



Numéro national de thèse : 2016LYSEN045

**THESE DE DOCTORAT DE L'UNIVERSITE DE LYON**

opérée par

**l'Ecole Normale Supérieure de Lyon**

**Ecole Doctorale** N° 52

**Physique et Astrophysique de Lyon**

**Discipline : Sciences de la Terre et de l'Univers**

Soutenue publiquement le 25/11/2016, par :

**Marie BOCHER**

---

# Reconstitution de la convection du manteau terrestre par assimilation de données séquentielle

Reconstruction of Mantle Circulation Using Sequential Data Assimilation

---

**Après l'avis de :**

| | | |
|---|---|---|
| Fleitout, Luce | Directrice de recherche Ecole Normale Supérieure Paris | Rapporteure |
| Cosme, Emmanuel | Maître de conférences Université Grenoble Alpes | Rapporteur |

**et devant le jury composé de :**

| | | |
|---|---|---|
| Cosme, Emmanuel | Maître de conférences Université Grenoble Alpes | Rapporteur |
| Labrosse, Stéphane | Professeur Ecole Normale Supérieure de Lyon | Président |
| Stadler, Georg | Associate Professor New York University | Examinateur |
| Janjic-Pfander, Tijana | Head of data assimilation branch Ludwig Maximillians university | Examinatrice |
| Coltice, Nicolas | Professeur Ecole Normale Supérieure de Lyon | Directeur de thèse |
| Fournier, Alexandre | Professeur Institut de Physique du Globe Paris | Codirecteur de thèse |

# Résumé

Cette thèse vise à proposer de nouvelles méthodes permettant de reconstruire la circulation dans le manteau terrestre et l'évolution de la tectonique de surface pour les deux cents derniers millions d'années. Nous utilisons des modèles numériques de convection mantellique dans lesquels la dynamique de surface est comparable à la tectonique terrestre. En combinant ces modèles avec des reconstructions de la tectonique des plaques il est possible d'estimer la structure et l'évolution du champ de température dans le manteau. Jusqu'à présent, l'inclusion des reconstructions de la tectonique des plaques se faisait en imposant des conditions aux limites du modèle (équilibre des forces, vitesses imposées...). Ces techniques, bien que permettant de tester la validité de différents scénarios tectoniques alternatifs, n'autorisent pas de rétroaction dynamique de la convection mantellique sur la tectonique de surface.

Dans ce travail, nous avons développé des techniques d'assimilation de données permettant d'intégrer les reconstructions de la tectonique des plaques dans un modèle numérique tout en laissant se développer de manière auto-cohérente cette rétroaction. Les techniques développées permettent également de prendre en compte les incertitudes associées aux reconstructions de la tectonique des plaques et de calculer les erreurs sur l'estimation finale de la circulation mantellique.

Dans un premier temps, nous avons développé un filtre de Kalman suboptimal qui permet d'estimer la structure et l'évolution de la circulation mantellique la plus probable à partir d'un modèle numérique de convection et d'une série temporelle d'observations de surface, ainsi que de leurs incertitudes respectives.

Ce filtre a été testé sur des expériences synthétiques. Celles-ci consistent à tenter de retrouver une évolution témoin à partir d'une série temporelle de données issues de cette évolution. Ces expériences ont montré qu'il était possible, en principe, de reconstruire la structure et l'évolution de l'ensemble du manteau à partir d'observations de vitesses et de flux de chaleur à la surface.

Dans un second temps, nous avons développé un filtre de Kalman d'ensemble. Ce filtre

permet non seulement d'estimer de manière plus précise la géométrie des structures mantelliques, mais aussi les incertitudes sur cette estimation.

# Abstract

This dissertation focuses on the developpement of data assimilation methods to reconstruct the circulation of the Earth's mantle and the evolution of its surface tectonics for the last 200 Myrs. We use numerical models of mantle convection in which the surface dynamics is similar to the Earth's. By combining these models with plate tectonics reconstructions, it is possible to estimate the structure and evolution of the temperature field of the mantle. So far, the assimilation of plate tectonics reconstructions was done by imposing specific boundary conditions in the model (force balance, imposed velocities...). These techniques, although insightful to test the likeliness of alternative tectonic scenarios, do not allow the full expression of the dynamical feedback between mantle convection and surface tectonics. We develop sequential data assimilation techniques able to assimilate plate tectonics reconstructions in a numerical model while simultaneously letting this dynamical feedback develop self-consistently. Moreover, these techniques take into account errors in plate tectonics reconstructions, and compute the error on the final estimation of mantle circulation.

First, we develop a suboptimal Kalman filter. This filter estimates the most likely structure and evolution of mantle circulation from a numerical model of mantle convection, a time series of surface observations and the uncertainty on both. This filter was tested on synthetic experiments. The principle of a synthetic experiment is to apply the data assimilation algorithm to a set of synthetic observations obtained from a reference run, and to then compare the obtained estimation of the evolution with the reference evolution. The synthetic experiments we conducted showed that it was possible, in principle, to reconstruct the structure and evolution of the whole mantle from surface velocities and heat flux observations.

Second, we develop an Ensemble Kalman Filter. Instead of estimating the most likely evolution, an ensemble of possible evolutions are computed. This technique leads to a better estimation of the geometry of mantle structures and a more complete estimation of the uncertainties associated.

# Sommaire Français

Résumé iii

Résumé en Anglais v

Table des Matières (Anglais) ix

Introduction 1

Introduction Française 7

1 Le Problème direct de la Circulation Mantellique 13

1.1 Mécanique des fluides pour la convection Mantellique ..... 13

1.2 Equations de Conservation ..... 14

1.3 Conditions initiales et aux limites ..... 21

1.4 Rheologie et Génération de plaques ..... 22

1.5 Approximation Numérique ..... 28

2 Les données et le problème inverse de la circulation Mantellique 33

2.1 Un cadre d'assimilation de données pour le problème de la circulation mantellique 34

2.1.1 l'inconnue: l'évolution du vrai état du manteau ..... 35

2.1.2 Les données observées sur la circulation mantellique ..... 36

2.1.3 Le modèle Dynamique ..... 37

2.1.4 L'état de base du manteau ..... 38

2.2 Les données sur la circulation mantellique ..... 39

2.2.1 Le champ de température mantellique au temps présent ..... 39

2.2.2 L'histoire de la cinématique de surface ..... 44

2.3 Les reconstitutions dynamiques de la circulation mantellique ..... 59

2.3.1 Les méthodes directes ..... 59

2.3.2 Méthodes d'assimilation de données ..... 64

3 Application de l'assimilation de données séquentielle à la reconstitution jointe de la convection mantellique et de la tectonique de surface. 77

3.1 Introduction ..... 79
3.2 Le filtre de Kalman étendu ..... 81
3.2.1 Initialisation ..... 84
3.2.2 Itérations analyse/prédiction ..... 84
3.3 Le modèle de convection, le vecteur d'état, et les données tectoniques ..... 86
3.3.1 Modèle de convection avec tectonique de type plaque ..... 86
3.3.2 L'état du manteau ..... 88
3.3.3 Les données: le flux de chaleur et les vitesses de surface ..... 89
3.3.4 L'opérateur observation et le vecteur d'état augmenté ..... 90
3.4 Algorithme d'assimilation de données pour la convection mantellique ..... 90
3.4.1 Initialisation ..... 91
3.4.2 Analyse ..... 98
3.4.3 Prédiction ..... 99
3.5 Expériences synthetiques ..... 100
3.5.1 Organisation des expériences ..... 101
3.5.2 Qualité de l'estimation issue de l'assimilation de données ..... 102
3.6 Discussion ..... 109
3.7 Conclusion ..... 113
3.8 Résumé en français ..... 114

4 L'assimilation de données d'ensemble pour la circulation mantellique 117
4.1 Introduction ..... 119
4.2 Présentation du Problème ..... 122
4.2.1 Modèle de convection mantellique ..... 122
4.2.2 Observations de la Circulation mantellique ..... 125
4.2.3 Filtre de Kalman d'ensemble: notations ..... 127
4.3 Le filtre de Kalman d'ensemble avec localisation et inflation ..... 129
4.3.1 Initialisation: première analyse et génération de l'ensemble de départ ..... 130
4.3.2 prédiction ..... 131
4.3.3 Analyse ..... 133

4.3.4 Implémentation du Filtre de Kalman d'Ensemble ..... 134
4.4 Evaluation a posteriori de la méthode du filtre de Kalman d'ensemble ..... 135
4.4.1 Descriptions des expériences jumelles ..... 135
4.4.2 Robustesse de l'algorithme d'assimilation ..... 136
4.4.3 Effet des paramètres d'assimilation sur la qualité de l'estimation ..... 146
4.4.4 Justesse de la reconstitution des structures géodynamiques ..... 149
4.5 Discussion ..... 154
4.6 Conclusion ..... 160

Discussion et Conclusion (anglais) 163
Discussion et Conclusion (français) 169
Bibliographie 176

# Contents

**Résumé** **iii**

**Abstract** **v**

**Contents** **ix**

**Introduction** **1**

**Introduction Française** **7**

**1 The Mantle Circulation Forward Problem** **13**

1.1 Fluid Mechanics for Mantle Convection . . . . . 13

1.2 Conservation Equations . . . . . 14

1.3 Initial and Boundary Conditions . . . . . 21

1.4 Rheology and Self-Consistent Plate Generation . . . . . 22

1.5 Numerical Approximation . . . . . 28

**2 Data and the Mantle Circulation Inverse Problem** **33**

2.1 A Data Assimilation Framework for Mantle Circulation Problems . . . . 34

2.1.1 The Unknown: the Evolution of the True State of the Mantle . . . 35

2.1.2 Observed Data on Mantle Circulation . . . . . 36
2.1.3 The Dynamical Model . . . . . 37
2.1.4 The Background State of the Mantle . . . . . 38
2.2 Data on Mantle Circulation . . . . . 39
2.2.1 Mantle Temperature Field at Present . . . . . 39
2.2.2 Surface Kinematics History of the Mantle . . . . . 44
2.3 Dynamical Mantle Circulation Reconstructions . . . . . 59
2.3.1 Direct Methods . . . . . 59
2.3.2 Data Assimilation Methods . . . . . 64

**3 A sequential data assimilation approach for the joint reconstruction of mantle convection and surface tectonics 77**

3.1 Introduction . . . . . 79
3.2 The Extended Kalman Filter . . . . . 81
3.2.1 Initialization . . . . . 84
3.2.2 Analysis and forecast sequence . . . . . 84
3.3 Convection model, Mantle State Vector and Tectonic Data . . . . . 86
3.3.1 Convection Model with Plate-Like Behaviour . . . . . 86
3.3.2 The State of the mantle . . . . . 88
3.3.3 The Data: Surface Heat Flux and Surface velocities . . . . . 89
3.3.4 The Observation Operator and the Augmented State . . . . . 90
3.4 Sequential Data Assimilation Algorithm for Mantle Convection . . . . . 90
3.4.1 Initialization . . . . . 91
3.4.2 Analysis . . . . . 98
3.4.3 Forecast . . . . . 99
3.5 Synthetic Experiments . . . . . 100
3.5.1 Setup of the Experiments . . . . . 101
3.5.2 Quality of the data assimilation estimate . . . . . 102
3.6 Discussion . . . . . 109
3.7 Conclusion . . . . . 113

**4 Ensemble Data Assimilation For Mantle Circulation** **117**

4.1 Introduction . . . . . 119

4.2 Presentation of the Problem . . . . . 122

4.2.1 Mantle Convection Model . . . . . 122

4.2.2 Observations of Mantle Circulation . . . . . 125

4.2.3 Ensemble Kalman Filtering Framework: Notations . . . . . 127

4.3 Ensemble Kalman Filter with Localization and Inflation . . . . . 129

4.3.1 Initialization: First Analysis and Generation of the starting ensemble . . . . . 130

4.3.2 Forecast . . . . . 131

4.3.3 Analysis . . . . . 133

4.3.4 Implementation of the Ensemble Kalman Filter . . . . . 134

4.4 A posteriori Evaluation of the Ensemble Kalman Filter Method . . . . . 135

4.4.1 Twin Experiment Setup . . . . . 135

4.4.2 Robustness of the Assimilation Algorithm . . . . . 136

4.4.3 Effect of the data assimilation parameters on the quality of the estimation . . . . . 146

4.4.4 Accuracy of the Reconstruction of Geodynamic Structures . . . . 149

4.5 Discussion . . . . . 154

4.6 Conclusion . . . . . 160

**5 Discussion and Conclusion** **163**

**Discussion et Conclusion** **169**

**Bibliography** **176**

# Introduction

The Earth is an active planet, as evidenced by the regular occurrence of earthquakes, volcanic eruptions and the presence of large regions of highly deformed rocks. This activity is the surface expression of a dynamic system comprised of the two outermost layers of the Earth: the crust (of variable thickness between $0$ up to $60$ km) and the mantle (extending from the base of the crust to $2900$ km in depth). To date, the Earth science community uses two different theories to describe the dynamics of the mantle-crust system at a global scale: plate tectonics theory describes the motion at the surface of the Earth while mantle convection theory describes its internal dynamics.

The theory of plate tectonics is a kinematic theory. Its original formulation (McKenzie & Parker, 1967; Morgan, 1968; Le Pichon, 1968) states that the Earth's surface is made of a jigsaw of internally rigid blocks, the plates, each of them moving with respect to the others. The relative motion of these plates is accommodated by deformation along the lines separating plates, called plate boundaries. This theory is successful in predicting the pattern of surface velocities and explaining the repartition of mountain ranges, earthquakes and volcanism (Figure 1 shows for example the fit between seismicity and plate boundaries). It also gives a global framework to integrate various types of data (magnetic, bathymetric, geodetic, paleontologic, geologic...) and provide models

of the motions at the surface of the Earth at present (Geologically current plate motion models, e.g. DeMets et al., 2010), and in the past (classical plate reconstruction models, e.g. Torsvik et al., 2010; Seton et al., 2012; Müller & Wessel, 2015).

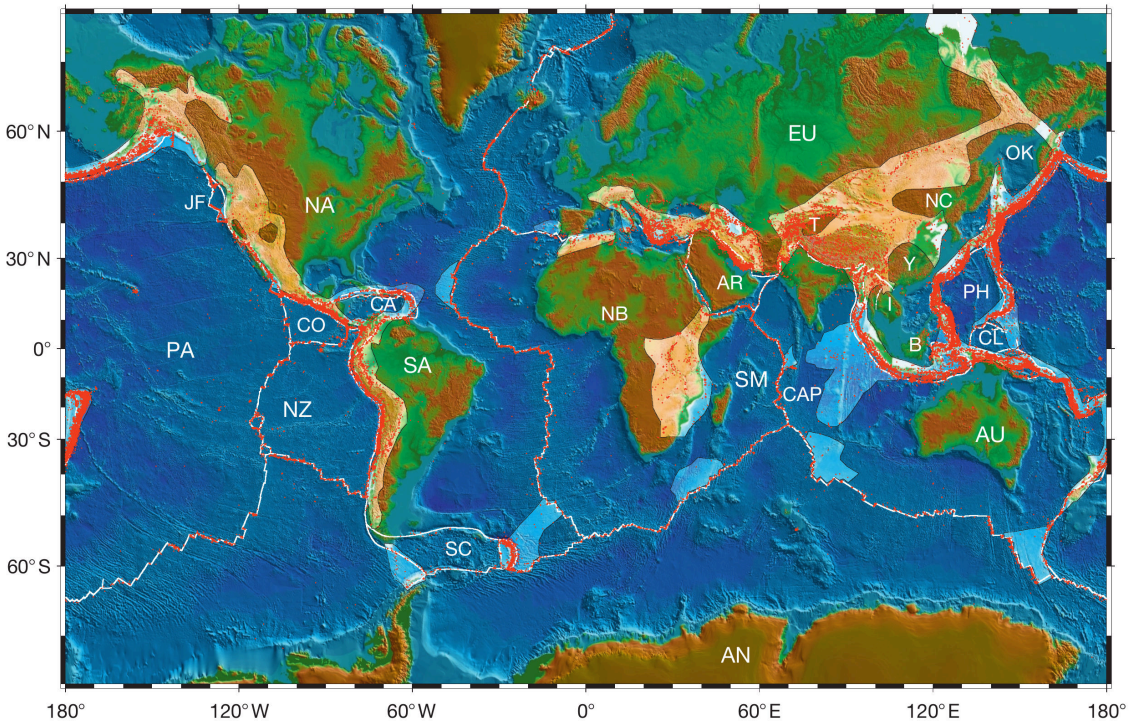


FIGURE 1: Synthetic world map taken from Müller & Wessel (2015). This map shows: in white, the plate boundaries from Coffin et al. (1997), in red dots the shallow seismicity ($0 - 70$ km depth) from Engdahl et al. (1998), and with areas of lighter illumination the regions of diffuse plate boundaries based on Figure 1 of Gordon (2000). Plate abbreviations: B, Borneo; AN, Antarctica; AR, Arabia; AU, Australia; CA, Caribbean; CAP, Capricorn; CL, Caroline; CO, Cocos; EU, Eurasia; I, Indo-China; IN, India; JF, Juan de Fuca; NA, North America; NB, Nubia; NC, North China; NZ, Nazca; OK, Okhotsk; PA, Pacific; SA, South America; SC, Scotia Sea; SM, Somalia; Y, Yangtze; T, Tarim Basin; PH, Philippine.

Mantle convection theory, on the contrary, is a physical theory. It is based on the observation that the mantle, although solid, behaves like a fluid at timescales greater than a few thousand years (Nansen, 1928; Haskell, 1937; Griggs, 1939a), and thus uses the principles and methods of fluid dynamics to study the mantle. Mantle convection theory states that the mantle is cooling down and that the principal mechanism for its transferring heat is convection: hot and thus lighter material in the interior rises while cold and denser material at the surface sinks, creating convecting currents. The evolution of mantle convection theory is tightly linked to the exploration of the surface

dynamics of the Earth. Originally, the possibility of mantle convection was studied (Holmes, 1931; Pekeris, 1935; Hales, 1936; Griggs, 1939b; Runcorn, 1962) so as to provide a viable driving mechanism for the theory of continental drift of Wegener (1924). The formulation of plate tectonics theory provided a more detailed description of the dynamics of the Earth's surface and it was soon thereafter recognized that the plates correspond to the thermal boundary layer in mantle convection theory (Turcotte & Oxburgh, 1972), or in short that plates are mantle convection. Since then, two of the central questions of mantle convection studies have been: how does plate tectonics arise from mantle convection, and what are the dynamical interactions between plates and the rest of the mantle?

The physical description of mantle convection is based on the derivation of partial differential equations governing the flow (Turcotte & Oxburgh, 1972). These equations are highly nonlinear, so the computation of a solution generally requires the development of numerical methods (McKenzie et al., 1974). The development of models of mantle convection has followed two complementary lines of inquiry to investigate the relationship between surface tectonics and mantle dynamics (see Lowman, 2011, for a review).

The first consist in using plate reconstruction models, either past or present, as boundary conditions of mantle convection models. In doing so, they provide estimations of the structure and evolution of the flow in the Earth's mantle. We will call these estimations mantle circulation reconstructions. Hager & O'Connell (1979) pioneered this type of studies and showed how the presence of plates at the surface affects the large scale flow of the mantle. Mantle circulation reconstructions have provided valuable insights into the dynamics and evolution of the deep earth mantle structures (Bunge et al., 1998; McNamara & Zhong, 2005; Davies et al., 2012; Bower et al., 2013), the evolution of mantle plumes and their relationship to hotspots (Hassan et al., 2016), changes in the Earth's rotation axis (Steinberger & O'Connell, 1997), sea-level (Moucha et al., 2008) or dynamic topography (Flament et al., 2013), for example.

The second line of research consist of implementing the physics necessary to generate

plates in convection codes, so that plate tectonics arises naturally from mantle convection models (Bercovici et al., 2015). Tackley (2000a) described this aim as the "Holy Grail" of mantle convection modellers: it would describe physically the dynamics of both the mantle and the surface, and, in doing so, unify the theories of plate tectonics and mantle convection. 16 years later, this aim is still not reached, but the progresses in this direction have led to the emergence of numerical models of mantle convection which display a surface dynamics similar to that of the Earth, at first order (Moresi & Solomatov, 1998; Moresi et al., 2000; Richards et al., 2001; Stein et al., 2004; Bercovici, 2003; Tackley, 2000b; Van Heck & Tackley, 2008). Although the mechanisms of plate generation are still not clearly understood, these models are realistic enough to study the dynamics of seafloor spreading (Coltice et al., 2012, 2013), the long-term motion of continents (Rolf et al., 2014; Yoshida, 2010) or the repartition of plate sizes and the dynamics of specific tectonic features (Mallard et al., 2016).

In this dissertation, we propose to take advantage of the predictive power of such models to develop a new generation of mantle circulation reconstructions, where observations of past and present plate motions are not imposed as boundary conditions, but assimilated in mantle convection models by an inverse method.

The problem of reconstructing mantle circulation is indeed an inverse problem: past and present motions at the surface of the Earth are the consequence of the structure and evolution of the flow in the mantle. Hence, we aim at determining the cause (the structure and evolution of the flow in the mantle) from its consequences (observations of surface motions). From the standpoint of inverse problem theory (see e.g. Wunsch, 2006, for a general introduction to inverse methods), the problem posed by mantle circulation reconstruction corresponds to the general question: how to estimate the state evolution of a system from a time series of data and a dynamical model? Methods solving this type of problem are grouped together under the generic term data assimilation methods.

Data assimilation methods were originally developed in the context of numerical weather prediction (Daley, 1993), and have since then found numerous applications, especially

in studies of geophysical fluids (i.e. the atmosphere Bengtsson et al., 1981; Daley, 1993; Kalnay, 2003, the ocean Chassignet & Verron, 2006 and the liquid outer core Fournier et al., 2010). These areas of research provide extensive resources on methods to solve data assimilation problems. We build on the experience and advances of data assimilation in those fields to design algorithms for the mantle circulation reconstruction problem.

This dissertation is organized as follows:

Chapter 1 presents the mantle circulation direct problem. We derive the partial differential equations governing mantle dynamics and the associated boundary and initial conditions. We also discuss constitutive laws chosen to describe mantle material, with a special focus on the rheology of the mantle, which plays a crucial role in the modelling of mantle convection with plate-like tectonics. Finally, we describe the numerical methods used to solve these equations.

Chapter 2 presents the mantle circulation inverse problem. We first introduce the general formalism of data assimilation and connect the mathematical objects that are defined to their meaning in the context of mantle circulation reconstruction. Then, we detail the data that is available to reconstruct mantle circulation. We finally review the reconstruction strategies that have so far been developed for mantle circulation.

Chapter 3 presents a proof of concept for the application of data assimilation methods to reconstruct mantle circulation from its surface tectonics history. We develop a suboptimal scheme derived from the Kalman filter, where surface velocities and seafloor age maps are used as data to assimilate. We test this algorithm on synthetic experiments using 2-D spherical annulus mantle evolutions. The results obtained show that it is possible, in principle, to reconstruct mantle circulation from the sole (and imperfect) knowledge of the history of surface tectonics.

In Chapter 4, we go one step further and apply the Ensemble Kalman Filter (EnKF) to this problem. The EnKF is a sequential Monte Carlo method particularly adapted to solve high dimensional data assimilation problems with a nonlinear dynamics. We test

the EnKF using synthetic observations, on a 2D-spherical annulus model and compare it with the method developed in Chapter 3. The EnKF performs on average better and is more stable than the former method. Moreover, the efficiency and scalability of the code we developed during this work opens the way to the application of the EnKF to realistic data assimilation with plate reconstructions and a 3D spherical mantle model.

Chapter 5 discusses the results obtained during this thesis and the possible future directions.

# Introduction Française

La Terre est une planète active, comme en témoignent l'occurence régulière de tremblements de terre, d'éruptions volcaniques et la présence de grandes régions fortement déformées. Cette activité est l'expression de surface d'un système dynamique constitué des deux couches les plus superficielles de la Terre: la croûte (d'épaisseur variable de $0$ à $60$ km) et le manteau (s'étendant de la base de la croûte à $2900$ km en profondeur). À ce jour, la communauté des sciences de la Terre utilise deux théories différentes pour décrire la dynamique du système manteau-croûte à une échelle globale: la théorie de la tectonique des plaques décrit les mouvements à la surface de la Terre alors que la théorie de la convection mantellique décrit sa dynamique interne.

La théorie de la tectonique des plaques est une théorie cinématique. Sa formulation originale (McKenzie & Parker, 1967; Morgan, 1968; Le Pichon, 1968) stipule que la surface de la Terre est faite d'un puzzle de blocs rigides, les plaques, qui se déplacent les unes par rapport aux autres. Le mouvement relatif de ces plaques est accomodé par la déformation le long des lignes séparant les plaques, appelées limites de plaques. Cette théorie réussit à prédire les vitesses de surface et à expliquer la répartition des chaînes de montagnes, des tremblements de terre et du volcanisme (la figure 2 montre par exemple l'adéquation entre la sismicité et les limites des plaques). Elle fournit

également un cadre global pour intégrer différents types de données (magnétiques, bathymétriques, géodésiques, paléontologiques, géologiques ...) et ainsi fournir des modèles des mouvements à la surface de la Terre à l'heure actuelle (modèles de mouvements de plaques géologiquement actuels, par exemple DeMets et al., 2010), et dans le passé (modèles classiques de reconstitution de la tectonique des plaques, par exemple Torsvik et al., 2010; Seton et al., 2012; Müller & Wessel, 2015).

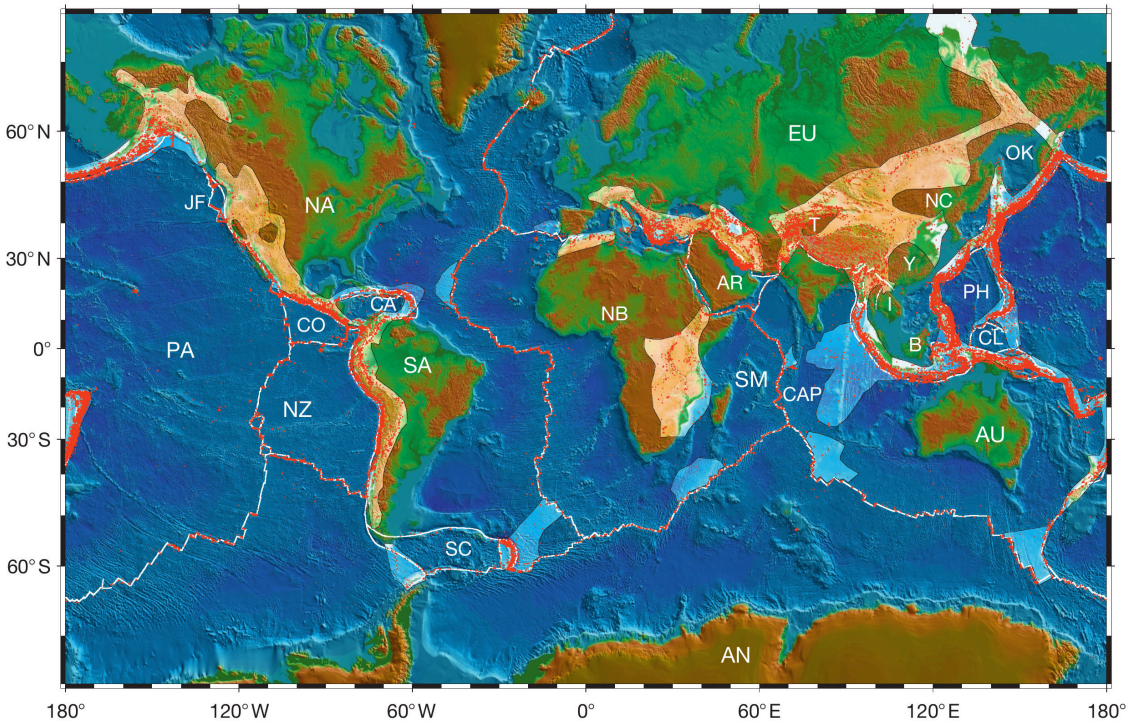


FIGURE 2: Carte synthétique du monde prise de Müller & Wessel (2015). Cette carte montre: en blanc, les limites de plaques de Coffin et al. (1997), en points rouges la sismicité peu profonde ($0 - 70$ km de profondeur) de Engdahl et al. (1998), et avec les zones illuminées les régions de limites de plaques diffuses, basées sur la figure 1 de Gordon (2000). Abréviations des plaques: B, Bornéo; AN, Antarctique; AR, Arabie; AU, Australie; CA, Caraïbes; CAP, Capricorne; CL, Caroline; CO, Cocos; UE, Eurasie; I, Indo-Chine; En Inde; JF, Juan de Fuca; NA, Amérique du Nord; NB, Nubia; NC, Chine du Nord; NZ, Nazca; OK, Okhotsk; PA, Pacifique; SA, Amérique du Sud; SC, Mer Scotia; SM, Somalie; Y, Yangtze; T, bassin de Tarim; PH, Philippine.

La théorie de la convection mantellique, au contraire, est une théorie physique. Elle est basée sur l'observation que le manteau, bien que solide, se comporte comme un fluide à des échelles de temps supérieures à quelques milliers d'années (Nansen, 1928; Haskell, 1937; Griggs, 1939a), et utilise les principes et les méthodes de la dynamique des fluides pour étudier le manteau. La théorie de la convection mantellique affirme

que le manteau se refroidit et que le mécanisme principal de transfert de chaleur est la convection: le matériel chaud et donc plus léger, à l'intérieur du manteau, s'élève tandis que le matériel froid et plus dense, à la surface, coule, créant ainsi des courants de convection. L'évolution de la théorie de la convection mantellique est étroitement liée à l'exploration de la dynamique de surface de la Terre. A l'origine, on a étudié la possibilité de la convection mantellique (Holmes, 1931; Pekeris, 1935; Hales, 1936; Griggs, 1939b; Runcorn, 1962) dans le but de fournir un mécanisme d'entraînement viable pour la théorie de la dérive des continents de Wegener (1924). La formulation de la théorie de la tectonique des plaques a fourni une description plus détaillée de la dynamique de la surface de la Terre et il a été bientôt reconnu que les plaques tectoniques correspondent à la couche limite thermique de surface dans la théorie de la convection mantellique (Turcotte & Oxburgh, 1972), ou autrement dit, que les plaques font partie de la convection mantellique. Depuis, deux des questions centrales des études de la convection mantellique ont été: Comment la tectonique des plaques découle-t-elle de la convection mantellique et quelles sont les interactions entre les plaques et le reste du manteau?

La description physique de la convection mantellique est basée sur la dérivation d'équations aux dérivées partielles régissant l'écoulement (Turcotte & Oxburgh, 1972). Ces équations sont très non linéaires, donc le calcul d'une solution nécessite généralement le développement de méthodes numériques (McKenzie et al., 1974). Le développement de modèles de convection mantellique a suivi deux lignes de recherche complémentaires pour étudier la relation entre la tectonique de surface et la dynamique du manteau (voir Lowman, 2011, pour une revue).

La première consiste à utiliser des modèles de reconstitution de plaques, passés ou présents, comme conditions aux limites des modèles de convection mantellique. On obtient ainsi des estimations de la structure et de l'évolution de l'écoulement dans le manteau terrestre. Nous appellerons ces estimations les reconstitutions de la circulation mantellique. Hager & O'Connell (1979) ont été les pionniers de ce type d'études et ont montré que la présence de plaques à la surface affecte l'écoulement à grande

échelle du manteau. Les reconstitutions de la circulation mantellique ont permis une meilleure compréhension de la dynamique et l'évolution des structures du manteau en profondeur (Bunge et al., 1998; McNamara & Zhong, 2005; Davies et al., 2012; Bower et al., 2013), de l'évolution des panaches mantelliques et de leur relation avec les points chauds (Hassan et al., 2016), des changements dans l'axe de rotation de la Terre (Steinberger & O'Connell, 1997), dans le niveau marin (Moucha et al., 2008) ou la topographie dynamique (Flament et al., 2013), par exemple.

