^3
ρ_0	masse volumique aux conditions STP	kg/m^3
$\rho_{0,Fe}$	dépendance de la masse volumique par rapport à la teneur en fer aux conditions STP	kg/m^3
σ	écart-type de la fonction gaussienne g	
σ_{ij}	tenseur des contraintes	Pa
ϕ	anisotropie radiale des ondes P	
Ω_1	domaine associé à la croûte et au manteau supérieur (0- 350 km)	
Ω_2	domaine associé à la zone de transition (350- 1000 km)	
ω	pulsation ou fréquence angulaire	rad/s

Annexe B

A Bayesian approach to infer radial models of temperature and anisotropy in the transition zone from surface wave dispersion curves

M. Drilleau¹, É. Beucler¹, A. Mocquet¹, O. Verhoeven¹, G. Moebs², G. Burgos³, J.-P. Montagner³ and P. Vacher¹

¹ Université de Nantes, Laboratoire de Planétologie et de Géodynamique, UMR 6112, Nantes, France.

² Université de Nantes, Laboratoire de Mathématiques Jean Leray, UMR 6629, Nantes, France.

³ Institut de Physique du Globe de Paris, Jussieu, Laboratoire de Sismologie, Paris, France.

Submitted to *Geophys. J. Int.* : april 2013

Accepted with moderate revisions : may 2013

The following version corresponds to the submitted one. No corrections have been applied.

Summary

Mineralogical transformations and material transfers within the Earth's mantle make the 350- 1000 km depth range (namely here the transition zone) highly heterogeneous and anisotropic. Most of the 3-D global tomographic models are anchored on small perturbations from 1-D models such as PREM, and are secondly interpreted in terms of temperature and composition distributions. However, the degree of heterogeneity in the transition zone can be strong enough that the concept of a 1-D reference seismic model might be addressed. To avoid the use of any seismic reference model, we present in this article a Markov chain Monte Carlo algorithm to directly interpret surface wave dispersion curves in terms of temperature and radial anisotropy distributions. These interpretations are based on laboratory measurements of elastic moduli and Birch-Murnaghan equation of state. In order to reduce the computing time and to generate a self-adaptating parameter space exploration, C1-Bézier curves are used to interpolate randomly chosen values for depth, temperature and anisotropy. This parameterization is able to explore both smoothly varying models and first-order discontinuities. Thanks to a Bayesian exploration, the probability distributions on temperature and anisotropy are governed by uncertainties on the data set. The method is applied to both synthetic data and real dispersion curves. Though surface wave data are not sensitive to sharp seismic discontinuities, the interpretation of the temperature distribution is highly related to the modeling of mineralogical phase transformations. Surface wave measurements along the Vanuatu-California path suggest a strong anisotropy above 400 km depth which disappears below, and a monotonous temperature distribution between 350 and 1000 km depth.

Key words : non-linear inversion, surface waves, probability distributions, transition zone, phase velocity.

B.1 Introduction

Seismic tomography is one of the most powerful way to provide information on the internal structure and dynamics of the Earth's deep interior. Seismic images reveal features at both lateral and radial resolutions which are continuously being improved. Various approaches for solving the inverse problem of describing the Earth's deep structure using seismological datasets have been implemented (*e. g.* [Montagner & Tanimoto, 1990](#); [van der Hilst et al., 1997](#); [Fukao et al., 2001](#); [Romanowicz, 2003](#)). Although high resolution Earth's models are more and more numerous, a quantitative comparison is difficult due to the amount of various techniques providing different kind of uncertainties.

Recent three-dimensional seismic models ([Fichtner et al., 2009](#)) are derived from sophisticated inverse theories and use accurate numerical methods for seismic wave propagation. While they improve the lateral resolution of the smooth 3-D starting models ([Ritsema et al., 1999](#); [Méglin & Romanowicz, 2000](#)), they are intrinsically rooted to first-order perturbations from a laterally homogeneous reference models such as PREM ([Dziewonski & Anderson, 1981](#)) or AK135 ([Kennett et al., 1995](#)). However, from a geodynamical point of view, in regions such as the transition zone, the degree of heterogeneity induced by mineralogical transformations, convective motions, upwelling and downwelling materials, and anisotropy, can be strong enough that the concept of a one-dimensional reference seismic model might be addressed.

In addition, while several studies (*e. g.* [Boschi & Dziewonski, 2000](#); [Trampert & Woodhouse, 2003](#); [Panning & Romanowicz, 2006](#); [Beghein et al., 2008](#); [Kustowski et al., 2008](#); [Visser et al., 2008a,b](#)) indicate radial anisotropic variations throughout the mantle, they only agree at long wavelengths ([Kustowski et al., 2008](#); [Panning et al., 2010](#)), even when higher mode surface waves are used to investigate the transition zone depth range, hereafter defined between 350 and 1000 km. These large uncertainties may be explained by the different inverse and regularization methods, the choice of the parameterization, and the data set.

When going beyond the tomographic methods and addressing the question of the interpretation of the seismological information in terms of mantle temperature and composition, two approaches may be followed : either from previously derived seismic models ([Vacher et al., 1996, 1998](#); [Goes et al., 2000](#); [Trampert et al., 2004](#); [Cammarano & Romanowicz, 2007](#)) or directly from the observed travel times, surface wave velocities, or eigenfrequencies ([Cammarano et al., 2009](#); [Khan et al., 2009](#)). One of the main advantages of the latter approach is to directly connect the posterior uncertainties on the geodynamical parameters to the variances on observations. However, the non-linear nature of the problem requires to use time consuming Monte Carlo like inverse methods. This difficulty has previously been handled by testing a limited amount of selected models ([Cammarano](#)

et al., 2009), or by exploring tightly constrained parameter spaces (Khan et al., 2009, 2011; Khan & Shankland, 2012).

In this article, we choose to address these questions by inferring the radial temperature field and anisotropy distribution for a given mineralogy from surface wave data. This allows to relax the priors and to sample a large range of possible parameter values.

Bayesian approaches allow to go beyond the classical computation of the unique best fitting model by providing a quantitative probabilistic measure of the model resolution, uncertainties and non-unicity (*e. g.* Mosegaard & Tarantola, 1995). Though these methods such as the well-known Markov chain Monte Carlo (MCMC) are popular in Geophysics, their use in global seismological studies and interpretation of seismological models in terms of mantle temperature and composition is recent. This is mainly due to the large amount of inverted parameters and to the huge required computing time (*e. g.* Verhoeven et al., 2009; Khan et al., 2009, 2011; Hauser et al., 2011; Bodin et al., 2012; Mosca et al., 2012). In a Bayesian framework, the known prior information on the parameters is combined with the observed data to generate the *a posteriori* distribution of the model parameters. MCMC methods perform a non-linear guided search by sampling the parameter space according to the posterior probabilities. One important advantage of this method is the complete independence from the choice of the initial model.

In this study, we develop a non-linear inverse approach to directly interpret Love and Rayleigh surface waves dispersion curves, in terms of temperature and radial anisotropy distributions. Starting from random realizations of the temperature field and anisotropy distribution, the seismic velocity profiles are computed, using a Birch-Murnaghan equation of state. The synthetic dispersion curves are then obtained by normal-mode summation mode and compared to the data through a Markov chain Monte Carlo method.

The forward problem is described in section B.2, while the Bayesian inversion is presented in section B.3. In section B.4, we detail the practical and numerical implementation of this procedure. Synthetic test results are exhibited in section B.5 and we illustrate the efficiency and the resolving power of the method by inverting phase velocities along the Vanuatu-California path, in terms of temperature and anisotropy of the transition zone.

B.2 Forward problem

This section describes the parameterization and modeling hypothesis, used to compute synthetic data for given temperature and anisotropy distributions, and the corresponding radial seismic models. In order to generate a self-adapting model space, with respect to the resolution power of the data and to slightly reduce the computing time, polynomial Bézier curves are chosen for the parameterization. This choice enables us to remove the effects of a regular depth-discretization, while allowing to generate a huge set of radial models.

B.2.1 Physical parameters

B.2.1.1 Isotropic parameters

A standard procedure to compute the density and elastic *moduli*, at a given depth, relies in extrapolating their values from Standard Temperature and Pressure (STP) conditions, using thermodynamic laws (*e. g.* [Jackson, 1998](#); [Stixrude & Lithgow-Bertelloni, 2010](#)). Laboratory experiments provide these STP values, their derivatives with respect to the temperature and pressure, and the thermodynamic parameters from which they can be derived. The elastic behavior of mantle materials may be computed starting either from their chemical or mineralogical compositions.

Here, the depth-dependent values of the density and the elastic *moduli* are first computed for each mineral using the approach described in [Vacher et al. \(1998\)](#), and the laboratory experiment values compiled by [Cammarano et al. \(2003\)](#) and [Verhoeven et al. \(2005\)](#). This method relies on a Birch-Murnaghan EOS, up to third-order in strain, associated to a Grüneisen correction for the temperature. The elastic properties of the mineralogical assemblages are computed using the Hashin-Shtrikman estimate ([Hashin & Shtrikman, 1963](#); [Watt et al., 1976](#)).

While the temperature and pressure increase, the mantle minerals undergo mineralogical transformations, including phase transitions. The volume fractions of the mineral phases are computed using phase diagrams evaluated from laboratory measurements. Phase diagrams are taken from [Vacher et al. \(1998\)](#). Here, two independent subsystems are considered : the phase diagram of olivine and its high pressure phases, and a phase diagram for all other components. Subsystems are computed using a Gibbs free-energy minimization method based on experimentally-constrained phase boundaries ([Ita & Stixrude, 1992](#)).

The corresponding radial density and seismic velocity models are then used to compute surface wave velocity dispersion curves, using the MINEOS package from CIG¹, based on the pioneer work of [Gilbert & Dziewonski \(1975\)](#) updated by [Woodhouse \(1988\)](#) and rewritten by G. Masters.

One may note that the standard seismic parameters (V_P , V_S and ρ) are not directly inverted, but are used, as an intermediate stage, to link the temperature and the anisotropy (for a given mineralogy) to the surface wave data.

B.2.1.2 Anisotropy

The other parameters of our models is the depth-distribution of the radial anisotropy, which is constrained by the Rayleigh-Love discrepancy ([Anderson, 1961](#)). In a

¹Computational Infrastructure for Geodynamics, <http://www.geodynamics.org/cig/software/mineos>

transversally isotropic medium, the number of independent elements of the fourth-order elastic tensor reduces to the five Love coefficients (Love, 1927), $A = \rho V_{P_H}^2$, $C = \rho V_{P_V}^2$, $N = \rho V_{S_H}^2$, $L = \rho V_{S_V}^2$, and F ; where ρ is the density, V_{P_H} , V_{P_V} , V_{S_H} , V_{S_V} , are the velocities of the horizontally and vertically propagating P waves, and the horizontally and vertically polarized S waves, respectively. Since the behaviour of these anisotropic *moduli*, as a function of temperature and pressure, is still unknown for most mantle minerals, we follow the procedure of Babuska & Cara (1991), Panning & Romanowicz (2006) and Khan et al. (2009), and define an isotropic shear wave velocity V_S by the Voigt average

$$V_S^2 = \frac{2V_{S_V}^2 + V_{S_H}^2}{3} = \frac{2L + N}{3\rho}. \quad (\text{B.1})$$

The anisotropy parameters are

$$\xi = \frac{V_{S_H}^2}{V_{S_V}^2} = \frac{N}{L}, \quad (\text{B.2})$$

and

$$\eta = \frac{F}{A - 2L}. \quad (\text{B.3})$$

Since the Rayleigh-Love discrepancy is mostly sensitive to the shear wave anisotropy and mostly insensitive to P wave anisotropy and η (Babuska & Cara, 1991), we discard P wave anisotropy and set $\eta = 1$. Equations (B.1) and (B.2) are combined to compute V_{S_H} and V_{S_V} as a function of V_S and ξ ,

$$V_{S_H} = V_S \sqrt{\frac{3\xi}{2 + \xi}}, \quad (\text{B.4})$$

$$V_{S_V} = V_S \sqrt{\frac{3}{2 + \xi}}. \quad (\text{B.5})$$

B.2.2 Model parameterization

The temperature and anisotropy distributions are described, as a function of depth, using C^1 Bézier polynomials (Bézier, 1966, 1967), based on randomly chosen control (or anchor) points. This procedure offers several advantages : (1) it does not impose a regularly spaced discretization of the models nor *a priori* on layer thicknesses and location of the seismic discontinuities ; (2) it can be used to describe both smooth (*e. g.* temperature gradients) and sharp (*e. g.* thin thermal boundary layers) variations with a minimum number of parameters ; and (3) it may be optimized during the iterative processes of the MCMC algorithm by adapting the number of curves that are necessary to describe a given structure to the resolving power of the observations.

The parameterization of a temperature or anisotropy distribution is sketched in Fig. B.1. The overall distribution is described by a series of N elementary cubic polynomials $B(t)$ of the form

$$B_j(t) = \sum_{i=0}^3 b_{i,3}(t) \mathcal{P}_{ji}, \quad t \in [0, 1], j = [1 \dots N], \quad (\text{B.6})$$

where the $\mathcal{P}_{ji}$, $i = 0, \dots, 3$ are the control points and where

$$b_{i,k}(t) = \binom{k}{i} t^i (1-t)^{k-i}, \quad i = 0, \dots, k \quad (\text{B.7})$$

are the Bernstein basis polynomials of degree k . At the ending points $\mathcal{P}_{j0}$ and $\mathcal{P}_{j3}$, the curvature is defined by the norm of the local tangent vectors $\mathcal{P}_{j0}\mathcal{P}_{j1}$ and $\mathcal{P}_{j2}\mathcal{P}_{j3}$, respectively. The C^1 class requires that the upward and downward derivatives are identical at each point joining two consecutive polynomials.

B.3 McMC method

Our inverse problem consists in computing the radial distributions of temperature and anisotropy from surface wave dispersion curves. This section outlines the fundamentals of the Bayesian inversion, based on the Markov chain Monte Carlo method, detailed in Mosegaard & Tarantola (1995) and Mosegaard (1998), and successfully applied to constrain the Earth's mantle structure using geophysical data (*e. g.* Khan et al., 2006; Verhoeven et al., 2009).

Let us denote by $\mathbf{p}$ and $\mathbf{d}$ the parameters of our model and the seismological data, respectively. The data are linked to the parameters through the equation, $\mathbf{d} = A(\mathbf{p})$, where the non-analytic and non-linear operator A represents the forward problem discussed in section B.2. Explicitly, the parameters are the control points of the cubic Bézier curves, namely depth (z), temperature (T) and anisotropy (ξ) at these particular points. In the Bayesian framework, a set of parameters is randomly chosen at each iteration. The corresponding radial distributions of seismic velocities and densities are used to compute Rayleigh and Love phase velocity dispersion curves.

The solutions of the inverse problem are described by the posterior probabilities $P(\mathbf{p}|\mathbf{d})$ that the parameters are in a configuration $\mathbf{p}$ given the data are in a configuration $\mathbf{d}$. The parameter space is sampled according to $P(\mathbf{p}|\mathbf{d})$. Bayes' theorem links the prior

distribution $P(\mathbf{p})$ and the posterior distribution $P(\mathbf{p}|\mathbf{d})$,

$$P(\mathbf{p}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{p})P(\mathbf{p})}{\sum_{\mathbf{p} \in \mathcal{M}} P(\mathbf{d}|\mathbf{p})P(\mathbf{p})}, \quad (\text{B.8})$$

where $\mathcal{M}$ denotes all the configurations in the parameter space. It exists a set of rules named priors $P(\mathbf{p})$ which defines the prior distribution, *i. e.* the set of possible models which reduce the configuration space and represents our state of knowledge. For example we do not know the temperature profile in the transition zone, but we can set upper and lower bounds on its expected value. Only models that belong to the prior distribution are proposed and tested against the data.

The probability distribution $P(\mathbf{d}|\mathbf{p})$ is a function of the misfit $S(\mathbf{d}, A(\mathbf{p}))$, which measures the difference between the observed data $\mathbf{d}$ and the computed synthetic data $A(\mathbf{p})$,

$$P(\mathbf{d}|\mathbf{p}) \propto \exp(-S(\mathbf{d}, A(\mathbf{p}))). \quad (\text{B.9})$$

The misfit function takes into account the overall variance that we assign to the data. We assume a Gaussian noise which implies that the observations are mutually independent. The misfit function is evaluated with a L_2 norm :

$$S(\mathbf{d}, A(\mathbf{p})) = \sum_i \frac{(d_i - A(p)_i)^2}{2\sigma_i^2} \quad (\text{B.10})$$

where d_i and $A(p)_i$ represent the i^{th} component of the observed data vector $\mathbf{d}$ and of the synthetic data, respectively, and where σ_i is the standard deviation associated to the i^{th} observation. The data variance directly determines the form of the posterior probability distribution and hence the posterior samples generated from it. In equation (B.10), the sum runs over all considered surface wave modes and frequencies.

Bayes' formula provides an estimate of the marginal probability $P(p_i = x|\mathbf{d})$, which is obtained by summing all the probabilities of the $\mathbf{p}$ configurations where the i^{th} parameter takes the value x :

$$P(p_i = x|\mathbf{d}) = \frac{\sum_{\mathbf{p} \in \mathcal{M}_x^i} P(\mathbf{d}|\mathbf{p})P(\mathbf{p})}{\sum_{\mathbf{p} \in \mathcal{M}} P(\mathbf{d}|\mathbf{p})P(\mathbf{p})}, \quad (\text{B.11})$$

where $\mathcal{M}_x^i$ represents all the configurations whose i^{th} component is x . The evaluation of the marginal probability is impossible to calculate in practice, because the sum in the denominator runs over the huge number of configurations. This number is equal to

$\prod_{i=1}^m |\mathcal{M}_i|$, where $|\mathcal{M}_i|$ is the number of values that the parameter i can take. We therefore choose a Markov chain Monte Carlo method (McMC), which samples the parameter space through a random walk, according to the probability $P(\mathbf{p}|\mathbf{d})$. The method is derived from the Metropolis-Hastings Algorithm (*e. g.* [Metropolis et al., 1953](#); [Hastings, 1970](#)), which generates samples according to the unknown posterior distribution. This step is done using a randomized decision rule which accepts or rejects the proposed models according to their fit to the data and the prior. After an initial exploration, the misfit function guides the research to choose another set of parameters, and the whole procedure is iterated.