La deuxième ligne de recherche consiste à implémenter la physique nécessaire pour générer des plaques dans les codes de convection, de sorte que la tectonique des plaques émerge naturellement dans les modèles de convection mantellique (Bercovici et al., 2015). Tackley (2000a) décrit ce but comme le "Saint Graal" des modélisateurs de la convection mantellique: on pourrait décrire physiquement la dynamique du manteau et de la surface, et, ce faisant, unifier les théories de la tectonique des plaques et de la convection mantellique. 16 ans plus tard, ce but n'est toujours pas atteint, mais les progrès dans ce sens ont conduit à l'émergence de modèles numériques de convection mantellique qui produisent une dynamique de surface semblable à celle de la Terre, au premier ordre (Moresi & Solomatov, 1998; Moresi et al., 2000; Richards et al., 2001; Stein et al., 2004; Bercovici, 2003; Tackley, 2000b; Van Heck & Tackley, 2008). Bien que les mécanismes de génération des plaques ne soient pas encore clairement compris, ces modèles sont suffisamment réalistes pour étudier la dynamique de formation des fonds océaniques (Coltice et al., 2012, 2013), le mouvement à long terme des continents (Rolf et al., 2014; Yoshida, 2010) ou encore la répartition des tailles des plaques et la dynamique de certains contextes tectoniques spécifiques (Mallard et al., 2016).

Dans cette dissertation, nous proposons de tirer parti du pouvoir prédictif de tels modèles pour développer une nouvelle génération de reconstitutions de la circulation mantellique, dans laquelle les observations des mouvements de plaques passées et présentes ne sont pas imposées comme des conditions aux limites, mais assimilées dans les modèles de convection mantellique par une méthode inverse.

Le problème de la reconstitution de la circulation mantellique est en effet un problème

inverse: les mouvements passés et présents à la surface de la Terre sont la conséquence de la structure et de l'évolution de l'écoulement dans le manteau. Par conséquent, nous cherchons à déterminer la cause (la structure et l'évolution de l'écoulement dans le manteau) à partir de ses conséquences (observations des mouvements de surface). Du point de vue de la théorie des problèmes inverses (voir, par exemple Wunsch, 2006, pour une introduction générale aux méthodes inverses), Le problème posé par la reconstitution de la circulation mantellique correspond à la question générale: Comment estimer l'évolution de l'état d'un système à partir d'une série chronologique de données et d'un modèle dynamique? Les méthodes de résolution de ce type de problème sont regroupées sous le terme générique de méthodes d'assimilation de données.

Les méthodes d'assimilation de données ont été initialement développées dans le contexte de la météorologie (Daley, 1993), et ont depuis lors trouvé de nombreuses applications, en particulier dans l'étude des fluides géophysiques (l'atmosphère Bengtsson et al., 1981; Daley, 1993; Kalnay, 2003, l'océan Chassignet & Verron, 2006 et le noyau externe liquide Fournier et al., 2010). Ces domaines de recherche fournissent des ressources étendues sur les méthodes permettant de résoudre les problèmes d'assimilation des données. Nous nous appuyons sur l'expérience et les progrès de l'assimilation de données dans ces domaines pour concevoir des algorithmes appliqués au problème de la reconstitution de la circulation mantellique.

Cette thèse est organisée comme suit:

Le chapitre 1 présente le problème direct de la circulation mantellique. Nous dérivons les équations aux dérivées partielles régissant la dynamique du manteau ainsi que les conditions aux limites et les conditions initiales. Nous discutons également les lois constitutives choisies pour décrire les propriétés mécaniques du manteau, en mettant l'accent sur la rhéologie du manteau, qui joue un rôle crucial dans la modélisation de la convection mantellique avec une tectonique de type plaque. Enfin, nous décrivons les méthodes numériques nécessaires pour résoudre ces équations.

Le chapitre 2 présente le problème inverse de la circulation mantellique. Nous introduisons d'abord le formalisme général de l'assimilation des données et relions les

objets mathématiques définis à leur signification dans le contexte de la reconstitution de la circulation mantellique. Ensuite, nous détaillons les données disponibles pour reconstruire la circulation mantellique. Nous examinons enfin les stratégies de reconstitution qui ont été développées jusqu'ici pour la circulation mantellique.

Le chapitre 3 présente une preuve du concept pour l'application de méthodes d'assimilation de données pour reconstruire la circulation mantellique à partir de l'histoire de sa tectonique de surface. Nous développons un schéma sous-optimal dérivé du filtre de Kalman, dans lequel les vitesses de surface et les cartes d'âge du fond marin sont utilisées comme données à assimiler. Nous testons cet algorithme sur des expériences synthétiques, en utilisant des évolutions 2D-sphériques annulaires. Les résultats obtenus montrent qu'il est possible, en principe, de reconstituer la circulation mantellique à partir de la connaissance unique (et imparfaite) de l'histoire de la tectonique de surface.

Dans le chapitre 4, nous allons plus loin et appliquons le filtre de Kalman d'Ensemble (EnKF) à ce problème. L'EnKF est une méthode séquentielle de Monte Carlo particulièrement adaptée pour résoudre l'assimilation de données pour des problèmes de grande dimension avec une dynamique non linéaire. Nous testons l'EnKF en utilisant des observations synthétiques, sur un modèle 2D-sphérique annulaire et comparons les résultats avec la méthode développée dans le chapitre 3. L'EnKF fonctionne en moyenne mieux et est plus stable que l'ancienne méthode. De plus, l'efficacité et la scalabilité du code que nous avons développé au cours de ce travail ouvrent la voie à l'application de l'EnKF à une assimilation avec des données provenant des reconstitutions des plaques terrestres et un modèle de manteau sphérique en 3D.

Le chapitre 5 traite des résultats obtenus au cours de cette thèse et des orientations futures possibles.

# CHAPTER 1

# The Mantle Circulation Forward Problem

Solving the direct problem of mantle circulation consists of specifying how to compute an evolution of mantle circulation from a given initial state. To do so, we use fluid dynamics principles. After justifying the use of fluid dynamics to compute mantle circulation, we derive the partial differential equations governing the flow, then specify the initial and boundary conditions necessary to solve these equations. An additional section is dedicated to the rheology of the mantle, which plays a major role in its dynamics. Finally, we discuss the numerical approximation of such a problem.

## 1.1 Fluid Mechanics for Mantle Convection

The study of mantle convection focuses on the dynamics of the mantle at length scales greater than one kilometer. This enables the modelling of the mantle as a continuous medium, although mantle material is polycrystalline and affected by brittle deformation near the surface.

Moreover, as in many solids, the response of mantle material to stress depends on the frequency content of the forcing. At shorter timescales (typically $< 100$ years), the mantle behaves like an elastic body, able to propagate seismic shear waves for example. However when submitted to stress over a longer period of time, mantle material deforms permanently. The Maxwell time $\tau_m$ gives a timescale for this transition from elastic to viscous behavior: $\tau_m = \mu_0/G$ where $\mu_0$ is the dynamic viscosity and $G$ is the shear modulus (elasticity). For the mantle, the shear modulus is of the order of $10^{11}$ Pa. If we take a viscosity of $10^{26}$ Pa.s for the lithosphere, and a viscosity of $10^{21}$ Pa.s for the rest of the mantle, we obtain a Maxwell time of $30$ Myr and $300$ yr respectively. So, at timescales of $1$ Myr or greater, the bulk of the mantle can be modelled as a viscous fluid. In the following, we shall make the same hypothesis for the lithosphere, although elasticity should be taken into account (Thielmann et al., 2015; Moresi et al., 2003).

Geochemical and seismological evidence suggests that the mantle is chemically heterogeneous at all length scales (see for example Meibom & Anderson (2004); Gurnis (1986a) for geochemistry and Ricard et al. (2014) for seismology). Although chemical heterogeneities play an important role for mantle convection (see for example Tackley, 2015, for an introduction to mantle geochemical geodynamics), we will consider only thermal heterogeneities in this development.

## 1.2 Conservation Equations

General derivations of equations governing fluid dynamics can be found in classical textbooks of fluid dynamics such as Batchelor (2000); Chorin et al. (1990). Here, we present the series of approximations classically used to model mantle convection, as developed in more details in Turcotte & Oxburgh (1972), Schubert et al. (2001) or Ricard (2015), for example.

## General Equations for the Conservation of Mass, Momentum and Energy

The conservation of mass implies that the variation in density $\rho$ of an open volume is compensated by a flow of mass $\rho\mathbf{u}$ (where $\mathbf{u}$ is the velocity) across its boundary. In differential form, it gives

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{1.1}$$

with $t$ the time. The variation of density with time is linked to sound wave propagation and is therefore negligible for the study of mantle convection.

Newton's second law (Newton, 1642–1727) states that the rate of change of momentum of a particle of fluid is equal to the sum of surface and body forces applied to this particle of fluid. Combined with the equation of conservation of mass and the divergence theorem, this leads to the differential equation

$$\rho \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = \nabla \cdot \boldsymbol{\sigma} + \mathbf{f}, \tag{1.2}$$

where $\boldsymbol{\sigma}$ is the stress tensor and $\mathbf{f}$ is the sum of body forces acting on the fluid per unit volume. The only body force acting on mantle material is the gravitational body force, so $\mathbf{f} = \rho\mathbf{g}$ with $\mathbf{g}$ the gravitational acceleration (see for example Ricard (2015), section 7.02.2.3.2 for a discussion on the effect of non-inertial forces on the mantle). The gravitational acceleration $\mathbf{g}$ depends on the repartition of mass in the mantle according to

$$\mathbf{g} = -\nabla \Psi \qquad \text{and} \qquad \nabla^2 \Psi = 4\pi G \rho, \tag{1.3}$$

where $\Psi$ is the gravitational potential and $G$ is the gravitational constant.

The stress tensor is decomposed into an isotropic thermodynamic pressure $P$ and a deviatoric velocity driven stress tensor $\boldsymbol{\tau}$, such that $\boldsymbol{\sigma} = -P\mathbf{I} + \boldsymbol{\tau}$, with $\mathbf{I}$ the identity matrix. As discussed previously, we suppose that mantle material behaves viscously. Although mantle anisotropy is established (Karato, 1998), and viscous anisotropy could affect the dynamics of the flow (Pouilloux et al., 2007; Mühlhaus et al., 2004), we assume, for simplicity, that the mantle viscosity is isotropic. The viscosity is then described with a bulk viscosity $\mu^v$ and a shear viscosity $\mu$. This leads to a velocity driven stress tensor $\boldsymbol{\tau}$ linked to velocity through

$$\boldsymbol{\tau} = \left[\left(\mu^v - \frac{2}{3}\mu\right)\nabla\cdot\mathbf{u}\right]\mathbf{I} + \mu\left(\nabla\mathbf{u} + [\nabla\mathbf{u}]^T\right). \tag{1.4}$$

The balance of energy gives

$$\rho\frac{\mathrm{D}\mathcal{U}}{\mathrm{D}t} = -\nabla\cdot\mathbf{q} - P\nabla\cdot\mathbf{u} + \nabla\cdot(\boldsymbol{\tau}\cdot\mathbf{u}) + \rho H. \tag{1.5}$$

where $\mathcal{U}$ is the internal energy per unit mass, $H$ the rate of energy production (radiogenic heating) per unit mass and $\mathbf{q}$ the heat flux.

We apply Fourier's law (Fourier, 1768–1830) to link the heat flux $\mathbf{q}$ to temperature $\mathbf{q} = -k\nabla T$ with $k$ the thermal conductivity.

By applying the first law of thermodynamics and supposing local thermodynamic equilibrium, we have

$$\frac{\mathrm{D}\mathcal{U}}{Dt} = C_p\frac{\mathrm{D}T}{\mathrm{D}t} - \frac{\alpha T}{\rho}\frac{\mathrm{D}P}{\mathrm{D}t} + P\frac{\nabla\cdot\mathbf{u}}{\rho} \tag{1.6}$$

where $\alpha$ is the thermal expansivity and $C_p$ is the heat capacity at constant pressure per unit volume.

We obtain the set of partial differential equations

$$\nabla \cdot (\rho \mathbf{u}) = 0, \tag{1.7a}$$

$$\rho \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla P + \nabla \cdot \boldsymbol{\tau} + \rho \mathbf{g}, \tag{1.7b}$$

$$\text{with} \quad \mathbf{g} = -\nabla \Psi \qquad \text{and} \qquad \nabla^2 \Psi = 4\pi G \rho,$$

$$\boldsymbol{\tau} = \left[ \left( \mu^v - \frac{2}{3}\mu \right) \nabla \cdot \mathbf{u} \right] \mathbf{I} + \mu \left( \nabla \mathbf{u} + [\nabla \mathbf{u}]^T \right),$$

$$\rho C_p \frac{\mathrm{D}T}{\mathrm{D}t} = \nabla \cdot (k \nabla T) + \alpha T \frac{\mathrm{D}P}{\mathrm{D}t} - \nabla \cdot (\boldsymbol{\tau} \cdot \mathbf{u}) + \rho H. \tag{1.7c}$$

For mantle convection, this system of equation is simplified either by using the anelastic liquid approximation or the Boussinesq approximation (Joseph Valentin Boussinesq, 1842–1929). The anelastic liquid approximation takes into account compressibility and the depth-dependence of density. Derivation of anelastic liquid equation for the mantle can be found in Jarvis & McKenzie (1980); Glatzmaier (1988); Bercovici et al. (1992); Alboussière & Ricard (2013). Here, we will derive the simpler Boussinesq equations.

## The Boussinesq Equations

To further simplify the system of equations, all the variables are written as the sum of a reference static solution and a perturbation. In the Boussinesq approximation, the reference state is chosen hydrostatic with a constant temperature $T_0$ and a constant density $\rho_0$. We write

$$P = \overline{P} + P' \qquad , \qquad T = T_0 + T' \qquad \text{and} \qquad \rho = \rho_0 + \rho'. \tag{1.8}$$

with $\overline{P}$ the hydrostatic pressure and $P'$, $T'$ and $\rho'$ the perturbations for pressure, temperature and density. The gravity can also be decomposed into a one-dimensional (1D)

profile $\overline{\mathbf{g}}$ and a perturbation $\mathbf{g}'$: $\mathbf{g} = \overline{\mathbf{g}} + \mathbf{g}'$. Following Equation 1.3, we have:

$$\overline{\mathbf{g}} = -\nabla\overline{\Psi} \qquad \text{and} \qquad \nabla^2\overline{\Psi} = 4\pi G\rho_0, \tag{1.9}$$

$$\mathbf{g}' = -\nabla\Psi' \qquad \text{and} \qquad \nabla^2\Psi' = 4\pi G\rho'. \tag{1.10}$$

The hydrostatic pressure $\overline{P}$ is

$$\nabla\overline{P} = \rho_0\overline{\mathbf{g}}. \tag{1.11}$$

We further suppose that the density perturbation depends only on temperature. In the hypothesis that $\rho' \ll \rho_0$, we obtain the equation of state

$$\rho' = -\rho_0\alpha T'. \tag{1.12}$$

with $\alpha$ the constant expansivity.

Compared to the anelastic liquid approximation, the Boussinesq approximation does not take into account compressibility and its main effect on temperature: the presence of a vertical adiabatic temperature gradient. Although the adiabatic temperature difference across the mantle is not negligible, the nonadiabatic temperature difference is high enough to be responsible for most of mantle dynamics. For the energy conservation equation (Eq. 1.7c), it means that the terms $\alpha T\frac{\mathrm{D}P}{\mathrm{D}t}$ and $\nabla\cdot(\boldsymbol{\tau}\cdot\mathbf{u})$ are negligible in the Boussinesq approximation (Bercovici et al., 1992; Glatzmaier, 1988; Jarvis & McKenzie, 1980; Tackley, 1996).

This leads to the Boussinesq equations

$$\nabla \cdot \mathbf{u} = 0 \tag{1.13a}$$

$$\rho_0 \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla P' + \nabla \cdot \left[\mu \left(\nabla \mathbf{u} + [\nabla \mathbf{u}]^T\right)\right] - \rho_0 \alpha T' \overline{\mathbf{g}} + \mathbf{g}' \rho_0, \tag{1.13b}$$

$$\text{with} \quad \overline{\mathbf{g}} = -\nabla \overline{\Psi} \qquad \text{and} \qquad \nabla^2 \overline{\Psi} = 4\pi G \rho_0,$$

$$\mathbf{g}' = -\nabla \Psi' \qquad \text{and} \qquad \nabla^2 \Psi' = 4\pi G \rho',$$

$$\rho_0 C_p \frac{\mathrm{D}T'}{\mathrm{D}t} = \nabla \cdot (k \nabla T') + \rho_0 H. \tag{1.13c}$$

The term $\mathbf{g}'\rho_0$, related to self-gravitation was shown to only affect the computation of gravity, but not the geometry of the flow (Zhong et al., 2008): we will omit it in the following.

## Further Simplifications and Nondimensionalization

We suppose $\overline{\mathbf{g}}$, $C_p$, $k$ and $H$ constant. The gravity is written $\overline{\mathbf{g}} = -g\mathbf{e}_r$ with $\mathbf{e}_r$ the radial unit vector.

We use four characteristic scales for the nondimensionalization:

- the length $D = 2900\ km$, corresponding to mantle thickness,
- the temperature difference $\Delta T = 1600\ K$, the temperature drop across the boundary layer (the lithosphere),
- the thermal diffusion time $\tau_D = \frac{D^2 \rho_0 C_p}{k} = 265$ billion years, i.e. the time scale of diffusion across the whole mantle,
- the reference viscosity $\mu_0 = 10^{23}\ Pa.s^{-1}$, which correspond to the viscosity at the mid-depth in the mantle,

and perform the following change of variables in the Boussinesq equations

$$T' = T^* \Delta T \ , \ t = t^* \tau_D \ , \ \mathbf{u} = \mathbf{u}^* D/\tau_D, \tag{1.14}$$

$$\mu = \mu^* \mu_0 \ , \ P' = P^* \mu_0/\tau_D \text{ and } \nabla = \nabla^*/D. \tag{1.15}$$

This gives the nondimensional set of equations

$$\nabla^* \cdot \mathbf{u}^* = 0, \tag{1.16a}$$

$$\frac{1}{\mathrm{Pr}} \frac{\mathrm{D}\mathbf{u}^*}{\mathrm{D}t^*} = -\nabla^* P^* + \nabla^* \cdot \left[\mu^* \left(\nabla^* \mathbf{u}^* + [\nabla^* \mathbf{u}^*]^T\right)\right] + \mathrm{Ra}\mathbf{e}_r T^*, \tag{1.16b}$$

$$\frac{\mathrm{D}T^*}{\mathrm{D}t^*} = \nabla^{*2} T^* + \mathrm{R}_h, \tag{1.16c}$$

where $\mathrm{Ra}$ and $\mathrm{Pr}$ are the Rayleigh and Prandtl numbers (John William Strutt, 3rd Baron Rayleigh, 1842-1919; Ludwig Prandtl, 1875-1953), defined as

$$\mathrm{Ra} = \frac{\rho_0^2 C_p g \alpha \Delta T D^3}{\mu_0 k} \qquad \text{and} \qquad \mathrm{Pr} = \frac{\mu_0 C_p}{k}. \tag{1.17}$$

For the mantle, $\mathrm{Ra}$ is estimated to be of the order of $10^7$ and $\mathrm{Pr}$ of the order $10^{24}$. The very high Prandtl number implies that the left-hand side of equation 1.16b is almost zero: the mantle does not have any inertia.

$\mathrm{R}_h$ is the nondimensional rate of energy production

$$\mathrm{R}_h = \frac{\rho_0 H D^2}{k \Delta T}. \tag{1.18}$$

Finally, we obtain the nondimensional equations of conservation of mass, momentum and energy that is used in this thesis to model mantle convection:

$$\nabla \cdot \mathbf{u} = 0, \tag{1.19a}$$

$$\nabla P - \nabla \cdot \left[\mu \left(\nabla \mathbf{u} + [\nabla \mathbf{u}]^T\right)\right] = \mathrm{Ra} T \mathbf{e}_r, \tag{1.19b}$$

$$\frac{\mathrm{D}T}{\mathrm{D}t} = \nabla^2 T + \mathrm{R}_h, \tag{1.19c}$$

To solve this set of equation, we need to specify the domain on which they are to be solved as well as the initial and boundary conditions. We also need to specify a constitutive equation for viscosity. These are the topics of the following two sections.

## 1.3 Initial and Boundary Conditions

The volume of the mantle, noted $\Omega$, is at first order a spherical shell of inner radius $3480$ km and outer radius $6371$ km (if we include the crust). We denote its outer boundary with $\Gamma_a$ and its inner boundary with $\Gamma_b$. The natural system to locate a point on such a geometry is spherical coordinates. We use the physical convention: the position vector $\mathbf{x}$ is defined by its colatitude $\phi$, its longitude $\theta$ and its radius $r$ and $\mathbf{e}_\phi$, $\mathbf{e}_\theta$, $\mathbf{e}_r$ are the corresponding unit vectors.

The upper boundary $\Gamma_a$ is in contact with air and water (dynamic viscosity $\eta = 10^{-5}$ and $10^{-3}$ Pa.s, respectively) and the lower boundary $\Gamma_b$ is in contact with the liquid core (dynamic viscosity of the order $10^{-3}$ Pa.s). The drastic drop of viscosity along these boundaries implies free surface conditions, with a deformable surface. The topography of the boundaries are very small compared to the length scale of convection so it is generally neglected in global mantle convection models. Mechanical boundary conditions are then free-slip with rigid, perfectly spherical and impermeable boundaries $\Gamma_a$ and $\Gamma_b$

$$\forall t,\ \forall \mathbf{x} \in \Gamma_b,\quad \mathbf{u}(\mathbf{x},t)\cdot\mathbf{e}_r = 0 \qquad \text{and} \qquad \boldsymbol{\sigma}(\mathbf{x},t)\cdot\mathbf{e}_r = \mathbf{0} \tag{1.20a}$$

$$\forall t,\ \forall \mathbf{x} \in \Gamma_a,\quad \mathbf{u}(\mathbf{x},t)\cdot\mathbf{e}_r = 0 \qquad \text{and} \qquad \boldsymbol{\sigma}(\mathbf{x},t)\cdot\mathbf{e}_r = \mathbf{0} \tag{1.20b}$$

However, regional studies have shown the importance of taking into account the deformation of the surface to obtain a surface dynamics closer to the Earth, especially in the context of subduction (Gerya, 2011). Recently, a global mantle convection model with a deformable surface has been developed (Crameri et al., 2012; Crameri & Tackley, 2014, 2015) but its high numerical cost limits its application for now.

An alternative mechanical boundary condition would be free slip at $\Gamma_b$ and imposed velocities at $\Gamma_a$

$$\forall t,\ \forall \mathbf{x} \in \Gamma_b,\ \ \mathbf{u}(\mathbf{x},t)\cdot\mathbf{e}_r = 0 \qquad \text{and} \qquad \boldsymbol{\sigma}(\mathbf{x},t)\mathbf{e}_r = \mathbf{0}, \tag{1.21a}$$

$$\forall t,\ \forall \mathbf{x} \in \Gamma_a,\ \ \mathbf{u}(\mathbf{x},t)\cdot\mathbf{e}_r = 0 \qquad \text{and} \qquad \mathbf{u}_H(\mathbf{x},t) = \mathbf{u}_H^a(\mathbf{x},t), \tag{1.21b}$$

with $\mathbf{u}_H$ the horizontal velocity. This boundary condition is used to force the surface dynamics, and we will discuss it in the next chapter on mantle circulation reconstructions.

Both boundaries $\Gamma_a$ and $\Gamma_b$ are supposed isothermal with temperature $T_a$ and $T_b$, respectively

$$\forall \mathbf{x} \in \Gamma_b,\ \ T(\mathbf{x}) = T_b, \tag{1.22a}$$

$$\forall \mathbf{x} \in \Gamma_a,\ \ T(\mathbf{x}) = T_a. \tag{1.22b}$$

The only prognostic equation is the energy conservation Eq.1.19c so the initial condition is entirely defined by an initial temperature field $T_i$

$$\forall \mathbf{x} \in \Omega,\ \ T(\mathbf{x}, t=0) = T_i(\mathbf{x}). \tag{1.23}$$

## 1.4 Rheology and Self-Consistent Plate Generation

Mantle material is polycrystalline. Its response to stress depends on various mechanisms of deformation at the crystal scale (Karato, 2013; Kohlstedt, 2015). The combination of all these mechanisms results in a complex and nonlinear rheology which is key to understanding mantle dynamics (Turcotte & Oxburgh, 1972; Davies & Richards, 1992; Bercovici, 2015).

In section 1.1, we supposed a relationship between the deviatoric stress tensor $\boldsymbol{\tau}$ and the strain rate tensor $\dot{\boldsymbol{\epsilon}}$ of the form

$$\boldsymbol{\tau} = 2\mu\dot{\boldsymbol{\epsilon}}, \tag{1.24}$$

with $\mu$ the viscosity and $\dot{\boldsymbol{\epsilon}} = \left(\nabla\mathbf{u} + [\nabla\mathbf{u}]^T\right)/2$. This formulation can take into account a wide range of deformation mechanisms by considering a composite rheology, where the effective viscosity $\mu$ is

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + ... \tag{1.25}$$

with $\mu_1$, $\mu_2$... the effective viscosities of each mechanism, potentially varying with temperature, pressure, strain rate, grain size and composition.
In this dissertation, we focus on the relationship between surface tectonics and mantle convection, so we limit the discussion on rheology to the question:

**what type of viscosity law generates Earth-like tectonics at the surface of convection models?**

One of the most important characteristics of mantle rheology is the dichotomy between a strong lithosphere and the rest of the mantle, which is softer. This is primarily due to the thermal activation of crystal deformation (Gordon, 1967; Weertman & Weertman, 1975), modelled by a viscosity following an Arrhenius law

$$\mu \propto \exp\left(\frac{E_A + PV_A}{T}\right), \tag{1.26}$$

with $E_A$ an activation energy and $V_A$ an activation volume. Both parameters set the amplitude of viscosity variations in the mantle and affect the style of convection. Systematic studies of convection with a temperature dependent rheology show three possible convective regimes, depending on the viscosity contrast between the surface and the bottom of the system (Christensen, 1984; Solomatov, 1995; Ratcliff et al., 1997). Low viscosity contrasts produce mobile-lid convection in which the flow resembles isoviscous convection. With a high viscosity contrast (greater than $10^4$), the convection is

limited to the lower part of the system, and the surface is frozen: this regime is called stagnant-lid convection. An intermediate regime exists, where the surface is sluggish.

Experimental studies, crystal deformation theory, and geophysical observations indicate an average viscosity profile with a strong lithosphere ($\mu > 10^{25}$ Pa.s), and a softer underlying mantle with a viscosity ranging from $10^{19}$-$10^{21}$ Pa.s in its upper part to $10^{22}$-$10^{23}$ Pa.s in its lower part (King, 2016). Consequently, the Earth's mantle convection regime should be stagnant-lid, if its viscosity was only temperature and pressure dependent. The presence of mobile plates at the surface of the Earth, and the high deformation observed at plate boundaries proves otherwise.

It is possible to obtain plate-like dynamics by artificially introducing weak zones in the top boundary layer of a stagnant-lid convective system (Christensen & Yuen, 1984; Gurnis & Hager, 1988; Puster et al., 1995; Zhong et al., 2000). Zhong et al. (1998) identify two diagnostics to evaluate how Earth-like the surface dynamics of a model is. Plateness (first defined in 2-D by Weinstein & Olson, 1992) quantifies the localization of deformation at the surface of a dynamic system. The plateness is maximum (generally set to one) when the surface is made of perfectly rigid blocks moving along infinitely small boundaries that concentrate surface deformation. Conversely, the plateness is low when the deformation is distributed uniformly at the surface. The plateness of the Earth is very high: Gordon (2000) estimates that the deformation at the surface of the Earth is concentrated in less than $15\%$ of its total area. The second diagnostic is the ratio of surface toroidal to poloidal motion. Any divergence free 3-D velocity field can be decomposed into a toroidal and a poloidal flow. At the surface, the poloidal flow creates divergence or convergence motions whereas the toroidal flow is associated with plate spin and strike-slip motion. On Earth, the ratio of surface toroidal to poloidal flow varies between $0.25$ to $0.5$ for the last $120$ Myr (Olson & Bercovici, 1991; Lithgow-Bertelloni et al., 1993). Models with imposed weak zones are a valuable tool to study the evolution of preexisting plates and plate boundaries, but they cannot create new plate boundaries.

To generate weak zones at the surface of models self-consistently, it is necessary to take

into account a nonlinear relationship between stress and strain rate. However, it is still not clear which relationship best describes lithosphere rheology, so various approaches coexist (see reviews of Tackley, 2000a; Lowman, 2011; Bercovici, 2003; Bercovici et al., 2015). Figure 1.1 summarizes the different relationships tested over the past 40 years. The first models proposed a power-law relationship between stress and strain rate

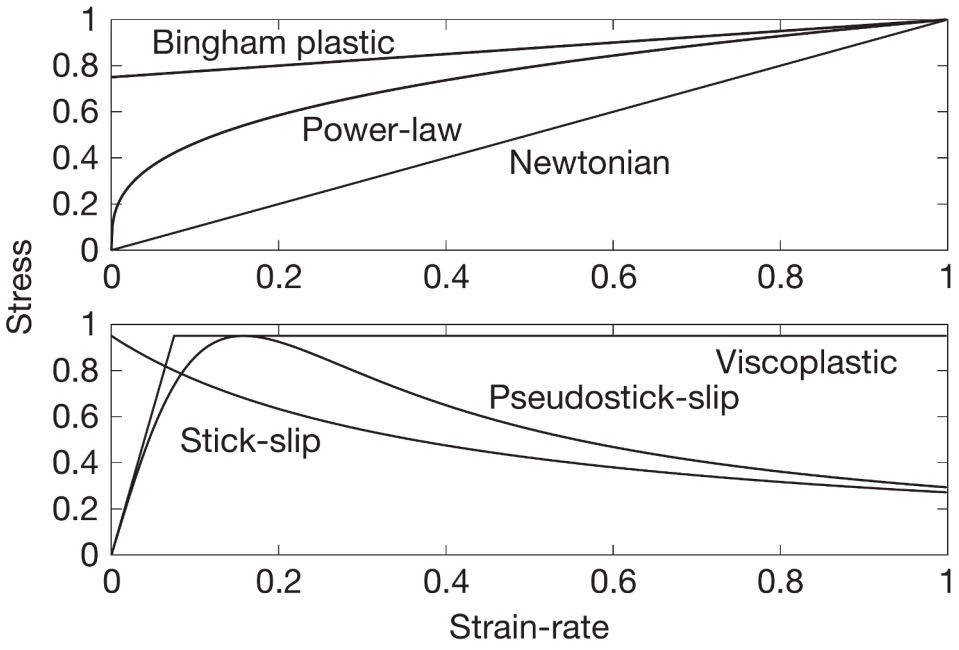


FIGURE 1.1: Curves of stress versus strain-rate for Newtonian, non-Newtonian power-law rheology (n=3), Bingham plastic, visco-plastic, stick-slip and pseudo-stick-slip rheologies. From Bercovici et al. (2015).

(Parmentier et al., 1976; Christensen, 1984). However, they produce very little plateness when the power-law index is 3, which is typically what is measured for mantle rocks (Weertman & Weertman, 1975; Evans & Kohlstedt, 1995). An increase of the power-law index leads to better localization of the deformation in narrow weak zones (Weinstein & Olson, 1992), but 3-D models yield only moderate toroidal motion (Christensen & Harder, 1991; Bercovici, 1993, 1995; Čadek et al., 1993). To enhance the localization of deformation, Bercovici (1993, 1995) proposed a power-law rheology with index -1 (pseudo-stick-slip in Figure 1.1), which would be the steady-state equivalent of a more complex, time dependent rheology (Tackley, 2000c). This type of rheology produces strain-rate weakening: at low strain-rates, the material behaves like a very viscous fluid but when the strain rate increases above a given threshold, the viscosity decreases so

drastically that it causes a decrease in stress. 3-D tests produce encouraging plateness and toroidal to poloidal ratio (Tackley, 1998a). However, such a rheology is difficult to implement in 3-D mantle convection models because of its high nonlinearity (Tackley, 2000c). Viscoplastic rheology (see figure 1.1) is generally preferred. In this formulation, the material behaves like a Newtonian viscous material when the stress is low. When the stress reaches a given threshold, called the yield stress $\sigma_{\text{Y}}$, the viscosity follows

$$\mu = \frac{\sigma_{\text{Y}}}{2\dot{\epsilon}_{\text{II}}}, \tag{1.27}$$

where $\dot{\epsilon}_{\text{II}} = \sqrt{\dot{\epsilon}_{ij}\dot{\epsilon}_{ij}}$ (in Einstein notation) is the second invariant of the strain rate tensor. This type of rheology aims at modelling the brittle and ductile deformation processes that take place in the lithosphere (Moresi & Solomatov, 1998; Trompert & Hansen, 1998). When combined with an asthenosphere of reduced viscosity, they produce plateness and toroidal to poloidal ratios comparable to that of the Earth (Tackley, 2000c; Richards et al., 2001; Stein et al., 2004). The progresses of numerical methods (e.g. Yoshida & Kageyama, 2004; Tackley, 2008; Burstedde et al., 2013) and the increase of computational power led to a series of models in 3-D spherical shell geometry self-consistently generating plate-like tectonics at their surface (Richards et al., 2001; Walzer & Hendel, 2008; Van Heck & Tackley, 2008; Yoshida, 2008; Foley & Becker, 2009). Figure 1.2 shows a snapshot of the surface and inner structure of a typical plate-like evolution produced by this type of models. Although these models display Earth-like plateness and toroidal:poloidal ratio, they have several major flaws. Firstly, the range of yield stresses ($\sigma_{\text{Y}}$) that produces plate-like tectonics is very limited (Foley & Becker, 2009) and does not correspond to estimated values of the strength of the lithosphere (Kohlstedt et al., 1995). However, the introduction of continents (Rolf & Tackley, 2011), melting (Lourenço et al., 2016) and a free surface (Crameri, 2013) tends to increase the yield stress at which plate tectonics appears on models. Secondly, during subduction, both sides dive into the mantle, whereas on Earth, subduction is exclusively single-sided (Gerya et al., 2008). Free surface boundary conditions give encouraging results for the development of single-sided subduction in such models (Crameri & Tackley,

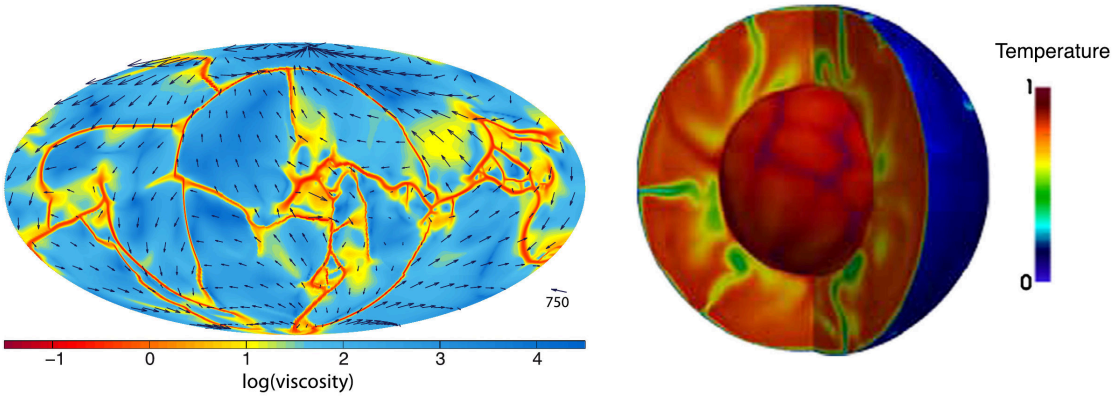


FIGURE 1.2: Left: surface velocity and viscosity fields snapshots of a free (i.e. unconstrained by data) mantle convection evolution with composite viscoplastic and temperature-pressure Newtonian rheology. Right: corresponding temperature field of the interior of the system. From Bello et al. (2015)

2014), especially combined with a weak-crust layer (Crameri & Tackley, 2015). Thirdly, the strike-slip boundaries are very diffuse, contrary to the Earth.

Several further developments have been proposed to obtain a surface dynamics closer to the Earth, such as taking into account the viscosity variations with the concentration of fluids in rocks, grain size and the deformation history (Bercovici et al., 2015). However, for the time being, a composite rheology with a Newtonian component of the form of equation 1.26, combined with a plastic behavior (Eq.1.27) remains the best option to study global convection with Earth-like tectonics. Recent applications of such models include comparison of seafloor age distributions and the time scale of spreading fluctuations in models and on Earth (Coltice et al., 2012, 2013), studies on the supercontinent cycle (Rolf et al., 2014), and on the distribution of plate sizes (Mallard et al., 2016).

## 1.5 Numerical Approximation

The final system of equations to solve is composed of three conservation equations

$$\nabla \cdot \mathbf{u} = 0, \tag{1.28}$$

$$-\nabla P + \nabla \cdot \left[\mu \left(\nabla \mathbf{u} + [\nabla \mathbf{u}]^T\right)\right] = \mathrm{Ra} T \mathbf{e}_r, \tag{1.29}$$

$$\frac{\mathrm{D}T}{\mathrm{D}t} = \nabla^2 T + \mathrm{R}_h, \tag{1.30}$$

a constitutive law for viscosity, which is a nonlinear function of temperature, pressure and stress

$$\mu = \left(\frac{1}{\mu_n} + \frac{1}{\mu_y}\right)^{-1}, \tag{1.31}$$

$$\text{with } \mu_n \propto \exp\left(\frac{E_A + PV_a}{T}\right), \tag{1.32}$$

$$\text{and } \mu_y = \frac{\sigma_Y}{2\dot{\epsilon}_{\mathrm{II}}}, \tag{1.33}$$

and the initial and boundary conditions

$$\forall \mathbf{x} \in \Omega, \;\; T(\mathbf{x}, t=0) = T_i(\mathbf{x}), \tag{1.34}$$

$$\forall \mathbf{x} \in \Gamma_a \cup \Gamma_b, \;\; \mathbf{u}_r(\mathbf{x}) = \mathbf{0} \qquad \text{and} \qquad \boldsymbol{\sigma}\mathbf{e}_r = \mathbf{0}, \tag{1.35}$$

$$\forall \mathbf{x} \in \Gamma_a, \;\; T(\mathbf{x}, t) = T_a \qquad \text{and} \qquad \forall \mathbf{x} \in \Gamma_b \;\; T(\mathbf{x}, t) = T_b. \tag{1.36}$$

This set of equations (or a variation of it, discussed in the previous sections) is discretized in time using finite differences and in space using finite difference (Kageyama & Sato, 2004), finite volume (Harder & Hansen, 2005; Tackley, 2008) or finite element (Baumgardner, 1985; Zhong et al., 2008; Burstedde et al., 2013; Kronbichler et al., 2012) methods. Those grid-based methods are preferred to spectral methods, which, although in principle more accurate and more efficient, have so far not been able to treat efficiently large horizontal viscosity variations (Balachandar et al., 1995; Zhang & Yuen, 1996).

Regardless of the technique, the evolution of a solution is computed by repeating two stages:

1. **Flow solution**: at a given time step, the velocity and pressure fields are computed by solving Equations 1.28 and 1.29, for a given temperature field.
2. **Time stepping**: the temperature field at the next time step is computed by solving Equation 1.30, using the former temperature and velocity solution.

(Ismail-Zadeh & Tackley, 2010 provide a detailed description of numerical methods for geodynamic modelling, and Zhong et al., 2015 review the recent advances of mantle convection numerical modelling).

In this dissertation, we use the finite volume code STAGYY (Tackley, 2008), which is an extension of STAG3D, continuously developed since 1992 (Tackley et al., 1993; Tackley, 1996, 1998b, 2002; Tackley & Xie, 2003). Recent developments of this code include the implementation of continents (Rolf & Tackley, 2011), free surface with the sticky air method (Crameri et al., 2012) and crust production by melting (Lourenço et al., 2016). This code has been extensively tested with the mantle convection community benchmarks from Blankenbach et al., 1989 (2-D), Busse et al., 1994 (3D Cartesian) , Van Keken et al., 1997 (thermo-chemical convection), Stemmer et al., 2006 (spherical geometry), Tosi et al., 2015 (viscoplasticity) and showed good agreement with those (Tackley et al., 1993; Tackley, 1996; Tackley & King, 2003; Tackley, 2008). Additionally, Tackley (2000c) performed a convergence test for models with self-consistent generation of plates. STAGYY is parallelized by domain decomposition, using the MPI message-passing library, and shows good scalability (Tackley, 2008).

The capabilities of this code go far beyond the simple model setup that we use in this dissertation. We chose a relatively simple setup to focus on the development of methods of reconstruction of mantle circulation, but it was important to develop a mantle circulation reconstruction framework on a code which allows more complexities, for future applications.

## Conclusion

In this chapter, we presented the solution for the direct problem of mantle circulation: we showed how we compute an evolution from a mantle convection model featuring plate-like tectonics at its surface and an initial temperature field. However, the initial temperature field of the Earth's mantle is unknown. Knowledge of past and present mantle circulation come from indirect sources of information such as seismic tomography, for the 3D current structure of the mantle and plate tectonics reconstructions, for past mantle circulation. This makes the reconstruction of mantle circulation an inverse problem. The aim of the next chapter is to describe this inverse problem, review methods used to tackle it, and present data assimilation methods, that we apply in chapter 4 and 5.

## Résumé du Chapitre

Résoudre le problème direct de la circulation dans le manteau consiste à spécifier comment calculer une évolution de la circulation dans le manteau à partir d'un état initial donné. Pour ce faire, nous utilisons les principes de dynamique des fluides. Après avoir justifié l'utilisation de la dynamique des fluides pour calculer la circulation dans le manteau, nous dérivons les équations aux dérivées partielles qui régissent l'écoulement, puis nous spécifions les conditions initiales et aux limites nécessaires pour résoudre ces équations. Une section supplémentaire est consacrée à la rhéologie du manteau, qui joue un rôle majeur dans sa dynamique. Enfin, nous discutons de l'approximation numérique de ce problème.

Dans ce chapitre, nous avons présenté la solution pour le problème direct de la circulation dans le manteau: nous avons montré comment nous calculons une évolution à partir d'un modèle de convection du manteau présentant une tectonique de type plaque à sa surface et un champ de température initial. Cependant, le champ de température initial du manteau terrestre est inconnu. La connaissance de la circulation dans le manteau passé et présent provient de sources indirectes d'information telles que la tomographie sismique, pour la structure de l'écoulement en 3D au présent et les reconstructions de la tectonique des plaques, pour la circulation passée dans le manteau. Ceci fait de la reconstitution de la circulation mantellique un problème inverse. Le but du chapitre suivant est de décrire ce problème inverse, d'examiner les méthodes utilisées pour le résoudre et de présenter les méthodes d'assimilation de données, que nous appliquons aux chapitres 4 et 5.

# CHAPTER 2

# Data and the Mantle Circulation Inverse Problem

The aim of a mantle circulation reconstruction is to estimate the past and present 3-D flow in the mantle, using data on mantle circulation in combination with a dynamical model of mantle convection. Data on mantle circulation is available both at present, with the constraints provided by seismic tomography for example, and in the past, with constraints on the surface tectonics history, for example. Inverse problems dealing with asynchronous data fall in the category of data assimilation problems. This chapter aims at presenting the data assimilation problem for mantle circulation. We start by formulating the mantle circulation problem in the data assimilation framework (first section). Then we detail the data available on the history of mantle circulation (second section). The third section finally presents the applications to the reconstruction of past mantle flow.

## 2.1 A Data Assimilation Framework for Mantle Circulation Problems

Data assimilation methods have originally been developed as a way to estimate the state of the atmosphere at present, so as to obtain an initial state for numerical weather predictions (see for example Ghil, 1989, for a short account of early data assimilation developments in weather forecasting). The improvement of both numerical weather models and the observational network, along with the increase in computational power, led to the evolution from subjective analysis of sparse observations, to objective analysis of a denser network of observations and finally to the assimilation of data not only spatially, but also through time, i.e. 4-D data assimilation (Charney et al., 1969; Smagorinsky et al., 1970; Rutherford, 1972). Atmospheric data assimilation borrows many concepts from control and estimation theory (see for example Ghil et al., 1981; Le Dimet & Talagrand, 1986). The main difference is in the larger size of the problem and dataset used in atmospheric data assimilation, which pushed the development of specific techniques. The application of data assimilation methods has been progressively opening to various geophysical systems, following the development of observational networks and numerical models in the corresponding disciplines. Current applications include oceanography (Chassignet & Verron, 2006), geomagnetism (Fournier et al., 2010), oil and gas reservoir (Evensen, 2009a, Chapter 17), and glaciology (Bonan et al., 2014), for example. In parallel, the goals of data assimilation methods have diversified: the original aim was to estimate the current state of a system to forecast future evolutions, but other applications have emerged, such as reanalysis of past data (hindcasting), sensitivity analysis of models to different parameters and initial and boundary conditions, assessment of the numerical model, and parameter estimation. Concerning mantle convection, data assimilation techniques provide powerful tools to investigate fundamental questions such as:

- what is the predictive power of state of the art mantle convection models, or how well do we model mantle dynamics?

- what was the circulation of the mantle in the past?
- how sensitive are mantle circulation models to the variation of key parameters, such as rheological ones?

The integration of data in dynamical models is a longstanding problem in global mantle convection studies (Hager & O'Connell, 1979) and in geodynamics in general (see for example the introductory book by Ismail-Zadeh et al., 2016). One of the ways to solve this problem has been to apply data assimilation methods to mantle circulation reconstructions. This approach was pioneered by Bunge et al. (2003) and Ismail-Zadeh et al. (2004) who developed variational data assimilation methods for mantle circulation reconstructions. In these articles, they present the data assimilation problem of mantle circulation reconstruction in its continuous form (using a set of partial differential equations similar to the one described in Section 1.2 and associated continuous fields). Here, we choose a different approach and formulate the data assimilation problem directly on the discretized version of mantle flow equations. We use the standard data assimilation formulation advised by Ide et al. (1997) and define the unknown (the state of the mantle), the sources of information (observations, background state and dynamical model) and how each of these components is related to the others.

### 2.1.1 The Unknown: the Evolution of the True State of the Mantle

The true state gathers all the variables that are necessary and sufficient to compute the 3-D flow of the Earth's mantle at a given time.

In the continuum approximation (see section 1.1), the evolution of mantle flow is described by a set of fields varying continuously in space and time. However, the computation of a dynamical evolution of those fields requires their discretization (see section 1.5). So, in practice, we aim at estimating the evolution of the discretized state of the mantle. In this case, the true state can take the form of a vector containing the values of the discretized fields at a given time step $k$. This vector is named the true state vector and noted $\mathbf{x}_k^t$. Then, we write the evolution of the true state as a time series

of true state vectors

$$\left\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_K^t\right\}, \tag{2.1}$$

with indices $\{1, ..., K\}$ referring to the numbers of the time steps.

The type and number of discretized fields included in the true state depend on the model chosen for mantle convection. For the model described in Chapter 1, the 3-D flow of the mantle at a given time is entirely defined from the sole knowledge of the temperature field. In this case, the true state vector contains the temperature values of the mantle at points defined on the numerical grid. For a model taking into account the variations of composition in the mantle, we would have to add the discretized compositional fields, for example.

### 2.1.2 Observed Data on Mantle Circulation

The set of observed data on mantle circulation is organized in the form of a series of observed data vectors

$$\{\mathbf{y}_1^o, \mathbf{y}_2^o, ..., \mathbf{y}_K^o\} \tag{2.2}$$

where $\mathbf{y}_k^o$ contains the data observed at the time step $k \in \{1, 2, ..., K\}$.

At each time step $k$, there is a function that associates a data vector to the set of continuous fields defining the state of the system. This function is approximated by a mapping $\mathcal{H}_k$ between the discretized state space and the data space. $\mathcal{H}_k$ is called the observation operator. For any state vector $\mathbf{x}_k$, the corresponding data vector $\mathbf{y}_k$ is given by

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \boldsymbol{\epsilon}_k^h \tag{2.3}$$

where $\boldsymbol{\epsilon}_k^h$ represents the errors caused by the hypotheses used to define $\mathcal{H}_k$ (e.g. discretization and description of the physics of the measurement).

Moreover, the observed data are affected by measurement errors, so the relation between the true state and the observed data at time step $k$ is

$$\mathbf{y}_k^o = \mathcal{H}_k(\mathbf{x}_k^t) + \boldsymbol{\epsilon}_k^o + \boldsymbol{\epsilon}_k^h \tag{2.4}$$

where $\boldsymbol{\epsilon}_k^o$ is the error vector of observed data at time step $k$.

In mantle circulation problems, the term "observation" is to be understood in a broad sense. It is common to integrate raw observations into a model to obtain global estimates of variables that are directly linked to mantle circulation. We then use these global estimates as "observations" in the data assimilation procedure. We will discuss this point in the following sections.

### 2.1.3 The Dynamical Model

The dynamical model $\mathcal{M}$ predicts the evolution of the state vector from one time step to the next. An example of $\mathcal{M}$ for mantle convection has been derived in chapter 1. As highlighted and discussed throughout chapter 1, the dynamical model $\mathcal{M}$ is a numerical approximation of an imperfect physical model describing the evolution of the continuous state of the system through time. From a state vector $\mathbf{x}_k$, we can write the state vector at time step $k+1$ as

$$\mathbf{x}_{k+1} = \mathcal{M}(\mathbf{x}_k) + \boldsymbol{\eta}_k, \tag{2.5}$$

with $\boldsymbol{\eta}_k$ the modelling errors due to discretization and simplifications in the physical modelling.

During this thesis, we used the numerical code STAGYY (see section 1.5) with a set of options and parameters described in the chapters 3 and 4 as the model $\mathcal{M}$.

### 2.1.4 The Background State of the Mantle

The background state of the mantle corresponds to the estimation of the state of the mantle before the observed data have been integrated. This a priori state comes from our knowledge of the physics of mantle convection, or the integration of another set of observed data, for example. The background state is linked to the first true state vector through

$$\mathbf{x}^b = \mathbf{x}_1^t + \boldsymbol{\epsilon}^b, \tag{2.6}$$

where $\boldsymbol{\epsilon}^b$ is the error vector of the background state vector.

In this work, we used the results of a very long evolution computed with the dynamical model $\mathcal{M}$ to infer the average temperature field in the mantle and evaluate the amplitude of the background error vector (see section 3.4.1 for more details).

## Summary

We have described the components of the general problem of mantle circulation reconstructions:

- the unknown: a time series of true state vectors $\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_K^t\}$
- three sources of information, all affected by errors:
  - the background state $\mathbf{x}^b$
  - the time series of observed data $\{\mathbf{y}_1^o, \mathbf{y}_2^o, ..., \mathbf{y}_K^o\}$ and their associated observation operators $\{\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}_K\}$
  - the dynamical model $\mathcal{M}$.

In the next section, we focus on the description of the observed data currently used to reconstruct mantle circulation. Then, in Section 2.3 we describe the algorithms used to solve the mantle circulation inverse problem.

## 2.2 Data on Mantle Circulation

Currently, two types of "observed data" are used in dynamical mantle circulation reconstructions:

- a temperature field of the mantle at present, which is derived from seismic tomography,
- a time series of surface velocities, which is the synthesis of a wide range of observations on the history of the surface of the Earth.

In the following, we detail the procedure used to obtain these data, and discuss the possible errors that affect them.

### 2.2.1 Mantle Temperature Field at Present

If we consider purely thermal convection, the density at a given depth is solely dependent on temperature. For example, in the model developed in chapter 1, the density anomaly $\rho'$ is linked to the temperature anomaly $T'$ through

$$\rho' = -\rho_0 \alpha T'. \tag{2.7}$$

with $\alpha$ the expansivity and $\rho_0$ the reference density. Hence, to obtain an estimate of the temperature field in the mantle at present, we need:

- a reference temperature profile (In the Boussinesq approximation, it is simply a constant temperature, see Section 1.2)
- a 3-D model of density anomaly

This model of density anomaly is derived from wave speed anomaly models inferred from seismic tomography.

**Seismic Tomography**

Dziewonski & Anderson (1981) computed the first widely accepted reference profile of the elastic properties of the Earth: the Preliminary Reference Earth Model (PREM). In particular, the PREM provides estimates of the velocity of compressional and shear waves (P- and S-waves), and density in the Earth as a function of depth. With these estimates it is possible to predict the travel time of any seismic wave for a given earthquake and a given seismic station. In practice, the predicted travel times do not fit exactly the observed travel times. Such discrepancies are due to the lateral heterogeneity of Earth's material elastic properties, interpreted as variations of either temperature or composition: the S-wave velocity $v_s$ and the P-wave velocity $v_p$ are

$$v_s = \sqrt{\frac{G}{\rho}} \qquad \text{and} \qquad v_p = \sqrt{\frac{K + \frac{4}{3}G}{\rho}} \tag{2.8}$$

where $K$ , $G$ and $\rho$ are the bulk modulus, the shear modulus and the density. At a given depth, each of them varies with temperature and composition.

Seismic tomography refers to inverse methods that aim at mapping in 2-D or 3-D the lateral heterogeneity of the elastic properties of the Earth, using seismograms as data, and a forward model specifying how seismic waves propagate in the Earth. The three components of the technique, i.e. the inverse method, the forward model, and the type of data extracted from the seismograms vary depending on the target region of mapping, the type of physical properties that are to be determined, the restrictions on computational power, and the availability of data. A full review of seismic tomography techniques is beyond the scope of this dissertation, and we refer the reader to Thurber & Ritsema (2015) for an introduction to the field and to Rawlinson et al. (2010); Trampert & Fichtner (2013); Liu & Gu (2012) for recent reviews.

Seismic tomography studies of the whole mantle have focused on estimating 3-D models of S-wave speed anomalies (e.g. Gu et al., 2001; Panning & Romanowicz, 2006; Kustowski et al., 2008; Ritsema et al., 2011; Auer et al., 2014; Chang et al., 2015), P-wave speed anomalies (e.g. Boschi & Dziewonski, 2000; Li et al., 2008), or both (e.g. Su &

Dziewonski, 1997; Masters et al., 2000; Antolik et al., 2003; Houser et al., 2008; Koelemeijer et al., 2016). The resulting models of velocity anomalies differ between studies (see for example Figure 2.1 which shows shear wave velocity anomalies for different models and at different depths). Becker & Boschi (2002) compared $4$ P-wave, $7$ S-wave and $3$ S- and P-wave tomography models. After a spherical harmonic decomposition of these models, they computed correlations between each pair of models, depending on the depth and the spherical harmonic degree. For the first 8 spherical harmonic degrees, they found an average correlation for pairs of P-wave models between $0.69$ and $0.85$, and slightly higher correlations for pairs of S-wave models. The correlations tended to decrease with increasing spherical harmonic degree.

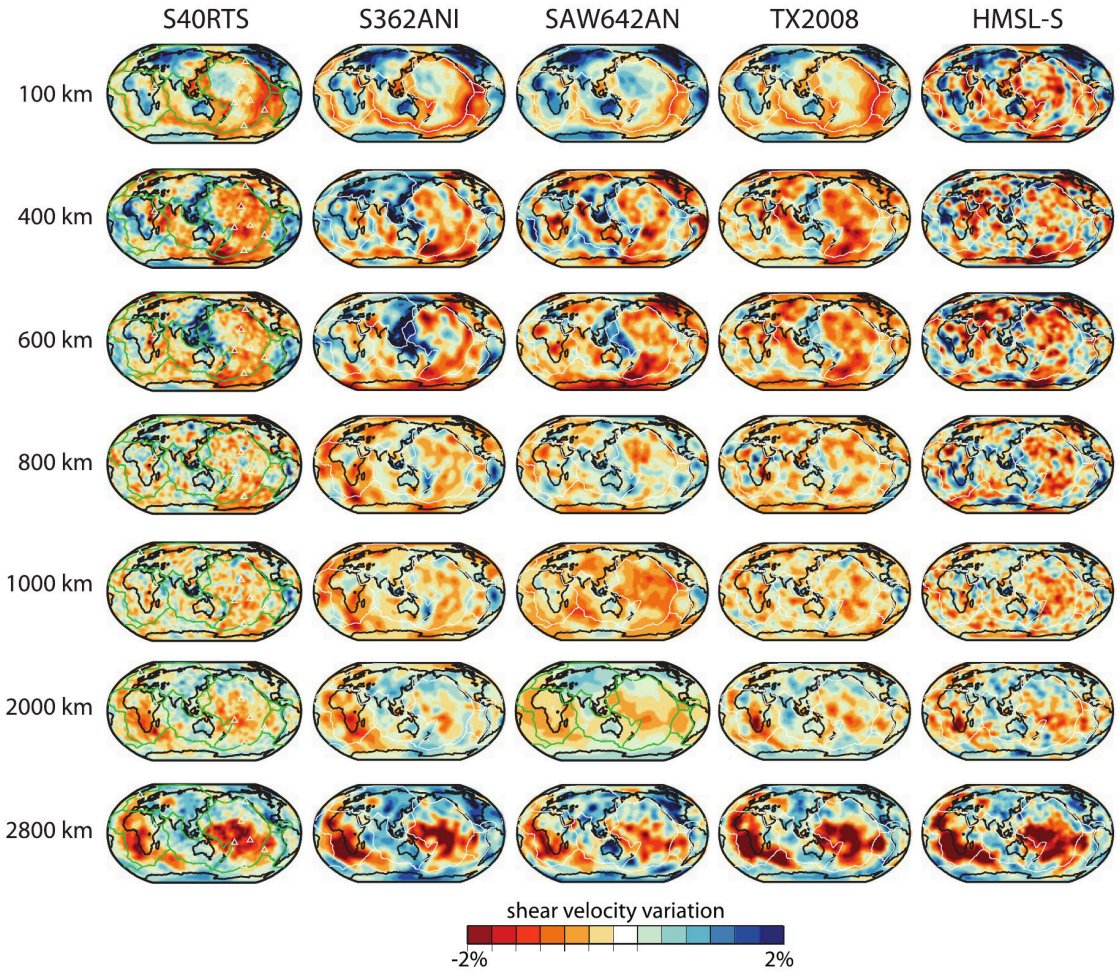


FIGURE 2.1: Maps of shear wave velocity anomalies at depth $100$, $400$, $600$, $800$, $1000$, $2000$ and $2800$ km of tomographic models (from left to right):S40RTS (Ritsema et al., 2011), S362ANI (Kustowski et al., 2008), SAW642AN (Panning & Romanowicz, 2006), TX2008 (Simmons et al., 2009), HMSL-S (Houser et al., 2008). Taken from Ritsema et al. (2011).

The differences in tomographic models are attributed to the variations in datasets, but also to the non-uniqueness of the solution to the inverse problem, which is dealt with differently depending on studies, using a wide range of inverse methods and parametrization of the forward model. The seismological community has developed several techniques to assess the reliability of tomographic models (see Rawlinson et al., 2014, for a review). A direct computation of the a posteriori error covariance matrix associated to the estimated S- and/or P-wave anomaly models would provide a much needed way to evaluate the information that could be extracted from those, and avoid geological and geodynamical misinterpretations (see Foulger et al., 2013, for a review on the pitfalls of seismic tomography interpretation for non-specialists). However, the size of the whole mantle seismic tomography problem makes the computation of this matrix very costly (and even intractable for some models). In practice, partial resolution tests are preferred (see examples in the previously cited works). Such tests do not quantify directly the absolute amplitude of errors on tomographic models. A project of 3-D Reference Earth Model for S- and P-wave speed, density, and anisotropy with the estimation of uncertainties is under construction (Moulik et al., 2016). Such model would provide a sound basis for geodynamical interpretation.

Meanwhile, several studies have adopted a more pragmatic approach to assess the robustness of the structures illuminated by seismic tomography. The idea is to compare various tomographic models in order to extract their common features, supposing that if tomographic models developed using different parametrizations, observations and inverse methods have common structures, then it is likely that these structures are not artefacts but reflect true heterogeneities of Earth's material properties. For example, Lekic et al. (2012) followed this approach and performed a cluster analysis of tomographic models of the lower mantle to infer its large scale structure. For the whole mantle, Becker & Boschi (2002) propose tomography-based benchmarks for geodynamics by computing a weighted average of available P-wave and S-wave models (called Pmean and Smean). They regularly update these models taking into account new tomographic studies (Becker & Boschi, 2016).

**Conversion of Wave Speed Anomaly Models to Temperature Fields**

Equation 2.8 shows that S- and P-wave speeds depend on three material properties: the bulk modulus $K$, the shear modulus $G$ and the density $\rho$. These material properties vary with temperature, pressure and composition. The first order relationship between a lateral relative density anomaly $\delta\rho/\rho = \delta \ln \rho$ and a lateral relative S- or P-wave speed anomaly $\delta v_s/v_s = \delta \ln v_s$ or $\delta v_p/v_p = \delta \ln v_p$ can be written as

$$\delta \ln \rho = \mathrm{R}_{\rho/s} \delta \ln v_s \qquad \text{and} \qquad \delta \ln \rho = \mathrm{R}_{\rho/p} \delta \ln v_p \tag{2.9}$$

with $\mathrm{R}_{\rho/s}$ and $\mathrm{R}_{\rho/p}$ the partial derivatives $\partial\rho/\partial v_s$ and $\partial\rho/\partial v_p$. At constant depth, $\delta \ln \rho$, $\delta \ln v_s$, $\delta \ln v_p$, $\mathrm{R}_{\rho/s}$ and $\mathrm{R}_{\rho/p}$ depend on temperature and composition, so the interpretation of wave speed anomaly models in terms of density variations of purely thermal origin is not rigorously exact (Mattern et al., 2005; Stixrude & Lithgow-Bertelloni, 2005, 2011). However, it is generally the assumption taken in mantle circulation reconstructions (e.g. Conrad & Gurnis, 2003; Steinberger & O'Connell, 1997; Spasojevic et al., 2009), where the density anomaly of the mantle below the lithosphere is computed with either constant or depth-dependent $\mathrm{R}_{\rho/s}$ or $\mathrm{R}_{\rho/p}$, estimated using mineral physics constraints (e.g. Karato & Karki, 2001; Cammarano et al., 2003). The S-wave velocity anomalies are generally preferred in these computations due to their better resolution of upper mantle heterogeneities.

Instantaneous models of mantle dynamics have also played a role in justifying such assumption and providing a scaling between wave speed anomalies and density anomalies (see Forte et al., 2015, for a review). Instantaneous models of mantle circulation estimate the flow of the mantle at present from:

- data such as seismic tomography models, the surface kinematics, the geometry of the geoid, the value of gravity anomalies, topography...
- a physical model of instantaneous mantle flow (Richards & Hager, 1984; Ricard et al., 1984, 1993), combined with models of the material properties of the mantle.

Hager et al. (1985) used the first models of seismic tomography to infer a density field and compute the geoid from it. In doing so, they were able to relate to first order seismic and geodynamical observables. By complementing the tomography-derived density model with a model of subducted slabs, Hager & Richards (1989) were able to explain $90\%$ of the current geoid. The improvement of both observations and models has produced a variety of instantaneous mantle circulation models which aim at reconciling seismic and geodynamical observations. For example, Simmons et al. (2006, 2009, 2010) perform a joint inversion of geodynamical and seismic observable to determine the optimum $\mathrm{R}_{\rho/s}$ profile and infer the temperature and compositional contributions to the variations of densities. Stadler et al. (2010), Ghosh et al. (2010) and Ghosh & Holt (2012) combine tomography-derived density anomalies and models of slabs to determine the state of stress in the lithosphere and Wang et al. (2015) build on this work to evaluate the capacity of various tomography models to produce accurate geodynamic observables.

So far, dynamical mantle circulation models have only used estimated temperature fields from tomography, or from an integration of other variables at present from joint geodynamical and seismic inversions (e.g. Glišović & Forte, 2016). However, the continuous developments in both instantaneous and dynamical mantle circulation models bring these areas of research ever closer and opens the way to a direct integration of present-day raw observations in a dynamical model.

### 2.2.2 Surface Kinematics History of the Mantle

Observations on the global horizontal motions at the surface of the Earth are generally sparse and indirect and, for a significant part, are of qualitative nature. The concepts and methods associated to plate tectonics theory enable us to produce a simple time series of quantitative data that can be integrated in a data assimilation framework as described in Section 2.1.