Because each new model is chosen to be in a specific neighbourhood of the previous model, each new step depends only on the previous step. Let us consider the Markov chain associated with the i^{th} parameter with value x at the iteration t . If the prior is considered to be uniform in a given interval, the marginal probability $P(p_i = x|\mathbf{d})$ is directly equal to

$$P(p_i = x|\mathbf{d}) = \frac{\sum_{\mathbf{p} \in \mathcal{M}_x} P(\mathbf{d}|\mathbf{p})}{\sum_{\mathbf{p} \in \mathcal{M}} P(\mathbf{d}|\mathbf{p})}. \quad (\text{B.12})$$

In this case, given a current model whose parameters are p_i^t and a proposed model p_i^{t+1} , a decision has to be made whether or not the new model is accepted or rejected.

- If $S(\mathbf{d}, P(p_i^{t+1})) \leq S(\mathbf{d}, P(p_i^t))$ the new model is accepted and added to the set of samples of the posterior distribution because it improves the fit to the data.
- If $S(\mathbf{d}, P(p_i^{t+1})) > S(\mathbf{d}, P(p_i^t))$ the new model is degrading the misfit function but is not necessarily rejected. The model is accepted when the ratio $P(p_i^{t+1}|\mathbf{d})/P(p_i^t|\mathbf{d})$, is larger than a random number taken from a uniform distribution between 0 and 1. Otherwise, the model is rejected.

One of the advantages of the Metropolis-Hastings algorithm is to prevent the random walk to be trapped into a local minimum. This feature can be of first importance when dealing with a non-linear problem. Another advantage is that only probability ratios are handled. Explicitly, the denominator of the equation (B.8) does not need to be computed.

If the amount of iterations can be large enough, the samples provide a good approximation of the posterior distribution for the model parameters, *i.e.* $P(\mathbf{p}|\mathbf{d})$. The posterior probability of the temperature and anisotropy at a given depth can be visualized by plotting the distribution of the selected values for the whole set of solutions.

B.4 Practical implementation

This section details the technical aspects of the developed algorithm and its implementation to obtain the invert synthetic and real data results shown in sections B.5 and B.6.

B.4.1 Boundary conditions

When studying the Earth's mantle, it is well known that Rayleigh and Love surface wave data are both sensitive to the crustal structure (*e. g.* [Montagner & Jobert, 1988](#); [Bozdag & Trampert, 2007](#)). Improper corrections may lead to the spurious mapping of crustal seismic signatures into the underlying mantle ([Boschi & Ekström, 2002](#); [Ferreira et al., 2010](#); [Panning et al., 2010](#)). Some studies ([Shapiro & Ritzwoller, 2002](#); [Visser et al., 2008b](#); [Khan et al., 2009](#)) attempt to circumvent this difficulty by allowing slight velocity and Moho depth variations with respect to an *a priori* crustal model.

Our region of interest is the transition zone (say between 350- and 1000 km depth), and an original strategy is followed for the crust and the uppermost mantle. Due to the heterogeneous mineralogical content of the 0- 350 km depth region and the inadequacy of EOS applied at lithospheric P - T conditions, only V_S and ξ are jointly inverted above 350 km, as many tomographic studies do. This means that the stage which consists in using temperature and Birch-Murnaghan EOS and anisotropy to secondly interpret seismic models, is not implemented in this depth range. Doing so, the reference model of the crust and the uppermost mantle only constrains the parameter range and its choice does not constrain the inverse scheme such as in the inversions based on perturbation theory. In order to assure a scale invariant parameterization ([Tarantola, 2005](#)), we adopt a $\log(V_S/V_{S_0})$ parameterization, between 0 and 350 km depth. The reference velocity, V_{S_0} , can take any arbitrary value. We use scaling relations based on the experimental study of [Isaak \(1992\)](#) to compute V_P and ρ profiles.

From 1500 km depth down to the center of the Earth, the seismic values are those of the PREM, but the deepest point in the inverted region is set to 1000 km. This defines a 500 km thick region within which the continuity between the seismic profile in the transition zone and in the PREM below is ensured. This means the temperature at 1500 km depth has to be set but various numerical experiments show that this value does not influence the results in the region of interest (350- 1000 km).

B.4.2 Prior information

The Bayesian formulation enables to account for *a priori* knowledge, provided that this information can be expressed as a probability distribution $P(\mathbf{p})$. One drawback of the

Bayesian inversion is the possible tuning of the prior sampling. Here we choose minimal prior information, which consists of uniform probability distributions in wide realistic parameter spaces. Following this statement, the priors on the parameters are uniformly distributed over wide domains (Fig. B.2).

For the temperature, where priors need to be set only between 350- and 1000 km depth, the low-temperature bound corresponds to the thermal state in a subducting lithosphere (Tagawa et al., 2007; Ganguly et al., 2009), while the high-temperature bound is defined by the dry solidus of a peridotite (Hirshmann, 2000; Hirshmann et al., 2009). For the anisotropy, inferred between 0- and 1000 km depth, the domain bounds are set to 0.9 and 1.1. These bounds are chosen to bracket the range of values obtained in recent surface wave studies (Panning & Romanowicz, 2006; Kustowski et al., 2008; Visser et al., 2008b; Khan et al., 2009, 2011). Even when radial anisotropy might decrease with increasing depth, we do not shrink the domain range in the deepest part of the model space. Between the surface and 350 km depth, the seismic structure is randomly sampled within the bounds of +10% and -10% of the PREM (Dziewonski & Anderson, 1981). The seismic value at 350 km depth is not randomly sampled, but is computed from the temperature value (since it is also the shallowest point of the transition zone), owing to the thermodynamical law. To prevent from an artificial discontinuity between the uppermost mantle and the transition zone, the point located just above 350 km depth is sampled in such a way that the V_S gradient is between $\pm 2.5 \times 10^{-3} \text{ s}^{-1}$.

Figs. B.3a→c represent the uniform prior distributions on the Bézier control points (see Table B.8 for details), and Figs. B.3d→f enhance the resulting prior probability distributions on the temperature, and anisotropy and uppermost V_S models, respectively, following eq.(B.6) and eq. (B.7). Note that the continuity of the Bézier curves modifies the former uniform sampling of the parameter space using discrete control points.

In order to be consistent with the data resolution power, minimal distances are set between two consecutively randomly sampled Bézier control points. This leads to unsampled regions, near the depth bounds (red lines), visible in Figs. B.3a→c.

B.4.3 The McMC algorithm within a three-step exploration

In the McMC algorithms, new models are proposed by randomly perturbing the previously accepted model. Here, the sampling of the parameter space is performed using a continuous proposal function and the stepsizes of the exploration depend on each parameter (V_S , T , ξ , z_{V_S} , z_T and z_ξ). For instance, defining the t^{th} and the $t + 1^{\text{th}}$ value of a parameter p , as p_i^t and p_i^{t+1} respectively, then the subsequent step may be defined as $p_i^{t+1} = p_i^t + w_i^t$, where w_i^t is the t^{th} stepsize, randomly sampled from a normal distribution

with zero mean. A Gaussian probability density distribution, centered at p_i^t ,

$$q(p_i^{t+1}|p_i^t) = \frac{1}{\theta_i \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{w_i^t}{\theta_i} \right)^2 \right], \quad (\text{B.13})$$

is classically used to randomly sample the p_i^{t+1} , where θ_i is the standard deviation of the Gaussian distribution. If p_i^{t+1} is out of the prior bounds, then the random walk reflects on the edge to respect the equirepartition of the proposal distribution. A proper choice for the θ_i value mostly affects the efficiency of the exploration and therefore the computing time, but it does not influence significantly the posterior probability distribution (MacKay, 2003). For normal distributions, an optimal jumping rule produces an acceptance rate of about 23% for multidimensional problems (Gelman et al., 2003).

In order to perform a comprehensive and efficient sampling of the parameter space within a reasonable computation time, the parallel predisposition of the algorithm (Rosen-thal, 2000) enables us to propose an unusual split of the inversion scheme into three steps. The two first steps look for a family of the best-misfit configurations of the parameters, and the statistics are performed during the third step, named *stationary period*. During the first step, the standard deviations (θ_i) are chosen to be large enough to explore a large part of the configuration space. A great number of chains (1600 in the presented inversions) are placed on a parallel computer system, and hence each chain runs on an independent processor. The chains start from a different state and follow different paths in the parameter space. At each iteration, the synthetic data computed from the given configuration is tested against the real data by computing the misfit (eq. B.10). First, the previously described procedure (see section B.3) is used to simultaneously explore all models providing that, at each iteration, a whole set of randomly sampled trial models is tested, up to a uniform sampling of the parameter space. Various amount of control points, which construct temperature Bézier curves (see section B.2.2), are used in order to not *a priori* fix the sampling of the parameter spaces. This step is stopped after 6×10^3 iterations and the best-misfit configuration is then determined for each independent chain.

One of the major issue when using MCMC exploration is to know which iteration rank is to estimate the number of iterations needed to achieve statistical convergence of the procedure. Practically, it can be however disabling, with respect to the computation time, to stop the iterations at any time. For this reason, the second step restarts with the parameter configuration which gives the best misfit during the first step.

Only the 20 % best-misfit configurations of all the chains are selected for the second step to refine the results. The second step is done, again using the same MCMC procedure during 6×10^5 iterations, using thinner Gaussian distributions for the parameter sampling but, this time, modifying only one couple of parameters at each iteration. The strategy of perturbing only one couple at a time ensures to preserve most of the characteristics of the

current model, which may have resulted in a good data fit. This induces an increase of the acceptance rate ($\sim 20\text{-}25\%$) while keeping a large-scale exploration of the models. The choice of the perturbed model, (V_S, z_{V_S}) , (T, z_T) or (ξ, z_ξ) , for the next iteration, is randomly done. This second step could be assimilated to the so-called *burn-in* period, but here it is not necessarily the configuration corresponding to the last iteration that is conserved.

The *stationary period* starts after the computation of the best-misfit configurations during the second step. Only 5 % of these configurations, with respect to the 1600 chains considered for the first step, are engaged in the third step. Each chain runs for 6×10^5 iterations in total. The amount of Bézier control points is the same as the one randomly chosen at the beginning of the second step, and since many Markov chains are started in parallel with various amount of Bézier points, the amount of parameters for temperature is also a parameter itself, which gives some common features with the reversible jump approach (Green, 1995; Green & Mira, 2001; Bodin & Sambridge, 2009; Bodin et al., 2012). Only the 4.8 millions of models sampled during the *stationary period* are used to compute the pdfs. This procedure should guarantee that the samples are less correlated as well as leading to a better coverage of the probability distribution. Another advantage, since the chains are preconditioned, thanks to the two first steps, the beginning of the *stationary period* is well defined and does not need any statistical assessment.

B.5 Inversion of synthetic data

In this section we show how we test our method, using synthetic data. Tests on synthetic data are useful to evaluate the robustness of the method and to highlight some features which could be found on real data, considering real uncertainties on dispersion curves. We also discuss the question of how to interpret the results by the analysis of the posterior distributions. The results demonstrate the potential of the inversion method to detect and decorrelate first-order variations of temperature in the transition zone, shear wave anisotropy between the surface and 1000 km depth, and shear wave velocity in the crust and the upper mantle.

B.5.1 Input models

Synthetic data are computed for known temperature and anisotropy models. A pyroclitic composition is considered because it is the most widely accepted model of mantle composition, and generally compatible with a large range of petrologic, geochemical and seismic observations (Ringwood, 1975). The iron number ($x_{\text{Fe}} = \text{Fe}/(\text{Fe}+\text{Mg})$, where Fe

and Mg indicate the bulk molar abundances) is assumed to be constant throughout the whole transition zone and is set to 0.11.

To illustrate the different features of the algorithm, we choose two synthetic models represented in Figs. B.4 and B.5. The 1-D profiles are plotted with black lines and the corresponding Bézier points with black stars. We choose temperature models far from an adiabatic profile (dashed lines in Figs. B.4h and B.5h), to demonstrate the ability of the algorithm to detect strong temperature variations. The model 1 is created in such a way that the perturbation on phase velocities produced by temperature variations compared to an adiabatic profile is two times larger than the perturbation caused by anisotropy. Second, we investigate with model 2 the question of whether or not the anisotropy in the transition zone can be detected by putting a shear wave anisotropy anomaly at 500 km depth.

The number of parameters used and the constraints for the Bézier curves computation are shown in Table B.8. To fully investigate the robustness of the method, a relative freedom is given for the placing of temperature parameters, setting the maximal gradient between two consecutive points to 5 K/km. We sample 4.8 millions of models during the *stationary period*. To eliminate dependent samples in the ensemble solution because only one couple of parameters is modified at each iteration, every 40th visited model is selected for analysis.

Synthetic data are Love and Rayleigh surface waves dispersion curves of the fundamental mode and the first three overtones. They are computed from the input models with the forward method described in section B.2. We use in average 35 periods per mode, covering the 45 - 273 s period range. Realistic uncertainties taken from Beucler et al. (2003) are used and they vary between 0.04 and 2.4 % of the synthetic phase velocity values.

B.5.2 Characterization of posterior results

Once the posterior probability density function is thought as the solution to our inverse problem, our main concern is now the analysis of the distributions. We have observed in Fig. B.3 that the pdfs of *a priori* 1-D profiles are not uniform. Consequently, results must be taken with care and other representations, in addition to color density plots, are necessary to fully investigate the results. There is no unique way on visualizing the system under study. Thus, the results presented here reflect the particular parameterization chosen.

Before discussing the results shown in Figs. B.4 and B.5, the four different sorts of representations used to analyze the distributions are detailed.

- The distribution of the Bézier points and its pdf (Figs. B.4a, e, i, for instance). This figure reveals the regions of the model space where the parameters are spread or narrowed. The depth of the Bézier points is not *a priori* fixed, so this representation is useful

to investigate if the points are concentrated around some specific depths. The pdf gives an additional information about the most sampled regions. The drawback is that we cannot see which points are linked together to form a 1-D profile.

- The distribution of the Bézier points and the misfit values (Figs. B.4b, f, j). The information provided by the misfit is usually disregarded in Bayesian studies, as the theory relies on probabilities based on the repetition (or not) of certain values of the parameters. But the misfit is the only quantitative expression indicating if the tested model is closed to the expected one. This figure allows to be sure that the algorithm was not trapped in a local minimum. Theoretically, the regions of low misfit values should match with regions of high probabilities.
- Some models randomly taken in the ensemble models and the four best-misfit models (Figs. B.4c, g, k). Mosegaard & Tarantola (1995) encourage this kind of representation. Here we consider a very small subset of models (20 models) and it cannot be used to infer statistical properties, but it is useful to visualize the diversity of the sampled models and to detect the depths which are the best resolved. In general, the 1-D models show different features, but all give a good fit to the data. For the synthetic tests, the profiles corresponding to the four best misfits are in general very close to expected model.
- The pdf of the Bézier curves (Figs. B.4d, h, l). This representation is the most widely employed. Using continuous Bézier curves, 1-D marginals of T , ξ and V_S can be computed at each depth. The juxtaposition of the histograms gives a 2-D density probability map. This figure provides an overview of the most frequently sampled paths, but has a tendency to smooth the results.

B.5.3 Results

The results of the probabilistic inversion of model 1 and 2 are shown in Figs. B.4 and B.5, respectively. Labels are placed on figures to make easier the visualization of some particular features discussed in the following.

B.5.3.1 Temperature

Concerning temperature, both Bézier curves distributions contain the input models. Bézier points are scattered around the profiles, with high density values of points in the vicinity of the profiles. The misfit values are low for the Bézier points located along the profiles to retrieved. The highest density values coincide with temperature extrema (labels 1)). The pdfs are sharper between 350 and 600 km depth than in the rest of the transition zone. The marginal probabilities of model 2 at 410 and 660 km depth are compared in Fig. B.6. At 410 km depth, the temperature to retrieve is, without ambiguity, the most

frequently sampled, while it is not so visible at 660 km depth. In Figs. B.4g and B.5g, the best-misfit profiles are all relatively close to the true models. The profiles randomly taken in the ensemble solution are oscillating around the input models between 600 and 800 km depth (labels [2]) and around 900 km depth (labels [3]). At these depths, the Bézier points are the most scattered. The distributions of Bézier points are narrower just above 1000 km depth due to the *a priori* PREM connection condition at 1500 km.

We performed an inversion with the same models but considering that the Earth is only composed of pure olivine, without phase transition, and we found a narrower distribution in the middle of transition zone. In this case, the pdf takes nearly the same value down to 800 km depth, suggesting that the relative broader distribution just below 600 km depth for models 1 and 2 are due to mineralogical transformations in both olivine and pyroxenes-garnet subsystems. In return, the oscillations below 800 km remain like in the case of a pyrolitic model, and are attributed to the decrease of surface waves sensitivity.

Fig. B.7 emphasizes that models with different attributes are able to describe the same data, accounting for data uncertainties. This figure represents the pdfs of the phase velocity distributions of all the sampled models for the inversion of model 2. Fig. B.7 clearly enhances that all the sampled models fit the synthetic data within their uncertainty bounds. Only the fundamental mode and overtones of Love waves are shown, but Rayleigh branches are seen to be fit within uncertainties too. Models on Figs. B.4g and B.5g individually highlight different temperature structures of the mid-transition zone, but nonetheless are models that all produce a good fit to the synthetic data. This good agreement between synthetic and tested data reinforces the method employed here of elucidating the transition zone interior from a big number of models, rather than just adopting the 1-D profile belonging to the region of highest probability as the prevalent one (Figs. B.4h and B.5h), and inferring the transition zone structure straight from this.

The number of Bézier points selected for each chain of the *stationary period* is summarized in Table B.8. Model 1 and model 2 are composed of 6 and 7 Bézier points, respectively. The right number of points is not the most frequently used. There is not a number of points that is significantly sampled rather another, which means that the number of points is not decisive because the parameters' depth is not *a priori* fixed. When the number of points is too high than the expected one, Bézier points are placed along the Bézier curve to retrieve and the form of the 1-D profile expected is preserved.

B.5.3.2 Radial anisotropy

Figs. B.4l and B.5l show that the distributions of Bézier curves are in very good agreement with the expected profiles. Both anisotropy perturbations at 200 and 500 km depth, for models 1 and 2, respectively, are retrieved at the right depths (labels [4]).

Figs. B.4j and B.5j show two clusters of low misfit values at these depths, which means that dispersion curves are sensitive to anisotropy variations in the upper mantle as well as in the transition zone.