The modern formulation of plate tectonics theory is the result of a collective effort of the geoscience community to integrate at a global scale observations on the horizontal motions of the Earth's surface from a wide range of disciplines (see for example Oreskes, 2003, for a historical description of the origins of plate tectonics theory and an account of its development by some of the researchers who took part in this scientific revolution). Here, we summarize the evolution of ideas and highlight key observations that led to the formulation of plate tectonics theory in the late $60$'s by McKenzie & Parker (1967), Morgan (1968) and Le Pichon (1968). We then describe the current models for past surface motions, which are to a large extent still based on these principles and observations.

In the beginning of the $20$th century, Alfred Wegener (1880-1930) proposed the continental drift hypothesis: the continents once formed a single landmass (called Pangaea) and have been drifting apart since then. Throughout his career, he collected observations supporting his theory and documented them in the successive editions of his book, "Die Entstehung der Kontinente und Ozeane" (see for example Wegener, 1924, the English translation of the third edition). Wegener bases his theory on the complementary shape of continents and the continuity of geological formations, paleoclimates and fossil repartition once continents are reunited.

Although the continental drift theory was rejected by a large part of the community (Oreskes, 1999), it gained a few early supporters. For example, Du Toit (1937) collected further geological observations supporting the theory of continental drift. Holmes (1931, 1944) proposed mantle convection as a viable mechanism for continental drift. Geodesist Felix Andries Vening-Meinesz pioneered the oceanic geodetic studies and interpreted the negative gravity anomalies at oceanic trenches (the deepest regions of the ocean floor) as the locations of downwelling convective currents (see Vening-Meinesz, 1948, for a synthesis of his gravity expeditions between 1923 and 1938). Griggs (1939b) designed experimental models of continental drift driven by convection currents.

The $1960 - 1970$ decade saw the progressive formulation and acceptance of the plate

tectonics theory, starting with the sea-floor spreading theory (Dietz, 1961; Hess, 1962) and new evidence of large scale lateral motion of continents (Runcorn, 1961) and finishing with the "new global tectonics" (i.e. plate tectonics) (Morgan, 1968; Le Pichon, 1968).

The intense marine geophysical exploration of the previous decades (started for military purposes during WWII and continued from then on) had produced an extensive dataset of the bathymetry, heat flow, seismic properties, gravity and magnetic anomalies of the seafloor, giving a global view of the oceanic domain and requiring new interpretations. Parallelly, the progresses in seismology made possible the detection and precise localization of ever smaller and distant earthquakes, as well as a better determination of earthquake properties such as the focal mechanism, and the associated slip (see e.g. Isacks et al., 1968, for a synthesis of data interpreted in the light of plate tectonics theory).

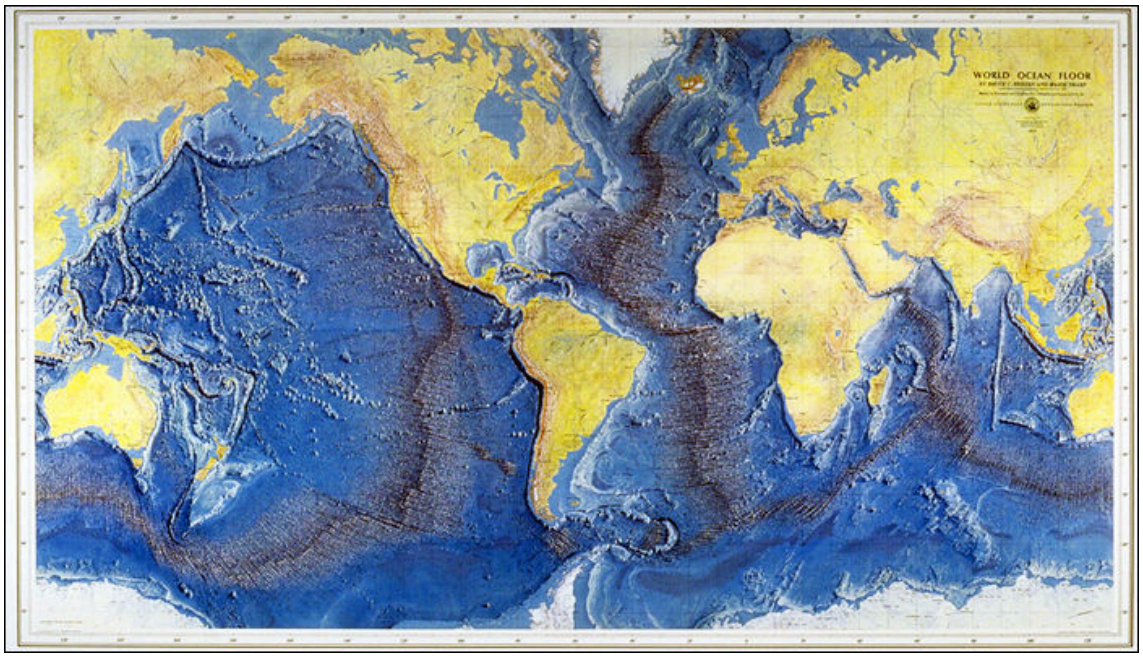


FIGURE 2.2: "World Ocean Floor Panorama" by Heezen & Tharp (1977), two pioneers of ocean exploration, who produced the first comprehensive maps of the ocean floor.

These studies led to the characterization of 5 main features of the ocean floor:

- **abyssal plains**, which represent most of the ocean floor, show very localized and limited deformation, and display "zebra patterns" of magnetic anomalies, i.e.

elongated stripes of alternating positive and negative magnetic anomalies (Mason & Draff, 1961). These zebra pattern are oriented parallely to the mid-ocean ridges.

- **mid-ocean ridges**, which form the largest mountain ranges on Earth and show remarkable continuity (Ewing & Heezen, 1956; Heezen et al., 1959, see also figure 2.2). They are associated with a high heat flow, volcanic activity, shallow earthquakes and a fracture zone resembling a rift at their center. Mid-oceanic ridges have been identified as spreading centers in Hess's and Dietz's theory of seafloor spreading.

- **deep trenches**, which are associated with mountain ranges on adjacent continents or island arcs, show intense seismic activity with deep earthquakes organized on a plane dipping towards the continents or the island arc (Benioff, 1949). Deep trenches are identified as zones of convergence where the oceanic floor dive into the mantle in Hess's and Dietz's theory of seafloor spreading, later labeled as subduction zones.

- **fracture zones**, linear features of several hundred to several thousand kilometers, perpendicular to mid-ocean ridges. These fracture zones are associated with an offset of the linear features they traverse (zebra pattern of magnetic anomalies and mid-ocean ridges) and hence appear as zones of shearing. Wilson (1965b) and Sykes (1967) showed these fractures correspond to transform faults which link divergent or convergent boundaries to form a consistent network of mobile belts dividing the surface of the Earth into rigid plates. Moreover, it was recognized that fracture zones are seismically active only on the sections that connect two other active boundaries, so that the rest of them are interpreted as fossil transform faults.

- **volcanic islands and seamounts** generally organized in linear chains. The radioactive dating of the volcanic rocks of pacific chains showed a regular age progression, with the younger islands being closer to the east pacific rise (i.e. the Pacific ocean ridge). Wilson (1963, 1965a) interpreted these chains as the result

of successive volcanic eruptions due to localized and fixed rising mantle currents below the moving plates. These hotspots were later connected to deep mantle plumes by Morgan (1971).

Simultaneously, the study of continental rocks brought key elements that led to the acceptance of plate tectonics theory. These continental arguments are the result of the interaction of studies on the remanent magnetism of rocks, radioactive dating and progresses in the understanding of the origin of the Earth's magnetic field, with the emergence of the geodynamo theory (Elsasser, 1946a,b).

First, a collection of paleomagnetic data of continental rocks indicated an apparent variation of the location of the magnetic pole through geological times. This apparent polar wander appeared to be different depending on continents. Runcorn (1961) showed that lateral motions of continents could reconcile the apparent polar wandering paths. Supposing that the magnetic pole has not moved over geological times, it is then possible to track the paleolatitude of continents with paleomagnetic data (the paleolongitude cannot be estimated because the Earth's magnetic field is radially symmetrical).

The second observation was the frequent reversed polarity measured in the remanent magnetism of rocks of various ages. Cox & Doell (1963) used the newly developed K-Ar dating method on samples coming from $6$ volcanic eruptions and were able to show the consistency of the polarity among rocks of the same age, which suggested the existence of global polarity reversals of the magnetic field. Additionally, they could date the reversal events, and obtained in this way the first geomagnetic polarity time scale, constantly updated since then (Ogg, 2012). Morley (1963) and Vine & Matthews (1963) connected these findings with the seafloor spreading theory, stating that if the ocean floor is produced at mid-ocean ridges and moving away from them with time, then the zebra patterns of magnetic anomalies on the seafloor correspond to the fossilization of the magnetic polarity through time and should be organized symmetrically around the mid-ocean ridges. This prediction was then confirmed by Heirtzler & Le Pichon (1965) and Heirtzler et al. (1966, 1968). Moreover, the identification of specific anomalies from

the center of the ridge outward, along with a magnetic polarity timescale provided a way to estimate the age of the ocean floor depending on the distance of the ridge, and thus to deduce the rate of seafloor spreading.

The final concept to formulate plate tectonics theory was found independently by McKenzie & Parker (1967) and Morgan (1968): since the Earth's surface is made of a jigsaw of internally rigid plates in constant motion, then the plate motions can be described with principles of solid body rotation on a sphere. More precisely, we can use Euler's rotation theorem (Leonhard Euler 1707-1783): a solid body motion on a sphere can always be described as a rotation around an axis passing through the center of the sphere, and the composition of two rotations is also a rotation. Consequently, absolute plate motions and relative motions of plate pairs are rotations, and are uniquely defined by their respective angular velocities and rotation poles (called the Euler poles and corresponding to one of the two points defined by the intersection of the rotation axis with the surface of the Earth). Morgan (1968) then deduced that transform faults could only develop along "small circles", i.e. circles at the surface of the Earth of constant latitude with respect to the Euler pole (see figure 2.3).

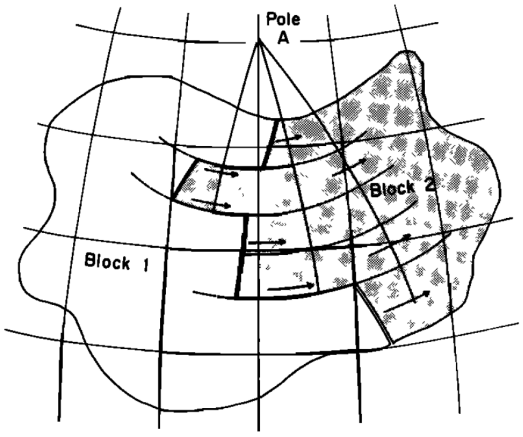


FIGURE 2.3: Diagram of two plates sharing a divergent plate boundary. The motion of plate 2 relative to plate 1 is described by the Euler pole A and an angular velocity. Transform faults develop along small circles of constant latitude with respect to the Euler pole A. Morgan (1968)

In this formulation, the plate tectonics jigsaw puzzle becomes a spherical geometry problem: we can integrate local observations on relative motions of plate pairs to estimate their Euler poles and angular velocities, and use geometrical constraints at triple

junctions (where 3 plates are in contact) to deduce the relative motions of plate pairs which are not constrained by data (see the description of the technique in McKenzie & Morgan, 1969). Following this principle, Le Pichon (1968) was the first to define a globally consistent plate circuit, i.e. a set of plates and plate boundaries which are related to each other in such a way that the relative motions of each plate pair can be either inferred from observations or deduced from other plate pairs relative motions (see Figure 2.4). He bases his analysis on data of seafloor spreading motions: the Euler poles are determined from the directions of fracture zones, and the spreading rates are inferred from the identification of magnetic anomalies around the ridges, the rest of relative plate motions are inferred geometrically. In his article, Le Pichon proposes also a partial reconstruction of plate motion throughout the Cenozoic ($65.5$ Myr to present) by applying successive rotations inferred by the same technique.

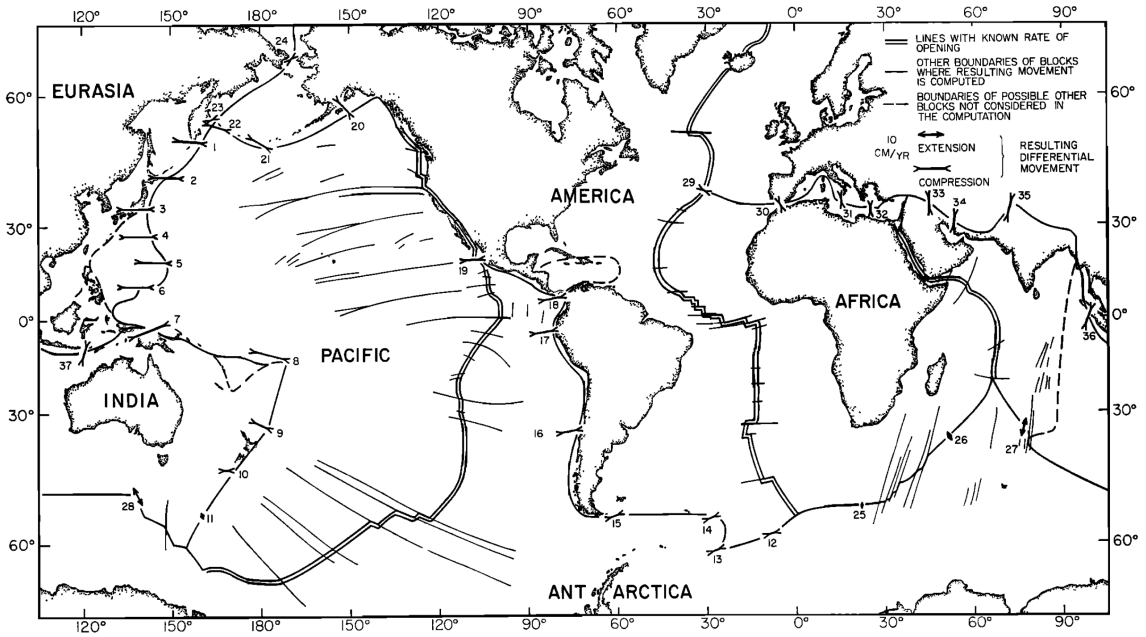


FIGURE 2.4: The first consistent global model of plate motions as published by Le Pichon (1968).

By the end of $1968$, Earth scientists had built a theory which at first order explains the surface dynamics of the Earth, and allows the integration of local observations of different nature, coming from almost all the disciplines of Earth sciences, into a consistent framework. The following decades saw the emergence of space geodesy (see Blewitt (2015) for a review), which made possible:

- the direct measurements of surface motions with techniques such as Very Long Baseline Interferometry, Satellite Laser Ranging, and the Global Positioning System. We can now measure surface motions and infer plate motions which could not be deduced with classical plate tectonics principles (e.g. DeMets et al., 2010). Those techniques also led to the production of instantaneous maps of plate motions (e.g. Argus et al., 2010; Kogan & Steblov, 2008), and the measurement of plate deformation (e.g. Kreemer et al., 2014).

- a better localization of ships during measurement campaigns at sea. This resulted in an improved mapping of magnetic anomalies identifications (e.g. Seton et al., 2014, see also Figure 2.5).

- the production of high resolution maps of gravity anomalies (e.g. Sandwell & Smith, 2009, see also Figure 2.6), which considerably increased the database of fracture zones, seamounts and other features of the ocean floor (e.g. Wessel, 2001; Matthews et al., 2011; Wessel et al., 2015).

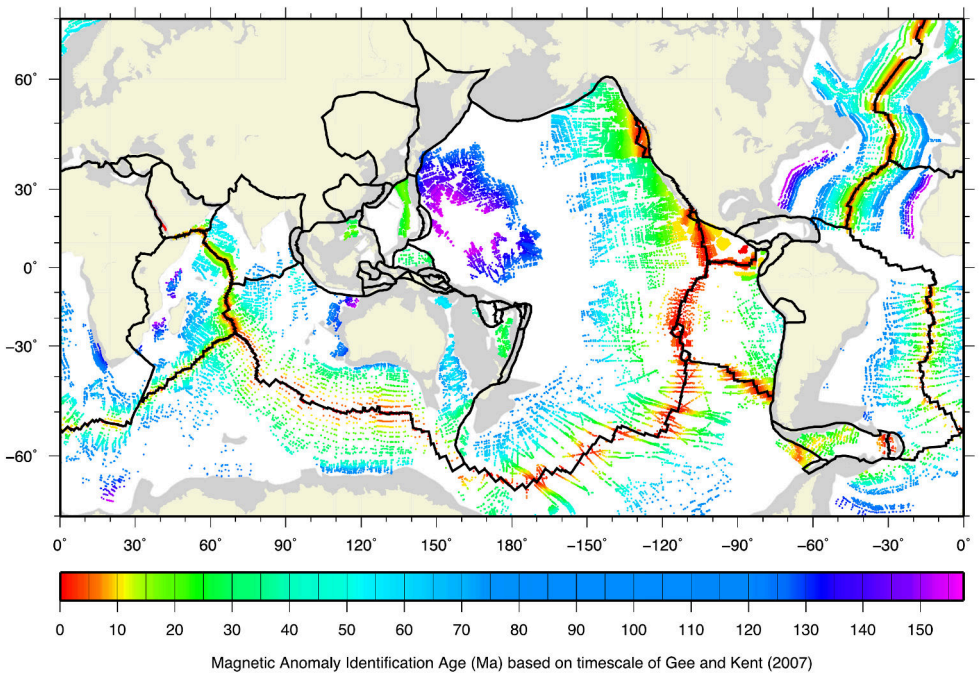


FIGURE 2.5: Map taken from Seton et al. (2014), which shows magnetic anomalies identifications (colored by age), recently collected to build a community database.

The maps of ocean magnetic anomaly identifications (Figure 2.5) and gravity anomaly (Figure 2.6) form the basis for the reconstruction of relative plate motions for the last

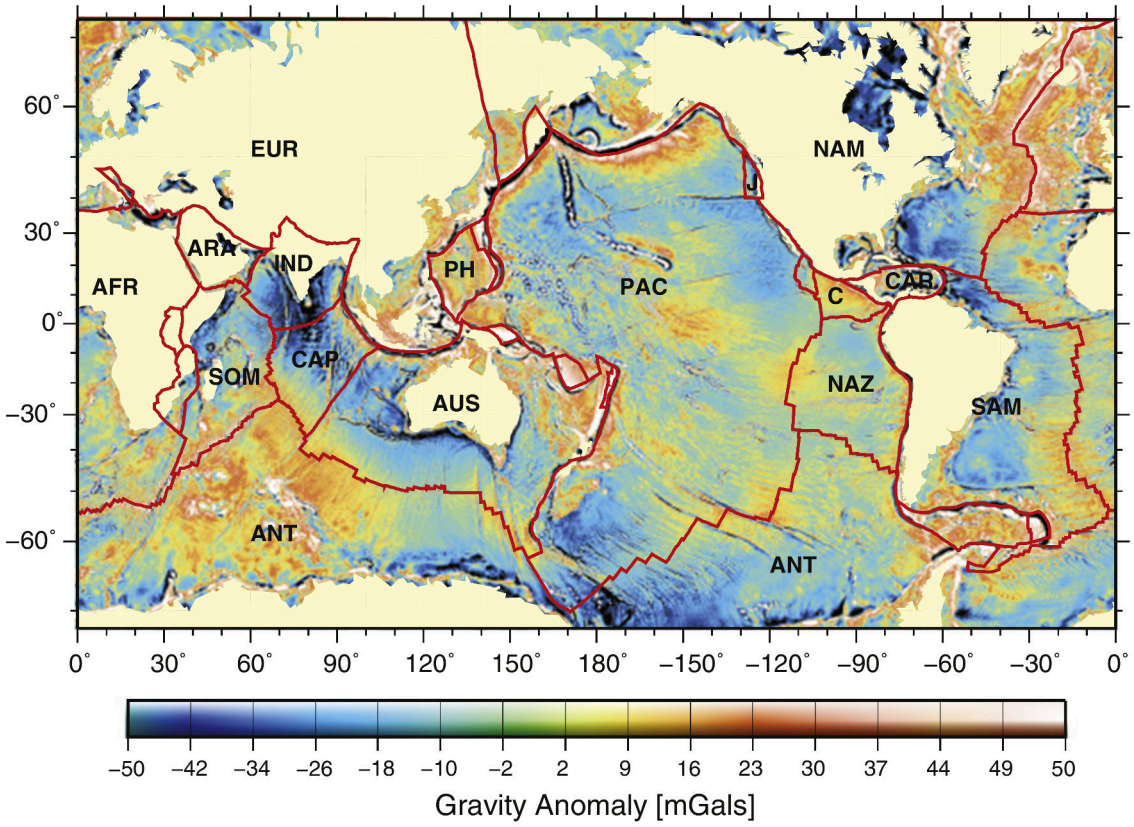


FIGURE 2.6: Map of gravity anomalies (Sandwell & Smith, 2009) and a model of plates at present. The name of major plates are: AFR: Africa, ANT: Antarctica, ARA: Arabia, AUS: Australia, C: Cocos, CAP: Capricorn, CAR: Caribbean, EUR: Eurasia, IND: India, NAM: North America, NAZ: Nazca, PAC: Pacific, PH: Philippine, SAM: South America, SOM: Somalia. Taken from Seton et al. (2012).

150 Myr: surface motions are reconstructed based on the spreading history of ocean basins, inferred from the age of the ocean floor (given by the magnetic anomaly identifications map), and indications of the direction of motion with active and fossil transform faults (identified on the gravity anomaly map). The general procedure consists in determining series of finite rotations defining the relative motions of plate pairs, starting at present and going back in time. The most common method to determine a finite rotation is the Hellinger method (Hellinger, 1981), summarized in Figure 2.7:

1. selection of magnetic anomalies of the same age from both sides of a ridge and identification of associated transform faults

2. construction of isochrons, i.e. lines of same age formed by magnetic anomaly identifications and transform faults. These isochrons are decomposed in great circle segments, and the conjugate great circle segment on each isochrons are

identified (numbers 1, 2, 3, 4 on the figure).

3. rotation of data of one plate to fit the data of the other plate. The common boundary is then reconstructed by a set of great circle segments, and the misfit to the data is computed by taking the weighted average of the distance of data points to their corresponding great circle segment.

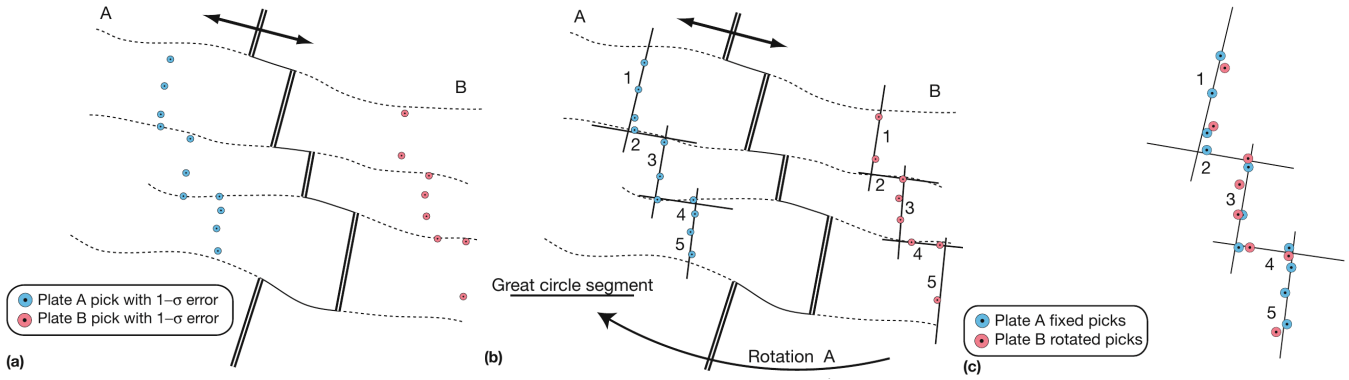


FIGURE 2.7: Diagram of the principle of the Hellinger method for relative plate motion reconstructions. Plates A and B are separated by a ridge (double lines), and active transform faults (single line). a) localization of magnetic anomaly identifications of the same age on both sides of the plate boundaries (pink and blue points) and of fossil transform faults (dashed lines) b) definition of isochrons and identification of conjugate great circle segments c) position of magnetic anomaly identifications after the finite rotation, and inferred past plate boundary. Taken from Wessel & Müller (2015)

By repeating this technique on isochrons of increasing age, we can construct a time series of finite rotations of plate pairs starting at present and going back in time. The extent of these time series is limited by the destruction of ocean floor at subduction zones. For example, the relative motion of the Pacific plate to the Nazca Plate (see Figure 2.6 for the localization of plates) can only be reconstructed directly for the last $\sim 50$ Myr since the older ocean floor of the Nazca plate has been subducted. Supposing symmetrical spreading rates, it is possible to estimate earlier finite rotations between these two plates by studying the sole pacific plate data (e.g. Seton et al., 2012), but this method introduces further uncertainties in the reconstruction. When both sides of the ridge have been destroyed, the reconstructions rely on the interpretations of onshore geology and extrapolations of the direction and the rate of spreading at more recent times (e.g. Rowley, 2008). In this case, the uncertainties on the reconstructions are high,

but difficult to quantify. Torsvik et al. (2010) defines the "world uncertainty" to quantify the increase of the global uncertainty of plate reconstructions through time, due to the absence of older ocean floor: the world uncertainty corresponds to the fraction of the Earth's lithosphere which has been subducted since a given time. For example, the oceanic floor older than $140$ Myr has almost entirely been subducted, which brings the world uncertainty at $60\%$ at that time. Besides, the time resolution of the finite rotation series is directly dependent on the frequency of polarity reversal of the magnetic field. One of the most important lack of data is the Cretaceous superchron (Granot et al., 2012): between $121$ and $83$ million years ago, the polarity of the Earth's magnetic field appears to have been stable, so a large portion of the ocean seafloor cannot be dated with precision using magnetic anomaly identifications, as is shown in Figure 2.5. Müller et al. (2008b) produced continuous maps of seafloor age uncertainty based on the distance from the nearest magnetic anomaly.

Once the motion of divergent plate boundaries is determined, it is possible to infer the relative motions of other boundaries using geometrical constraints. To determine the relative motions of all plate pairs, we need to construct a plate circuit such that each plate is connected to the plate circuit by at least one divergent boundary where the finite rotations are computable, as Le Pichon (1968) had constructed at present. However, it is not possible to build such a circuit for the last $200$ Myr. Some plates, such as the Philipine plate at present (Bird, 2003), have very limited portions of divergent boundaries and are thus almost isolated from the rest of the global circuit. Before $83$ Myr, the Pacific plate cannot be linked to the global circuit either (Seton et al., 2012). Moreover, the uncertainties on finite rotations propagate through the circuit and two alternative circuits can lead to different results (e.g. Steinberger et al., 2004). Finally, some parts of the surface of the Earth have undergone deformation, for example South East Asian plates. In these places, plate tectonics geometrical rules do not apply (e.g. Holt et al., 2000; Gordon, 2000). To produce a global model of plate reconstruction it is thus necessary to take into account other sources of information. These complementary sources of informations are of two types: additional observations on the relative motion of plates, and observations on the absolute motion of plates.

Observations on the relative motions of plates include for example paleontological, structural, stratigraphic and regional geology constraints (e.g. Hosseinpour et al., 2016; Barnett-Moore et al., 2016). The integration of such constraints into global reconstructions is not done in a systematic way. On the contrary, it generally requires choices and interpretations from the geoscientist who builds the reconstruction. The resulting uncertainties are then difficult to quantify.

Observations on the absolute plate motions primarily include data from island and seamount chains, and paleomagnetism. Figure 2.8 represents the linkage between the relative motion of a plate with its neighboring plates, and its absolute motion. All these motions are represented for a plate which is stationary for the last $80$ Myr, for illustrative purposes. The figure also depicts the predicted apparent polar wander path (i.e. the time evolution of the direction of the paleopole as recorded by the remanent magnetism of the rocks belonging to this plate) and the geometry and age progression of an island and seamount chain. These predictions are made with the following assumptions: 1) the Earth's magnetic field has always been primarily dipolar, and its poles have always remained stationary 2) the seamounts and ocean islands are created by static and localized hotspots below the plate (first hypothesized by Wilson, 1963). With these hypotheses, it is possible to infer the paleolatitude of the plate from APW paths (the paleolongitude cannot be inferred because the Earth magnetic field is radially symmetrical) and the absolute motion of the plate by considering several seamount and island chains and estimating finite rotation series, similarly to what is done for the relative plate motions from magnetic anomalies and fracture zones. The respective uncertainties on observations can also be estimated (see for example Torsvik et al., 2008 for APW uncertainty estimations, and O'Neill et al., 2005; Wessel et al., 2006 for uncertainty estimations on absolute plate motions from seamount and island chain motions).

In practice, the estimation of absolute plate motions is more complex. First, the study of APW paths reveals a significant component of whole mantle and crust rotation, called true polar wander (e.g. Goldreich & Toomre, 1969; Besse & Courtillot, 2002; Evans,

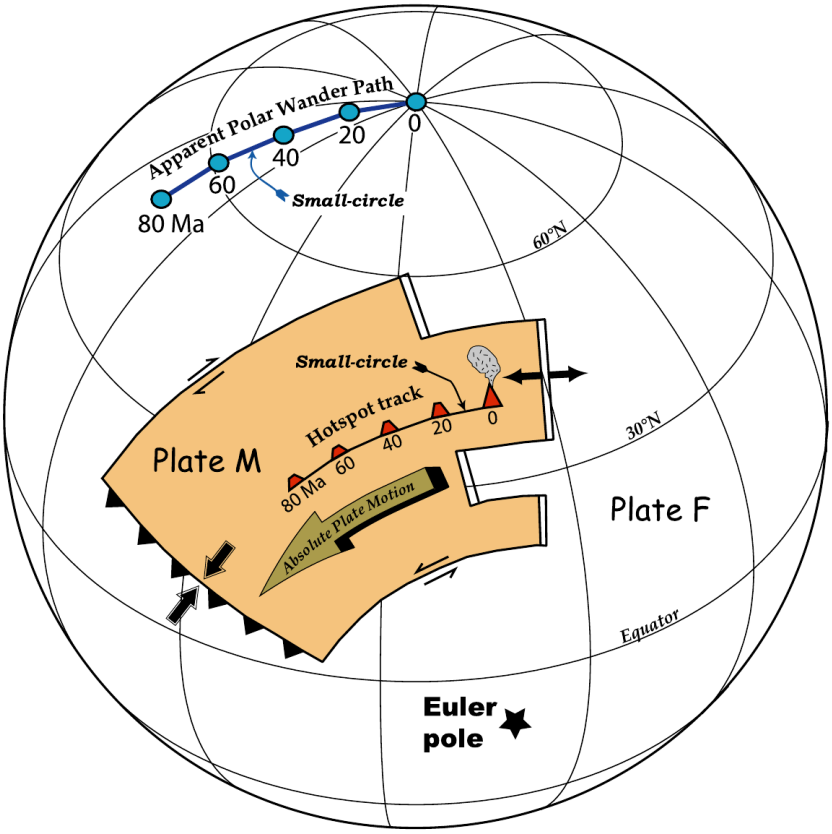


FIGURE 2.8: Diagram of the relationship between relative plate motions (shown on the plate boundaries), absolute plate motions and predicted observations of apparent polar wander and seamount and island chain geometry, if the magnetic pole of the Earth and the hotspots were static. Taken from Torsvik et al. (2008).

2003; Maloof et al., 2006; Steinberger & Torsvik, 2008). The true polar wander is due to the redistribution of masses in the convecting mantle, which changes the axis of maximum non-hydrostatic moment of inertia and leads to a whole mantle and crust rotation so as to realign this axis to the spin axis of the Earth. Second, the global study of seamount and ocean island chains shows that the hotspots are slightly moving through time, due to convective currents in the mantle (e.g. Molnar & Stock, 1987; Tarduno & Gee, 1995; Tarduno & Cottrell, 1997; DiVenere & Kent, 1999; Tarduno et al., 2003). As a result, the APW paths and seamount and ocean island chains actually give indications on the relative motions of plates with respect to the magnetic pole of the Earth and the moving mantle, respectively. To integrate these observations together with relative plate motions in a global model of plate tectonics reconstruction, it is necessary to define a common reference frame and relate all the data to it.