The subset of 1-D models is nearly similar to the input models down to 750 km depth for model 2. Below this depth, some 1-D models are oscillating around the value of one (labels [5]) and the Bézier points fill nearly the entire parameter space. Even though some points are far from the Bézier curves to retrieve, the corresponding misfit values can be relatively low (label [6]). The same behaviour is observed for model 1, where the points are getting wider as a function of depth symmetrically around $\xi = 1$. According to sensitivity kernel theory, the fundamental mode and the first overtone are very sensitive to the 200 km of the Earth. Higher modes enable to enhance the resolution down to 1000 km depth, but to a lesser extent, which explains the broader distribution of Bézier points below 700 km depth, as for temperature. Our results demonstrate that the probabilistic inversion method employed here presents a clear potential to detect radial anisotropy in the transition zone, if any.

B.5.3.3 V_S in the uppermost mantle

Figs. B.4d and B.5d show that V_S profiles in the uppermost mantle are well retrieved for both cases, even when the input profile is close to the prior bounds, like for model 2 (label [7]). Strong changes of V_S gradient are well characterised by clusters of Bézier points. At 50 km (model 1) and 80 km depth (model2), where the slope change is the strongest, we observe a correlation between the depth of the Bézier points and V_S values (labels [8]). This trade-off means unsurprisingly, that the data are fit equally well when the discontinuity is deeper and V_S is higher, and *vice-versa*. This illustrates the limitation of the resolving power of the data. The other points are located along the Bézier curves to retrieve.

The points are more concentrated than for T and ξ results and are less numerous, which means that V_S parameters were already close to the expected values at the end of the second step of the three-step exploration (see section B.4.3), and that the algorithm did not need to change them too much during the *stationary period*. The well constrained V_S profiles are in good agreement with the fundamental mode sensitivity to the shallowest part of the Earth. Fig. B.7 shows that fundamental mode error bounds are very thin (between 0.04 and 0.065 %) within the 50 to 100 s period range, which means that the misfit is significantly improved when the V_S profile is close to the true one.

B.6 Real data

Our non-linear algorithm is used to determine the average temperature and anisotropy distributions under the Vanuatu-California path, which has been the focus of several previous studies (e. g. [Cara, 1979](#); [van Heijst & Woodhouse, 1997](#); [Beucler et al., 2003](#)).

B.6.1 Data and *a priori* conditions

The data set we analyze is made of Love and Rayleigh phase velocities up to the third order of [Visser et al. \(2008a\)](#). Phase velocity measurements were obtained using a model search approach ([Visser et al., 2007](#)), yielding consistent uncertainties between all the measurements. We use a total of 107 distinct phase velocities as function of period that we invert jointly for a thermal and radial anisotropy structure along a great circle. The period range investigated is 35-175 s and uncertainties are between 0.29 and 0.77 % of phase velocities. The bulk and shear dissipations are specified as in PREM. Pressure profile remains fixed to the PREM values, and a pyrolite composition is considered. The expected temperature gradient in the mantle is 0.4 K/km ([Katsura et al., 2010](#)), but we let a large freedom in the location of the Bézier points by fixing the value of ± 1.5 K/km to the maximal gradient between two consecutive points.

Prior boundaries are slightly different of the ones used with synthetic tests. Considering previous studies, a strong positive radial anisotropy is expected under the Pacific plate ([Ekström & Dziewonski, 1998](#); [Kustowski et al., 2008](#); [Nettles & Dziewonski, 2008](#); [Khan et al., 2009](#)). Consequently, we extend the upper bound of ξ to 1.3 instead of 1.1. Because the path under study is oceanic, high seismic values are expected in shallow layers and we allow a wider range of seismic velocities for the first 25 km. Setting relatively large bounds prevents from compensation effects between the different sets of parameters because several parameters cannot reach adequate values.

B.6.2 Results

B.6.2.1 Temperature distribution

Figs. [B.8e](#)→[h](#) show that the temperature distribution is globally close to an adiabat. As for synthetic tests, the Bézier points are spread at all depths and the profiles do not depend on the number of points used (see Table [B.8](#)). The temperature is particularly well defined between 350 and 550 km depth, where the pdfs are the highest (Figs. [B.8e](#) and [h](#)). The distribution is more scattered in the mid-transition zone, between 550 and 750 km depth.

In Fig. B.8g, three kinds of profiles can be distinguished. The first family (referred as profiles 1 in red) is nearly adiabatic and is very close from the median of the distribution (in blue). The temperature gradient of the median profile is about 0.22 K/km, which is slightly lower than the supposed one in the mantle. When the temperature values are extrapolated to the surface, a potential temperature of about 1600 K is obtained. The second (in green) and third (in purple) families of models are characterised by an increase and a decrease of temperature, respectively, with respect to the median profile.

Although the pdf of Bézier curves appears to be smooth and unimodal between 550 and 750 km depth (Fig. B.8h), the statistic results on Bézier points (Fig. B.8e) enhance two distinct clusters (labels 1 and 2), also visible with misfit values (Fig. B.8f). They can be assimilated to the families 2 and 3. All the results gathered together, it seems that both high and low temperatures in the mid-transition zone are able to fit the data within uncertainties, as shown in Fig. B.9.

B.6.2.2 Radial anisotropic structure

Between the surface and 350 km depth, a strong positive ($V_{SH} > V_{SV}$) anisotropic signal nearly equal to 1.05 is revealed. It seems to be a robust feature because Bézier points are all concentrated around this value (Figs. B.8i and j). This result is different from the PREM but the shear wave anisotropy structure retrieved here concurs with upper mantle results obtained from recent whole mantle seismic tomographic studies. The origin of this anomaly is not well understood, though several authors have proposed that this region of anisotropy corresponds to a flow channel at the boundary between the lithosphere and asthenosphere (Gaboret et al., 2003; Gung et al., 2003). A stronger anisotropy is depicted near the surface. Between 350 and 450 km depth, the signal amplitude of anisotropy is decreasing as a continuous fashion as the transition zone approached. The distribution is centered around a value of one down to 1000 km depth. As we observed with synthetic tests, the distribution of Bézier points is broader below 700 km depth, due the decrease of surface waves sensitivity.

Several models (Fig. B.8k) show slight oscillations around $\xi = 1$ at depths deeper than 400 km (labels 3 in Fig. B.8k and 4 in Fig. B.8j). The marginal distributions of ξ at this depth is represented for Bézier points and Bézier curves in Fig. B.10. Both histograms clearly show that $\xi \sim 1$ is the most probable value. The case for a slight (~ 0.97 at most) negative shear wave anisotropy ($V_{SH} < V_{SV}$) below 400 km depth, as observed by Panning & Romanowicz (2006) and Visser et al. (2008b), is not substantiated here and our results indicate that the transition zone is nearly isotropic.

B.6.2.3 Seismic velocities and density

We describe as in the synthetic tests, the results of the primary parameters of the inversion scheme, *i. e.* V_S between the surface and 350 km depth (Figs. B.8a→d). Moreover, the derived physical properties from the temperature distribution are discussed, in the form of S and P waves velocities and density distributions of the transition zone (Fig. B.11), since these parameters are also determined during the whole direct problem.

V_S in the uppermost mantle

At the surface, V_S values are globally large compared to PREM and are consistent with an oceanic path. Unlike synthetic tests, all depths are sampled with Bézier points (Figs. B.8a and b). This is due to larger data uncertainties, which allow the algorithm to accept wider configurations of Bézier points.

Between 50 and 150 km depth, V_S is smoothly decreasing, and can be considered to be a low-velocity zone extensively described before under oceans (*e. g.* Gaboret et al., 2003; Gung et al., 2003; Nettles & Dziewonski, 2008). A change slope that we assimilate to the Lehmann discontinuity occurs between 200 and 300 km depth. This feature is clearly visible on Bézier points, where they split into two branches (label 5). Precise agreement is not warranted here as surface wave data are not as sensitive to the exact location of discontinuities as are converted or reflected phases, for example.

V_S , V_P and ρ in the transition zone

The V_S , V_P and ρ distributions plotted in Fig. B.11 are computed using the forward problem as detailed in section B.2.1.

In the upper half of the transition zone (between 400 and 520 km depth), P and S wave velocities distributions are found to be larger than the PREM, while in the lower part P and S waves velocities are slightly lower, implying an overall steeper velocity gradient in this region. The density distribution agrees with PREM between the transformations ol→wad and wad→ring. At the latter, density increases and become more than PREM.

The depth where the phase transition ol→wad occurs is similar to the one of PREM, roughly at 400 km depth. The temperature the most frequently sampled at this depth is ~ 1650 K, within the range that is expected for this transition from high pressure experiments, *i. e.* $\sim 1750 \pm 100$ K (Ito & Takahashi, 1989). A striking feature is the magnitude of the seismic jump near 400 km depth of our V_S and V_P distributions, relative to PREM. This discrepancy between the results based on elastic moduli taken from laboratory experiments and global models was already remarked by Stixrude (1997), Cammarano et al. (2003) or Katsura et al. (2004) for a pyrolitic composition. The amplitude of the discontinuity could be reduce considering a less olivine-rich mineralogical model (Duffy et al.,

1995). Instead, the amplitude of the seismic discontinuities may be underestimated in the global seismic models (Shearer, 2000), which may be due to finite frequency effect of seismic waves (Jackson, 1998). However, as the data considered here are less sensitive to the location and size of the discontinuities, we leave it for further study to investigate this discrepancy.

The phase transition wad→ring (the 520-km discontinuity) is not present in PREM nor in the V_S distribution, but appears in the density distribution, where the jump occurs between 490 and 520 km depth. The V_P distribution shows a slightly low velocity zone around 500 km depth. This behaviour is linked to the particular values chosen for the wadsleyite and ringwoodite elastic moduli and their derivatives with respect to temperature and pressure, and was already noted by Cammarano et al. (2005a).

Between 600 and 700 km depth, the distributions are more complicated. The inverse problem is highly non-linear, and hence the posterior is far from being an unimodal Gaussian distribution. To illustrate this, we plot in Fig. B.12 the marginal distributions for V_S at 655, 660 and 665 km depth. The 660 km histogram has two maxima, which means that the marginal distribution is influenced by velocity values above and below the discontinuity. Looking at Fig. B.11a, the mineralogical transformation of ringwoodite to perovskite and magnesiowustite occurs over a wide depth range above the PREM discontinuity. In Fig. B.11a are plotted the V_S profiles corresponding to the three families of temperature profiles shown in Fig. B.8g. Because the ring→pv + mgw transformation is endothermic (e. g. Ito & Takahashi, 1989), the coldest model (profile 3) has the deepest discontinuity, and the hottest (profile 2) has the shallowest one.

Just below 700 km depth, an additional transition to higher velocities are observed on several models in Fig. B.11a, due to other phase transformations in the pyroxenes-garnet subsystem. Deeper, S and P wave velocities are equal or slightly higher than PREM, but the density is higher of about 1.7 %.

B.6.3 Effect of the Clapeyron slope at 660 km depth on the temperature distribution

The distribution of the Bézier points defines three possible profiles of temperature. They are displayed in Fig. B.13, along with stability range of the olivine polymorphs (in black) and Mg-perovskite and magnesiowustite (in pink). The limit between the latter domains clearly coincides with the Clapeyron slope (dP/dT) used in this study, $\Gamma_{660} = -2.8$ MPa/K (Ito & Takahashi, 1989).

In order to illustrate the covariance between temperature and composition, we further investigate the effect of the Clapeyron slope on the temperature distribution, by setting

its value to $\Gamma_{660} = 0$ MPa/K. Then, the transformation occurs instantaneously when the temperature profile cross the limit of 24 GPa (Fig. B.13). The results for V_S and ξ distributions are very similar to the ones obtained with $\Gamma_{660} = -2.8$ MPa/K (not shown here), and no trade-offs between V_S , T and ξ were obtained. The temperature results are displayed in Fig. B.14. Although the distribution is similar to the one of Figs. B.8e→h down to 500 km depth, a $\Gamma_{660} = 0$ MPa/K implies a temperature decrease in the mid-transition zone. If the best-misfit profiles (Fig. B.14c) are compared to the three families of profiles (Fig. B.13), we observe that only family 3 is preserved. The results of this test indicate that the temperature values that are inverted close to seismic discontinuities depend on the priors on the composition and thermodynamical parameters.

B.7 Conclusion

In this paper, we present a new method combining a thermodynamical law with a non-linear inversion scheme to infer radial distributions of temperature in the transition zone, shear wave anisotropy in the mantle and shear wave velocity in the crust and in the upper mantle from seismic surface wave data. This method was used to invert fundamental and higher-order Love and Rayleigh dispersion curves (up to the third overtone). This was done with a MCMC algorithm, that is, by performing a random walk in a multidimensional model space that combines prior information with misfit to the data from model parameters. Input to the algorithm were random models generated according to the prior distribution and the likelihood function. As output we assimilated random realizations of the posterior distribution. We chose to invoke as few prior constraints as possible in order to gauge which particular feature is most probable given data and *a priori* informations. In our approach, size and location of discontinuities in seismic properties are modeled in a physically realistic manner, as their variations depend on composition and physical conditions of the particular model being considered. The use of Bézier points as models' parameters, whose depths are not *a priori* fixed, constitutes an adaptive parameterization that allows a great flexibility in the exploration of the model space.

We tested our algorithm with synthetic and real phase averaged velocity dispersion curves. We furthermore give some examples of how to investigate the sampled models. The results show that the data are sensitive to temperature and shear wave anisotropy in the transition zone. Regarding anisotropy, the picture that emerges of the Visser et al. (2008a)'s data is in overall accordance with previous studies, that is a strong shear wave anisotropy in the upper mantle and a roughly isotropic behaviour in the transition zone. Improvements on phase velocities uncertainties are surely welcome to discriminate whether or not a significant sign change really occurs below 400 km depth. The temperature distribution is nearly adiabatic with a gradient of 0.22 K/km and a potential temperature

of 1600 K, but is relatively complex between 600 and 700 km depth, which means that the thermic structure of the transition zone cannot be reduce to an unique 1-D profile. In addition, we observed that the temperature distribution is very sensitive to the choice of the Clapeyron slope of the ring $\rightarrow$ pv + mgw transformation.

The range of reported Clapeyron slopes for the perovskite-forming reaction is wide : -3.0 (Ito & Takahashi, 1989) to -0.2 MPa/K (Litasov et al., 2005). A further study would be to evaluate the effects of the Clapeyron slope on our inferred temperature distribution by considering Γ_{660} as a parameter in itself. We tested for a joint inversion of temperature and mineralogy for synthetic data, but posterior distributions of volume fractions of mantle minerals was always very close to the prior distributions. Olivine and pyroxenes exhibit a nearly identical mechanical behaviour, and cannot be distinguished on the basis of seismic data only. Given the generality of our formulation, additional independent data sets such as electrical conductivity and anelasticity properties, can easily be incorporated to separate the effects of temperature and mineralogy (*e. g.* Verhoeven et al., 2009).

MCMC methods in seismology are promising, and offer a nice alternative to deterministic methods which tend to find a model which lies close to its start. The main restriction is the computational cost. In spite of this limitation, since the proposed algorithm is a direct parameter search, the forward calculations are separate routines independent of the main algorithm, and hence they can be easily replaced by alternative algorithms.

B.8 Figures and tables

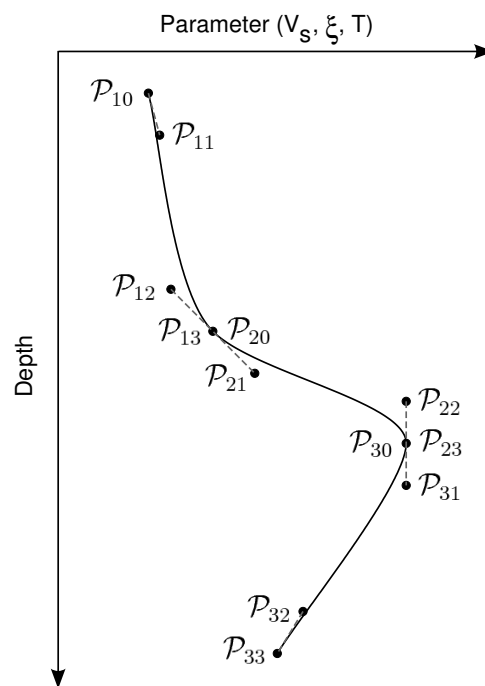


FIG. B.1 : Schematic representation of the model parameterization using a continuous set of C^1 Bézier curves. Four control points define each polynomial and its derivative. Continuity requires that the upward and downward derivatives are identical at each point that is common to two consecutive polynomials. It works for any kind of parameter.

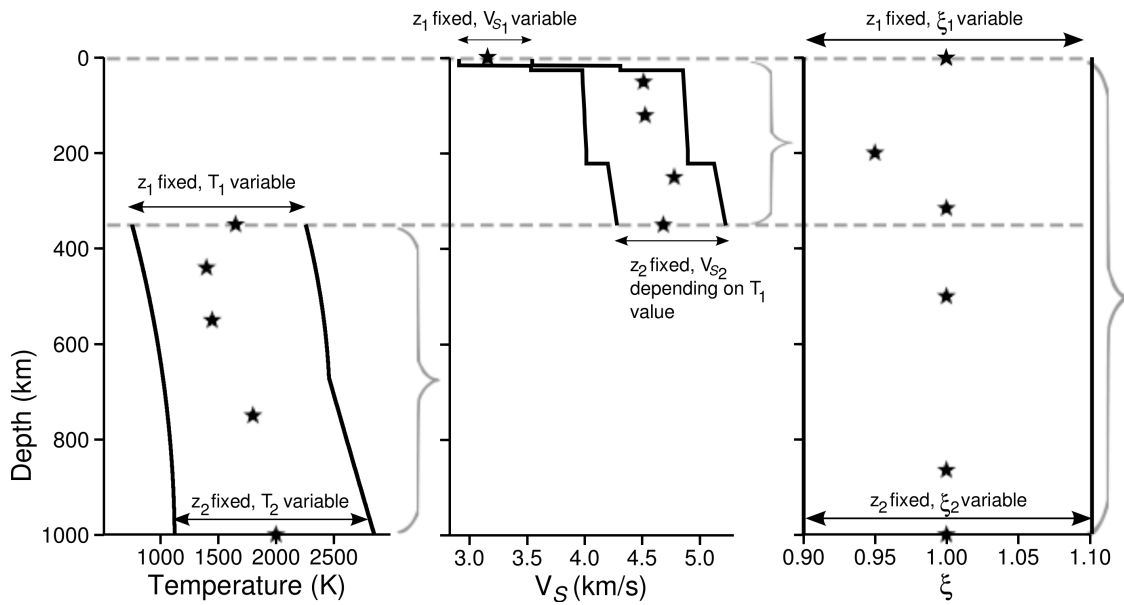


FIG. B.2 : Illustration of *a priori* setting on the Bézier points used in both synthetic and real data inversions for temperature (T), shear wave velocity (V_S) and anisotropy (ξ). Black stars represent the Bézier points. Dotted gray lines define 0 and 350 km depths. Black lines are the bounds of the model space. Embraces indicate the depth ranges where the points are randomly sampled. The depth ranges for T , V_S and ξ profiles are 350- 1000, 0- 350, and 0- 1000 km depth, respectively. The shear wave velocity V_{S2} at 350 km depth is computed according to the temperature T_1 .

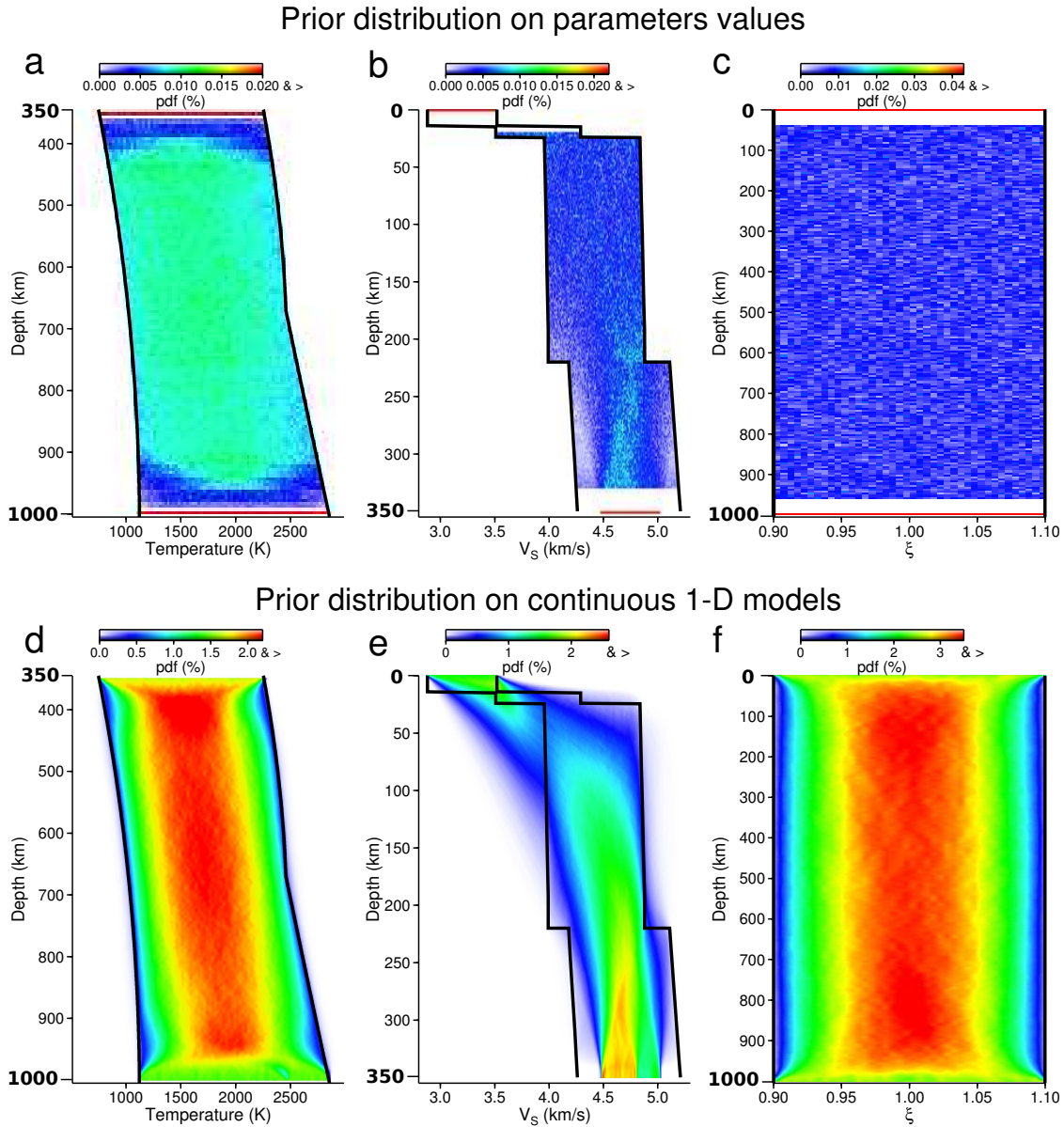


FIG. B.3 : Marginal probability density functions (pdf) for temperature (left), shear wave velocity (middle) and anisotropy (right), considering that all the sampled models which are in good agreement with *a priori* informations are accepted. The results are shown for the discrete values of the parameters (top row) and for the 1-D models (bottom row). Red and blue colors show high and low probabilities, respectively. Black lines represent the minimum and maximum parameter values allowed. The pdf values for the parameters are computed by counting the number of Bézier points in each cases. The sizes of the cases for T , V_S and ξ are $25 \text{ K} \times 5 \text{ km}$, $0.01 \text{ km/s} \times 2 \text{ km}$ and $0.005 \times 5 \text{ km}$, respectively. The pdf values for the 1-D models are computed by counting the number of curves in each parameter interval at each depth. For a given depth, the sum of the pdf over all the parameter intervals is equal to 100 %

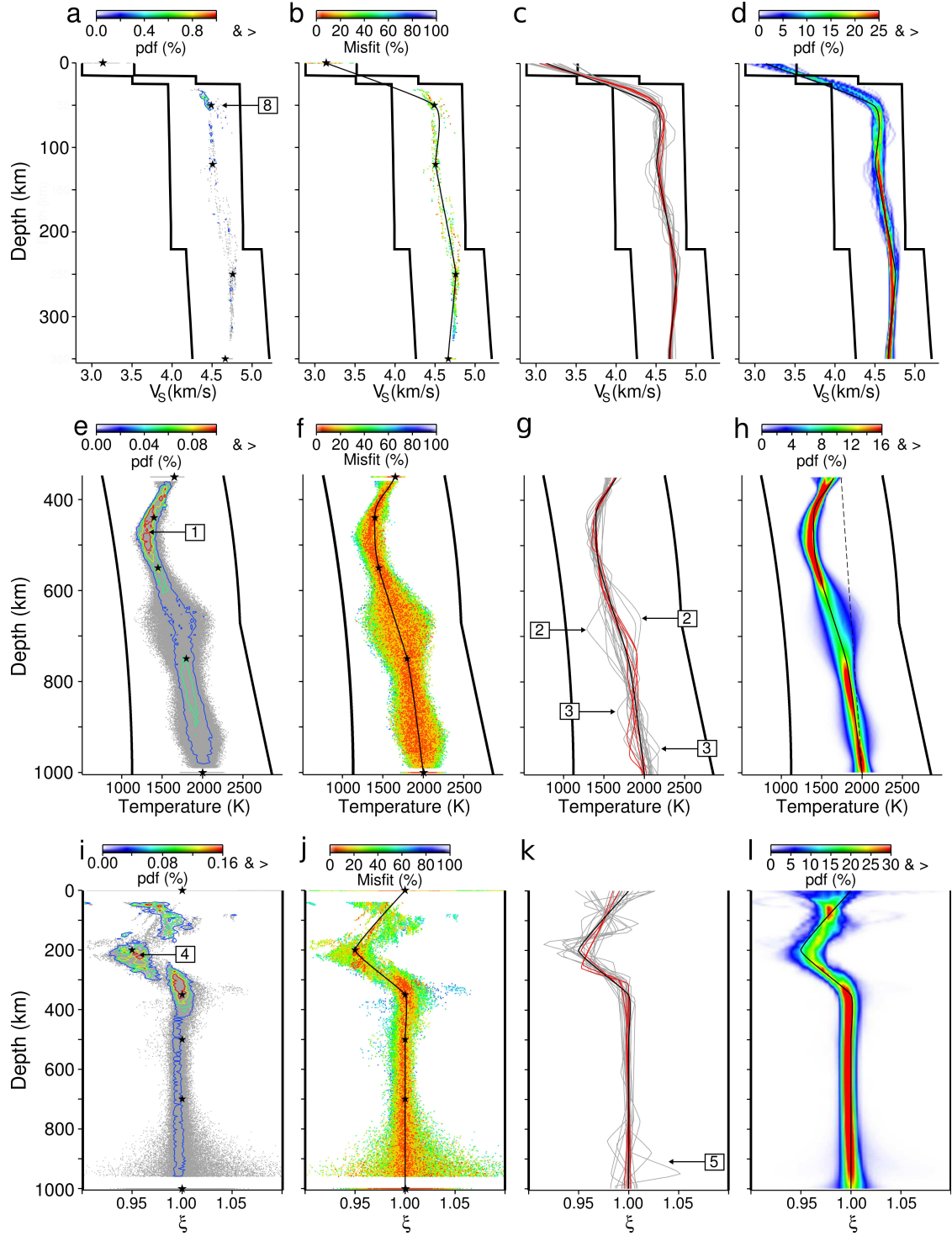


FIG. B.4 : Results of the probabilistic inversion for the synthetic model 1 for V_S (top), T (middle) and ξ (bottom). Black lines and black stars show the models to retrieve and the corresponding Bézier points. Thick black lines are the prior bounds. (a), (e) and (i) show the Bézier points accepted during the MCMC exploration. Color lines represent the corresponding pdf. (b), (f) and (j) display the Bézier points and the misfit values. Blue and red colors are high and low misfit, respectively. Misfit values are scaled according to the highest and the lowest values encountered. (c), (g) and (k) show in gray a random subset of 20 models taken in the ensemble solution, which contains 120 000 models. Red lines are the profiles corresponding to the four best misfits. (d), (h) and (l) are color density plots of 1-D profiles. Dashed lines in (h) show the 1600 K adiabat considering a temperature gradient of 0.4 K/km. The squared labels refer to special comment, see section B.5.3 for details.

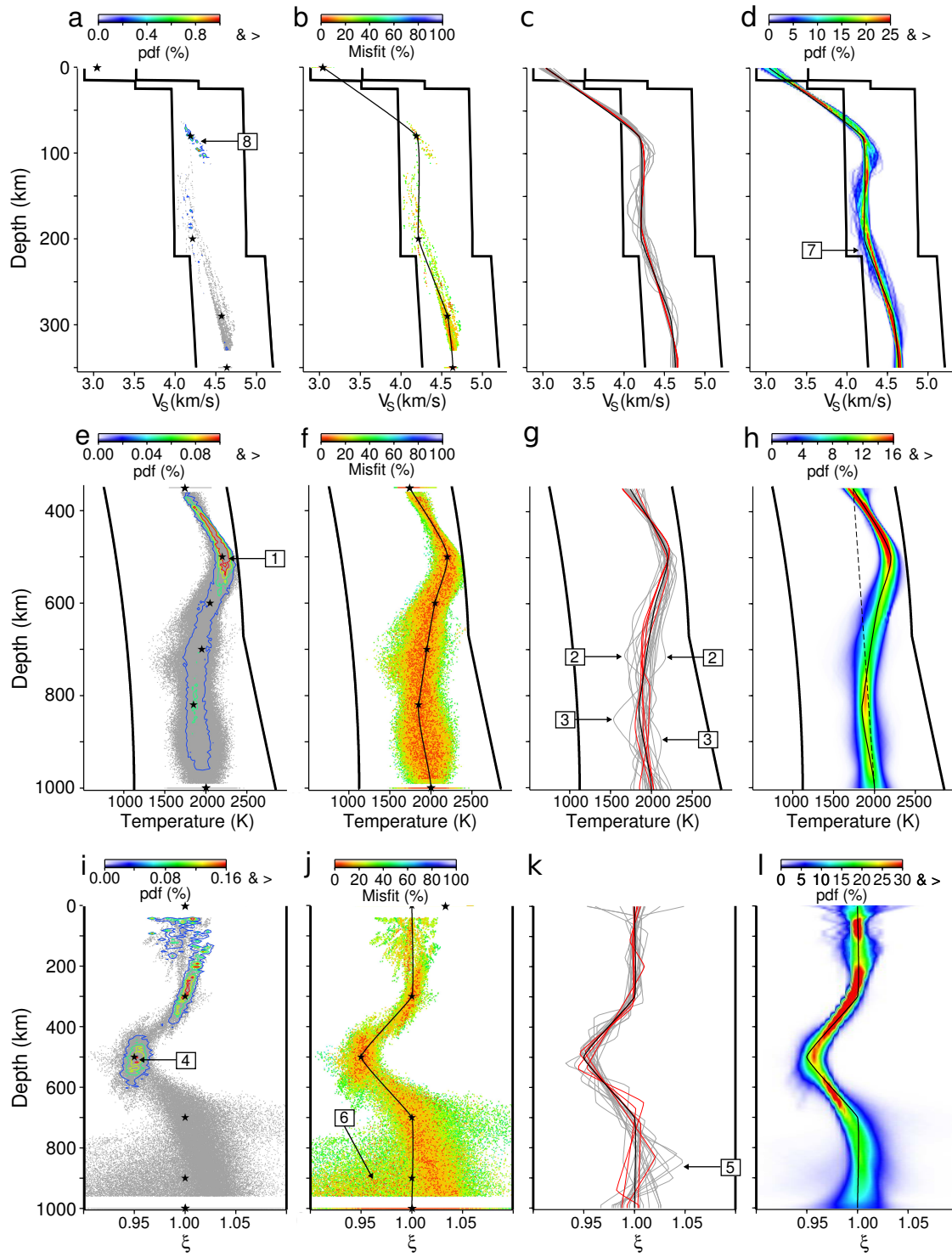


FIG. B.5 : Same as Fig. B.4 but for model 2.

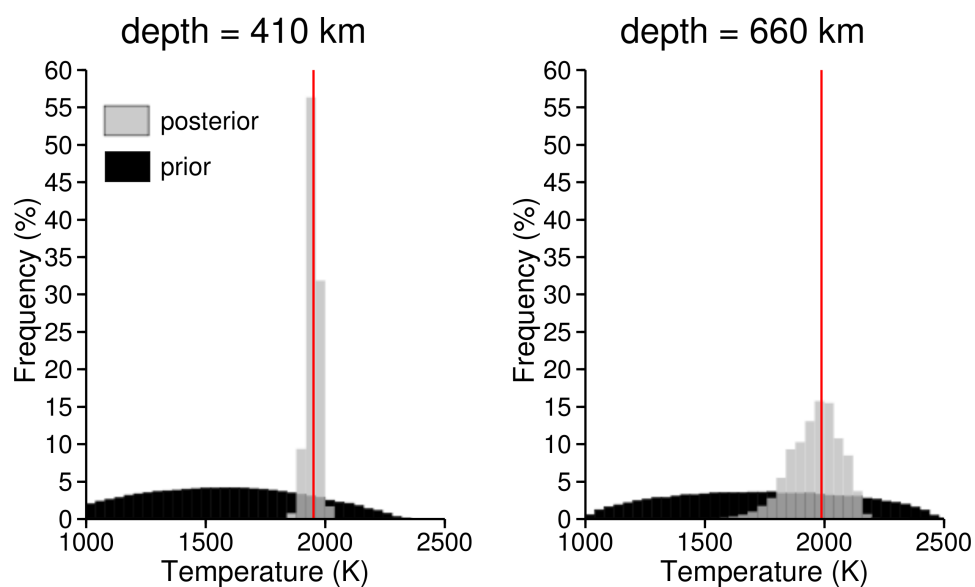


FIG. B.6 : Prior (in black) and posterior (in gray) marginal probabilities of Bézier curves on temperature at 410 and 660 km depth for model 2. Input values are shown in red (1951 K at 410 km and 1987 K at 660 km). Intervals are 20 K wide.

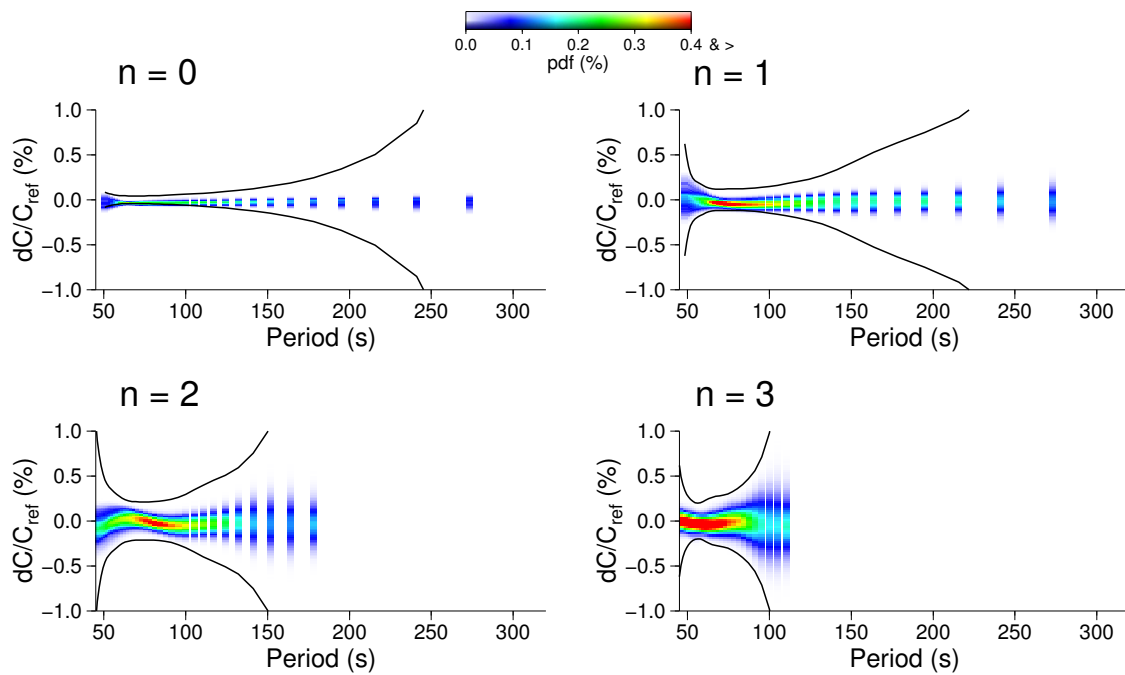


FIG. B.7 : Posterior phase velocity distributions of Love fundamental mode and overtones of model 2. Results are shown in percentage of phase velocities (C) compared with phase velocities of model 2 (C_{ref}). Black curves are uncertainties on phase velocities. For long periods, uncertainties are large and are not displayed.