Reference frames vary depending on studies. Torsvik et al. (2008) review and compare four types of reference frames generally used in plate tectonics reconstructions: paleomagnetic, African fixed hotspot, African moving hotspot, and global moving hotspot. The paleomagnetic reference frame uses APW paths corrected from their estimated true polar wander to determine the variations in paleolatitudes of continents. Since the longitude is not constrained by APW paths, the reference frame also considers one plate as fixed in longitude (generally, Africa is chosen due to its central place in the Pangaea, and its supposed small amplitude of motion). The paleomagnetic reference frame is the only reference frame available before $130$ Myr. The three other reference frames are variations on hotspot-based reference frames using different approximations. The African fixed hotspot reference frame considers that the hotspots in the Atlantic and Indian oceans are totally fixed (e.g. Müller et al., 1993), and anchors relative plate motions in a frame using data coming from those hotspots. On the contrary, the African moving hotspot, and global moving hotspot reference frames estimate the motions of hotspots, and propose corrections to construct a reference frame. The estimation of the motion of hotspots is done using mantle circulation models (e.g. O'Neill et al., 2005; Doubrovine et al., 2012), that we will discuss in the next section.

Recent plate tectonics reconstructions use a hybrid reference frame, as proposed by Torsvik et al. (2008), where a moving hotspot reference frame is used for recent times (up to $100$ Ma) and is linked to a paleomagnetic reference frame for earlier times (e.g. Torsvik et al., 2010; Seton et al., 2012; Müller et al., 2016).

Additionally, mantle structures estimated from tomography models have been included as absolute plate motion constraints in some plate tectonics reconstruction models. Van Der Meer et al. (2010) identified paleoslabs (remnants of subducted lithosphere in the mantle) using seismic tomography, and opened the way to the introduction of such features into plate tectonic reconstructions (e.g. Williams et al., 2015). Torsvik et al. (2014) used the location of deep mantle structures (Large Low Shear Velocity Provinces, LLSVP), that are thought to be responsible for the production of Large igneous provinces, to estimate the longitude of continents in the Paleozoic.

The final result of these plate tectonics reconstructions is a time-series of plate layouts associated with pairs of rotation poles and angular velocities, which describe the history of the motions at the surface of the Earth. Modelling techniques and software developments have recently allowed the production of plate tectonics reconstructions continuous in time (e.g. Gurnis et al., 2012; Müller et al., 2016), which made their integration into mantle convection models both easier and more consistent (see next section). Given the amount of observations used for these reconstructions, and the variety of specific techniques used to integrate those observations, the quantification of uncertainties in the final result is difficult to estimate. To our knowledge, a rigorous quantification of errors in the global plate tectonics reconstructions has not been performed yet. The progress made in successive plate reconstruction models are generally evaluated by comparing them with previous ones on their ability to explain observations (Müller et al., 2016) and the plausibility of the dynamics they predict (Williams et al., 2015).

## Conclusions on Observed Data

In this section, we described the data currently used in mantle circulation models: the temperature field at present and the history of surface kinematics. These data are not raw observations, but the result of models that integrate a wide range of observations. This setup is not ideal, since the addition of successive layers of models and/or inversions make the link between raw observations and the final result (the estimated mantle circulation) complex and the uncertainties difficult to assess. In this sense, the ultimate goal of data assimilation for mantle convection is to develop a framework able to integrate directly raw observations. The construction of community databases gathering raw data in a unified format (Seton et al., 2014; Wessel et al., 2015; Moulik et al., 2016), the progresses made in each of the disciplines to model and understand mantle structure and dynamics, and the techniques developed to assimilate data into models all contribute to this aim.

## 2.3 Dynamical Mantle Circulation Reconstructions

Three alternative strategies have been proposed to reconstruct the history of mantle circulation: backward advection, semi-empirical sequential methods and variational data assimilation. In the present dissertation, we propose a fourth one: sequential data assimilation, that we develop in Chapters 3 and 4. Backward advection and empirical sequential methods have already produced numerous estimations of global mantle circulation evolutions in the last decades. The variational data assimilation methods are more recent, and have so far been applied to regional reconstructions (Liu & Gurnis, 2008; Spasojevic et al., 2009) and just recently to global mantle circulation models (Glišović & Forte, 2014; Horbach et al., 2014; Glišović & Forte, 2016).

Backward advection and empirical sequential methods are direct methods: they integrate data in mantle convection models as boundary or initial conditions, without taking into account the uncertainties on the data and/or the model. We present these methods in the first part of this section. On the contrary, variational and sequential data assimilation consider mantle circulation reconstruction as an inverse problem, and explore (at least partially) the mantle state space to determine the evolution that best fits the observed data, given uncertainties on both data and the convection model. We describe these techniques in the second part of this section.

### 2.3.1 Direct Methods

By the end of the 90's, the increase in computational power and the advances in numerical modelling had led to the emergence of numerical models able to solve the governing equations of mantle convection in 3D spherical geometry, with radial viscosity variations and at a Rayleigh number 1 to 2 orders of magnitude lower than that of the Earth's mantle (Tackley et al., 1993; Bunge et al., 1996). Simultaneously, the resolution of mantle tomographic models had improved, revealing the distribution and geometry of subducted slabs (Su & Dziewonski, 1997; Masters et al., 1996; Li & Romanowicz, 1996; Grand et al., 1997; Van der Hilst et al., 1997). Finally, semi-dynamical

models of mantle density evolution derived from surface kinematic history had shown their ability to reproduce the current geoid and seismic wave speed anomalies, and to predict plate velocities in the past (e.g. Ricard et al., 1993; Zhang & Christensen, 1993; Deparis et al., 1995; Lithgow-Bertelloni & Richards, 1998). The Backward advection (Steinberger & O'Connell, 1997) and semi-empirical sequential (Bunge et al., 1998) methods emerged from these progresses, as a way to investigate the dynamics of the mantle and its interaction with surface tectonics.

**Backward advection**

Backward advection consists in reversing the time in mantle convection equations, to compute the evolution of mantle circulation from its present state (derived from seismic tomography) backward in time, using plate tectonics reconstructions as surface boundary conditions. As described in section 1.3, the conditions at the upper boundary $\Gamma_a$ of the model are then

$$\forall t,\ \forall \mathbf{x} \in \Gamma_a,\ \ \mathbf{u}(\mathbf{x},t) \cdot \mathbf{e}_r = 0 \qquad \text{and} \qquad \mathbf{u}_H(\mathbf{x},t) = \mathbf{u}_H^a(\mathbf{x},t), \tag{2.10}$$

with $\mathbf{u}$ the total velocity and $\mathbf{u}_H$ the horizontal velocity. $t$, $\mathbf{x}$ and $\mathbf{e}_r$ are the time, the position vector and the unit radial vector. $\mathbf{u}_H^a$ is the surface velocity field given by plate tectonics reconstructions. In this formulation, it is necessary to have plate tectonics reconstructions which are continuous in time.

For mantle convection, the only prognostic equation is the equation of energy conservation. In the Boussinesq approximation, the nondimensional form of this equation is

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \nabla^2 T + \mathrm{R}_h, \tag{2.11}$$

where $T$ and $\mathrm{R}_h$ are the nondimensional temperature and rate of energy production. The backward heat advection-diffusion equation should be:

$$\frac{\partial T}{\partial t} - \mathbf{u} \cdot \nabla T = -\nabla^2 T - \mathrm{R}_h. \tag{2.12}$$

This problem is an ill-posed problem: the backward diffusion term $-\nabla^2 T$ will transform a slight positive temperature anomaly at present into a very large and localized temperature anomaly in the past. It follows that the final solution (the temperature field at some time in the past) does not depend continuously on the initial condition (the temperature field at present): the problem is unstable, and following Hadamard (1902) definition, ill-posed (see Ismail-Zadeh et al., 2016, section 1.1 for a simple 1-D example of this behavior). In practice, the diffusion term is dropped. Then, the backward rate of energy production $-\mathrm{R}_h$ also needs to be dropped, for global energy balance reasons. The backward conservation energy equation becomes a simple backward advection equation

$$\frac{\partial T}{\partial t} = \mathbf{u} \cdot \nabla T. \tag{2.13}$$

The backward advection method should be valid for a timespan which is short compared to the diffusion time of temperature heterogeneities in the mantle. If we consider a temperature anomaly of characteristic size $100$ km (5 times smaller than the current resolution of global tomography models, see e.g. Ritsema et al., 2011), we obtain a characteristic diffusion time of

$$\tau_d = \frac{d^2}{\kappa} = 300 \text{ Myr}, \tag{2.14}$$

where $\kappa = 10^{-6} m^2.s^{-1}$ is the coefficient of thermal diffusivity. Conrad & Gurnis (2003) add to the backward advection algorithm an analytic solution describing the thermal boundary layers (where diffusion is the principal mechanism of heat transfer) to model more accurately mantle circulation. They test the reliability of the backward advection method by first computing the state of the mantle in the past by backward advection,

and then computing a forward evolution from this initial state. Finally, they compare the predicted state of the mantle at present with the state of the mantle derived from seismic tomography. With this setup, they estimate that backward advection is reliable to reconstruct mantle circulation for the last $50$ to $75$ Myr. O'Neill et al. (2005) push the reconstructions of mantle circulation to $120$ Myr, using a different technique for the boundary layers. From their mantle circulation models, they estimate the motions of hotspots through time, in order to build a moving hotspot reference frame (see previous section). They take into account uncertainties in the plate tectonics reconstructions and convection models by successively running backward advection evolutions and modifying the parameters of the model and plate tectonics reconstructions, to converge to a result consistent with both the physics and the data. Doubrovine et al. (2012) also use backward advection to estimate a moving hotspot reference frame for the last $130$ Myr. However, they stop the backward advection at $70$ Myr, and estimate the motion of hotspots between $70$ and $130$ Myr using the velocity field at $70$ Myr. Williams et al. (2015) compare the moving hotspot reference frame of O'Neill et al. (2005); Doubrovine et al. (2012) to the absolute reference frame recently derived from the identification of paleoslabs in tomography models (Van Der Meer et al., 2010). Williams et al. (2015) find discrepancies between reference frames for times older than $70$ Myr, further confirming the results of Conrad & Gurnis (2003): backward advection models are probably not reliable beyond $70$ Myr.

Furthermore, uncertainties on the state of the mantle at present (derived from seismic tomography) and on plate tectonics reconstructions add to the intrinsic time limit of retrodictions for the backward advection method.

**Semi-empirical Sequential Methods**

The semi-empirical sequential methods estimate mantle circulation by integrating plate tectonics reconstructions chronologically into a mantle convection model. Plate tectonics reconstructions are either introduced as velocity boundary conditions (following Equation 2.10), as first described by Bunge et al. (1998), or with a more sophisticated method, by blending a convection solution with thermal and kinematic models of plates and slabs (Bower et al., 2015). Contrary to backward advection, the timespan of the reconstruction is not limited by the assimilation method, but the availability of plate tectonics reconstructions. This makes possible the computation of mantle circulation models for much longer timespans (so far, up to 450 Myr, as done by Zhang et al., 2010). Such long term evolutions allow the investigation of the effect of surface tectonics on the deep mantle structure (e.g. McNamara & Zhong, 2005; Schuberth et al., 2009b,a; Zhang & Zhong, 2011; Davies et al., 2012; Bower et al., 2013; Bull et al., 2014; Hassan et al., 2015), and the investigation of alternative scenarios of absolute plate motions beyond 70 Myr (e.g. Shephard et al., 2012; Nerlich et al., 2016; Zahirovic et al., in press).

One important issue of semi-empirical sequential methods is that the initial condition is unknown. A first guess is generally obtained by supposing a simple 1D temperature profile (e.g. McNamara & Zhong, 2005), by constructing a mantle structure consistent with the first plate layout used (e.g. Flament et al., 2014), or by computing a solution with surface boundary conditions corresponding to the first plate layout for tens of millions of years or more (e.g. Bunge et al., 2002). McNamara & Zhong (2005) studied the sensitivity of mantle circulation models to initial conditions. They considered three initial conditions: a first case with a 1-D temperature profile and two other cases where they imposed the velocity boundary conditions from the reconstructed plate tectonics 119 Myr ago for 60 Myr and 120 Myr, respectively. Figure 2.9 shows the estimated temperature field at the base of the mantle circulation models after 119 Myr of reconstruction. Although the global pattern is similar for the three mantle circulation models, the details of the temperature field vary between models.

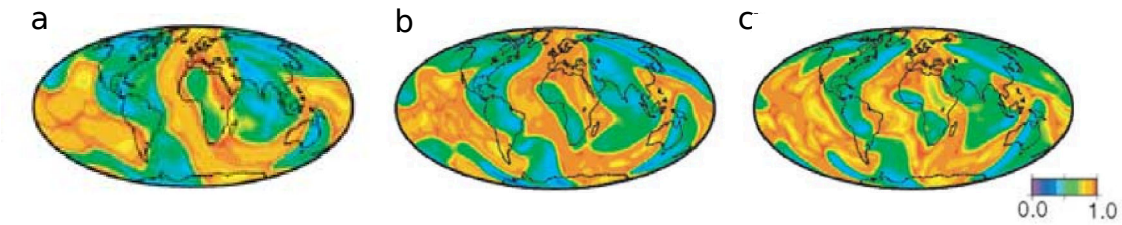


FIGURE 2.9: Maps of the estimated nondimensional temperature at present at 2811 km depth for three mantle circulation models computed from three different initial states, all other conditions being equal (thermochemical mantle model, 119 Myr of assimilation of plate tectonics reconstructions). Initial states are: a) 1-D temperature profile b) 60 Myr of imposed boundary conditions corresponding to the velocity and layout of plates 119 Myr ago c) same as b) but with 120 Myr. Taken from McNamara & Zhong (2005).

Both semi-empirical sequential and backward advection methods have shown their sensitivity to variations in mantle convection models (e.g. O'Neill et al., 2005; Bello et al., 2015) and in observed data (e.g. Doubrovine et al., 2012; Shephard et al., 2012). The limits of these methods motivate the development of methods to assimilate data on mantle circulation into mantle convection models while taking into account the uncertainties on both the model and the observed data. We describe these techniques in the next section.

### 2.3.2 Data Assimilation Methods

In Section 2.1, we described the problem of mantle circulation reconstruction as as a time-dependent state estimation problem, following the classical notations of data assimilation. Two approaches exist to find a solution to this problem: sequential data assimilation and 4-D variational data assimilation.

So far, data assimilation methods for mantle circulation have been developed either as simplified schemes directly derived from 4-D variational data assimilation (e.g. Bunge et al., 2003), or as techniques derived from optimal control theory, bearing similarities with 4-D variational data assimilation (e.g. Ismail-Zadeh et al., 2003b,a, 2007). In this dissertation, we present two applications of sequential data assimilation methods to the mantle circulation problem. To ease the comparison between techniques, we keep

the general formulation of data assimilation derived from the discretized state and model, as in Section 2.1. We first describe the common objective of data assimilation techniques. Then we describe the principle of 4-D variational data assimilation and the applications to mantle circulation. Finally, we describe the general principle of sequential data assimilation.

In the following, we consider only the second order statistics of the errors affecting the states and data. Other formulations exist, taking into account a more thorough description of error statistics (see e.g. Bocquet et al., 2010, for a review). More detailed descriptions of data assimilation techniques can be found in classical textbooks of atmospheric sciences and numerical weather prediction (e.g. Bengtsson et al., 1981; Daley, 1993; Kalnay, 2003), oceanography (e.g. Chassignet & Verron, 2006) or data assimilation in general (e.g. Evensen, 2009a; Swinbank et al., 2012; Blayo et al., 2014).

**Objective of Data Assimilation**

As described in Section 2.1, the unknown is defined as a time series of true state vectors $\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_K^t\}$ and three sources of information are available, all affected by errors:

- the background state $\mathbf{x}^b$, which is related to the first true state by

$$\mathbf{x}^b = \mathbf{x}_1^t + \boldsymbol{\epsilon}^b \tag{2.15}$$

  with $\boldsymbol{\epsilon}^b$ the error on the background state. The error is modelled as a random vector of $\mathbf{0}$ mean (the background state is supposed unbiased) and covariance matrix $\mathbf{P}^b$.

- the time series of observed data $\{\mathbf{y}_1^o, \mathbf{y}_2^o, ..., \mathbf{y}_K^o\}$ and their associated observation operators $\{\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}_K\}$. At a given time step $k$ the observed data vector $\mathbf{y}_k^o$ is related to the true state $\mathbf{x}_k^t$ by

$$\mathbf{y}_k^o = \mathcal{H}_k(\mathbf{x}_k^t) + \boldsymbol{\epsilon}_k^o + \boldsymbol{\epsilon}_k^h \tag{2.16}$$

with $\boldsymbol{\epsilon}_k^o$ the measurement errors and $\boldsymbol{\epsilon}_k^h$ the errors linked to the observation operator. The sum of errors $\boldsymbol{\epsilon}_k^{o,h} = \boldsymbol{\epsilon}_k^o + \boldsymbol{\epsilon}_k^h$ is modelled as a random vector of $\mathbf{0}$ mean (observations are supposed unbiased) and covariance matrix $\mathbf{R}_k$.

- the dynamical model $\mathcal{M}$, which relate a given state $\mathbf{x}_{k-1}$ at time step $k-1$ to the state vector $\mathbf{x}_k$ at time step $k$ by

$$\mathbf{x}_k = \mathcal{M}(\mathbf{x}_{k-1}) + \boldsymbol{\eta}_k \tag{2.17}$$

with $\boldsymbol{\eta}_k$ the errors linked to modelling approximations. The error is modelled as a random vector of $\mathbf{0}$ mean (the model is supposed unbiased) and covariance matrix $\mathbf{Q}_k$.

We cannot compute directly the error made by a given evolution $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K\}$ with respect to the true state evolution $\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_K^t\}$, since it is unknown. However, we can compute the distances between the estimated evolution and the constraints given by the different sources of information. More precisely, we compute the Mahalanobis distances (Mahalanobis, 1936). The Mahalanobis distance $\Delta_{\mathbf{a}}(\mathbf{x})$ is a measure of the distance of a vector $\mathbf{x}$ to the multivariate continuous random vector $\mathbf{a}$ of mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

$$\Delta_{\mathbf{a}}(\mathbf{x}) = \sqrt{(\mathbf{x}-\boldsymbol{\mu})^T (\boldsymbol{\Sigma})^{-1} (\mathbf{x}-\boldsymbol{\mu})}. \tag{2.18}$$

For example, when the covariance matrix $\boldsymbol{\Sigma}$ is the identity matrix, this means that the variance of each component of the random vector $\mathbf{a}$ is one and the components of $\mathbf{a}$ are uncorrelated. In this case, the Mahalanobis distance corresponds to the Euclidean distance between the vector $\mathbf{x}$ and the mean $\boldsymbol{\mu}$ (i.e. the L2 norm of $\mathbf{x} - \boldsymbol{\mu}$).

For our problem, we compute the following Mahalanobis distances:

$$\left[\Delta_{\boldsymbol{\epsilon}^b}(\mathbf{x}_1 - \mathbf{x}^b)\right]^2 = (\mathbf{x}_1 - \mathbf{x}^b)^T \left(\mathbf{P}^b\right)^{-1} (\mathbf{x}_1 - \mathbf{x}^b) \tag{2.19a}$$

$$\forall k \in \{1, .., K\},$$

$$\left[\Delta_{\boldsymbol{\epsilon}_k^{o,h}}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k^o)\right]^2 = (\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k^o)^T (\mathbf{R}_k)^{-1} (\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k^o) \tag{2.19b}$$

$$\forall k \in \{2, .., K\},$$

$$\left[\Delta_{\boldsymbol{\eta}_k}(\mathcal{M}(\mathbf{x}_{k-1}) - \mathbf{x}^k)\right]^2 = (\mathcal{M}(\mathbf{x}_{k-1}) - \mathbf{x}_k)^T (\mathbf{Q})^{-1} (\mathcal{M}(\mathbf{x}_{k-1}) - \mathbf{x}_k). \tag{2.19c}$$

The global aim of data assimilation methods is to find a state evolution which minimizes these distances. 4-D variational and sequential data assimilation methods differ by the strategies adopted to determine this evolution.

#### 4-D Variational Data Assimilation

4-D variational data assimilation methods aim at providing a time series of the estimated evolution of the state of the system such that the objective function $\mathcal{J}$ defined by

$$\begin{aligned}\mathcal{J}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K) =& (\mathbf{x}_1 - \mathbf{x}^b)^T \left(\mathbf{P}^b\right)^{-1} (\mathbf{x}_1 - \mathbf{x}^b) \\ &+ \sum_{k=1}^{K} (\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k^o)^T (\mathbf{R}_k)^{-1} (\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k^o) \\ &+ \sum_{k=2}^{K} (\mathbf{x}_k - \mathcal{M}(\mathbf{x}_{k-1}))^T (\mathbf{Q}_k)^{-1} (\mathbf{x}_k - \mathcal{M}(\mathbf{x}_{k-1}))\end{aligned} \tag{2.20}$$

is minimized. The minimization of this function is called weak constraint four dimensional variational assimilation or weak constraint 4D-var, following the terminology of Sasaki (1970). The term "weak constraint" comes from the fact that we take into account modelling errors.

Bunge et al. (2003) writes the equivalent of the objective function of Equation 2.20 in the continuous case, using the classical conservation equations governing mantle flow,

as described in Chapter 1. However, the mantle model they consider differs from the model used in this dissertation on two major points.

First, in these models, the viscosity depends neither on temperature, nor on the state of stress, but only on depth. This has important consequences for the development of data assimilation techniques. As discussed in Section 1.4, the temperature and stress dependence of the viscosity is necessary to obtain mantle convection models which display plate-like tectonics at their surface. If we use a mantle model which does not have this type of viscosity law, then it is not possible to integrate the data derived from plate reconstructions as is done in classical data assimilation schemes, i.e. by defining an operator $\mathcal{H}$ that links a state computed from the model to surface velocities. In this case, the only data which is usable as proper "observed data", as defined previously, is the temperature field at present, estimated from seismic tomography. The time series of observed data $\{\mathbf{y}_1^o, \mathbf{y}_2^o, ..., \mathbf{y}_K^o\}$ shrinks to $\{\mathbf{y}_K^o\}$. The vector $\mathbf{y}_K^o$ contains the values of the discretized temperature field derived from seismic tomography. The associated observation operator $\mathcal{H}_K$ corresponds to the interpolation of the temperature field from the grid defined in mantle convection models to the different discretization used in seismic tomography models. The equivalent discrete objective function becomes:

$$\begin{aligned}\mathcal{J}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K) =& (\mathbf{x}_1 - \mathbf{x}^b)^T \left(\mathbf{P}^b\right)^{-1} (\mathbf{x}_1 - \mathbf{x}^b) \\ &+ (\mathcal{H}_K(\mathbf{x}_K) - \mathbf{y}_K^o)^T (\mathbf{R}_K)^{-1} (\mathcal{H}_K(\mathbf{x}_K) - \mathbf{y}_K^o) \\ &+ \sum_{k=2}^{K} (\mathbf{x}_k - \mathcal{M}(\mathbf{x}_{k-1}))^T (\mathbf{Q}_k)^{-1} (\mathbf{x}_k - \mathcal{M}(\mathbf{x}_{k-1}))\end{aligned} \tag{2.21}$$

The second difference is the mechanical boundary conditions at the surface of the model. Since the mantle convection model cannot produce plate-like tectonics at the surface self-consistently, plate tectonics reconstructions have to be imposed as surface boundary conditions (Equation 2.10) to obtain an "Earth-like" evolution. In this formulation, the mantle model $\mathcal{M}$ includes the plate tectonics reconstructions.

In this specific case, Bunge et al. (2003) derives Euler-Lagrange equations associated to the continuous objective function. In principle, the solution to the Euler-Lagrange

equations provides the state evolution which minimizes the objective function. However, in practice, further assumptions are made before computing the solution to this problem. Bunge et al. (2003) suppose no error on the model (which means in their formulation also no error on plate tectonics reconstructions). If we neglect modelling errors, then the evolution of the state of the system can be determined solely from the first state:

$$\forall k \in \{2, ..., K\}, \quad \mathbf{x}_k = \mathcal{M}(\mathbf{x}_{k-1}) \tag{2.22}$$

Under this assumption, the objective function $\mathcal{J}$ depends only on the initial state $\mathbf{x}_1$ and becomes

$$\begin{aligned}\mathcal{J}(\mathbf{x}_1) =&(\mathbf{x}_1 - \mathbf{x}^b)^T \left(\mathbf{P}^b\right)^{-1} (\mathbf{x}_1 - \mathbf{x}^b) \\ &+ (\mathcal{H}_K(\mathbf{x}_K) - \mathbf{y}_K^o)^T (\mathbf{R}_K)^{-1} (\mathcal{H}_K(\mathbf{x}_K) - \mathbf{y}_K^o).\end{aligned} \tag{2.23}$$

We consider further that the covariance matrices associated to the background state error $\mathbf{P}^b$ and the observed data error $\mathbf{R}_K$ are of the form:

$$\mathbf{P}^b = V^b \mathbf{I} \qquad \text{and} \qquad \mathbf{R}_K = V^o \mathbf{I} \tag{2.24}$$

where $\mathbf{I}$ is the identity matrix and $V^b$ and $V^o$ are the variances associated to each component of the background state and observation error random vectors.

We obtain the cost function

$$\begin{aligned}\mathcal{J}(\mathbf{x}_1) =&\frac{1}{V^b}(\mathbf{x}_1 - \mathbf{x}^b)^T(\mathbf{x}_1 - \mathbf{x}^b) \\ &+ \frac{1}{V^o} (\mathcal{H}_K(\mathbf{x}_K) - \mathbf{y}_K^o)^T (\mathcal{H}_K(\mathbf{x}_K) - \mathbf{y}_K^o) .\end{aligned} \tag{2.25}$$

In the continuous form, and considering that the states and data contain only temperature fields, this translates to

$$\mathcal{J}_c(\mathcal{T}_i) = \frac{1}{V^b} \int_\Omega \left[\mathcal{T}_i(\mathbf{a}) - \mathcal{T}^b(\mathbf{a})\right]^2 \, \mathrm{d}\mathbf{a} + \frac{1}{V^o} \int_\Omega \left[\mathcal{T}_f(\mathbf{a}) - \mathcal{T}^o(\mathbf{a})\right]^2 \, \mathrm{d}\mathbf{a}. \tag{2.26}$$

Where $\mathcal{T}_i$ is the initial temperature field to be determined, $\mathcal{T}_f$ the final temperature, computed from the evolution of $\mathcal{T}_i$ using the forward model of mantle convection. $\mathcal{T}^b$ is the a priori initial temperature field, and $\mathcal{T}^o$ is the temperature field at present, derived from seismic tomography. $\Omega$ represents the volume of the mantle and $\mathbf{a}$ is the position vector.

The function $\mathcal{J}_c$, (or a further simplification, taking $V^b \to \infty$ and $V^o = 1$), along with a mantle convection model, form the basis of the development of current data assimilation methods applied to the mantle circulation inverse problem. In practice, iterative methods for the minimization of the function $\mathcal{J}_c$ are developed in the continuous form, and the resulting equations are then discretized to solve the system.

The general organization of algorithms is:

1. Initialization with a discretized temperature field $T_i^{(1)}$.
2. Iteration until convergence of the solution. At the $n$th iteration
   (a) Computation of the state evolution using the forward code of mantle convection, from a given discretized initial state $T_i^{(n)}$ we obtain a final state $T_f^{(n)}$.
   (b) Computation of the correction $\delta T^{(n)}$ to add to the initial state to minimize the cost function. The new initial temperature field is then $T_i^{(n+1)} = T_i^{(n)} + \delta T^{(n)}$

Algorithms differ by the techniques used to estimate the correction $\delta T^{(n)}$ in stage 2b.

One solution, adopted by Bunge et al. (2003), and further investigated by Horbach et al. (2014), Vynnytska & Bunge (2015) and Ghelichkhan & Bunge (2016) is to use the adjoint method to obtain a first order estimation of the gradient of the objective function $\nabla \mathcal{J}_c$ around a given initial temperature field. This method requires the derivation of a set of partial differential adjoint equations associated with the set of partial differential equations defining the forward model of mantle convection. The stage 2b corresponds to the backward computation of the discretized adjoint equations. The advantage of this method is that the partial differential adjoint equations are very similar to the forward differential equations, which allows the use of the same code to perform stage 2a

and stage 2b. However, this also means that the computational time required for stage 2a is equivalent as the one needed for stage 2b. Moreover, the computation of stage 2b requires to store the whole set of temperature and velocity evolutions computed in stage 2a. Bunge et al. (2003) performs a preliminary test in 3 dimensions for a $100$ Myr evolution and obtain convergence of the objective function value after $100$ iterations. Horbach et al. (2014) apply this method to the reconstruction of mantle circulation for the past $40$ Myr. They test the sensitivity of the method to the initial temperature field chosen in stage 1 of the algorithm, and show that the solutions, after several iterations, all converge to the same evolution. Vynnytska & Bunge (2015) further test this variational algorithm using synthetic experiments in 2D cylindrical geometry. First, they compute a reference evolution using the forward model, with a free-slip boundary condition. The surface velocities and final temperature fields are extracted from the reference evolution. Then the variational algorithm is applied to estimate an evolution from these two constraints: surface velocities are used as boundary conditions, and the final temperature field is the data to fit. The background state is a spherically symmetric temperature field. They limit the timespan of the evolution to $50$ Myr. Within this period, the evolution inferred from the variational data assimilation algorithm converges towards the reference solution. They also perform synthetic experiments where the surface boundary conditions are not imposed during the assimilation. In this case, the assimilation algorithm produces an evolution whose final temperature field converges towards the reference final temperature field, but the whole estimated evolution diverges from the reference evolution, showing the presence of local minima of the objective function in Equation 2.26 if velocities are not imposed at the surface of the mantle convection model. More generally, these results show the importance of surface velocity constraints to estimate the evolution of the temperature field in the past.

Ismail-Zadeh et al. (2003a,b, 2004) also use the adjoint method but derive the adjoint equations considering a model restricted to the equations of conservation of energy (and associated boundary conditions), supposing no effect of the temperature variation on the velocity field. In step 2a, they consider the complete mantle model with the

equations of conservation of mass, momentum and energy and the associated boundary conditions, arguing that the iteration of step 2a and step 2b will then lead to an estimation of the evolution consistent with the physics of mantle convection. The computational cost of this technique is lighter than the former method and this formulation also allows the implementation of temperature-dependent viscosity. Liu & Gurnis (2008) perform synthetic regional tests using this method and show that the result of the assimilation depends strongly on the chosen initial state in stage 1 of the algorithm. To obtain convergence towards the reference evolution, it is necessary to compute the initial temperature field (stage 1 of the algorithm) from a simple backward advection, as done in section 2.3.1.

Finally, Ismail-Zadeh et al. (2007) propose an algorithm based on the regularization of the backward heat equation (Equation 2.12) for stage 2b. This method is called the quasi-reversibility method (Lattés & Lions, 1969). With synthetic tests, Ismail-Zadeh et al. (2007) finds a lower accuracy of the quasi-reversibility method compared to adjoint methods. However, the quasi-reversibility method is computationally lighter (see Ismail-Zadeh et al., 2016, Chapter 7 for a comparison between backward advection, adjoint and quasi-reversibility methods). Glišović & Forte (2014, 2016) recently applied the quasi-reversibility method to the reconstruction of mantle circulation for the last $65$ Myr.

All the techniques described previously, when applied to global reconstructions of mantle circulation, impose plate tectonic reconstructions as mechanical boundary conditions, without taking into account the uncertainties on these reconstructions. As shown for example by the synthetic experiments of Vynnytska & Bunge (2015), the imposition of surface constraints during the assimilation is decisive to obtain solutions that converge to the true evolution (also highlighted by Glišović & Forte, 2016). Given the uncertainties of plate tectonics reconstructions as we go back in time (see section 2.2.2), it then appears important to be able to take them into account in the data assimilation algorithms developed.