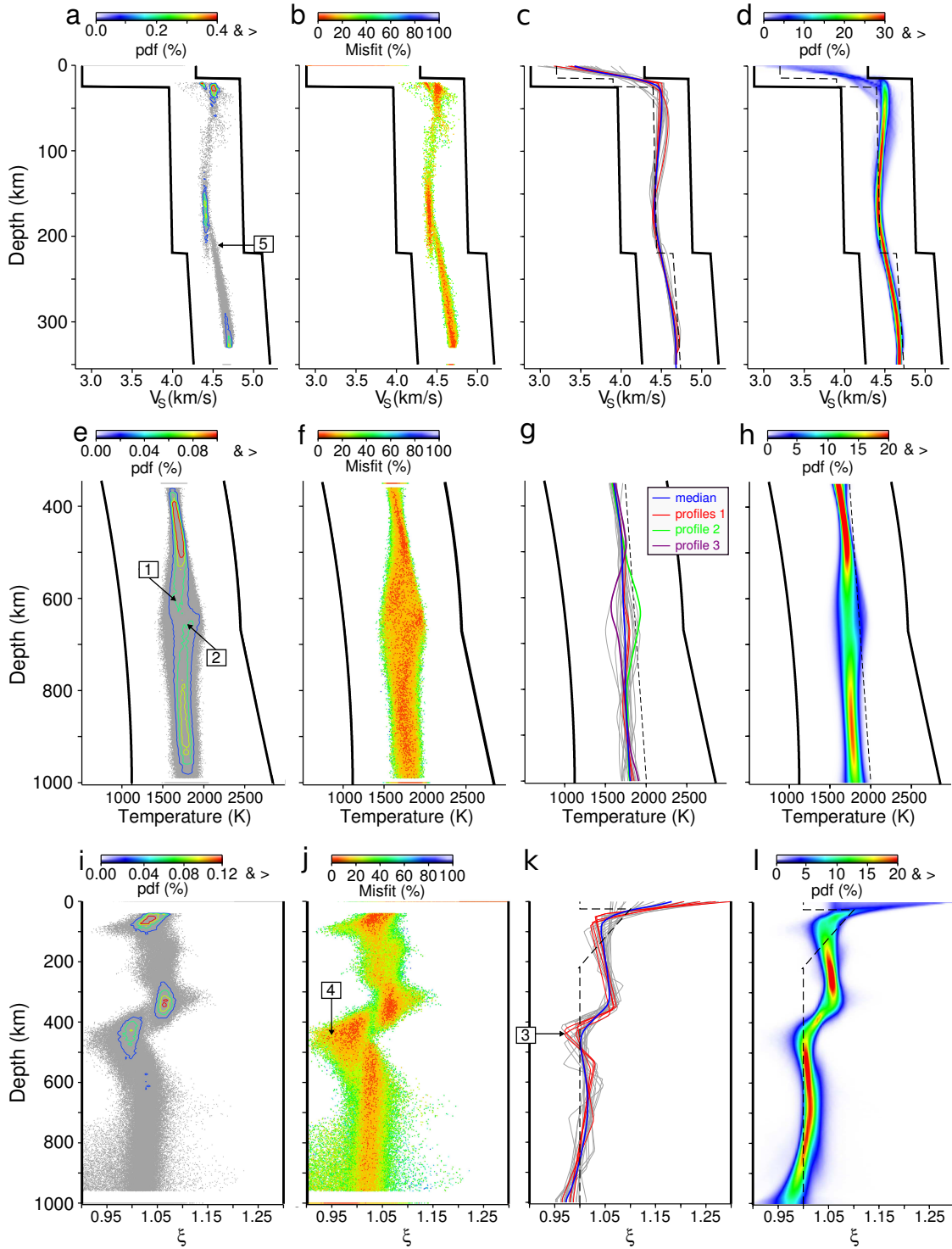


FIG. B.8 : Same as Fig. B.4 but for the Vanuatu-California data. Dashed lines in (c), (d), (k), and (l) are PREM values. Dashed lines in (g) and (h) show the 1600 K adiabat considering a temperature gradient of 0.4 K/km. Blue lines in (c), (g) and (k) are the median profiles of the distributions.

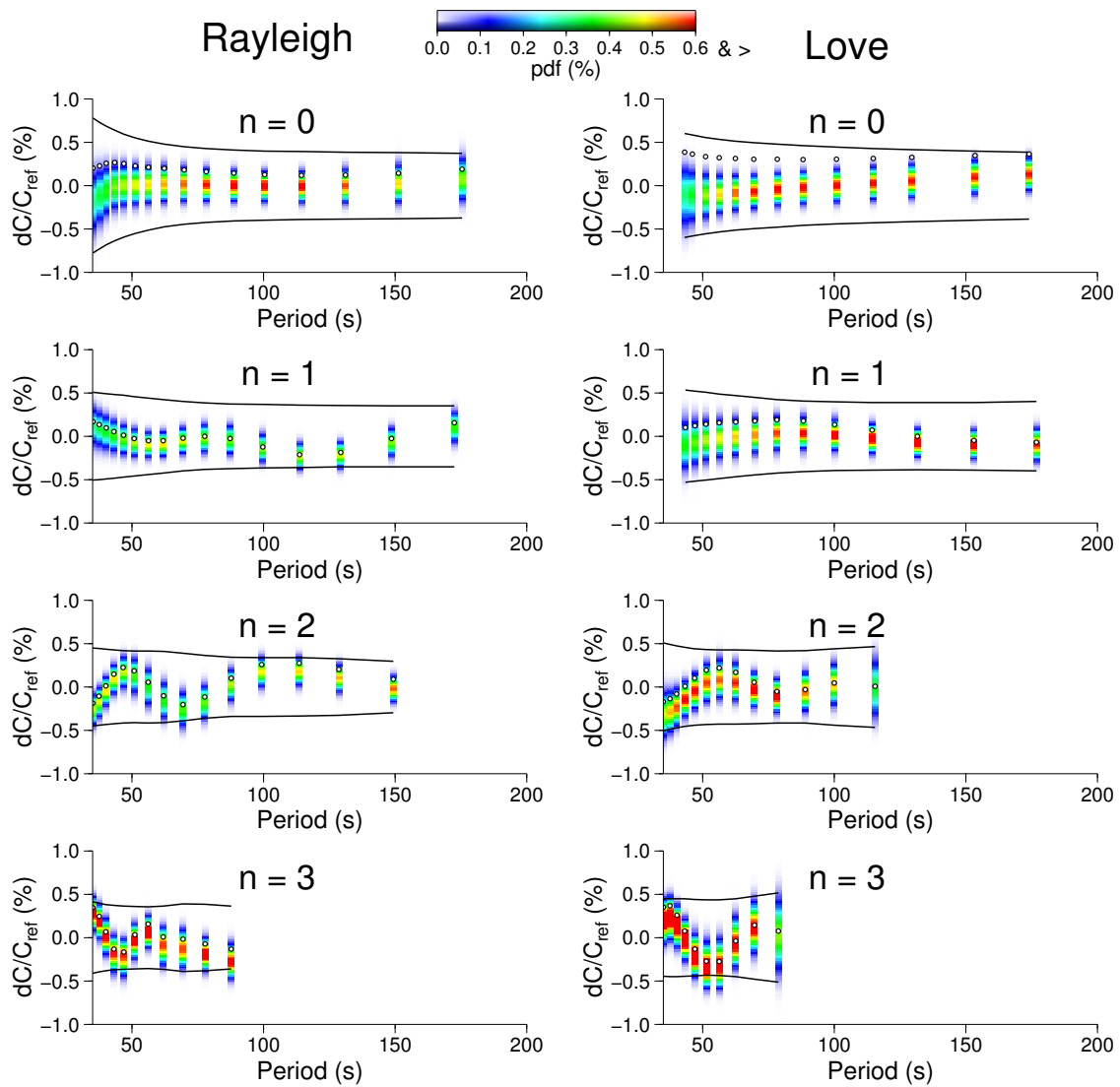


FIG. B.9 : Same as Fig. B.7 but for the Vanuatu-California data. White dots are phase velocities of the median profiles plotted in Fig. B.8.

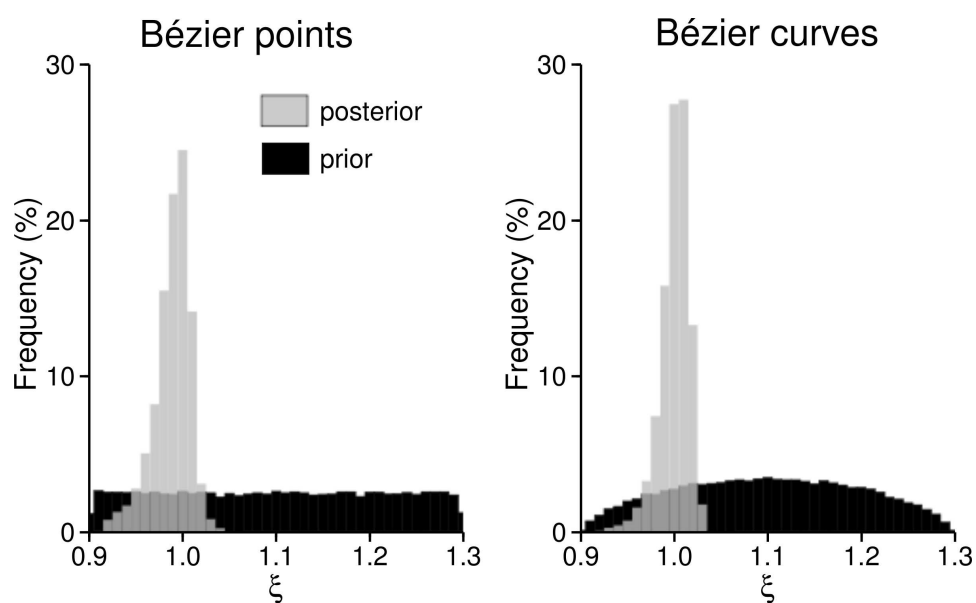


FIG. B.10 : Marginal prior (in black) and posterior (in gray) distributions of Bézier points and Bézier curves for shear wave anisotropy at 450 km depth. It concerns the Vanuatu-California path. Intervals are 0.01 wide.

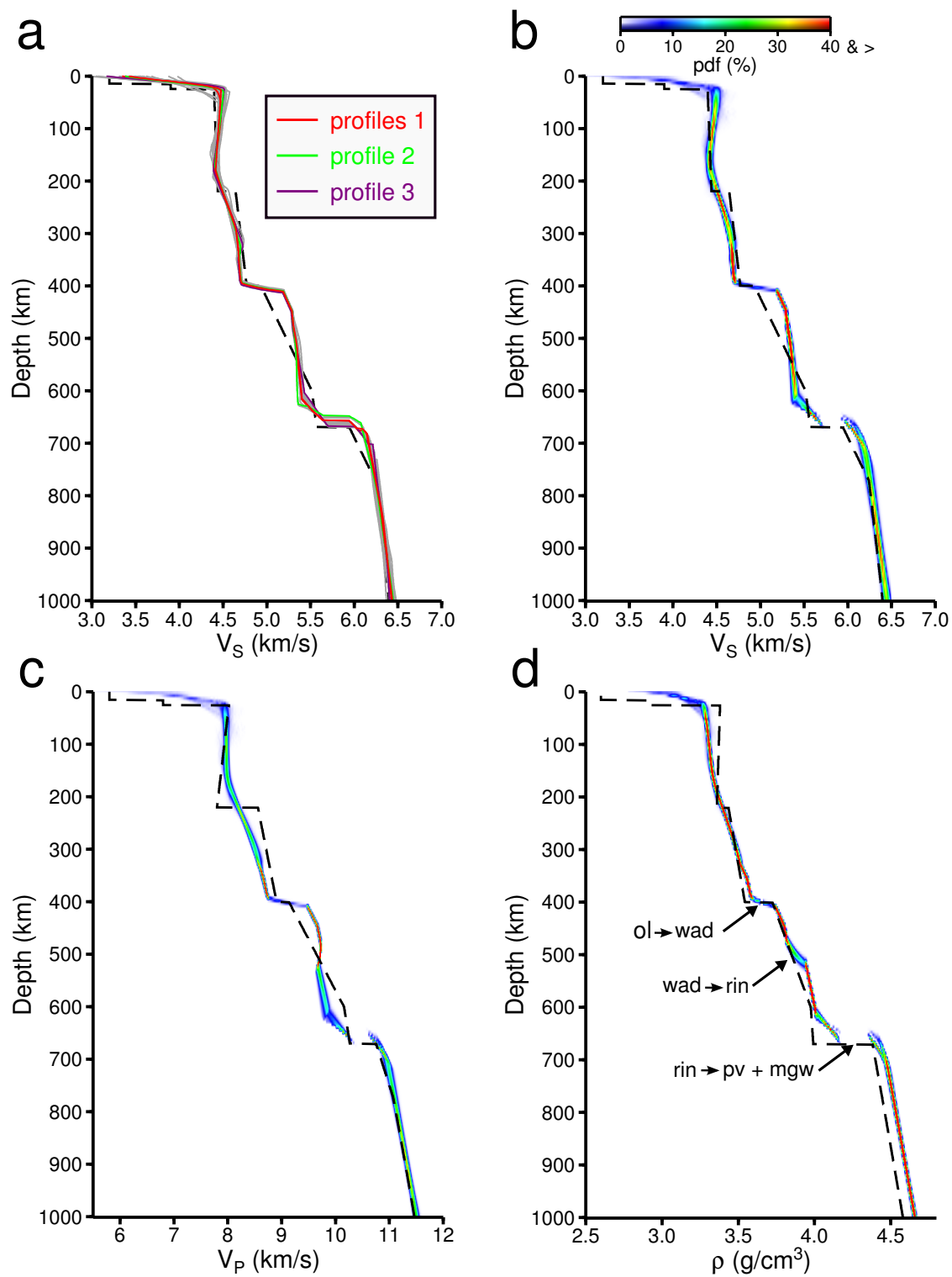


FIG. B.11 : Probability maps of (b) V_S , (c) V_P and (d) ρ as a function of depth for the Vanuatu-California data. The abbreviations in (d) stand for ol : olivine, wad : wadsleyite, rin : ringwoodite, pv : perovskite, mgw : magnesiowustite. (a) shows in gray a random subset of 20 models in the ensemble solution and in color the four models corresponding to the best misfits. Dashed lines are PREM values.

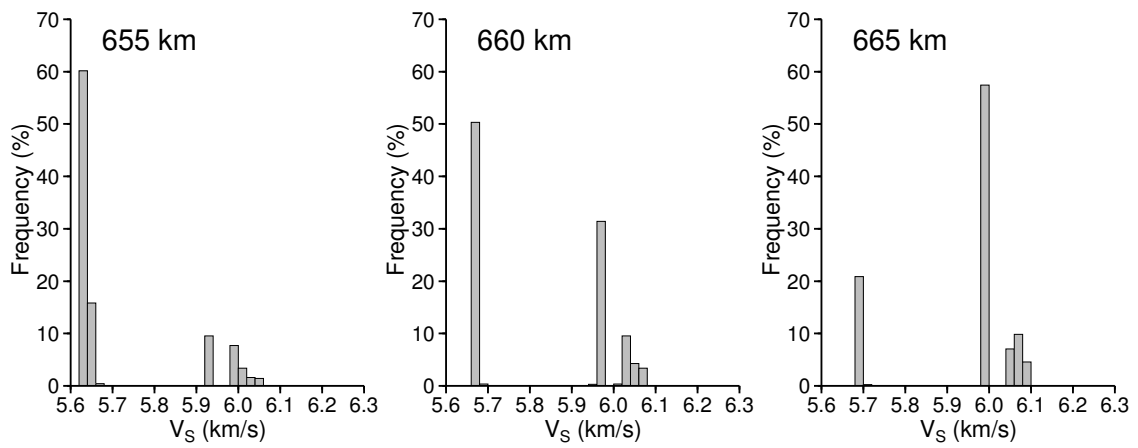


FIG. B.12 : Marginal posterior distributions for V_S at 655, 660 and 665 km depth (*i.e.*, around the transformation of ringwoodite into perovskite and magnesiowustite). Intervals are 0.02 km/s wide. The distribution is clearly influenced by both V_S values taken above and under the discontinuity.

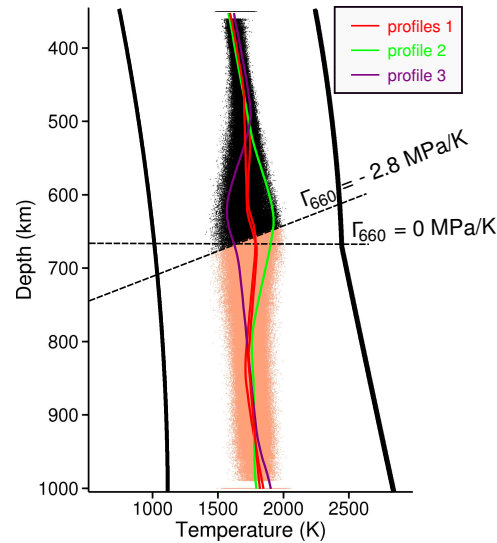


FIG. B.13 : Distribution of Bézier points for temperature. When the thermodynamic conditions favor the olivine polymorphs, the points are plotted in black, and when perovskite and magnesiowustite are created, the points are shown in pink. Γ_{660} refers to the Clapeyron slope value.

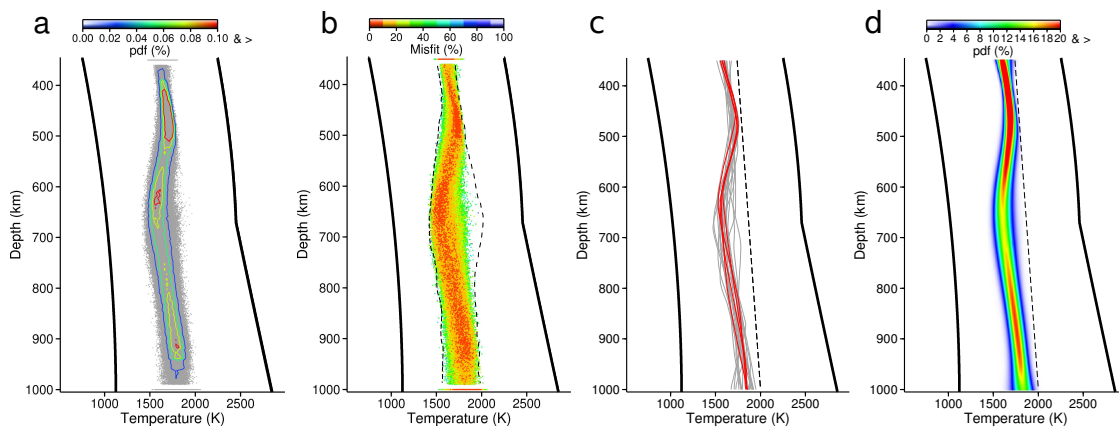


FIG. B.14 : Same as Fig. B.8 but for the Vanuatu-California data, considering a Clapeyron slope equal to zero for the ringwoodite to perovskite and magnesiowustite transformation. Dashed lines in (b) are the contours of the Bézier points distribution displayed in Fig. B.8e and f.