In this dissertation, we focus on the development of data assimilation methods that

consider the history of plate tectonics not as a boundary condition that is imposed at the surface of the model, but as observed data, which are assimilated in a consistent fashion in mantle circulation models. To do so, we use direct models of mantle convection that are able to produce plate-like tectonics at their surface self-consistently. These models have viscosity laws that strongly vary with both temperature and stress, as is discussed in Section 1.4. This introduces strong nonlinearities in the flow equations, which complicates the development of 4-D variational data assimilation methods. Hence, we turn to sequential data assimilation methods, the principle of which we describe in the next section.

**Sequential Data Assimilation**

The exact formulation of the sequential data assimilation algorithms that were developed during this thesis are described in the next two chapters. Here, we simply highlight the principle of the method. Although sequential methods find their roots in estimation theory more than optimal control theory (see the following Chapters), we present here the different estimations as the minimization of objective functions, so as to ease the comparison with the previous 4-D variational data assimilation methods.

Sequential methods divide the problem into several consecutive stages:

1. The best guess for the initial state $\mathbf{x}_1$ is computed by minimizing the objective function:

$$\begin{aligned}\mathcal{J}(\mathbf{x}_1) =&(\mathbf{x}_1 - \mathbf{x}^b)^T \left(\mathbf{P}^b\right)^{-1} (\mathbf{x}_1 - \mathbf{x}^b) \\ &+ (\mathcal{H}_1(\mathbf{x}_1) - \mathbf{y}_1^o)^T (\mathbf{R}_1)^{-1} (\mathcal{H}_1(\mathbf{x}_1) - \mathbf{y}_1^o)\end{aligned} \tag{2.27}$$

the result of this minimization is called the first analyzed state $\mathbf{x}_1^a$, which is linked to the first true state by $\mathbf{x}_1^a = \mathbf{x}_1^t + \boldsymbol{\epsilon}_1^a$, with $\boldsymbol{\epsilon}_1^a$ the error on the first analyzed state. $\boldsymbol{\epsilon}_1^a$ is modelled as a random vector, of mean $\mathbf{0}$ and covariance matrix $\mathbf{P}_1^a$ (called the analyzed error covariance matrix). $\mathbf{P}_1^a$ is also evaluated during this stage by

taking into account the errors on both the background state and the first observed data.

2. Two steps are repeated, until all observations have been taken into account. At timestep $k \in \{2, 3, ..., K\}$:

   (a) the **Forecast** state $\mathbf{x}_k^f$ is computed using the model $\mathcal{M}$:

   $$\mathbf{x}_k^f = \mathcal{M}(\mathbf{x}_{k-1}^a). \tag{2.28}$$

   The forecast state is linked to the $k$th true state by $\mathbf{x}_k^f = \mathbf{x}_k^t + \boldsymbol{\epsilon}_k^f$, with $\boldsymbol{\epsilon}_k^f$ the error on the $k$th forecast state. $\boldsymbol{\epsilon}_k^f$ is modelled as a random vector, of mean $\mathbf{0}$ and covariance matrix $\mathbf{P}_k^f$ (called the forecast error covariance matrix). $\mathbf{P}_k^f$ is also evaluated during this stage, by taking into account the errors on the analyzed state and the modelling errors.

   (b) the **Analyzed** state $\mathbf{x}_k^a$ is determined by minimizing the objective function

   $$\begin{aligned}\mathcal{J}(\mathbf{x}_k) =& (\mathbf{x}_k - \mathbf{x}_k^f)^T \left(\mathbf{P}_k^f\right)^{-1} (\mathbf{x}_k - \mathbf{x}_k^f) \\ &+ (\mathcal{H}_1(\mathbf{x}_k) - \mathbf{y}_k^o)^T (\mathbf{R}_k)^{-1} (\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k^o).\end{aligned} \tag{2.29}$$

   $\mathbf{x}_k^a$ is linked to the first true state by $\mathbf{x}_k^a = \mathbf{x}_k^t + \boldsymbol{\epsilon}_k^a$, with $\boldsymbol{\epsilon}_k^a$ the error on the $k$th analyzed state. $\boldsymbol{\epsilon}_k^a$ is modelled as a random vector, of mean $\mathbf{0}$ and covariance matrix $\mathbf{P}_k^a$ (called the analyzed error covariance matrix). $\mathbf{P}_k^a$ is also evaluated during this stage, by taking into account errors on the $k$th forecast state and errors on the $k$th observed data.

In this scheme, the observations are integrated sequentially: the evolution of the estimated state is computed forward in time, and the observed data, as they become available, are assimilated into the estimated state evolution. The estimated state evolution should then converge towards the true evolution, as more and more data are assimilated.

The following chapters 3 and 4 propose two algorithms based on this principle to reconstruct mantle circulation. We performed synthetic experiments on both these algorithms to determine whether it is possible to reconstruct mantle circulation from the sole and imperfect knowledge of the surface tectonics history, and under which conditions a satisfactory reconstruction is possible. The first one is a suboptimal scheme derived from the classical Kalman Filter (see e.g. Todling & Cohn, 1994, for a review on suboptimal schemes based on Kalman filters). This scheme provided encouraging results and drove us to develop a more efficient sequential data assimilation algorithm for the mantle circulation problem. This is what we do in Chapter 4 with the implementation of an Ensemble Kalman Filter for this problem.

## Résumé du chapitre

Le but d'une reconstitution de la circulation mantellique est d'estimer l'écoulement 3-D dans le manteau, au présent et dans le passé. Pour atteindre ce but, on combine des données sur la circulation mantellique avec un modèle dynamique de convection mantellique. On dispose de données sur la circulation mantellique au présent, avec par exemple les contraintes fournies par la tomographie sismique, et dans le passé, avec par exemple les contraintes sur l'histoire de la tectonique de surface. Les problèmes inverses s'intéressant à des données asynchrones appartiennent à la catégorie des problèmes d'assimilation de données. Ce chapitre vise à présenter le problème de l'assimilation de données pour la circulation mantellique. Nous commençons par formuler le problème de la circulation mantellique dans le cadre de l'assimilation de données. Nous décrivons ensuite les données actuellement utilisées dans les modèles de circulation mantellique: le champ de température actuel et l'histoire de la cinématique de surface. Finalement on présente les applications à la reconstruction de la circulation mantellique. Trois stratégies alternatives ont été proposées pour reconstruire l'histoire de la circulation mantellique: l'advection rétrograde, les méthodes séquentielles semi-empiriques et l'assimilation de données variationnelle. Dans la présente thèse, nous proposons une quatrième méthode: l'assimilation de données séquentielle, que nous développons dans chapitres 3 et 4. Ces quatre méthodes sont présentées dans la troisième section de ce chapitre.

# CHAPTER 3

## A sequential data assimilation approach for the joint reconstruction of mantle convection and surface tectonics

The following Chapter has been published in january 2016 as:

Bocher, M., Coltice, N., Fournier, A., & Tackley, P., 2016. A sequential data assimilation approach for the joint reconstruction of mantle convection and surface tectonics, *Geophysical Journal International*, **204**(1), 200–214

# Abstract

With the progress of mantle convection modelling over the last decade, it now becomes possible to solve for the dynamics of the interior flow and the surface tectonics to first order. We show here that tectonic data (like surface kinematics and seafloor age distribution) and mantle convection models with plate-like behaviour can in principle be combined to reconstruct mantle convection. We present a sequential data assimilation method, based on suboptimal schemes derived from the Kalman filter, where surface velocities and seafloor age maps are not used as boundary conditions for the flow, but as data to assimilate. Two stages (a forecast followed by an analysis) are repeated sequentially to take into account data observed at different times. Whenever observations are available, an analysis infers the most probable state of the mantle at this time, considering a prior guess (supplied by the forecast) and the new observations at hand, using the classical best linear unbiased estimate. Between two observation times, the evolution of the mantle is governed by the forward model of mantle convection. This method is applied to synthetic 2-D spherical annulus mantle cases to evaluate its efficiency. We compare the reference evolutions to the estimations obtained by data assimilation. Two parameters control the behaviour of the scheme: the time between two analyses, and the amplitude of noise in the synthetic observations. Our technique proves to be efficient in retrieving temperature field evolutions provided the time between two analyses is 10 Myr. If the amplitude of the a priori error on the observations is large (30 per cent), our method provides a better estimate of surface tectonics than the observations, taking advantage of the information within the physics of convection.

## 3.1 Introduction

Global tectonics is a surface expression of mantle convection (Bercovici, 2003): the motions of continents and seafloor are generated by forces acting within the mantle and the lithosphere. For example, Ricard et al. (1989) and Alisic et al. (2012) obtained a consistent description of surface kinematics by converting the long wavelength heterogeneities of seismic velocity into buoyancy forces. Mantle convection studies also attest to this link, showing for example that a downwelling in a context of large-scale convection draws continents to aggregate (Zhong, 2001; Rolf et al., 2014).

Reconstructing the convective history of the Earth's mantle is a long-standing problem in geosciences (Bunge et al., 1998). Indeed, an accurate determination of mantle temperature and velocity fields evolution only for the last $200$ My would already lead to a better understanding of ore deposits formation, water resources distribution, the evolution of the geodynamo on geological time scales or deep material properties. The geodynamics community has taken advantage of the link between mantle convection and tectonics to build a first approach to reconstruct the convection history of the past $100$ My. For this approach, plate tectonics theory is used to describe surface kinematics. By driving convection currents at the surface with plate kinematic reconstructions, it is possible to propose a scenario of the evolution of mantle heterogeneities. Results so obtained proved to be relatively consistent with long wavelength seismic tomography or/and the geoid (Bunge & Grand, 2000; Flament et al., 2013). These 3D spherical convection models are called mantle circulation models (Bunge et al., 2002), as a reference to oceanography. They require very fine parametrization in order to generate internal structures consistent with both the physics of convection and geophysical observations (Bower et al., 2015). To go one step further, several groups used tomographic models as data, in addition to plate reconstructions (Bunge et al., 2003; Ismail-Zadeh et al., 2007; Liu & Gurnis, 2008). A temperature field of the present-day Earth's mantle is generated from these tomographic models, and the past mantle circulation is retrieved following various methods such as backward advection, variational methods and quasi-reversibility methods (see Ismail-Zadeh & Tackley (2010) for a full report on

these methods). There are two drawbacks in this methodology, both difficult to overcome. First, the convection models used so far do not naturally produce Earth-like tectonics: they generate either sluggish tectonics at the surface or a stagnant lid when larger viscosity contrasts are prescribed (Solomatov, 1995). Second, although tectonic data are the most important source of information (they are time-dependent, unlike seismic data), they are used only to drive the flow, not as data to match. To drive the flow, the kinematics needs to be known at every time-step of the calculation. This is not the case for plate tectonic reconstructions, even though Gurnis et al. (2012) designed a procedure to interpolate the plate tectonics geometry between times at which it is known.

Significant progress has been made in the past 10 years on modelling convection that produces more realistic surface tectonics. Convection with a pseudo-plastic rheology generates surface tectonics with a plate-like behaviour (Moresi & Solomatov, 1998; Moresi et al., 2000; Stein et al., 2004; Bercovici, 2003; Tackley, 2000b; Van Heck & Tackley, 2008; Bercovici & Ricard, 2014). Recent models display seafloor spreading and continental drift comparable to that of the Earth to first order: seafloor age distributions and the time scale of spreading fluctuations are consistent with what has been inferred for the Earth for the last 200 My (Coltice et al., 2012, 2013). This opens the way to producing a mantle circulation model using convection models with plate-like behaviour, under the observational constraint provided by tectonic data. This constraint is not to be enforced through kinematic boundary conditions imposed at every time step. Instead, it should be taken into account in a statistically consistent fashion, whereby the prediction due to the model and the observations are combined in a way that respects the uncertainties affecting both.

We explore here this possibility using a controlled case, and present a proof of concept. We have developed a sequential data assimilation method which is a suboptimal scheme based on the Kalman filter (see Todling & Cohn (1994) for a review of suboptimal schemes based on the Kalman filter). Such data assimilation methodology

provides the best linear unbiased estimate of the temperature field and surface velocities at times data are available, and a prediction of the evolution of the system state between those times. Using synthetic tests in spherical annulus geometry (Hernlund & Tackley, 2008), we show that our data assimilation method is efficient for reconstructions over $200$ My with data gaps of $5 - 10$ My and relative uncertainties on surface observations below $30\%$. This method also allows the refinement of surface tectonics, taking advantage of the information within the physics of convection.

In the following, we first describe the general approach to Kalman Filtering in Section 2. Section 3 provides details on the definition of the different operators and vectors forming the backbone of a sequential data assimilation framework: the physical model, the data and the state for mantle convection, as well as the observation operator. Section 4 presents our sequential data assimilation algorithm and the adjustments used for the reconstruction of mantle convection. Then, we present in Section 5 an evaluation of the method with a series of synthetic experiments, which leads to the discussion of section 6.

## 3.2 The Extended Kalman Filter

We introduce in this section the Extended Kalman Filter that we adapt and apply to the convection reconstruction problem. We use here classical notations, taken from Ide et al. (1997). Our goal is to estimate the evolution of the mantle temperature field, using information from a mantle convection model and tectonic data including global kinematics and seafloor ages. The following is a rather general description of the Extended Kalman Filter; similar developments can be found in Ghil & Malanotte-Rizzoli (1991); Kalnay (2003); Wunsch (2006) for meteorology and oceanography and in Fournier et al. (2010) for geomagnetism, for example.

Data assimilation aims at estimating the state of a dynamical system and its evolution for a given period of time, combining information from observations and a physical model. The evolution of the state of the system is described by the set of true state

vectors, storing the chosen scalar fields that are discretized:

$$\left\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_{n_t}^t\right\}, \tag{3.1}$$

the subscripts $\{1, 2, ..., n_t\}$ being the discrete time steps.

Likewise, the observations are defined as a set of column vectors

$$\left\{\mathbf{y}_{t_1}^o, \mathbf{y}_{t_2}^o, ..., \mathbf{y}_{t_n}^o\right\} \text{ with } \{t_1, t_2, ..., t_n\} \subset \{1, 2, ..., n_t\}. \tag{3.2}$$

The true state and the observations are linked linearly by the observation operator $\mathbf{H}$

$$\forall t_i \in \{t_1, t_2, ..., t_n\}, \quad \mathbf{y}_{t_i}^o = \mathbf{H}\mathbf{x}_{t_i}^t + \boldsymbol{\epsilon}_{t_i}^o, \tag{3.3}$$

where $\boldsymbol{\epsilon}_{t_i}^o$ is the observation error at the timestep $t_i$. The precise value of $\boldsymbol{\epsilon}_{t_i}^o$ is unknown. However, it is possible to model it as a random vector and estimate its probability density function (PDF). We assume $\boldsymbol{\epsilon}_{t_i}^o$ follows a centred Gaussian statistics, and its associated covariance matrix

$$\mathbf{R}_{t_i} = \mathrm{Covar}(\boldsymbol{\epsilon}_{t_i}^o) = \langle \boldsymbol{\epsilon}_{t_i}^o (\boldsymbol{\epsilon}_{t_i}^o)^T \rangle \tag{3.4}$$

is called the observation error covariance matrix. The operator $[]^T$ means transpose. The Extended Kalman Filter belongs to the Sequential Data Assimilation techniques. It consists of integrating the observation vectors one after the other into a numerical model, when they become available. Fig. 3.1 gives an overview of this procedure. The initialization, at timestep $t_1$, provides an a priori guess of the state of the system. Then two steps called analysis and forecast are performed sequentially, every time a new observation vector becomes available, until all the observations have been taken into account. Hence the name of the technique: sequential data assimilation. The analysis corrects the current estimate of the system state by considering the new observations. The forecast provides an estimation of the evolution of the system until the next time new observations are available. The evolution of the estimated state should converge

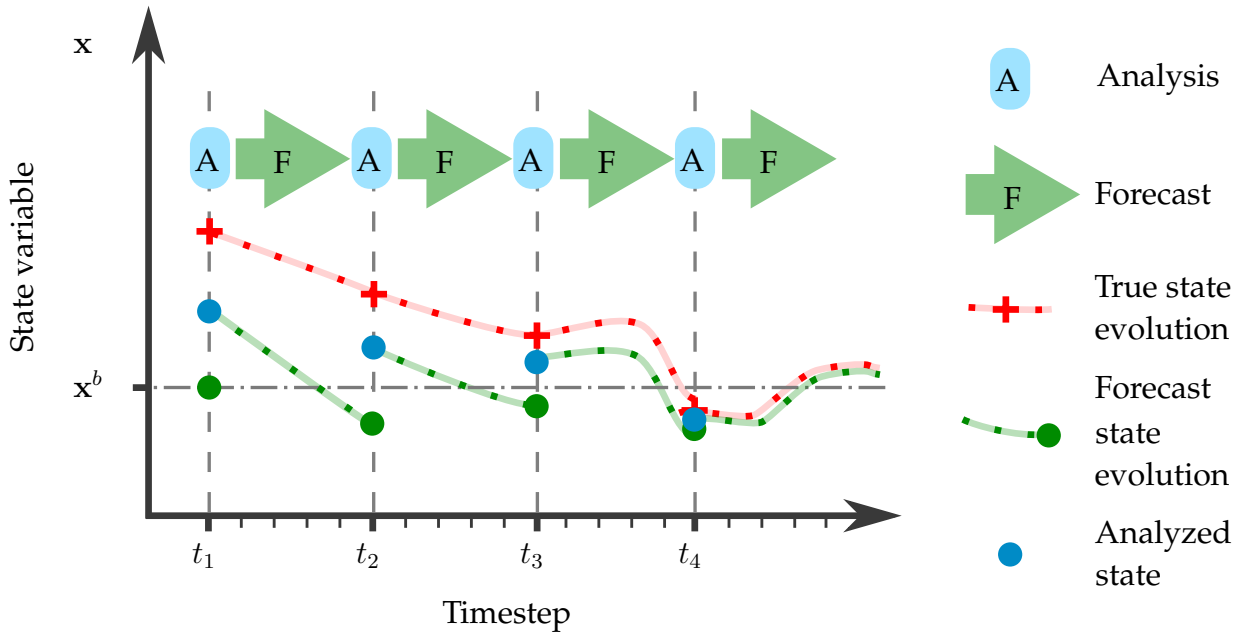


FIGURE 3.1: Principle of Sequential Data Assimilation. The initialization at time $t_1$ sets the background state $\mathbf{x}^b$ (green dot at $t_1$). $\mathbf{x}^b$ is corrected by taking into account new observations, leading to the first analyzed state $\mathbf{x}^a_{t_1}$ (blue dot at $t_1$), which is closer to the true state $\mathbf{x}^t_{t_1}$ (red cross at $t_1$). The evolution of the state (green line) is then computed by the model $\mathcal{M}$ until $t_2$, leading to a new forecast state $\mathbf{x}^f_{t_2}$ (green dot). A new analyzed state $\mathbf{x}^a_{t_2}$ is computed considering new observations. The sequence is repeated until all observations have been taken into account. The aim of data assimilation is to get as close as possible to the true state evolution (red line).

towards the true state as more observations are assimilated.

### 3.2.1 Initialization

The Extended Kalman Filter algorithm starts at the timestep $t_1$, when the first observations are available. Before taking into account the first observations, an a priori state of the system has to be estimated. It is called the background state $\mathbf{x}^b$. The relationship between the background state and the true state at $t_1$ is

$$\mathbf{x}_b = \mathbf{x}^t_{t_1} + \boldsymbol{\epsilon}^b, \tag{3.5}$$

where $\boldsymbol{\epsilon}^b$ is the background state error. $\boldsymbol{\epsilon}^b$ is a random vector with $\langle \boldsymbol{\epsilon}^b \rangle = \mathbf{0}$ and $\mathrm{Covar}(\boldsymbol{\epsilon}^b) = \mathbf{P}^b$.

The background state vector is considered as the forecast state vector at timestep $t_1$, written $\mathbf{x}^f_{t_1}$ and its covariance matrix $\mathbf{P}^b$ is the forecast error covariance matrix at $t_1$, written $\mathbf{P}^f_{t_1}$.

### 3.2.2 Analysis and forecast sequence

At $t_i$, a new estimate of the state of the system is calculated considering $\mathbf{y}^o_{t_i}$, $\mathbf{R}_{t_i}$, $\mathbf{x}^f_{t_i}$ and $\mathbf{P}^f_{t_i}$. This step is called analysis, the new estimate being the analyzed state $\mathbf{x}^a_{t_i}$ . The relationship between the analyzed state and the true state is

$$\mathbf{x}^a_{t_i} = \mathbf{x}^t_{t_i} + \boldsymbol{\epsilon}^a_{t_i}, \tag{3.6}$$

where $\boldsymbol{\epsilon}^a_{t_i}$ is the analysis error at $t_i$. $\boldsymbol{\epsilon}^a_{t_i}$ is modelled as a random vector following a Gaussian law, and its covariance $\mathbf{P}^a_{t_i}$ is the analysis error covariance matrix. The analyzed state $\mathbf{x}^a_{t_i}$ is chosen so that

- $\mathbf{x}^a_{t_i}$ is a linear combination of $\mathbf{x}^f_{t_i}$ and $\mathbf{y}^o_{t_i}$,
- $\mathrm{Tr}(\mathbf{P}^a_n)$ is minimised,

- $\langle \boldsymbol{\epsilon}_n^a \rangle = \mathbf{0}$.

These conditions lead to the classical Best Linear Unbiased Estimate (see Ghil & Malanotte-Rizzoli (1991) for example)

$$\mathbf{x}_{t_i}^a = \mathbf{x}_{t_i}^f + \mathbf{K}_{t_i} \left[ \mathbf{y}_{t_i}^o - \mathbf{H}\mathbf{x}_{t_i}^f \right], \tag{3.7}$$

$$\mathbf{K}_{t_i} = (\mathbf{H}\mathbf{P}_{t_i}^f)^T \left[ \mathbf{H}(\mathbf{H}\mathbf{P}_{t_i}^f)^T + \mathbf{R} \right]^{-1}, \tag{3.8}$$

$$\mathbf{P}_{t_i}^a = (\mathbf{I} - \mathbf{K}_{t_i}\mathbf{H})\mathbf{P}_{t_i}^f, \tag{3.9}$$

where $\mathbf{K}_{t_i}$ is the Kalman gain at $t_i$. If the probability density functions of the errors on the forecast state and the observed data are Gaussian and the observation operator is linear, then $\mathbf{x}_{t_i}^a$ is not only the estimate of minimum variance but as well the most likely state: in this case, the analysis will be optimal.

The evolution of the estimated state of the system from $t_i$ to $t_{i+1}$, the next timestep observations are available, is computed using the direct numerical model $\mathcal{M}$:

$$\mathbf{x}_{t_{i+1}}^f = \mathcal{M}\left(\mathbf{x}_{t_i}^a\right) + \boldsymbol{\eta}_{t_i}. \tag{3.10}$$

Its associated covariance matrix is

$$\mathbf{P}_{t_{i+1}}^f = \mathbf{M}\mathbf{P}_{t_i}^a\mathbf{M}^T + \mathbf{Q}_{t_i}, \tag{3.11}$$

where $\mathbf{M}$ is the tangent linear model, i.e. the linearised version of the operator $\mathcal{M}$ and $\mathbf{Q}_{t_i}$ is the model error covariance matrix. Equations 3.10 and 3.11 describe the forecast step. The estimated state of the system at a timestep $t_{i+1}$ $\mathbf{x}_{t_{i+1}}^f$ is called the forecast state and $\mathbf{P}_{t_{i+1}}^f$ is the forecast error covariance matrix at timestep $t_{i+1}$. $\mathbf{P}_{t_{i+1}}^f$ takes into account both the propagation of errors of the analyzed state at timestep $t_i$, $\boldsymbol{\epsilon}_{t_i}^a$ and the model errors $\boldsymbol{\eta}_{t_i}$. We assume a perfect model in this study, so that

$$\forall t_i, \quad \boldsymbol{\eta}_{t_i} = \mathbf{0}. \tag{3.12}$$

## 3.3 Convection model, Mantle State Vector and Tectonic Data

In this section, we describe in detail what $\mathcal{M}$, $\mathbf{H}$, $\mathbf{x}$ and $\mathbf{y}$ refer to in our mantle convection context.

### 3.3.1 Convection Model with Plate-Like Behaviour

The forward model $\mathcal{M}$ is our source of prior information. It solves the equations of conservation of mass, momentum and energy with classical simplifications for mantle convection: infinite Prandtl number and Boussinesq approximation. We further assume an isochemical mantle, and non-dimensionalize the equations to thermal diffusion scales (for a full development of the equations, see Ricard (2015) for example). We obtain

$$\nabla \cdot \mathbf{u} = 0, \tag{3.13}$$

$$\nabla \cdot \left[ \mu \left( \nabla \mathbf{u} + (\nabla \mathbf{u})^T \right) \right] - \nabla p + \mathrm{Ra} T \mathbf{e}_r = 0, \tag{3.14}$$

$$\frac{DT}{Dt} = \nabla^2 T + H, \tag{3.15}$$

where $\mathbf{u}$, $p$, $T$ and $t$ are the non-dimensional velocity, dynamic pressure, temperature and time, respectively. We work in spherical coordinates $(r, \theta, \phi)$ of unit vectors $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi)$. $\mathrm{Ra}$ is the Rayleigh number and $H$ is the non-dimensional internal heating rate. The models presented here have $10\%$ basal heating and $90\%$ internal heating.

The temperatures at the top and bottom boundaries are set to $T_a$ and $T_b$. The surface and the base of the model are shear-stress free.

The dynamic viscosity $\mu$ varies with temperature and stress following the equation

$$\mu = \left( \mu_T^{-1} + \mu_y^{-1} \right)^{-1}, \tag{3.16}$$

$\mu_T$ decreasing exponentially with temperature (according to Arrhenius law), and divided by $\beta$ when reaching the solidus:

$$\mu_T = \exp\left(\frac{E_A}{T+T_1} - \frac{E_A}{2T_1}\right) \quad if \quad T < T_s, \tag{3.17}$$

$$\mu_T = \beta^{-1}\exp\left(\frac{E_A}{T+T_1} - \frac{E_A}{2T_1}\right) \quad if \quad T > T_s. \tag{3.18}$$

$T_1$ is the temperature for which $\mu_T = 1$, $E_A$ is the activation energy and $T_s = T_{s_0} + \nabla_r T_s(r_a - r)$ with $r_a$ the surface value of $r$. $T_s$ models the variation of solidus with depth and is tuned so that the viscosity drop is located at the base of the top boundary layer. This results in a weaker asthenosphere and favours plate-like behaviour (Richards et al., 2001; Tackley, 2000b).

$\mu_y$ is defined by

$$\sigma_{yield} = \sigma_Y + (r_a - r).\nabla_r\sigma_Y, \quad \mu_y = \frac{\sigma_{yield}}{2\dot{\epsilon}}, \tag{3.19}$$

with $\sigma_Y$, $\nabla_r\sigma_Y$ and $\dot{\epsilon}$ being the yield stress at the surface, the depth-dependence of the yield stress and the second invariant of the strain rate tensor respectively. The strain rate tensor is linked to the velocity by

$$\dot{\boldsymbol{\epsilon}} = \frac{1}{2}\left[\boldsymbol{\nabla}\mathbf{u} + (\boldsymbol{\nabla}\mathbf{u})^T\right]. \tag{3.20}$$

Solutions are computed using StagYY (Tackley et al., 1993), a finite-volume, multigrid convection code. We use a spherical annulus grid which provides results closer to the spherical grid than a cylindrical geometry (Hernlund & Tackley, 2008). The grid is refined in the radial direction near the upper boundary of the model. In the following, the longitudinal coordinate of a point is written $\phi_m$, with $m \in \{1, 2, ..., M\}$ and its radial coordinate is written $r_n$ with $n \in \{1, 2, ..., N\}$, $r$ varying from $r_b$ to $r_a$. The value of the parameters used for this work are given in Table 3.1.

Since this work is a proof of concept, we chose a fairly simple model, with equation

TABLE 3.1: Values of the parameters of the forward models

| Symbol | Meaning | model 1 | model 2 | model 3 |
|---|---|---|---|---|
| $Ra$ | Rayleigh number | $10^6$ | $10^6$ | $10^7$ |
| $H$ | Non-dimensional internal heating rate | 20.5 | 13.5 | 20.5 |
| $M$ | Resolution of the grid in the longitude direction | 384 | 384 | 512 |
| $N$ | Resolution of the grid in the radial direction | 48 | 48 | 64 |
| $r_a$ | Radius of the top of the domain | 2.2 | 2.2 | 2.2 |
| $r_b$ | Radius of the bottom of the domain | 1.2 | 1.2 | 1.2 |
| $T_a$ | Temperature at the top of the domain | 0 | 0 | 0 |
| $T_b$ | Temperature at the bottom of the domain | 0.9 | 1.3 | 1 |
| $E_A$ | Activation Energy | 23.03 | 23.03 | 23.03 |
| $T_1$ | Temperature at which $\mu_T = 1$ | 1 | 1 | 1 |
| $\beta$ | Factor of viscosity reduction for partial melting | 10 | 10 | 10 |
| $T_{s_0}$ | Solidus Temperature at $r = r_a$ | 0.6 | 0.6 | 0.6 |
| $\nabla_r T_s$ | Radial gradient of the solidus temperature | 2 | 2 | 2 |
| $\sigma_Y$ | Yield Stress | $1.10^4$ | $1.10^4$ | $4.10^4$ |
| $\nabla_r \sigma_Y$ | Radial gradient of the yield stress | $2.10^5$ | $2.10^5$ | $2.10^5$ |

governing the flow relying on strong assumptions (incompressible, isochemical mantle). The Rayleigh number is $10^6$ which is one or two orders of magnitude lower than that of the Earth, but high enough to ensure chaotic convection. The viscosity law self-consistently generates plates at the surface, as it can be seen from the surface velocity in Fig.3.10. Ignoring many more complexities does not mean they are not fundamental for reconstructing Earth mantle evolution, but we focus in this manuscript on the data assimilation methodology.

### 3.3.2 The State of the mantle

Given Equations 3.13 to 3.15 above, the field variables describing the state of the Earth's mantle $\mathbf{x}$ are velocity, temperature and pressure. Inertial forces are negligible for mantle convection, which means that the velocity and pressure fields can be deduced from the temperature field at any time given the viscosity law and other parameters, as shown in the diagnostic Equations 3.13 and 3.14. Since it is possible to compute $\mathbf{u}$ and $p$ from the sole knowledge of $T$, $\mathbf{x}$ should be restricted to $T$ alone.

### 3.3.3 The Data: Surface Heat Flux and Surface velocities

As a first approach, the data $\mathbf{y}^o$ we use are not direct measurements per se, but plate reconstruction models. For instance, Seton et al. (2012) or Shephard et al. (2013) proposed plate tectonics reconstructions for the last $200$ My, using the continuously closing plates methodology (Gurnis et al., 2012) so that a reconstruction can be numerically computed for any time between a given $t_i$ and the subsequent $t_{i+1}$. These reconstructions integrate paleomagnetic, paleobiological and geological data to provide continuous maps of surface velocity and seafloor age as well as the position and geometry of continents. It is this type of data that is used today in convection reconstructions with imposed boundary conditions. One fundamental difference between these methods and our sequential assimilation method is that the latter naturally takes into account uncertainties in the reconstructions. A second difference is that we do not need the surface data to be known at all times.

Plate reconstructions provide estimates of the velocity at any location on the surface of the Earth in the approximation of the plate tectonics theory, as well as the age of the seafloor. In the model we use in this manuscript, surface heat flux is an excellent proxy for the age of the seafloor (Coltice et al., 2012). Consequently, we propose to consider surface heat flux and surface velocity as the data to assimilate. However, with more sophisticated models, small scale convection would require an explicit computation of the age of the seafloor.

At each time $t_i$, the data vector for the present study will be

$$\mathbf{y} = \left[q_s(\phi_1), \ldots, q_s(\phi_N), V_s(\phi_1), \ldots, V_s(\phi_N)\right]^T, \tag{3.21}$$

where $q_s(\phi_m)$ and $V_s(\phi_m)$ are the surface heat flux and tangential velocity values at longitude $\phi_m$, respectively. Hence, for the model parameters set of Table 3.1, the data vector contains $2N = 768$ values.