	T	ξ	V_S
number of Bézier points	6 to 10	7	5
depth interval (km)	350- 1000	0- 1000	0- 350
minimal distance between 2 consecutive points (km)	80 to 10	40	20
norm of the local tangent vector (km)	40 to 5	20	10
maximal gradient between 2 consecutive points	5 K/km		

TAB. B.1 : Synthesis of the parameters constraining the localisation of Bézier points.

number of Bézier points	6	7	8	9	10
model 1 (6 points)	18.75	17.50	23.75	27.15	12.50
model 2 (7 points)	22.50	20.00	18.75	18.75	20.00
Vanuatu-California	22.50	16.25	25.00	18.75	17.50

TAB. B.2 : Number of Bézier points selected for the *stationary period* (%).

Références bibliographiques

- Adams, R. D., 1971. Reflections from discontinuities beneath Antarctica, *Bull. Seismol. Soc. Am.*, **61**, 1441–1451.
- Akaogi, M. & Ito, E., 1993. Refinement of enthalpy measurements of MgSiO_3 perovskite and negative pressure-temperature slopes for perovskite-forming reactions, *Geophys. Res. Lett.*, **20**, 1839–1842.
- Akaogi, M., Ito, E., & Navrotsky, A., 1989. Olivine-modified spinel-spinel transitions in the system $\text{Mg}_2\text{SiO}_4\text{-Fe}_2\text{SiO}_4$: calorimetric measurements, thermochemical calculation, and geophysical application, *J. Geophys. Res.*, **94**, 15671–15685.
- Akaogi, M., Takayama, H., Kijitani, H., Kawaji, H., & Atake, T., 2007. Low-temperature heat capacities, entropies and enthalpies of Mg_2SiO_4 polymorphs, and α - β - γ and post-spinel phase relations at high pressure, *Phys. Chem. Minerals*, **34**, 169–198.
- Aki, K. & Richards, P. G., 1980. *Quantitative seismology, theory and methods*, vol 1, Freeman, San-Francisco, 557p.
- Anderson, D. L., 1961. Elastic wave propagation in layered anisotropic media, *J. Geophys. Res.*, **66**, 2953–2963.
- Anderson, D. L., 1979. Evidence supporting the approximation $\gamma\rho = \text{constant}$ for the grüneisen parameter of the Earth lower mantle, *J. Geophys. Res.*, **84**, 3537–3542.
- Anderson, D. L., 1988. Temperature and pressure derivatives of elastic constants with applications to the mantle, *J. Geophys. Res.*, **93**, 4688–4700.
- Andrieu, C. & Moulines, E., 2005. On the ergodicity properties of some adaptative MCMC algorithms, *Ann. Appl. Probab.*, **16**, 1462–1505.
- Babuska, V. & Cara, M., 1991. *Seismic anisotropy in the Earth*, Kluwer academic publishers, Boston, 271p.
- Backus, G. E. & Gilbert, J. F., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. Geophys. Res.*, **67**, 4427–4440.

- Backus, G. E. & Gilbert, J. F., 1968. The resolving power of gross Earth data, *Geophys. J. R. Astron. Soc.*, **16**, 169–205.
- Baker, M. B. & Beckett, J. R., 1999. The origin of abyssal peridotites : A reinterpretation of constraints based on primary bulk compositions, *Earth Planet. Sci. Lett.*, **171**, 49–61.
- Bass, J. D. & Anderson, D. L., 1984. Composition for the upper mantle : geophysical tests of two petrological models, *Geophys. Res. Lett.*, **11**, 237–240.
- Bassin, C., Laske, G., & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *EOS Trans. AGU*, **81**, F897.
- Becker, T. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.*, **3**, doi :10.1029/2001GC000168.
- Becker, T. W., Kellogg, J., Ekström, G., & O'Connell, R., 2003. Comparison of azimuthal seismic anisotropy from surface waves and finite strain from global mantle-circulation models, *Geophys. J. Int.*, **155**, 696–714.
- Beghein, C., 2010. Radial anisotropy and prior petrological constraints : A comparative study, *J. Geophys. Res.*, **115**, B03303, doi :10.1029/2008JB005842.
- Beghein, C., Resovsky, J., & van der Hilst, R. D., 2008. The signal of mantle anisotropy in the coupling of normal modes, *Geophys. J. Int.*, **175**, 1209–1234, doi : 10.1111/j.1365-246X.2008.03970.x.
- Bercovici, D. & Karato, S., 2003. Whole-mantle convection and the transition-zone water filter, *Nature*, **425**, 39–44.
- Besag, J., Green, P., Higdon, D., & Mengersen, K., 1995. Bayesian computation and stochastic systems (with discussion), *Statistical Science*, **10**, 3–66.
- Beucler, É., Stutzmann, É., & Montagner, J.-P., 2003. Surface-wave higher mode phase velocity measurements using a roller coaster type algorithm, *Geophys. J. Int.*, **155**, 289–307.
- Bézier, P., 1966. Définition numérique des courbes et surfaces I, *Automatisme*, pp. 625–632.
- Bézier, P., 1967. Définition numérique des courbes et surfaces II, *Automatisme*, pp. 625–632.
- Bijwaard, H., Spakman, W., & Engdahl, E. R., 1998. Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, **103**, 30055–30078.

- Bina, C. R. & Helffrich, G., 1994. Phase transition clapeyron slopes and transition zone seismic discontinuity topography, *J. Geophys. Res.*, **99**, 15853–15860.
- Birch, F., 1952. Elasticity and constitution of the Earth's interior, *J. Geophys. Res.*, **57**, 227–286.
- Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *J. Geophys. J. Int.*, **178**, 1411–1436.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., & Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface wave dispersion, *J. Geophys. Res.*, **117**, B02301, doi :10.1029/2011JB008560.
- Boschi, L. & Dziewonski, A. M., 2000. Whole Earth tomography from delay times of P, PcP, PKP phases : Lateral heterogeneities in the outer core, or radial anisotropy in the mantle ?, *J. Geophys. Res.*, **105**, 13675–13696, doi :10.1029/2000JB900059.
- Boschi, L. & Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. Geophys. Res.*, **107**, doi :10.1029/2000JB000059.
- Bozdag, E. & Trampert, J., 2007. On crustal corrections in surface wave tomography, *Geophys. J. Int.*, **172**, 1076–1088.
- Brooks, S. P. & Roberts, G. O., 1998. Diagnosing convergence of Markov chain Monte Carlo algorithms, *Statistics and Computing*, **8**, 319–335.
- Cammarano, F. & Romanowicz, B., 2007. Insights into the nature of the transition zone from physically constrained inversion of long-period seismic data, *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 9139–9144.
- Cammarano, F., Goes, S., Vacher, P., & Giardini, D., 2003. Inferring upper mantle temperatures from seismic velocities, *Phys. Earth Planet. Int.*, **138**, 197–222.
- Cammarano, F., Deuss, A., Goes, S., & Giardini, D., 2005a. One dimensional physical reference models for the upper mantle and transition zone : combining seismic and mineral physics constraints, *J. Geophys. Res.*, **110**, doi :10.1029/2004JB003272.
- Cammarano, F., Goes, S., & Giardini, D., 2005b. Is a pyrolitic mantle compatible with seismic data ?, *Earth Planet. Sci. Lett.*, **232**, 227–243.
- Cammarano, F., Romanowicz, B., Stixrude, L., Lithgow-Bertelloni, C., & Xu, W., 2009. Inferring the thermochemical structure of the upper mantle from seismic data, *Geophys. J. Int.*, **179**, 1169–1185.

- Capdeville, Y., Chaljub, E., Vilotte, J.-P., & Montagner, J.-P., 2003. Coupling the spectral element method with a modal solution for elastic wave propagation in global Earth models, *Geophys. J. Int.*, **152**, 34–67.
- Cara, M., 1979. Lateral variations of *S*-velocity in the upper mantle from higher Rayleigh modes, *Geophys. J. R. Astr. Soc.*, **57**, 649–670.
- Castillo, J., 2001. *Contribution à l'étude de la structure interne des planètes telluriques*, Ph.D. thesis, Laboratoire de Planétologie et de Géodynamique de Nantes.
- Chambers, K., Woodhouse, J. H., & Deuss, A., 2005. Topography of the 410-km discontinuity from *PP* and *SS* precursors, *Earth Planet. Sci. Lett.*, **235**, 610–622.
- Chen, W. & Brudzinski, M. R., 2003. Seismic anisotropy in the mantle transition zone beneath Fiji-Tonga, *Geophys. Res. Lett.*, **1682**, 30(13), doi :10.1029/2002GL016330.
- Christensen, U. R. & Hofmann, A. M., 1994. Segregation of subducted oceanic crust in the convecting mantle, *J. Geophys. Res.*, **99**, doi/10.1029/93JB03403.
- Cobden, L., Goes, S., Cammarano, F., & Connolly, A., 2008. Thermochemical interpretation of one-dimensional seismic reference models for the upper mantle : evidence for bias due to heterogeneity, *Geophys. J. Int.*, **175**, 627–648.
- Collier, J., Hellfrich, G., & Wood, B. J., 2001. Seismic discontinuities and subduction zones, *Phys. Earth planet. Int.*, **127**, 35–49.
- Collins, M. D. & Brown, J. M., 1998. Elasticity of a upper mantle clinopyroxene, *Phys. Chem. Min.*, **26**, 7–13.
- Cowles, M. & Carlin, B., 1995. Markov chain Monte Carlo convergence diagnostics : A comparative review, *Journal of the American Statistical Association*, **91**, 883–904.
- Dahlen, F. A. & Tromp, J., 1998. *Theoretical global seismology*, Princeton Univ. Press, Princeton, New Jersey, 1025p.
- Davies, G. F. & Dziewonski, A. M., 1975. Homogeneity and constitution of the Earth's lower mantle and outer core, *Phys. Earth and Planet. Int.*, **10**, 336–343.
- Demengel, G. & Pouget, J.-P., 1998. *Modèles de Bézier, des B-Splines et des NURBS*, ellipses, Paris, 288p.
- Deng, L. & Scales, J., 1999. Estimating the topography of multi-dimensional fitness functions, *Cent. for Wave Phenomena Tech. Rep.*, **208**.

- Deschamps, F. & Tackley, P., 2009. Searching for models of thermo-chemical convection that explain probabilistic tomography. II-Influence of physical and compositional parameters, *Phys. Earth Planet. Int.*, **176**, 1–18.
- Deschamps, F. & Trampert, J., 2003. Mantle tomography and its relation to temperature and composition, *Phys. Earth Planet. Int.*, **140**, 277–291.
- Deschamps, F., Snieder, R., & Trampert, J., 2001. The relative density-to-shear velocity scaling in the uppermost mantle, *Phys. Earth Planet. Int.*, **124**, 193–211.
- Deuss, A., 2009. Global observations of mantle discontinuities using *SS* and *PP* precursors, *Surveys in Geophysics*, **30**, 301–326.
- Deuss, A. & Woodhouse, J. H., 2001. Seismic observations of splitting of the mid-transition zone discontinuity in Earth’s mantle, *Science*, **294**, 354–357.
- Deuss, A. & Woodhouse, J. H., 2004. The nature of the Lehmann discontinuity from its seismological Clapeyron slopes, *Earth Planet. Sci. Lett.*, **225**, 295–304.
- Duffy, T. S. & Anderson, D. L., 1989. Seismic velocities in mantle minerals and the mineralogy of the Earth as determined from seismic profiles, *J. Geophys. Res.*, **94**, 1895–1912.
- Duffy, T. S., Zha, C., Downs, R. T., Mao, H. K., & Hemley, R. J., 1995. Elastic constants of forsterite Mg_2SiO_4 to 16 GPa, *Nature*, **378**, 170–173.
- Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary Reference Earth Model, *Phys. Earth Planet. Int.*, **25**, 297–356.
- Ekström, G. & Dziewonski, A. M., 1998. The unique anisotropy of the Pacific upper mantle, *Nature*, **394**, 168–172.
- Engdahl, E. R. & Flinn, E. A., 1969. Seismic waves reflected from discontinuities within the Earth’s upper mantle, *Science*, **163**, 177–179.
- Farber, D. L., Williams, Q., & Ryerson, F. J., 1994. Diffusion in Mg_2SiO_4 polymorphs and chemical heterogeneity in the mantle transition zone, *Nature*, **371**, 693–695.
- Fei, Y., van Orman, J., Li, J., van Westrenen, W., Sanloup, C., Minarik, W., Hirose, K., Komabayashi, T., Walter, M., & Funakoshi, K., 2004. Experimentally determined postspinel transformation boundary in Mg_2SiO_4 using MgO as an internal pressure standard and its geophysical implications, *J. Geophys. Res.*, **10**, B02305, doi :10.1029/2003JB002562.

- Ferreira, A., Woodhouse, J., Visser, K., & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography, *J. Geophys. Res.*, **115**, B04313, doi :10.1029/2009JB006716.
- Fichtner, A. & Trampert, J., 2011. Resolution analysis in full waveform inversion, *Geophys. J. Int.*, **187**, 1604–1624.
- Fichtner, A., Kennett, B., Igel, H., & Bunge, H., 2009. Full seismic waveform tomography for upper-mantle structure in the Australian region using adjoint methods, *Geophys. J. Int.*, **179**, 1703–1725.
- Fischer, K. M. & Wiens, D. A., 1996. The depth distribution of mantle anisotropy beneath the Tonga subduction zone, *Earth Planet. Sci. Lett.*, **142**, 253–260.
- Flanagan, M. P. & Shearer, P. M., 1998. Global mapping of topography on transition zone velocity discontinuities by stacking *SS* precursors, *J. Geophys. Res.*, **103**, 2673–2692.
- Flanagan, M. P. & Shearer, P. M., 1999. A map of topography on the 410-km discontinuity from *PP* precursors, *Geophys. Res. Lett.*, **26**, 549–552.
- Flesch, L. M., Li, B., & Liebermann, R. C., 1998. Sound velocities of polycrystalline MgSiO_3 orthopyroxene to 10 GPa at room temperature, *Am. Mineral.*, **83**, 444–450.
- Fouch, M. J. & Fisher, K. M., 1996. Mantle anisotropy beneath north-west Pacific subduction zones, *J. Geophys. Res.*, **101**, 15987–16002.
- Friederich, W. & Dalkolmo, J., 1995. Complete synthetic seismograms for a spherically symmetric Earth by a numerical computation of the Green's function in the frequency domain, *Geophys. J. Int.*, **122**, 537–550.
- Frisillo, A. L. & Barsch, G. R., 1972. Measurement of single-crystal elastic constants of bronzonite as a function of pressure and temperature, *J. Geophys. Res.*, **77**, 6360–6384.
- Fukao, Y., Obayashi, M., Inoue, H., & Nenbai, M., 1992. Subducting slab stagnant in the mantle transition zone, *J. Geophys. Res.*, **97**, 4809–4822.
- Fukao, Y., Widiyantoro, S., & Obayashi, M., 2001. Stagnant slabs in the upper and lower mantle transition region, *Rev. of Geophysics*, **39**, 291–323.
- Gaboret, C., Forte, A. M., & Montagner, J., 2003. The unique dynamics of the Pacific hemisphere mantle and its signature on seismic anisotropy, *Earth Planet. Sci. Lett.*, **208**, 219–233.

- Ganguly, J., Freed, A. M., & Saxena, S. K., 2009. Density profiles of oceanic slabs and surrounding mantle : Integrated thermodynamic and thermal modeling, and implications for the fate of slabs at the 660 km discontinuity, *Phys. Earth Planet. Int.*, **172**, 257–267.
- Gelman, A., Richardson, S., & D., S., 1996. *Markov Chain Monte Carlo in Practise*, Chapman & Hall/CRC, U. K., 486p.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B., 2003. *Bayesian Data Analysis*, Chapman & Hall/CRC, U. K., 696p.
- Gilbert, F. & Dziewonski, A. M., 1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Philos. Trans. R. Soc. London A.*, **278**, 187–269.
- Goes, S., Govers, R., & Vacher, P., 2000. Shallow mantle temperatures under Europe from *P* and *S* wave tomography, *J. Geophys. Res.*, **105**, 11153–11169.
- Gossler, J. & Kind, R., 1996. Seismic evidence for very deep roots of continents, *Earth Planet. Sci. Lett.*, **138**, 1–13.
- Grand, S. P., 1994. Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.*, **99**, 11591–11621.
- Green, P., 1995. Reversible jump MCMC computation and bayesian model selection, *Biometrika*, **88**, 711–732.
- Green, P. J. & Mira, A., 2001. Delayed rejection in reversible jump Metropolis-Hastings, *Biometrika*, **88**, 1035–1053.
- Gu, Y., Dziewonski, A. M., & Agee, C. B., 1998. Global de-correlation of the topography of transition zone discontinuities, *Earth Planet. Sci. Lett.*, **157**, 57–67.
- Gu, Y., Dziewonski, A. M., & Ekström, G., 2003. Simultaneous inversion for mantle shear wave velocity and topography of transition zone discontinuities, *Geophys. J. Int.*, **154**, 559–583.
- Gu, Y., Lerner-Lam, A., Dziewonski, A. M., & Ekström, G., 2005. Deep structure and seismic anisotropy beneath the East Pacific Rise, *Earth Planet. Sci. Lett.*, **232**, 259–272.
- Gung, Y., Panning, M., & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature*, **422**, 707–711.
- Hashin, Z. & Shtrikman, S., 1962. On some variational principles in anisotropic and nonhomogeneous elasticity, *J. Mech. Phys. Solids*, **10**, 335–342.