### 3.3.4 The Observation Operator and the Augmented State

As described by Equation 3.3, the link between the state of the system $\mathbf{x}$ and the data $\mathbf{y}$ has to be expressed in the form of a linear equation. This is straightforward for surface heat flux. At a given longitude $\phi_m$, the surface heat flux is approximated by a first order discretization of Fourier's law

$$\forall \phi_m \quad q(\phi_m) = -k \frac{T(\phi_m, r_N) - T(\phi_m, r_{N-1})}{r_N - r_{N-1}}, \tag{3.22}$$

where $k$ is the thermal conductivity, which is $1$ for our non-dimensional framework.

However, the link between the velocity field and the temperature field is highly non-linear, because of our choice of rheology. This is why we consider an augmented state of the mantle, containing both temperature and surface velocity fields:

$$\mathbf{x} = [T(\phi_1, r_1), \ldots, T(\phi_M, r_1), T(\phi_1, r_2), \ldots, \tag{3.23}$$

$$T(\phi_M, r_N), V(\phi_1), \ldots, V(\phi_M)], \tag{3.24}$$

where $T(\phi_m, r_n)$ is the temperature value at longitude $\phi_m$ and radius $r_n$ and $V(\phi_m)$ is the surface tangential velocity value at longitude $\phi_m$. For the model parameters set of Table 3.1, the state vector contains $NM + N = 18816$ values.

## 3.4 Sequential Data Assimilation Algorithm for Mantle Convection

We describe here the steps of our sequential algorithm, which are

- an initialization, which evaluates the background state $\mathbf{x}^b$ and its error covariance matrix $\mathbf{P}^b$,

- an analysis which estimates

$$\left\{\mathbf{x}_{t_i}^f, \mathbf{P}_{t_i}^f, \mathbf{y}_{t_i}^o, \mathbf{R}_{t_i}^o\right\} \rightarrow \left\{\mathbf{x}_{t_i}^a\right\},$$

- a forecast, which computes

$$\left\{\mathbf{x}_{t_i}^a\right\} \rightarrow \left\{\mathbf{x}_{t_{i+1}}^f, \mathbf{P}_{t_{i+1}}^f\right\}.$$

### 3.4.1 Initialization

The background state $\mathbf{x}^b$ and its associated covariance matrix $\mathbf{P}^b$ describe an estimate of the state of the system when no observation is yet available. $\mathbf{x}^b$ components correspond to the average temperature field and surface velocities of the mantle. $\mathbf{P}^b$ contains the two-point spatial correlation of the temperature field and surface velocity field, as described by Balachandar (1998).

To estimate $\mathbf{x}^b$ and $\mathbf{P}^b$, we use a set of $K = 200$ states of the system $\mathbf{x}_k$ with $k \in \{1, ..., K\}$, computed with the mantle convection code. These states are extracted from a free run of the convection model (i.e. unconstrained by data). To ensure decorrelation between the states, the time between two snapshots is equivalent to $5$ Lyapunov times. The Lyapunov time is the e-folding time for the growth of an initial error in a dynamic system. In the calculations we describe below, we obtain a typical Lyapunov time of $140$ My, similar to the Lyapunov times computed in 3D spherical convection models (Bello et al., 2014).

The background state is

$$\mathbf{x}^b = \langle \mathbf{x} \rangle = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_k \tag{3.25}$$

and the error covariance matrix is

$$\mathbf{P}^b = \mathrm{Covar}(\mathbf{x}) = \frac{1}{K-1}\sum_{k=1}^{K}\left(\mathbf{x}_k - \mathbf{x}^b\right)\left(\mathbf{x}_k - \mathbf{x}^b\right)^T. \tag{3.26}$$

The probability density functions of all the variables are also estimated, to compare them to Gaussian probability density functions. Indeed, the best linear unbiased estimate used during analysis is optimal only in the linear Gaussian case, as described in section 2.

As noted before, the initial setting of the model is spherically symmetric. As a consequence

$$\forall(\phi, r),\ \langle T(\phi, r)\rangle = \langle T(0, r)\rangle, \tag{3.27}$$

$$\forall(\phi_1, \phi_2, r_1, r_2),$$

$$\mathrm{Cov}(T(\phi_1, r_1), T(\phi_2, r_2)) = \mathrm{Cov}(T(0, r_1), T(\phi_1 - \phi_2, r_2)), \tag{3.28}$$

$$\forall(\phi_1, \phi_2, r_1, r_2),$$

$$\mathrm{Cov}(T(\phi_1, r_1), T(\phi_2, r_2)) = \mathrm{Cov}(T(0, r_1), T(\phi_2 - \phi_1, r_2)). \tag{3.29}$$

Likewise

$$\forall\phi,\ \langle V_s(\phi)\rangle = \langle V_s(0)\rangle, \tag{3.30}$$

$$\forall(\phi_1, \phi_2),$$

$$\mathrm{Cov}(V_s(\phi_1), V_s(\phi_2)) = \mathrm{Cov}(V_s(0), V_s(\phi_1 - \phi_2)), \tag{3.31}$$

$$\forall(\phi_1, \phi_2, r_1),$$

$$\mathrm{Cov}(T(\phi_1, r_1), V_s(\phi_2)) = \mathrm{Cov}(T(0, r_1), V_s(\phi_2 - \phi_1)), \tag{3.32}$$

$$\mathrm{Cov}(T(\phi_1, r_1), V_s(\phi_2)) = -\mathrm{Cov}(T(0, r_1), V_s(\phi_1 - \phi_2)). \tag{3.33}$$

These symmetries are enforced during the computation of $\mathbf{x}^b$ and $\mathbf{P}^b$, which decreases the number $K$ of states of the system needed to obtain converged statistics. For our

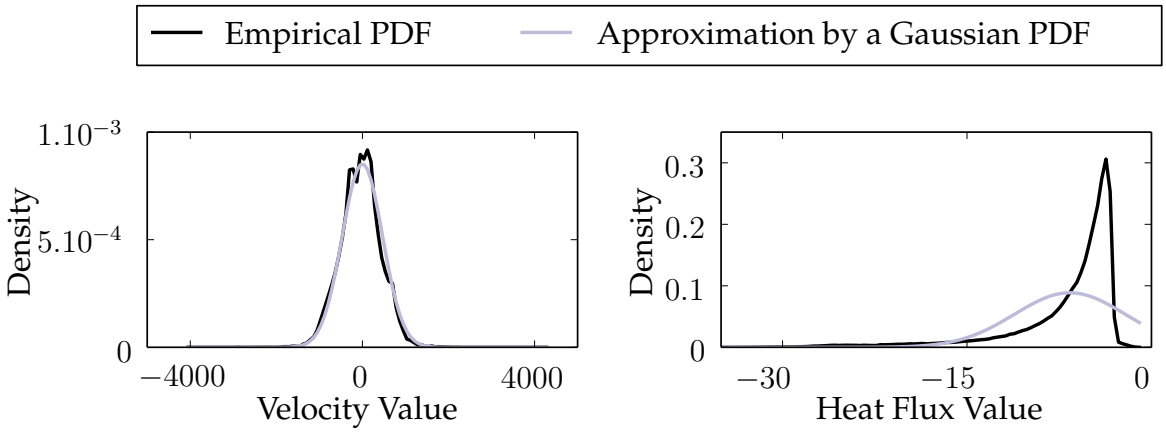


FIGURE 3.2: Probability density functions of background surface velocity (left) and background surface heat flux (right) compared to their Gaussian approximations (in light gray)

test case, the covariance matrix $\mathbf{P}^b$ contains $(NM+N)^2 = 18816^2 = 354,041,856$ components. The symmetries in the covariances dramatically reduce the number of independant components to $N/2(M+1)^2 = 3,557,400$.

From $\mathbf{x}^b$ and $\mathbf{P}^b$, a background data vector $\mathbf{y}^b$ and its associated covariance matrix $\mathbf{R}^b$ are defined as

$$\mathbf{y}^b = \mathbf{H}\mathbf{x}^b, \tag{3.34}$$

$$\mathbf{R}^b = \mathbf{H}\mathbf{P}^b\mathbf{H}^T. \tag{3.35}$$

Figs 3.2 and 3.3 represent the PDF of surface velocity, surface heat flux and temperature for different depths for our test model described in Section 3.1. Surface velocities (Fig. 3.2 left) have a Gaussian PDF which is consistent with the approximation we made. On the contrary, the PDFs of temperature (Fig. 3.3) are more complex. For temperature values close to the surface, the PDFs are highly skewed and have a strong kurtosis. This behaviour is due to the strongly nonlinear rheology at the surface of the model (kurtosis), and the isothermal boundary conditions (skewness). As a consequence, the estimation performed during analysis will not be optimal.

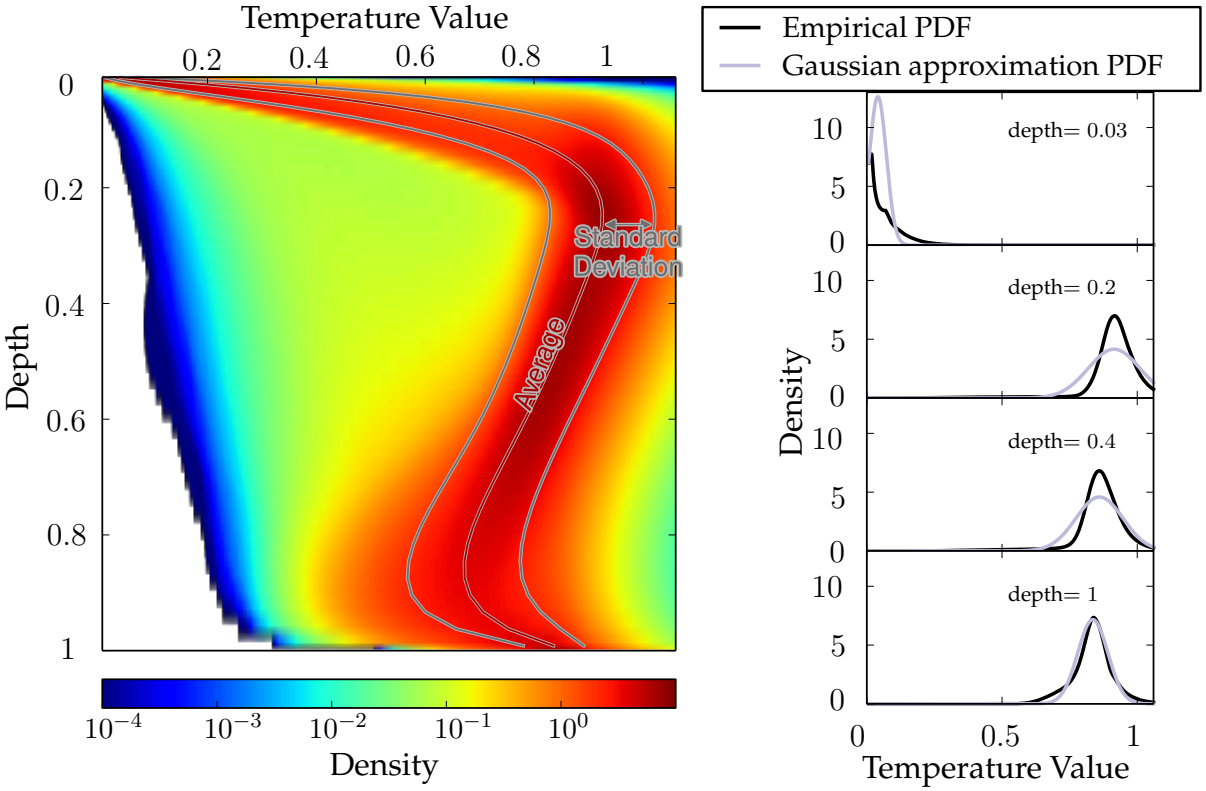


FIGURE 3.3: Map of the probability density functions of background temperature calculated at each depth (left). On the right, probability density functions for different depth are shown, so as to compare them with their Gaussian approximations used for data assimilation.

$\mathbf{x}^b$, $\mathbf{P}^b$, $\mathbf{y}^b$ and $\mathbf{R}^b$ are used to standardize the state $\mathbf{x}$ and data $\mathbf{y}$ for the analysis step. We define their corresponding standard score state $\widetilde{\boldsymbol{x}}$ and standard score data $\widetilde{\boldsymbol{y}}$ by

$$\widetilde{\boldsymbol{x}} = \left(\boldsymbol{S}^b\right)^{-1}\left(\mathbf{x} - \mathbf{x}^b\right), \tag{3.36}$$

$$\widetilde{\boldsymbol{y}} = \left(\boldsymbol{S}_y^b\right)^{-1}\left(\mathbf{y} - \mathbf{y}^b\right), \tag{3.37}$$

where $\boldsymbol{S}^b$ is the diagonal matrix containing the standard deviations of each component of the background state and $\boldsymbol{S}_y^b$ the diagonal matrix containing the standard deviations of each component of the background data,

$$\boldsymbol{S}^b = \mathrm{Diag}(\mathbf{P}^b)^{1/2}, \tag{3.38}$$

$$\boldsymbol{S}_y^b = \mathrm{Diag}(\mathbf{R}^b)^{1/2}. \tag{3.39}$$

The standard score observation operator $\widetilde{\boldsymbol{H}}$ is then

$$\widetilde{\boldsymbol{H}} = (\boldsymbol{S}_y^b)^{-1}\boldsymbol{H}\boldsymbol{S}^b. \tag{3.40}$$

The standardized background state $\widetilde{\boldsymbol{x}}^b$ and its associated covariance matrix $\widetilde{\boldsymbol{P}}^b$ are then

$$\widetilde{\boldsymbol{x}}^b = \mathbf{0}, \tag{3.41}$$

$$\widetilde{\boldsymbol{P}}^b = (\boldsymbol{S}^b)^{-1}\mathbf{P}^b(\boldsymbol{S}^b)^{-1}. \tag{3.42}$$

The covariance matrix $\widetilde{\boldsymbol{P}}^b$, associated with the standardized background state corresponds as well to the correlation matrix of the background state. Equations 3.28 to 3.33 are also valid for correlations values, i.e. for the components of $\widetilde{\boldsymbol{P}}^b$.

Equations 3.28 and 3.29 show that the correlation between any couple of temperature variables can be summed up by plotting only the correlations

$$\forall r_1, r_2 \in [r_b, r_a]^2, \forall \Delta\phi \in [0, \pi],\ \mathrm{Cor}(T(0, r_1), T(\Delta\phi, r_2)) \tag{3.43}$$

Hence, for a given $r_1$, the two points spatial temperature-temperature correlations can

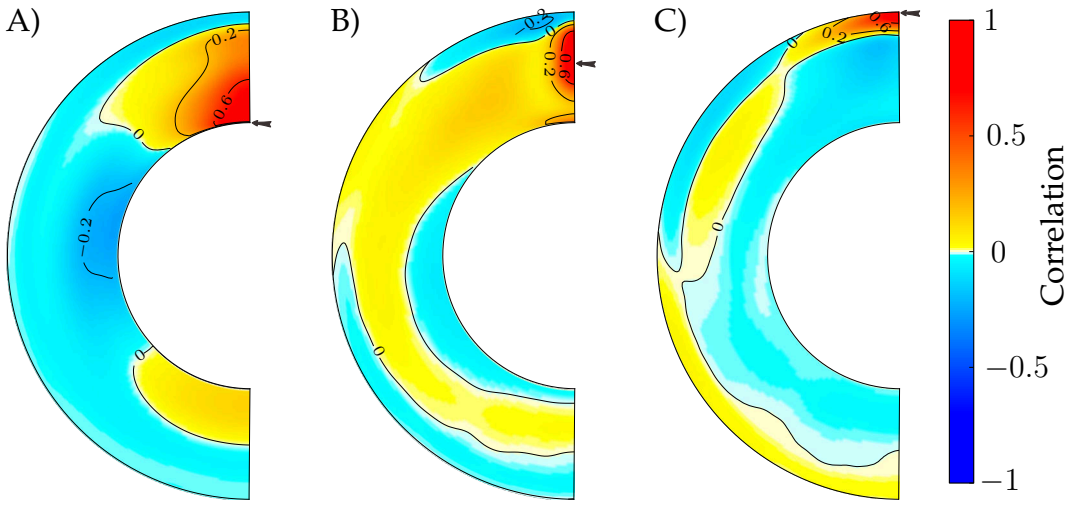


FIGURE 3.4: Maps of the two-point spatial correlation of temperature. The arrow indicates the location of the anchor point. The maps represents the correlation between the temperature at anchor points and the temperature points on the rest of the domain. Considering the symmetries of correlations described on equations 3.28 and 3.29, only the correlations $\forall \phi' \in [0, \pi], r' \in [0, 1], \mathrm{Cov}(T(0, r), T(\phi', r'))$ for $r = r_b$ (A), $r = (r_a + r_b)/2$ (B) and $r = r_a$ (C) are shown.

be mapped onto a half-spherical annulus. Fig. 3.4 represents these correlation maps for different values of $r_1$: $r_1 = r_b$ (Fig. 3.4(a)), $r_1 = (r_a + r_b)/2$ (Fig. 3.4(b)) and $r_1 = r_a$ (Fig. 3.4(c)). Fig. 3.4(a) shows that bottom temperatures have strong correlations with neighbouring temperatures, up to $0.6$ at a distance corresponding to a third of the domain depth. A fourth order convection pattern is also notable, with bottom temperature values anticorrelated with temperatures at a longitude of $+90°$and correlated with values at a longitude of $+180°$. The mid-mantle temperature (Fig. 3.4b) has an ellipsoidal zone of high correlation, which is the result of the dominant vertical motion in this part of the domain. Finally the top temperatures have a zone of correlation which extends towards the longitudinal direction, corresponding to the dominant horizontal motion at the surface of the domain. The anticorrelation between the top boundary layer and the rest of the domain is also remarkable.

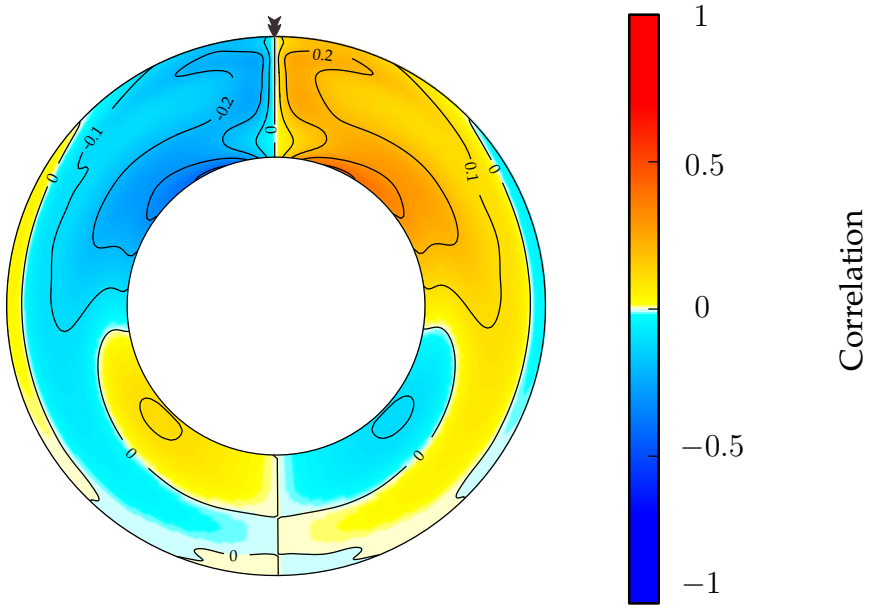


FIGURE 3.5: Map of the two-point spatial correlations of temperature and velocity. The arrow indicate the anchor point. The map represents the correlation between the surface velocity value at the anchor point and the temperature values on the rest of the domain. Considering the symmetries of correlations described on equations 3.32, only the correlations $\forall r \in [r_b, r_a], \forall \Delta\phi \in [0, 2\pi],\ \mathrm{Cor}(V_s(0), T(\Delta\phi, r))$ are shown.

Equation 3.32 shows that the correlation between any couple of temperature-surface velocity variables can be summed up by plotting only the correlations

$$\forall r \in [r_b, r_a], \forall \Delta\phi \in [0, 2\pi],\ \mathrm{Cor}(V_s(0), T(\Delta\phi, r)) \tag{3.44}$$

Hence, the two points spatial temperature-velocity correlations can be mapped onto a spherical annulus. Fig. 3.5 represents this correlation map. Fig. 3.5 shows that surface velocities are weakly correlated with temperature, with a maximum correlation value of $0.4$. Nevertheless, the correlations between surface velocity and temperature at the bottom are not negligible, and this is mainly due to the sinking of slabs to the bottom of the model.

Although we ignore the non-Gaussianity of some distributions, second order statistics succeed in summing up the main features of the convection model dynamics.

Furthermore, the matrix $\widetilde{\boldsymbol{P}}^b$ is eigendecomposed leading to

$$\widetilde{\boldsymbol{P}}^b = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T. \tag{3.45}$$

where $\boldsymbol{V}$ contains the eigenvectors of $\widetilde{\boldsymbol{P}}^b$ and $\boldsymbol{\Lambda}$ its eigenvalues. The background correlation matrix $\widetilde{\boldsymbol{P}}^b$ is reduced to

$$\widetilde{\boldsymbol{P}}^b_r = \boldsymbol{V}_r\boldsymbol{\Lambda}_r\boldsymbol{V}_r^T \tag{3.46}$$

where $\boldsymbol{\Lambda}_r$ contains only the 1928 largest eigenvalues, which account for $99.98\%$ of the cumulative variance of $\widetilde{\boldsymbol{P}}^b$. $\boldsymbol{V}_r$ is composed of the corresponding eigenvectors. The reduction of the correlation matrix is equivalent to assuming that the correlations associated with the lowest eigenvalues and corresponding eigenvectors are not meaningful. This technique, described by Cane et al. (1996), not only lightens the computational cost, but also corrects the state of the system only in those directions followed by the dynamical model.

We can now describe the first standard score forecast state as

$$\widetilde{\boldsymbol{x}}^f_{t_1} = \mathbf{0}, \tag{3.47}$$

$$\widetilde{\boldsymbol{P}}^f_{t_1} = \widetilde{\boldsymbol{P}}^b_r = \boldsymbol{V}_r\boldsymbol{\Lambda}_r\boldsymbol{V}_r^T. \tag{3.48}$$

### 3.4.2 Analysis

At every time $t_i$, $i \in \{1, ..., n\}$, new observed data $\mathbf{y}^o_{t_i}$ are available. These observations are assumed to be unbiased, which means $\langle \boldsymbol{\epsilon}^o_{t_i} \rangle = \mathbf{0}$. An error covariance matrix $\mathbf{R}_{t_i}$ is associated with these observations. The observation error covariance matrix $\mathbf{R}$ is assumed here to be the same for any time. $\mathbf{R}$ and $\mathbf{y}^o_{t_i}$ are standardized with respect to

$\mathbf{y}^b$ and $\boldsymbol{S}_y^b$

$$\widetilde{\boldsymbol{y}}_{t_i}^o = \left(\boldsymbol{S}_y^b\right)^{-1}\left(\mathbf{y}_{t_i}^o - \mathbf{y}^b\right), \tag{3.49}$$

$$\widetilde{\boldsymbol{R}} = \left(\boldsymbol{S}_y^b\right)^{-1}\mathbf{R}\left(\boldsymbol{S}_y^b\right)^{-1}. \tag{3.50}$$

Independently, a prior estimate of the state is computed. It is the forecast state $\mathbf{x}_{t_i}^f$, assumed to be unbiased ($\langle\boldsymbol{\epsilon}_{t_i}^f\rangle = \mathbf{0}$). Its associated error covariance matrix is $\mathbf{P}_{t_i}^f$. $\mathbf{P}_{t_i}^f$ and $\mathbf{x}_{t_i}^f$ are standardized with respect to $\mathbf{x}^b$ and $\boldsymbol{S}^b$

$$\widetilde{\boldsymbol{x}}_{t_i}^f = \left(\boldsymbol{S}^b\right)^{-1}\left(\mathbf{x}_{t_i}^f - \mathbf{x}^b\right), \tag{3.51}$$

$$\widetilde{\boldsymbol{P}}_{t_i}^f = \left(\boldsymbol{S}^b\right)^{-1}\mathbf{P}_{t_i}^f\left(\boldsymbol{S}^b\right)^{-1}. \tag{3.52}$$

To further simplify this problem, we assume that

$$\widetilde{\boldsymbol{P}}_{t_i}^f \approx \widetilde{\boldsymbol{P}}^b \approx \boldsymbol{V}_r\boldsymbol{\Lambda}_r\boldsymbol{V}_r^T. \tag{3.53}$$

Under this assumption, the reduced state Kalman filter is given by

$$\widetilde{\boldsymbol{K}}_r = (\widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{P}}_{t_i}^f)^T\left[\widetilde{\boldsymbol{H}}(\widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{P}}_{t_i}^f)^T + \widetilde{\boldsymbol{R}}\right]^{-1}. \tag{3.54}$$

The result of the analysis is finally rescaled using the forecast standard deviation

$$\mathbf{x}_{t_i}^a = \mathbf{x}_{t_i}^f + \boldsymbol{S}_{t_i}^f\widetilde{\boldsymbol{K}}_r\left[\widetilde{\boldsymbol{y}}_{t_i}^o - \widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{x}}_{t_i}^f\right]. \tag{3.55}$$

### 3.4.3 Forecast

The forecast state at time $t_{i+1}$, $\mathbf{x}_{t_{i+1}}^f$ is computed by the convection code STAGYY, taking $\mathbf{x}_{t_i}^a$ as starting state.

The associated covariance matrix $\mathbf{P}^f_{t_{i+1}}$ is estimated by assuming

$$\mathbf{P}^f_{t_{i+1}} = \boldsymbol{S}^f_{t_{i+1}} \boldsymbol{V}_r \boldsymbol{\Lambda}_r \boldsymbol{V}_r^T \boldsymbol{S}^f_{t_{i+1}} \tag{3.56}$$

$$\boldsymbol{S}^f_{t_{i+1}} = \alpha \boldsymbol{S}^f_{t_i}. \tag{3.57}$$

The coefficient $\alpha$ is estimated by considering errors on data and using the fact that the error on forecast data and on observed data are uncorrelated:

$$\alpha = \sqrt{\frac{\mathrm{Tr}(\mathrm{Covar}(\boldsymbol{H}\boldsymbol{x}^f_{t_{i+1}} - \mathbf{y}^o_{t_{i+1}}) - R)}{\mathrm{Tr}(\mathrm{Covar}(\boldsymbol{H}\boldsymbol{x}^f_{t_i} - \mathbf{y}^o_{t_i}) - R)}}. \tag{3.58}$$

## 3.5 Synthetic Experiments

We test here our sequential assimilation algorithm on synthetic cases. The parameters chosen for the model 1 are described in Table 3.1, along with the parameters of two additional models: model 2, which has a higher bottom heat flux to surface heat flux ratio ($37\%$), and model 3, which has a higher Rayleigh number ($10^7$). The time in the evolutions computed is rescaled as classically done, using the transit time of the mantle as the relevant scaling time (see Gurnis, 1986b, for example). The Earth's mantle transit time $t^E_t$ estimates the average time spent by a particle to move from the surface to the core-mantle boundary. For the Earth, it is evaluated by considering the thickness of the mantle $D^E$ and the root mean square of surface velocities $v^E_{rms}$

$$t^E_t = \frac{D^E}{v^E_{rms}}. \tag{3.59}$$

$v^E_{rms}$ is obtained from plate tectonics reconstructions (Shephard et al., 2013; Seton et al., 2012, for example). $t^M_t$, the transit time for the model can be computed by the same procedure. The rescaling of time for a computed evolution is

$$t = t^* \frac{t^E_t}{t^M_t}, \tag{3.60}$$

where $t^*$ is the non-dimensional time. This rescaling is designed so that the time of the evolution can compare to that of the Earth, even if the vigour of convection differs.

### 3.5.1 Setup of the Experiments

We test our data assimilation method on synthetic experiments. They consist of two phases.

First, we synthesize a true state evolution $\left\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_{n_t}^t\right\}$ and corresponding observed data sets $\left\{\mathbf{y}_{t_1}^o, \mathbf{y}_{t_2}^o, ..., \mathbf{y}_{t_n}^o\right\}$. This is done by computing a convection evolution spanning $600$ My and considering it as the true state evolution $\left\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_{n_t}^t\right\}$. From this evolution, we extract sets of surface heat flow and surface velocities at regular intervals $\Delta t$. These data are randomly noised to produce $\left\{\mathbf{y}_{t_1}^o, \mathbf{y}_{t_2}^o, ..., \mathbf{y}_{t_n}^o\right\}$. We set the average amplitude of the random noise as a fraction $\gamma$ of $q_{rms}$ and $V_{rms}$, the root mean squares of surface heat flux and velocity, respectively. The observed error covariance matrix is then

$$\mathbf{R} = \gamma^2 \begin{pmatrix} q_{rms}^2 & ... & 0 & 0 & ... & 0 \\ 0 & ... & 0 & 0 & ... & 0 \\ 0 & ... & q_{rms}^2 & 0 & ... & 0 \\ 0 & .. & 0 & V_{rms}^2 & ... & 0 \\ 0 & ... & 0 & 0 & ... & V_{rms}^2 \end{pmatrix}. \tag{3.61}$$

Second, we apply the assimilation algorithm using the synthetic data $\left\{\mathbf{y}_{t_1}^o, \mathbf{y}_{t_2}^o, ..., \mathbf{y}_{t_n}^o\right\}$. Then, we compare the result of the assimilation, $\left\{\mathbf{x}_1^f, \mathbf{x}_2^f, ..., \mathbf{x}_{n_t}^f\right\}$ with the true state $\left\{\mathbf{x}_1^t, \mathbf{x}_2^t, ..., \mathbf{x}_{n_t}^t\right\}$.

For each set of model parameters, we compute 21 mantle convection evolutions starting from different initial conditions to produce 21 possible true state evolutions. For model 1, we test the assimilation for each evolution with 10 different pairs of parameters of data assimilation $(\Delta t, \gamma)$. $\Delta t$ is varying from $1$ My to $50$ My and $\gamma$ from $5\%$ to $50\%$. For Model 2 and 3, we test the assimilation for each evolution with $\Delta t = 10$ My

TABLE 3.2: Values of the parameters of the assimilation for synthetic experiments.

| name of experiment | Rayleigh number | Heat flux ratio | $\Delta t$ (My) | $\gamma$ (%) | number of synthetic experiments |
|---|---|---|---|---|---|
| $t1\gamma10$ | $10^6$ | 10% | 1 | 10 | 21 |
| $t5\gamma10$ | $10^6$ | 10% | 5 | 10 | 21 |
| $t10\gamma5$ | $10^6$ | 10% | 10 | 5 | 21 |
| $t10\gamma10$ | $10^6$ | 10% | 10 | 10 | 21 |
| $t10\gamma30$ | $10^6$ | 10% | 10 | 30 | 21 |
| $t10\gamma50$ | $10^6$ | 10% | 10 | 50 | 21 |
| $t15\gamma10$ | $10^6$ | 10% | 15 | 10 | 21 |
| $t20\gamma10$ | $10^6$ | 10% | 20 | 10 | 21 |
| $t50\gamma10$ | $10^6$ | 10% | 50 | 10 | 21 |
| $t10\gamma10Q40$ | $10^6$ | 37% | 10 | 10 | 21 |
| $t10\gamma10Ra7$ | $10^7$ | 26% | 10 | 10 | 21 |

and $\gamma = 10\%$. The exact combinations of parameters tested are presented in Table 3.2. Table 3.2 defines as well the name of the different assimilations that we will use in the following.

### 3.5.2 Quality of the data assimilation estimate

We evaluate the quality of the data assimilation scheme on its ability to retrieve the true temperature fields and to match surface data.