- Hashin, Z. & Shtrikman, S., 1963. A variational approach to the theory of the elastic behaviour of multiphase materials, *J. Mech. Phys. Solids*, **11**, 12–140.
- Hastings, W. K., 1970. Monte carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, 97–109.
- Hauser, J., Dyer, K. M., Pasyanos, M. E., Bungum, H., Faleide, J. I., Clark, A., & Schweitzer, J., 2011. A probabilistic seismic model for the European Arctic, *J. Geophys. Res.*, **116**, B01303, doi :10.1029/2010JB007889.
- Hedlin, M. A. H., Shearer, P. M., & Earle, P. S., 1997. Seismic evidence for small-scale heterogeneity throughout the Earth's mantle, *Nature*, **387**, 145–150.
- Helffrich, G. R., 2006. Heterogeneity in the mantle : Its creation, evolution and destruction, *Tectonophysics*, **416**, 23, doi :10.1038/35087500.
- Helffrich, G. R. & Wood, B. J., 2001. The Earth's mantle, *Nature*, **412**, doi :10.1038/35087500.
- Hill, R., 1952. Elastic properties of reinforced solids : some theoretical principles, *Proc. of the Phys. Soc. of London Series*, **A65**, 349–354.
- Hirose, K., 2002. Phase transitions in pyrolitic mantle around 670-km depth : implications for upwelling plumes from the lower mantle, *J. Geophys. Res.*, **107**, doi :10.1029/2001JB000597.
- Hirshmann, M., 2000. Mantle solidus : Experimental constraints and the effects of peridotite composition, *Geochem. Geophys. Geosyst.*, **1**, GC000070.
- Hirshmann, M., Tenner, T., Aubaud, C., & Withers, A., 2009. Dehydration melting of nominally anhydrous mantle : The primacy of partitioning, *Phys. Earth Planet. Int.*, **176**, 54–68.
- Hofmann, A. W., 1997. Mantle geochemistry : The message from oceanic volcanism, *Nature*, **385**, doi :10.1038/385219a0.
- Houser, C. & Williams, Q., 2010. Reconciling Pacific 410 and 660 km discontinuity topography, transition zone shear velocity patterns, and mantle phase transitions, *Earth Planet. Sci. Lett.*, **296**, 255–266.
- Houser, C., Masters, G., Flanagan, M., & Shearer, P., 2008. Determination and analysis of long-wavelength transition zone structure using *SS* precursors, *Geophys. J. Int.*, **174**, 178–194.

- Ibrahim, A. K. & Nuttli, O. W., 1967. Travel-time curves and upper mantle structure from long-period *S* waves, *Bull. Seism. Soc. Am.*, **57**, 1063–1092.
- Inoue, E., Yurimoto, H., & Kudoh, Y., 1995. Hydrous modified spinel $\text{Mg}_{1.5}\text{SiH}_{0.5}\text{O}_4$: a new water reservoir in the mantle transition region, *Geophys. Res. Lett.*, **22**(2), 117–120.
- Irifune, T., 1987. An experimental investigation of the pyroxene-garnet transformation in a pyrolite composition and its bearing on the constitution of the mantle, *Earth. Planet. Sci. Lett.*, **45**, 324–336.
- Irifune, T., Koizumi, T., & Ando, J. I., 1996. An experimental study of the garnet-perovskite transformation in the system MgSiO_3 - $\text{Mg}_3\text{Al}_2\text{Si}_3\text{O}_{12}$, *Phys. Earth Planet. Int.*, **96**, 147–157.
- Irifune, T., Nishiyama, N., Kuroda, K., Inoue, T., Isshiki, M., Utsumi, W., Funakoshi, K., Urakawa, S., Uchida, T., Katsura, T., & Ohtaka, O., 1998. The postspinel phase boundary in Mg_2SiO_4 determined by in situ X-ray diffraction, *Science*, **279**, 1698–1700.
- Irifune, T., Higo, Y., Inoue, T., Kono, Y., Ohfuji, H., & Funakoshi, K., 2008. Sound velocities of majorite and the composition of the mantle transition region, *Nature*, **451**, 814–817.
- Isaak, D. G., 1992. High-temperature elasticity of iron-bearing olivines, *J. Geophys. Res.*, **97**, 1871–1885.
- Isaak, D. G., Anderson, O., Goto, T., & Suzuki, I., 1989. Elasticity of single crystal forsterite measured to 1700 K, *J. Geophys. Res.*, **97**, 5895–5906.
- Ita, J. & Stixrude, L., 1992. Petrology, elasticity and composition of the mantle transition zone, *J. Geophys. Res.*, **97**, 6849–6866.
- Ito, E., 1977. The absence of oxide mixture in high-pressure phases of Mg-silicates, *Geophys. Res. Lett.*, **4**, 72–74.
- Ito, E. & Takahashi, E., 1989. Postspinel transformations in the system Mg_2SiO_4 - Fe_2SiO_4 and some geophysical implications, *J. Geophys. Res.*, **94**, 10637–10646.
- Jackson, I., 1998. Elasticity, composition and temperature of the Earth's lower mantle : A reappraisal, *Geophys. J. Int.*, **134**, 291–311.
- Jackson, I., 2008. Properties of rocks and minerals - physical origins of anelasticity and attenuation in rock, in *Treatise on Geophysics : Mineral Physics*, pp. 493–525, ed. Schubert, G., Elsevier, Amsterdam.

- Jacobsen, S. D. & Smyth, J. R., 2006. Effect of water on the sound velocities of ringwoodite in the transition zone, in *Earth's Deep Water Cycle*, pp. 131–145, eds Jacobsen, S. D. & Lee, S. v. d., American Geophysical Union, Washington DC.
- Jacobsen, S. D., Jiang, F., Mao, Z., Duffy, T. S., Smyth, J. R., Holl, C. M., & Frost, D. J., 2008. Effects of hydration on the elastic properties of olivine, *Geophys. Res. Lett.*, **35**, L14303, doi :10.1029/2008GL034398.
- Jasbinsek, J. & Dueker, K. G., 2007. Ubiquitous low velocity-layer atop the 410-km discontinuity in the northern Rocky Mountains, *Geochem. Geophys. Geosyst.*, **8**, Q10004, doi :10.1029/2007GC001661.
- Jeanloz, R. & Thompson, A. B., 1983. Phase transitions and mantle discontinuities, *Rev. Geophys. Space Phys.*, **21**, 51–74.
- Julian, B. R. & Anderson, D. L., 1968. Travel times, apparent velocities and amplitudes of body waves, *Bull. Seism. Soc. Am.*, **58**, 339–366.
- Jung, U. & Karato, S., 2001. Water-induced fabric transitions in olivine, *Science*, **293**(5534), 1460–1463.
- Káráson, H. & van der Hilst, R. D., 2000. Constraints on mantle convection from seismic tomography, in *History and Dynamics of Plate Motion*, pp. 277–288, eds Richards, M. A., Gordon, R., & van der Hilst, R. D., American Geophysical Union, Washington DC.
- Karato, S., 1993. Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, **20**, 1623–1626.
- Karato, S., 2008. *Deformation of Earth Materials : Introduction to the Rheology of the Solid Earth*, Cambridge University Press, 463p.
- Karato, S., 2011. Water distribution across the mantle transition zone and its implications for global material circulation, *Earth Planet. Sci. Lett.*, **301**, 413–423.
- Karato, S. I., Forte, A. M., Liebermann, R. C., Masters, G., & Stixtude, L., 2000. *Earth's Deep Interior : Mineral Physics and Tomography From the Atomic to the Global Scale*, American Geophysical Union, Washington D. C., 289p.
- Katayama, I., Hirauchi, K., Michibayashi, K., & Ando, J., 2009. Trench-parallel anisotropy produced by serpentine deformation in the hydrated mantle wedge, *Nature*, **461**, 1114–1117.

- Katsura, T. & Ito, E., 1989. The system $\text{Mg}_2\text{SiO}_4\text{-Fe}_2\text{SiO}_4$ at high pressures and temperatures : precise determination of stabilities of olivine, modified spinel, and spinel, *J. Geophys. Res.*, **94**, 15663–15670.
- Katsura, T., Yamada, H., Shinmei, T., Kubo, A., Ono, S., Kanzaki, M., Yoneda, A., Walter, M. J., Ito, E., Urakawa, S., Funakoshi, K., & Utsumi, W., 2003. Post-spinel transition in the Mg_2SiO_4 determined by high P-T in situ X-ray diffraction, *Phys. Earth Planet. Int.*, **136**, 11–24.
- Katsura, T., Yamada, H., Nishikawa, O., Song, M., Kubo, A., Shinmei, T., Yoshino, S., Aizawa, Y., Yoshino, T., Walter, M., Ito, E., & Funakoshi, K., 2004. Olivine-Wadsleyite transition in the system $(\text{Mg,Fe})_2\text{SiO}_4$, *J. Geophys. Res.*, **109**, doi :10.1029/2003JB002438.
- Katsura, T., Yoneda, A., Yamazaki, D., Yoshino, T., & Ito, E., 2010. Adiabatic temperature profile in the mantle, *Phys. Earth Planet. Int.*, **183**, 212–218.
- Kawakatsu, H. & Yoshioka, S., 2011. Metastable olivine wedge and deep dry cold slab beneath southwest Japan, *Earth Planet. Sci. Lett.*, **303**, 1–10.
- Kellogg, J. B., Jacobsen, S. B., & O’Connell, R. J., 2002. Modeling the distribution of isotopic ratios in geochemical reservoirs, *Earth Planet. Sci. Lett.*, **204**, 183–202.
- Kennett, B. L. N., Engdahl, E. R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, **122**, 108–124.
- Khan, A. & Mosegaard, K., 2002. An inquiry into the lunar interior : A non-linear inversion of the Appolo lunar seismic data, *J. Geophys. Res.*, **107**, 5036, doi :10.1029/2001JE001658.
- Khan, A. & Shankland, T. J., 2012. A geophysical perspective on mantle water content and melting : Inverting electromagnetic sounding data using laboratory-based electrical conductivity profiles, *Earth Planet. Sci. Lett.*, **317**, 27–43.
- Khan, A., Connolly, A. D., & Olsen, A. N., 2006. Constraining the composition and thermal state of the mantle beneath Europe from inversion of long-period electromagnetic sounding data, *J. Geophys. Res.*, **111**, B10102, doi :10.1029/2006JB004270.
- Khan, A., Boschi, L., & Connolly, A. D., 2009. On mantle chemical and thermal heterogeneities and anisotropy as mapped by inversion of global surface wave data, *J. Geophys. Res.*, **114**, B09305, doi :10.1029/2009JB006399.

- Khan, A., Boschi, L., & Connolly, J. A. D., 2011. Mapping the Earth's thermochemical and anisotropic structure using global surface wave data, *J. Geophys. Res.*, **116**, doi :10.1029/2010JB007828.
- Kiefer, B., Stixrude, L., & Wentzcovitch, R., 1997. Calculated elastic constants and anisotropy of Mg_2SiO_4 spinel at high pressure, *Geophys. Res. Lett.*, **24**, 2841–2844.
- Kohlstedt, D. L., Keppler, H., & Rubie, D. C., 1996. Solubility of water in the α - β - γ phases of $(\text{Mg}, \text{Fe})_2\text{SiO}_4$, *Contrib. Mineral. Petrol.*, **1223**, 345–357.
- Kustowski, B., Ekström, G., & Dziewonski, A. M., 2008. Anisotropic shear wave velocity structure of the Earth's mantle : A global model, *J. Geophys. Res.*, **113**, B06306, doi :10.1029/2007JB005169.
- Lawrence, J. F. & Shearer, P. M., 2008. Imaging mantle transition zone thickness with *SdS-SS* finite-frequency sensitivity kernels, *Geophys. J. Int.*, **174**, 143–158.
- Lay, T. & Wallace, T. C., 1995. *Modern global seismology*, Academic press, San Diego, 521p.
- Li, C., van der Hilst, R. D., Engdahl, E. R., & Burdick, S., 2008. A new global model for *P* wave speed variations in Earth's mantle, *Geochem. Geophys. Geosys.*, **9**, Q05018, doi :10.1029/2007GC001806.
- Li, X. D. & Romanowicz, B., 1996. Global mantle shear-velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res.*, **101**, 22245–22272.
- Litasov, K., Ohtani, E., Sano, A., Suzuki, A., & Funakoshi, K., 2005. Wet subduction versus cold subduction, *Geophys. Res. Lett.*, **32**, L13312, doi :10.1029/2005GL022921.
- Liu, L., 1977. The system enstatite-pyroxene at high pressures and temperatures and the mineralogy of the Earth's mantle, *Earth Planet. Sci. Lett.*, **36**, 237–245.
- Liu, Q. & Gu, Y., 2012. Seismic imaging : from classical to adjoint tomography, *Tectonophysics*, **566**, 31–66.
- Love, A. E. H., 1927. *A treatise on the mathematical theory of elasticity*, Cambridge University Press, New York, 643p.
- Lukas, M. A., 2008. Strong robust generalized cross-validation for choosing the regularization parameter, *Inverse Probl.*, **24**, doi :10.1088/0266-5611/24/3/034006.
- MacKay, D. J. C., 2003. *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 640p.

- Mainprice, D., 2007. Seismic anisotropy of the deep Earth from a mineral and rock physics perspective, in *Treatise on Geophysics : Mineral Physics*, pp. 437–491, ed. Price, G. D., Elsevier, Amsterdam.
- Mao, Z., Jacobsen, S. D., Jiang, F., Smyth, J. R., Holl, C. M., & Duffy, T. S., 2008. Elasticity of hydrous wadsleyite to 12 GPa : Implications for Earth’s transition zone, *Geophys. Res. Lett.*, **35**, L21305, doi :10.1029/2008GL035618.
- Marone, F. & Romanowicz, B., 2007. Non-linear crustal corrections in high-resolution waveform seismic tomography, *Geophys. J. Int.*, **170**, 460–467.
- Marone, F., Gung, Y., & Romanowicz, B., 2007. Three-dimensional radial anisotropic structure of the North American upper mantle from inversion of surface waveform data, *Geophys. J. Int.*, **171**, 206–222.
- Masters, G., Johnson, S., Laske, G., & Bolton, H., 1996. A shear-velocity model of the mantle, *Phil. Trans. R. Soc. London A*, **354**, 1285–1411.
- Masters, G., Laske, G., Bolton, H., & Dziewonski, A. M., 2000. The relative behaviour of shear velocity, bulk sound speed, and compresionnal velocity in the mantle : Implications for chemical and thermal structure, in [Karato et al. \(2000\)](#), pp. 63–87.
- Mégnin, C. & Romanowicz, B., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophys. J. Int.*, **143**, 709–728.
- Meier, U., Curtis, A., & Trampert, J., 2007. Global crustal thickness from neural network inversion of surface wave data, *Geophys. J. Int.*, **169**, 706–722.
- Metropolis, N. & Ulam, S., 1949. The Monte Carlo method, *J. Am. Stat. Ass.*, **44**, 335–341.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21**, 1087–1091.
- Montagner, J., 2007. Upper mantle structure : Global isotropic and anisotropic tomography, in *Treatise in Geophysics*, pp. 559–590, eds Romanowicz, B. & Dziewonski, A. M., Elsevier, Amsterdam.
- Montagner, J.-P., 1998. Where can seismic anisotropy be detected in the Earth’s mantle ? in boundary layers ..., *Pure Appl. Geophys*, **58**, 223–256.

- Montagner, J.-P., 2002. Low anisotropy channels below the Pacific Plate, *Earth Planet. Sci. Lett.*, **6320**, 1–12.
- Montagner, J.-P. & Anderson, D. L., 1989. Petrological constraints on seismic anisotropy, *Phys. Earth Planet. Int.*, **54**, 82–105.
- Montagner, J.-P. & Jobert, N., 1988. Vectorial tomography-II. Application to the Indian ocean, *Geophys. Journal*, **94**, 309–344.
- Montagner, J.-P. & Kennett, B. L. N., 1996. How to reconcile body-wave and normal-mode reference Earth models, *Geophys. J. Int.*, **125**, 229–248.
- Montagner, J.-P. & Nataf, H. C., 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *J. Geophys. Res.*, **91**, 511–520.
- Montagner, J.-P. & Nataf, H. C., 1988. Vectorial Tomography. I : Theory, *Geophys. J.R. astr. Soc.*, **94**, 295–307.
- Montagner, J.-P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. Geophys. Res.*, **95**, 4797–4819.
- Mosca, I., Cobden, L., Deuss, A., Ritsema, J., & Trampert, J., 2012. Seismic and mineralogical structures of the lower mantle from probabilistic tomography, *J. Geophys. Res.*, **117**, B06304, doi :10.1029/2011JB008851.
- Mosegaard, K., 1998. Resolution analysis of general inverse problems through inverse Monte-Carlo sampling, *Inverse Problems*, **14**, 405–426.
- Mosegaard, K. & Tarantola, A., 1995. Monte-Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, **100**, 12431–12447.
- Murakami, M., Ohishi, Y., Hirao, N., & Hirose, K., 2009. Elasticity of MgO to 130 GPa : Implications for lower mantle mineralogy, *Earth and Planet. Sci. Lett.*, **277**, 123–129.
- Nakagawa, T., Tackley, P. J., Deschamps, F., & Connolly, J. A., 2009. Incorporating self-consistent calculated mineral physics into thermochemical mantle convection simulations in a 3-D spherical shell and its influence on seismic anomalies in Earth's mantle, *Geochim. Geophys. Geosyst.*, **10**, Q03004, doi :10.1029/2008GC002280.
- Nataf, H. C. & Ricard, Y., 1996. 3SMAC : An *a priori* tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth Planet. Int.*, **95**, 101–122.
- Neal, R. M., 2005. The short-cut metropolis method, doi : arXiv :math/0508060v1.

- Nettles, M. & Dziewonski, A. M., 2008. Radially anisotropic shear-velocity structure of the upper mantle globally and beneath North America, *J. Geophys. Res.*, **113**, B02303, doi : 10.1029/2006JB004819.
- Niazi, M. & Anderson, D. L., 1965. Upper mantle structure of western North America from apparent velocities of *P* waves, *J. Geophys. Res.*, **70**, 4633–4640.
- Nolet, G., 2008. *A breviary of Seismic Tomography : Imaging the Interior of the Earth and Sun*, Cambridge University Press, 344p.
- Ohuchi, T., Kawazoe, T., Nishihara, Y., Nishiyama, N., & Irifune, T., 2011. High pressure and temperature fabric transitions in olivine and variations in upper mantle seismic anisotropy, *Earth Planet. Sci. Lett.*, **304**, 55–63.
- Okino, K., Ando, M., Kaneshima, S., & Hirahara, K., 1989. The horizontally lying slab, *Geophys. Res. Lett.*, **16**, 1059–2062.
- Pacalo, R. E. G. & Weidner, D. J., 1997. Elasticity of majorite, MgSiO_3 tetragonal garnet, *Phys. Earth Planet. Int.*, **99**, 145–154.
- Panning, M. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, **167**, 361–379, doi :10.1111/j.1365-246X.2006.03100.x.
- Panning, M., Lekic, V., & Romanowicz, B., 2010. The importance of crustal corrections in the development of a new global model of radial anisotropy, *J. Geophys. Res.*, **115**, doi :10.1029/2010JB007520.
- Poirier, J. P., 2000. *Introduction to the Physics of the Earth's Interior*, Cambridge University Press, 328p.
- Rawlinson, N., Reading, A. M., & Kennett, B. L. N., 2006. Lithospheric structure of Tasmania from a novel form of teleseismic tomography, *J. Geophys. Res.*, **111**, doi :10.1029/2005JB003803.
- Rawlinson, N., Sambridge, M., & Saygin, E., 2008. A dynamic objective function technique for generating multiple solution models in seismic tomography, *Geophys. J. Int.*, **174**, 295–308.
- Rawlinson, N., Pozdag, S., & Fishwick, S., 2010. Seismic tomography : A window into deep Earth, *Phys. Earth Planet. Int.*, **178**, 101–135.
- Revenaugh, J. & Sipkin, S. A., 1994. Seismic evidence for silicate melt atop the 410-km mantle discontinuity, *Nature*, **369**, 474–476.

- Ringwood, A. E., 1975. *Composition and Petrology of the Earth's Mantle*, McGraw-Hill, New York, 618p.
- Ritsema, J., van Heijst, H. J., & Woodhouse, J. H., 1999. Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science*, **286**, 1925–1928.
- Ritsema, J., van Heijst, H. J., & Woodhouse, J. H., 2004. Global transition zone tomography, *J. Geophys. Res.*, **109**, B02302, doi :10.1029/2003JB002610.
- Ritsema, J., Xu, W., Stixrude, L., & Lithgow-Bertelloni, C., 2009. Estimates of the transition zone temperature in a mechanically mixed upper mantle, *Earth Planet. Sci. Lett.*, **244**, 244–252.
- Romanowicz, B., 2003. Global mantle tomography : progress status in the last 10 years, *Ann. Rev. of Earth and Planet. Sciences*, **31**, 303–328.
- Rosenthal, J., 2000. Parallel computing and Monte Carlo algorithms, *Far Est Journal of Theoretical Statistics*, **4**, 207–236.
- Saikia, A., Frost, D. J., & Rubie, D. C., 2008. Splitting of the 520-kilometer seismic discontinuity and chemical heterogeneity in the mantle, *Science*, **319**, 1515–1518.
- Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm-I. Searching a parameter space, *Geophys. J. Int.*, **138**, 479–494.
- Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm-II. Appraising the ensemble, *Geophys. J. Int.*, **138**, 727–746.
- Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, **40**, doi :10.1029/2000RG000089.
- Sawamoto, H., Weidner, D., Sasaki, S., & Kumazawa, M., 1984. Single-crystal elastic properties of the modified spinel (beta) phase of Mg_2SiO_4 , *Science*, **224**, 749–751.
- Shapiro, N. M. & Ritzwoller, M. H., 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J. Int.*, **151**, 88–105.
- Shearer, P. M., 1991. Constraints on upper mantle discontinuities from observations of long-period reflected and converted phases, *J. Geophys. Res.*, **96**, 18,147–18,182.
- Shearer, P. M., 2000. *Upper mantle seismic discontinuities*, pp. 115–131, in [Karato et al. \(2000\)](#).
- Shearer, P. M. & Masters, T. G., 1992. Global mapping of topography on the 660-km discontinuity, *Nature*, **355**, 791–796.

- Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., & Lin, F.-C., 2013. Joint inversion of surface wave dispersion and receiver functions : a bayesian monte-carlo approach, *Geophys. J. Int.*, **192**, 807–836.
- Simons, F. J., Forte, A. M., & Grand, S. P., 2009. Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity : Implications for the relative importance of thermal versus compositional heterogeneities, *Geophys. J. Int.*, **177**, 1284–1304.
- Sinogeikin, S. V., Katsura, T., & Bass, J. D., 1998. Sound velocities and elastic properties of Fe-bearing wadsleyite and ringwoodite, *J. Geophys. Res.*, **103**, 20819–20825.
- Sinogeikin, S. V., Zhang, J., & Bass, J., 2004. Elasticity of single crystal and polycrystalline MgSiO_3 perovskite by Brillouin spectroscopy, *Geophys. Res. Lett.*, **31**, L06620, doi :10.1029/2004GL019559.
- Smyth, J., Holl, C., Frost, D., Jacobsen, S., Langenhorst, F., & McCammon, C., 2003. Structural systematics of hydrous ringwoodite and water in the Earth's interior, *Am. Mineral.*, **88**, 1402–1407.
- Sotin, C. & Parmentier, E. M., 1989. On the stability of a fluid layer containing an univariant phase transition : application to planetary interiors, *Phys. Earth Planet. Int.*, **55**, 10–25.
- Stacey, F. D., 1992. *Physics of the Earth*, Brookfield Press, Brisbane, 513p.
- Stixrude, L., 1997. Structure and sharpness of phase transitions and mantle discontinuities, *J. Geophys. Res.*, **102**, 14835–14582.
- Stixrude, L. & Lithgow-Bertelloni, C., 2005a. Thermodynamics of mantle minerals, I : physical properties, *Geophys. J. Int.*, **162**, 610–632.
- Stixrude, L. & Lithgow-Bertelloni, C., 2005b. Mineralogy and elasticity of the oceanic upper mantle : Origin of the low-velocity zone., *J. Geophys. Res.*, **110**, B03204, doi :10.1029/2004JB002965.
- Stixrude, L. & Lithgow-Bertelloni, C., 2010. Thermodynamics of the Earth's mantle, *Rev. Min. Geoch.*, **71**, 465–484.
- Suzuki, A., Ohtani, E., Morishima, H., Kubo, T., Kanbe, Y., Kondo, T., Okada, T., Terasaki, H., Kato, T., & Kikegawa, T., 2000. In situ determination of the phase boundary between wadsleyite and ringwoodite in Mg_2SiO_4 , *Geophys. Res. Lett.*, **27**, 803–806.

- Tackley, P. J., 2008. Geodynamics : Layer cake or plum pudding ?, *Nat. Geosci.*, **157**, doi :10.1038/ngeo134.
- Tackley, P. J., Stevenson, D. J., Glatzmaier, G. A., & Schubert, G., 1984. Effects of multiple phase transitions in a three dimensional spherical model of convection in Earth's mantle, *J. Geophys. Res.*, **99**, 15977–15901.
- Tackley, P. J., Xie, S., Nakagawa, T., & Hernlund, J. W., 2005. Numerical and laboratory studies of mantle convection : Philosophy, accomplishments, and thermochemical structure and evolution, in *Earth's Deep Mantle : Structure, Composition, and Evolution*, pp. 83–99, eds van der Hilst, R. D., Bass, J. D., Matas, J., & Trampert, J., Geophys. Monogr. Ser., vol. 160, American Geophysical Union, Washington, D. C.
- Tagawa, M., Nakakuki, T., & Tajima, F., 2007. Dynamical modeling of trench retreat driven by the slab interaction with the mantle transition zone, *Earth, Planets and Space*, **59**, 65–74.
- Takeushi, H. & Saito, M., 1972. Seismic surface waves, in *Methods in Computational Physics*, vol. 11, pp. 217–295, ed. Bolt, A. B., Academic Press, New York.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, 342p.
- Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.*, **20**, 219–232.
- Tauzin, B., Debayle, E., & Wittlinger, G., 2010. Seismic evidence for a global low-velocity layer within the Earth's upper mantle, *Nature Geoscience*, **3**, 718–721.
- Tibi, R. & Wiens, D. A., 2005. Detailed structure and sharpness of upper mantle discontinuities in the Tonga subduction zone from regional broad-band arrays, *J. Geophys. Res.*, **110**, doi :10.1029/2004JB003433.
- Trampert, J. & van Heijst, H. J., 2002. Global azimuthal anisotropy in the transition zone, *Science*, **296**, 1297–1299.
- Trampert, J. & Woodhouse, J. H., 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophys. J. Int.*, **154**, 154–165.
- Trampert, J., Vacher, P., & Vlaar, N., 2001. Sensitivities of seismic velocities to temperature, pressure and composition in the lower mantle, *Phys. Earth Planet. Int.*, **124**, 255–267.

- Trampert, J., Deschamps, F., Resovsky, J., & Yuen, D., 2004. Probabilistic tomography maps chemical heterogeneities throughout the lower mantle, *Science*, **306**, 853–856.
- Turcotte, D. L. & Schubert, G., 1982. *Geodynamics*, John Wiley and Sons, New York, 456p.
- Turcotte, D. L., Schubert, G., & Olson, P., 2000. *Mantle Convection in the Earth and Planets*, Cambridge University Press, 940p.
- Vacher, P., Mocquet, A., & Sotin, C., 1996. Comparison between tomographic structures and models of convection in the upper mantle, *Geophys. J. Int.*, **124**, 45–56.
- Vacher, P., Mocquet, A., & Sotin, C., 1998. Computation of seismic profiles from mineral physics : The importance of the non-olivine components for explaining the 660 km depth discontinuity, *Phys. Earth Planet. Int.*, **106**, 275–298.
- Vacher, P., Spakman, W., & Wortel, M. J. R., 1999. Numerical tests on the seismic visibility of metastable minerals in subduction zones, *Earth Planet. Sci. Lett.*, **170**, 335–349.
- van der Hilst, R. D., Engdahl, E. R., Spakman, W., & Nolet, G., 1991. Tomographic imaging of subducted lithosphere below northwest Pacific island arcs, *Nature*, **353**, 37–43.
- van der Hilst, R. D., Widyanoro, S., & Engdahl, E. R., 1997. Evidence for deep mantle circulation from global tomography, *Nature*, **386**, 578–584.
- van Heijst, H. J. & Woodhouse, J. H., 1997. Measuring surface-wave overtone phase velocities using a mode-branch stripping technique, *Geophys. J. Int.*, **131**, 209–230.
- Verhoeven, O., Rivoldini, A., Vacher, P., Mocquet, A., Choblet, G., Menvielle, M., Dehant, V., Hoolst, T. V., Sleewaegen, J., Barriot, J.-P., & Lognonné, P., 2005. Interior structure of terrestrial planets : Modeling Mars’ mantle and its electromagnetic, geodetic, and seismic properties, *J. Geophys. Res.*, **110**, 10.1029/2008JB005678.
- Verhoeven, O., Mocquet, A., Vacher, P., Rivoldini, A., Menvielle, M., Arrial, P.-A., Choblet, G., Tarits, P., Dehant, V., & Hoolst, T. V., 2009. Constraints on thermal state and composition of the Earth’s lower mantle from electromagnetic impedances and seismic data, *J. Geophys. Res.*, **114**, 10.1029/2004JE002271.
- Visser, K., Lebedev, S., Trampert, J., & Kennett, B. L. N., 2007. Global Love wave overtone measurements, *Geophys. Res. Lett.*, **34**, L03302, doi :10.1029/2006GL028671.
- Visser, K., Trampert, J., & Kennett, B. L. N., 2008a. Global anisotropic phase-velocity maps for higher mode Love and Rayleigh waves, *Geophys. J. Int.*, **172**, 1016, doi :10.1111/j.1365-246X.2007.03685.x.

- Visser, K., Trampert, J., Lebedev, S., & Kennett, B. L. N., 2008b. Probability of radial anisotropy in the deep mantle, *Earth Planet. Sci. Lett.*, **270**, 241–250.
- Wang, Y., Uchida, T., Zhang, J., Rivers, M. L., & Sutton, S. R., 2004. Thermal equation of state of akimotoite MgSiO_3 and effects of the akimotoite-garnet transformation on seismic structure near the 660 km discontinuity, *Phys. Earth planet. Int.*, **143**, 57–80.
- Watt, J. P., Davies, G. F., & O’Connell, R. J., 1976. The elastic properties of composite materials, *Rev. Geophys. Space Phys.*, **14**, 541–563.
- Webb, S. L. & Jackson, I., 1993. The pressure dependence of the elastic moduli of single-crystal orthopyroxene ($\text{Mg}_{0.8}\text{Fe}_{0.2}$) SiO_3 , *Europ. J. Min.*, **5**, 1111–1119.
- Weber, Z., 2000. Seismic travelttime tomography : a simulated annealing approach, *Phys. Earth Planet. Int.*, **119**, 149–159.
- Weidner, D. J., 1985. A mineral physics test of a pyrolite mantle, *Geophys. Res. Lett.*, **12**, 449–455.
- Weidner, D. J. & Ito, E., 1985. Elasticity of MgSiO_3 in the ilmenite phase, *Phys. Earth Planet. Int.*, **40**, 65–70.
- Weidner, D. J. & Wang, Y., 1998. Chemical- and clapeyron-induced buoyancy at the 660 km discontinuity, *J. Geophys. Res.*, **103**, 7431–7441.
- Weidner, D. J. & Wang, Y., 2000. *Phase transformations : implications for mantle structure*, pp. 215–235, in [Karato et al. \(2000\)](#).
- Weidner, D. J., Wang, H., & Ito, J., 1978. Elasticity of orthoenstatite, *Phys. Earth Planet. Int.*, **17**, 7–13.
- Wood, B. J., 1995. The effect of H_2O on the 410-kilometer seismic discontinuity, *Science*, **268**, 74–76.
- Woodhouse, J., 1988. The calculation of eigenfrequencies and eigenfunctions of the free oscillations of the Earth and the Sun, in *Seismological Algorithms*, pp. 321–370, ed. Doornbos, D., Academic Press, London.
- Workman, R. K. & Hart, S. R., 2005. Major and trace element composition of the depleted MORB mantle (DMM), *Earth Planet. Sci. Lett.*, **231**, 53–72.
- Xu, W., Lithgow-Bertelloni, C., Stixrude, L., & Ritsema, J., 2008. The effect of bulk composition and temperature on mantle seismic structure, *Earth Planet. Sci. Lett.*, **275**, 70–79.

- Xu, Y., Shankland, T. J., & Poe, B. T., 2000. Laboratory-based electrical conductivity in the Earth's mantle, *J. Geophys. Res.*, **105**, 27865–27875.
- Yeganeh-Haeri, A., Weidner, D. J., & Ito, E., 1989. Elasticity of MgSiO_3 in the perovskite structure, *Science*, **243**, 787–789.
- Yuan, H., Romanowicz, B., Fischer, K. M., & Abt, D., 2011. 3-D shear wave radially and azimuthally anisotropic velocity model of the North American upper mantle, *Geophys. J. Int.*, **184**, 1237–1260.
- Zha, C.-S., Duffy, T. S., Downs, R. T., Mao, H.-K., & Hemley, R. J., 1998. Brillouin scattering and X-ray diffraction of San Carlos olivine : Direct pressure determination to 32 GPa, *Earth Planet. Sci. Lett.*, **159**, 25–33.
- Zhao, D., 2001. Seismological structure of subduction zones and its implications for arc magmatism and dynamics, *Phys. Earth Planet. Int.*, **127**, 197–214.
- Zhou, Y., Nolet, G., Dahlen, F. A., & Laske, G., 2006. Global upper mantle structure from finite frequency surface wave tomography, *J. Geophys. Res.*, **111**, B04304, doi :10.1029/2005JB003677.

Une approche bayésienne pour estimer les propriétés physiques dans la zone de transition à partir des ondes de surface

Résumé. Les nouvelles méthodes tomographiques, qui utilisent le calcul des noyaux de sensibilité 3D, emploient souvent comme modèles de référence des modèles V_P , V_S à grandes longueurs d'onde obtenus par des inversions linéarisées. Ces modèles sont basés sur de petites perturbations de modèles globaux 1D et sont utilisés dans un deuxième temps pour estimer les distributions de température et de composition. D'un point de vue sismologique, le degré d'hétérogénéité dans la zone de transition (350- 1000 km de profondeur), lié aux transitions de phase et aux mouvements de convection, peut être suffisamment important pour que le concept d'un modèle sismique 1D de référence soit remis en question. Un algorithme de Monte Carlo par chaînes de Markov a été implémenté. Il détermine directement de façon statistique l'état thermique et la structure anisotrope du manteau à partir des données de dispersion des ondes de surface de Love et de Rayleigh. Des courbes polynomiales de Bézier sont choisies pour la paramétrisation et produisent à la fois des modèles lisses et des discontinuités. La solution est décrite en termes de probabilités, ce qui permet de prendre en compte les incertitudes. La méthode est illustrée avec des courbes de dispersion synthétiques et réelles. Les résultats indiquent une distribution de température complexe au milieu de la zone de transition sous l'Océan Pacifique. La structure anisotrope détectée est en bon accord avec les études précédentes qui indiquent une anisotropie positive dans le manteau supérieur. En considérant peu de contraintes *a priori*, la zone de transition paraît isotrope, le long du trajet étudié.

Mots-clés : inversion non-linéaire, ondes de surface, distributions de probabilité, zone de transition, température, anisotropie radiale, vitesse de phase.

A Bayesian approach to infer the physical properties in the transition zone from surface waves

Abstract. The new tomographic methods involving 3-D kernel computations often use, as reference models, 3-D large wavelength V_P , V_S models obtained by linearized inversions. These models are based on small perturbations of 1-D global models and are secondly used to derive temperature and composition distributions. From a seismological point of view, the degree of heterogeneity in the transition zone (350- 1000 km depth), due to phase transitions and convective motions, can be strong enough that the concept of a 1-D reference seismic model might be addressed. A Markov chain Monte Carlo algorithm was implemented. This directly determines the statistical thermal state and anisotropic structure of the mantle from the dispersion data of Love and Rayleigh surface waves. Polynomial Bézier curves are chosen for the parameterization and are able to explore both smoothly varying models and first-order discontinuities. The solution is described in probabilistic terms, allowing uncertainties to be fully accounted for. The method is illustrated with both synthetic data and real dispersion curves. The results indicate a complex temperature distribution in the mid-transition zone beneath the Pacific Ocean. The retrieved anisotropy structure agrees with previous studies indicating positive uppermost mantle anisotropy. Considering few *a priori* conditions, the transition zone appears to be isotropic, along the investigated path.

Key words : non-linear inversion, surface waves, probability distributions, transition zone, temperature, radial anisotropy, phase velocity.