Fig. 3.6 shows examples of the final forecast state (second column) for evolutions with different parameters, after 300 My of data assimilation: $t10\gamma10$ (31 observation times, model 1), $t50\gamma10$ (7 observation times, model 1), $t10\gamma10Q40$ (31 observation times, model 2) and $t10\gamma10Ra7$ (31 observation times, model 3). The two first cases are done using the same model parameters, so we display the data assimilation results of the same evolution for better comparison. The true temperature fields for each case are displayed on the first column. The local error on the third column of Fig. 3.6 is the absolute value of the error at each coordinate $(\phi_m, r_n)$

$$\epsilon_T^f(\phi_m, r_n) = |T^f(\phi_m, r_n) - T^t(\phi_m, r_n)|. \tag{3.62}$$

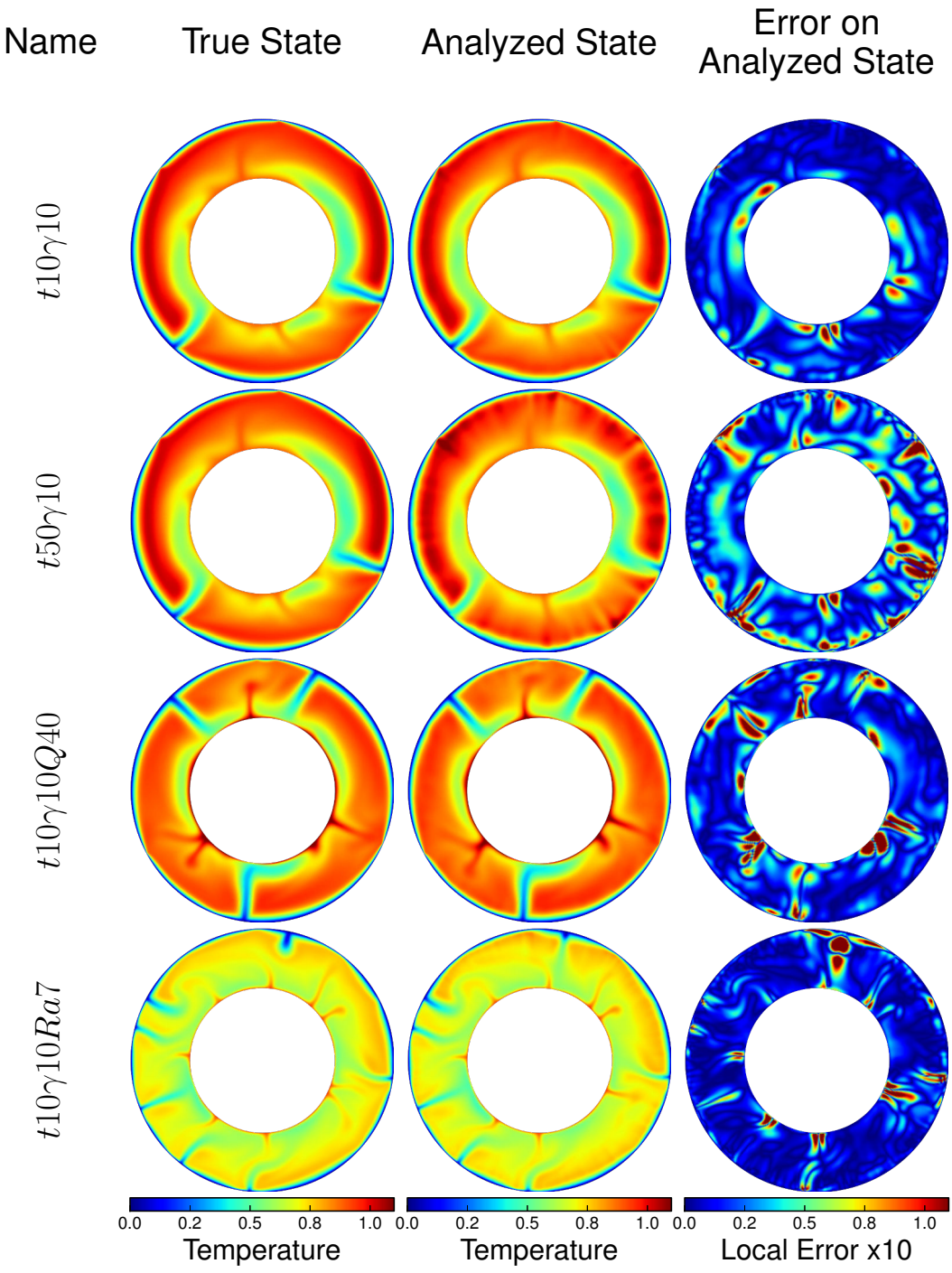


FIGURE 3.6: Comparison of the results of data assimilation after 300 My for four cases. First row: data assimilation with a $\Delta t = 10$ My (experiment $t10\gamma10$), second row: data assimilation with a $\Delta t = 50$ My (experiment $t50\gamma10$), third row: data assimilation with a $\Delta t = 10$ My (experiment $t10\gamma10Q40$) and fourth row: data assimilation with a $\Delta t = 10$ My (experiment $t10\gamma10Ra7$). The first column displays the true temperature field for each experiment, the second column shows the analyzed temperature field, and the third column is the associated error multiplied by 10 and represented using the same color scale than the field itself, in order to ease visual inspection.

Below the boundary layer, the locations of positive temperature anomalies are better retrieved for $t10\gamma10$ than for $t50\gamma10$. The general geometry of slabs at the bottom is representative of the true state, even though some details are missing. The geometry of upwellings is retrieved for $t10\gamma10$, but only their approximate location is well estimated for $t50\gamma10$. Some minor positive anomalies are missing at the bottom of the domain for both cases. The relative error shows that although the geometry of the temperature field is reasonably well recovered, the actual temperature values for down- and upwellings are not as accurately estimated. This tendency is especially noticeable for $t50\gamma10$. Overall, the convection pattern is correctly estimated, but for $t50\gamma10$ slabs are blurred. The structures are also well recovered for $t10\gamma10Q40$ and $t10\gamma10Ra7$. We can see that the general shape of the hot upwellings is well estimated for both cases, except on one upwelling, on the bottom right corner of the $t10\gamma10Ra7$ experiment, where the hot upwelling with two branches is estimated as only one bigger upwelling. The structure of slabs are also reasonnably well estimated. The value of the errors for both $t10\gamma10Q40$ and $t10\gamma10Ra7$ are higher than those of $t10\gamma10$, on some localized regions.

We compute the spatial average error of the temperature field for each timestep $i \in \{1, 2, ..., n_t\}$

$$\epsilon_T^f(i) = \sqrt{\frac{\sum_{m=1}^{M}\sum_{n=1}^{N}(T_i^f(\phi_m, r_n) - T_i^t(\phi_m, r_n))^2 \Delta V(\phi_m, r_n)}{\sum_{m=1}^{M}\sum_{n=1}^{N} T_i^t(\phi_m, r_n)^2 \Delta V(\phi_m, r_n)}}, \tag{3.63}$$

where $\Delta V(\phi_m, r_n)$ is the volume of the cell centred on $(\phi_m, r_n)$. For each combination of $\Delta t$ and $\gamma$, $\epsilon_T^f(i)$ is averaged over 21 synthetic experiments differing only by their initial conditions for the model 1. The synthetic experiments $t10\gamma10Q40$ (higher bottom to surface heat flux ratio) and $t10\gamma10Ra7$ (higher Rayleigh number) showed respectively $20$ and $19$ successes out of $21$. We identify a data assimilation experiment as failing if there is a time after which the global error $\epsilon_T^f$ is consistently increasing when analyses are performed. The three failure cases show a consistently increasing difference between observed data and forecast data, making them easy to identify even without the knowledge of the true state evolution. Picking another random noise of the

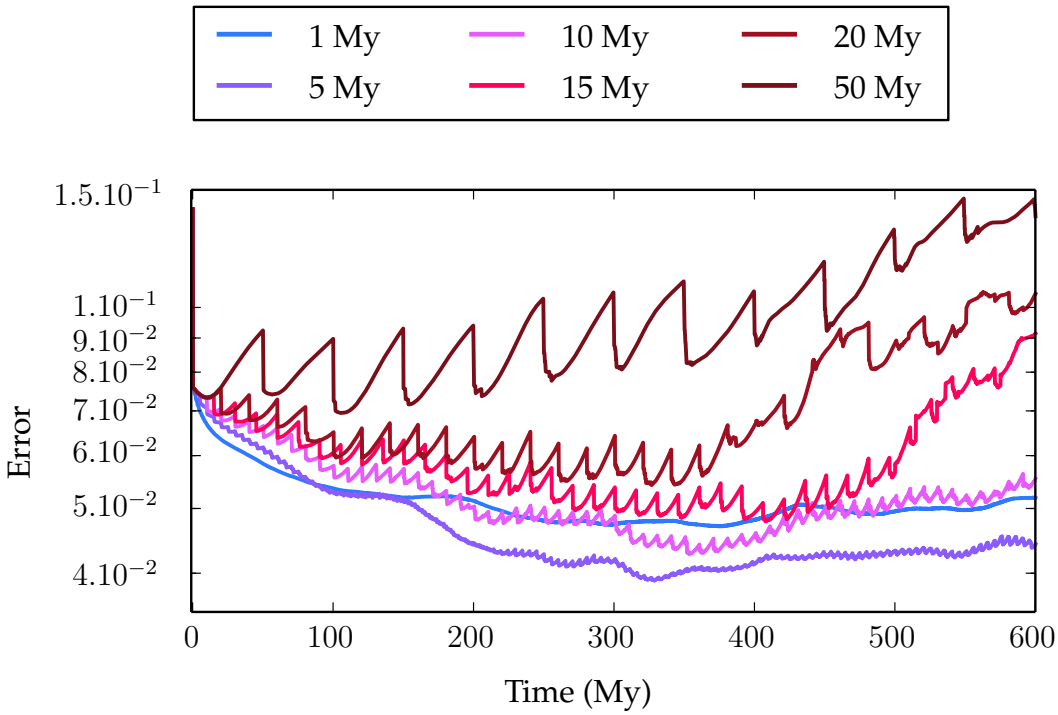


FIGURE 3.7: Evolution of the average relative error on the forecast temperature field with $\gamma = 10\%$ and $\Delta t = 1$, 5, 10, 15, 20 and 50 My. Note that the scale on the y-axis is logarithmic.

same magnitude for each of the three experiments that failed produced a successful data assimilation trial. Hence, peculiar random noise structure could prevent accurate retrieval by the methodology presented here. We compute the average evolution of errors for synthetic experiments $t10\gamma10Q40$ and $t10\gamma10Ra7$ taking into account only the successes, ie 20 experiments for $t10\gamma10Q40$ and 19 for $t10\gamma10Ra7$. Hence, the estimation of the error is not specific to a case that is easier to retrieve than others. On every curve, the analysis times are characterized by a sudden drop of the error (see Figs 3.7, 3.8 and 3.9). This is remarkable for the first analysis, at time 0, where the error is reduced by a factor of two for any set of parameters tested. These results show that taking into account only the second order statistics is efficient at correcting a forecast state and improving the estimate of the true state.

As observed in Figs 3.7, 3.8 and 3.9, the general evolution of errors with time follows three phases: an error reduction phase, which may be followed by an error stabilization phase around the lower value, and finally an error growth phase. The duration and existence of each phase depends on the chosen combination of $\Delta t$, $\gamma$ and the model

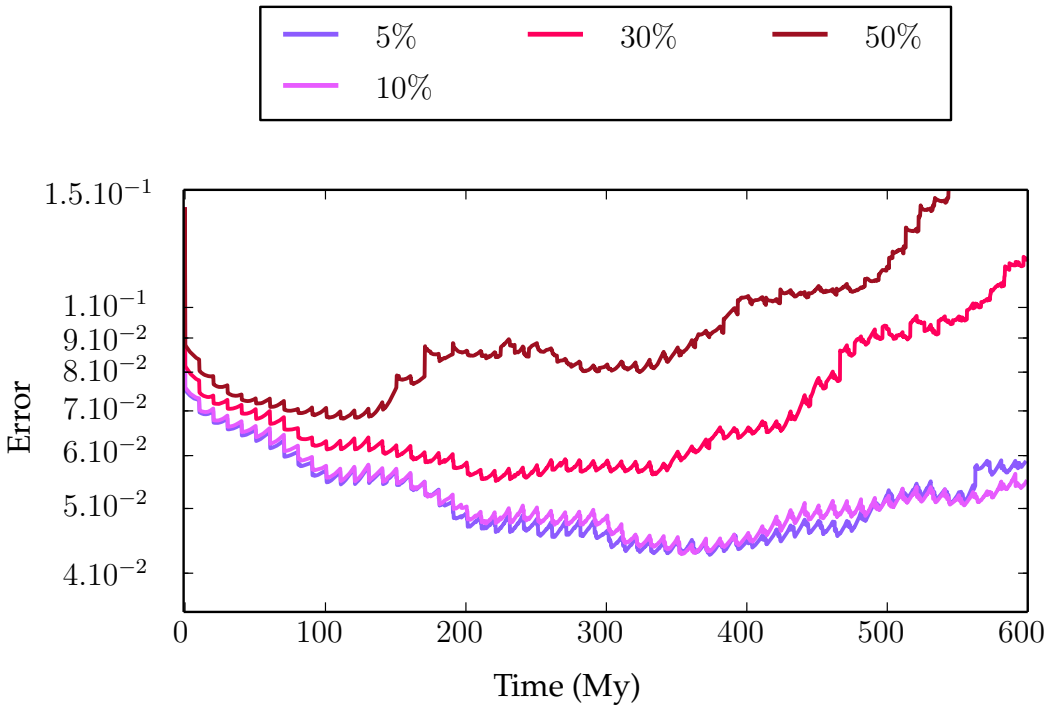


FIGURE 3.8: Evolution of the average relative error on the forecast temperature field with $\Delta t = 10$ My and $\gamma = 5$, 10, 30 and 50%. Note that the scale on the y-axis is logarithmic.

parameters.

Fig. 3.7 shows the evolutions of errors for data assimilations with noise on observations of 10% and different frequencies of analyses: $t1\gamma10$, $t5\gamma10$, $t10\gamma10$, $t15\gamma10$, $t20\gamma10$ and $t50\gamma10$. The error reduction phase lasts between 300 to 400 My (2 to 2.5 Lyapunov times) for all cases except for the extreme case of $t50\gamma10$ (50 My representing one third of the Lyapunov time). For the first 100 My of assimilation, the shorter the time between 2 analyses, the faster the errors decrease.

If analysis intervals are equal to or longer than 15 My, growth of the error happens within 600 My of assimilation. In contrast, for $t1\gamma10$, $t5\gamma10$ and $t10\gamma10$, no error growth occurs, and errors are stable at around 5% after 200 My. Throughout the assimilation, errors for $t1\gamma10$ become greater than the errors for $t5\gamma10$. This is due to the method used for updating forecast errors in our scheme. Indeed, their estimation is based on the comparison between observed and forecast surface data. A 1 My interval between analyses does not allow the possible errors made internally to be propagated onto the

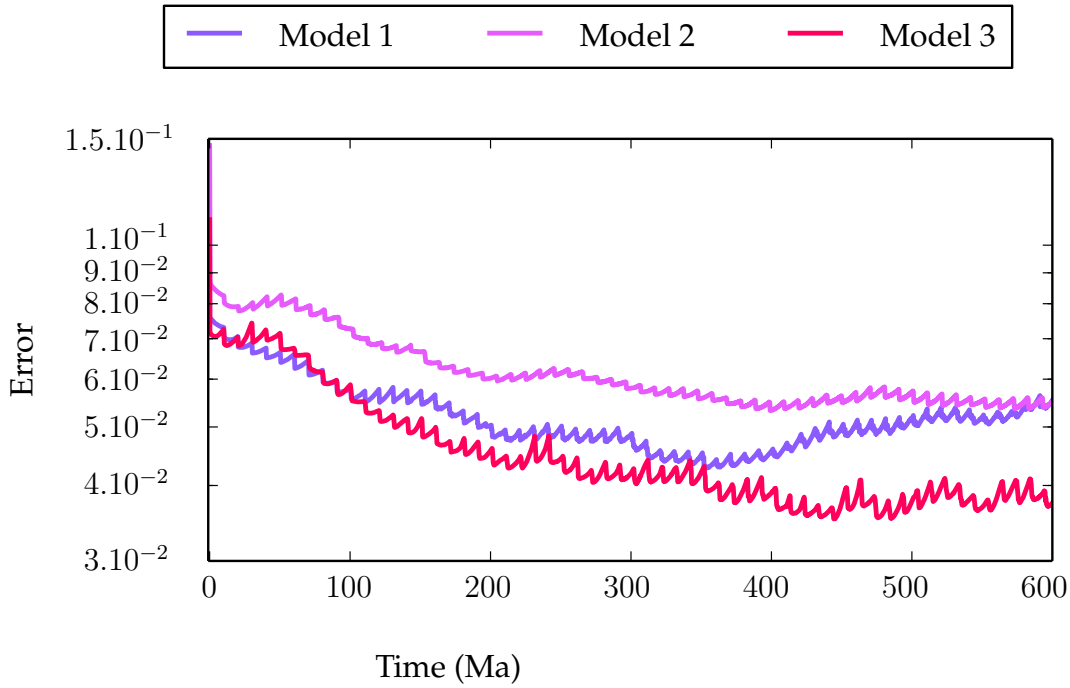


FIGURE 3.9: Evolution of the average relative error on the forecast temperature field with $\Delta t = 10$ My and $\gamma = 10\%$ for the model 1, model 2 and model 3. Note that the scale on the y-axis is logarithmic.

surface. It follows that the error on the forecast is underestimated for short intervals of analysis, leading to results that are not as good as for longer intervals of analysis.

Fig. 3.8 shows the evolutions of errors for data assimilations with analyses every $10$ My and different errors on observations: $t10\gamma5$, $t10\gamma10$, $t10\gamma30$ and $t10\gamma50$. For levels of noise in observation ranging from $5$ to $30\%$, the error reduction phase lasts more than $300$ My. Even when observations are noised as much as $50\%$, data assimilation keeps errors to values lower than $10\%$ on the estimate of the temperature field over the first $100$ My. Fig. 3.8 shows that low levels of noise in observations ($5$ to $10\%$) give results of similar quality.

The error evolutions for $t10\gamma10Ra7$ and $t10\gamma10Q40$ are shown in Fig. 3.9, along with the evolution of errors of $t10\gamma10$. The error for $t10\gamma10Ra7$ is on average smaller than the error for $t10\gamma10$. On the contrary, the error for $t10\gamma10Q40$ is on average bigger than the error of $t10\gamma10$. We note as well that, between two analyses, the errors are increasing much more for $t10\gamma10Ra7$ than for the other cases.

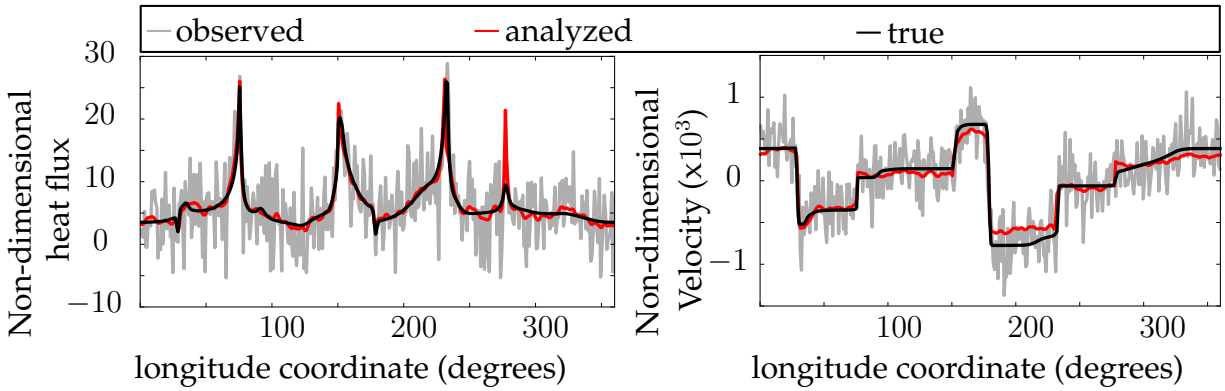


FIGURE 3.10: Comparison between true, observed and analyzed surface heat flux (left) and surface velocities (right) after 100 My of assimilation with the parameters $t10\gamma50$

Fig. 3.10 demonstrates that our data assimilation scheme provides a posteriori a better estimate of surface data than the a priori observation for extreme cases. Indeed, for $t10\gamma50$, after 100 My of data assimilation, the a posteriori retrieved surface heat flux and surface velocities are better estimates of the true values than the ones used as the observations.

We conducted here synthetic experiments, where the true state is known, and the result of data assimilation can be compared to this true state evolution. However, in a data assimilation using real observations, the true state of the system is unknown. In this case, it is necessary to use other diagnostics to evaluate the quality of the assimilation. One diagnostic is based on the study of the innovation, which is the difference between the observed data $\mathbf{y}_i^o$ and the forecast data $\boldsymbol{H}\boldsymbol{x}_i^f$,

$$\boldsymbol{d}_i = \mathbf{y}_i^o - \boldsymbol{H}\boldsymbol{x}_i^f. \tag{3.64}$$

The simplest evaluation is to consider the evolution of the cumulative mean innovation

$$d_k = \left\| \sum_{i=1}^{k} \boldsymbol{d}_i \right\| \tag{3.65}$$

which is supposed to converge to zero if the scheme is unbiased (Talagrand, 2003).

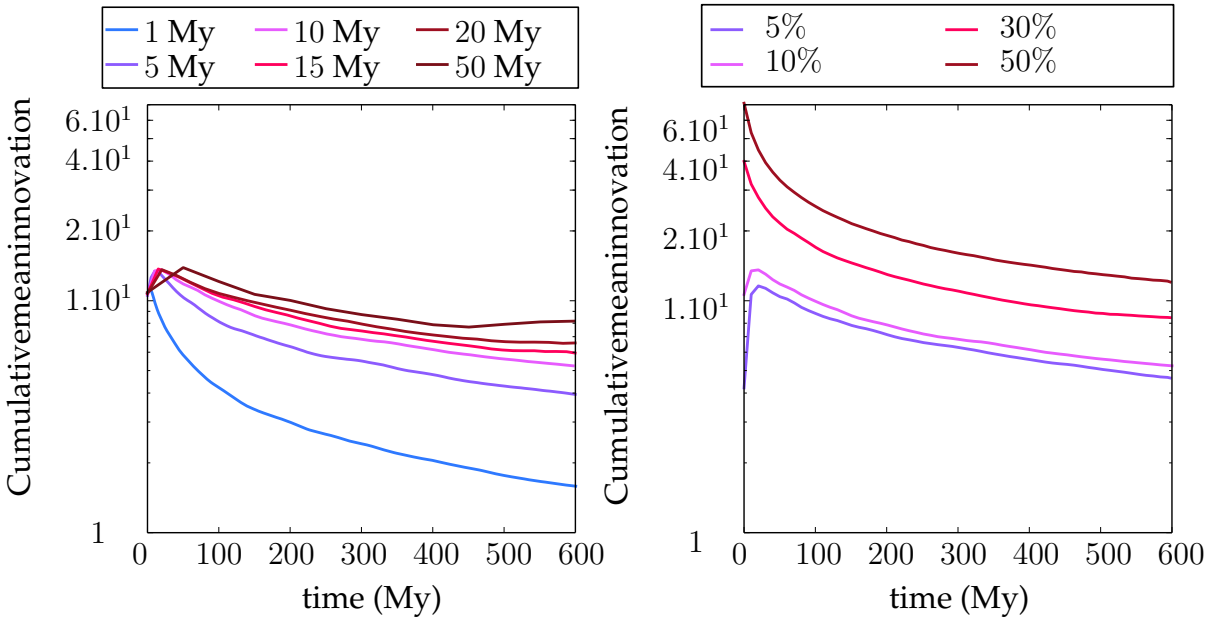


FIGURE 3.11: Evolution of the cumulative mean innovation as a function of the number of analyses. On the left: for model 1 with $\gamma = 10\%$ and $\Delta t = 1$, 5, 10, 15, 20 and 50 My; on the right: for Model 1 with $\Delta t = 10$ My and $\gamma = 5$, 10, 30 and 50%

Fig. 3.11 shows the cumulative mean innovation for the different parameters, as a function of the number of analyses. The cumulative mean innovation is decreasing through time for all cases, showing that our scheme is unbiased. This should not come as a surprise, since the observations and the model are governed exactly by the same physics.

## 3.6 Discussion

We chose a sequential method because of its relative ease of implementation compared to alternative methods such as variational data assimilation (see Talagrand, 1997; Evensen, 2007, for discussions on both methodologies). Moreover, sequential data assimilation has proved to be efficient in solving a wide range of geophysical problems (Aubert & Fournier, 2011; Hoteit & Pham, 2004, for example). We opted for a suboptimal scheme based on the Kalman filter instead of driving the flow with surface velocities so as to exploit the surface information a step further. The initial condition, or first guess, is given as a compromise between a 1D temperature profile and the inversion of the first data available. Then, the whole temperature field is updated

whenever observations are available. This means our models are less sensitive to initial conditions than mantle circulation models in which plate velocities are imposed as boundary conditions.

The specificity of sequential data assimilation is that the estimated state of the mantle improves through time, as more data are assimilated into the model (Fig. 3.1). This raises an issue since the information at hand for the Earth tends to improve in accuracy as it gets closer to present-day: classical sequential data assimilation schemes are not able to extend the recent information back in time. This difficulty can be overcome by implementing Kalman smoothers (Cohn et al., 1994; Cosme et al., 2010; Nerger et al., 2014) which use time correlations to update the previous states with new observations. Applying these techniques would also enable us to add tomographic models as observations in mantle circulation models, and to propagate this information back in time. However, one of the benefits of the method we developped would be to provide an image of the present state of the Earth mantle which is independant from seismic data, and would give the opportunity to work on the interpretation of tomographic models.

Our sequential assimilation technique gives encouraging results for our test cases: data assimilation gradually reduces the distance between the true models and the estimated ones for the first $300$ My, even for highly noised data (up to $30\%$ of noise). The analysis is based on a linear correction, the amplitude of which depends on the distance between the predicted data and the observed data. The linearity of the correction is an approximation that is efficient as soon as the prediction is close enough to the observation. When these two are too different, the analysis tends to over-correct the state, which leads to an unrealistic analyzed state. The algorithm is stable as long as the time between two observations is short enough: if it is $15$ My and greater the estimated temperature field will eventually diverge from the true temperature field. If the time between two analyses is too long, the linear approximation for correcting the state field is not valid any more. The temperature field is over-corrected and this leads to analyzed temperatures inconsistent with the physical model.

We evaluate hereafter how the required $\Delta t$ and $\gamma$ in our test cases compare to the

corresponding values available for Earth's data: time between two plate reconstruction stages and uncertainty on plate kinematics and seafloor age maps. The seafloor spreading isochrons used for tectonic reconstructions are based on the identification of magnetic anomalies (Müller et al., 2008a). This gives a strong constraint on the motion of plates at least every $10$ My, except for the Cretaceous superchron spanning between $83.5$ to $125$ Ma. Other geological arguments have been provided to infer plate evolutions for this period (Torsvik et al., 2009). For more recent times (the last $20$ My), more precise surveys of magnetic anomalies have led to regional reconstructions with a temporal resolution of around $1$ My (Merkouriev & DeMets, 2006). Moreover, Gurnis et al. (2012) developed a way to estimate kinematic states between two stages of reconstructions. With this method, having data to assimilate at least every $10$ My is ensured, although interpolated solutions could miss peculiar changes in plate motions, and errors on the reconstructed states propagate into the interpolated states.

Regarding these errors, noise in plate tectonic reconstructions is difficult to estimate since the process involves taking into account various types of data and human syntheses. Hence, uncertainties on plate reconstructions increase as we go back in time, since less data are available and more interpretation is required. The accuracy of reconstructions also varies in space. For instance the evolution of the Atlantic for more than $100$ My is well known, however, the eastern Pacific before $60$ My is mostly unknown because the oceanic seafloor has subducted. Efforts to estimate errors on plate tectonic reconstructions have been directed to the estimation of errors on present maps of oceanic sea floor (Müller et al., 2008a), and on the noise reduction in finite rotations deduced from them (Iaffaldano et al., 2014). Both of these works show that the error of age maps is smaller than $10\%$ for recent times ($< 50$ My). However, errors back in time are more difficult to estimate since accuracy for certain plates decreases (like those in the eastern Pacific as discussed before for instance), and an error on the angular velocity or the position of the Euler pole would constrain the errors on velocities over the whole plate domain (Molnar & Stock, 1985). These errors would have to be propagated to seafloor age maps.

As successful as it may appear, the method presented here has several limitations. First, our technique does not take into account non-linearities of the model physics, and the non-Gaussianity of the distribution of temperature, heat flux and velocity values. Indeed, we computed only second order statistics to estimate the link between temperatures and velocities. However, the method we present here already captures the essential features of mantle convection, and provides accurate convection reconstructions with the approximations made.

In this study, we assumed a perfect model (with $\eta = 0$ as stated in Equation 3.12). However, this would not be the case for applying the technique with Earth data. Model equations assume simplifications of the physics, detailed in Ricard (2007). Any parametrization we may use would have shortcomings. The most important, least known, ingredient is thought to be rheology, which is a strong limitation according to Worthen et al. (2014). Applying our method to Earth data would help evaluate the forecasting power of the state-of-the-art convection models. Explicitly introducing a model error term in the data assimilation scheme could be a way to improve the reconstructions (Evensen, 2007). Also, we used heat flux as a proxy for seafloor age because there is no small scale convection in our test cases. For Earth data, monitoring age with tracers, giving the exact estimate of seafloor age would be required. The state vector and observation matrix would differ slightly, but the technique would still apply. Complexifying the direct model used in data assimilation (ie having a 3D model at higher Rayleigh number, with a more complex rheology and a multi component system) will as well increase dramatically the size of the data assimilation problem. In particular, the state covariance matrix will reach a size at which the direct computations carried out in this paper will be extremely heavy. One solution would be to take advantage of the structure and properties of the covariance matrix (symmetries, periodicity) to design faster algorithms. The method could also be modified to avoid the computation of the covariance matrix and calculate directly the smaller matrix $\widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{P}}^f_{t_i}$.

## 3.7 Conclusion

We have applied sequential data assimilation methodology to reconstruct mantle convection using surface data (velocity and heat flux as a proxy for seafloor age). This work is a first approach to test the efficiency of assimilating surface data to recover the evolution of convecting structures. Our technique belongs to the suboptimal schemes of sequential data assimilation based on the Kalman filter. We modified the optimal interpolation method to update the forecast error covariance matrix.

Our scheme of sequential data assimilation proves to be efficient in recovering the temperature field of a convective system with plate-like tectonics at its surface over several $100$ My. The only observations used were surface heat fluxes and surface velocities. We tested the robustness of the method by conducting synthetic experiments in 2D spherical annulus geometry. We obtained accurate results, even for the location of plumes at the base of the models, for periods of at least $300$ My provided the time between analyses is shorter than $15$ My and the noise in observation is lower than $30\%$. These requirements are satisfactory, since current plate reconstruction models already provide estimates for velocities and ages at least every $10$ My and with a fine accuracy for the past $60$ My, and possibly beyond.

Application to the Earth would first involve the use of 3D-spherical sophisticated models, with high resolution, high convective vigour, continents and a more realistic rheology. In 3D the ratio of number of data points to unknowns would be the same as in 2D spherical annulus geometry by definition. However, increasing the resolution of our model would make this ratio deteriorate, which could be compensated by increasing the correlation between closer nodes. Moreover, although convection models can account for complex parametrizations, their limitations could introduce errors in the reconstructions.

These difficulties should not stop the exploitation of the method: if successful, the convection reconstruction would provide a new image of the mantle, alternative to seismic tomography, which would appear as an independent source of information and assess

the quality of the estimate of the state of the mantle at present day; if unsuccessful, it would provide a quantitative evaluation of forward models and help us decide how to improve them.

## acknowledgments

We like to thank the editor as well as the two anonymous reviewers for their comments that helped to improve the initial manuscript. The research leading to these results has received funding from the European Research Council within the framework of the SP2-Ideas Program ERC-2013-CoG, under ERC grant agreement 617588. Calculations were performed on LGLTPE Seisglob high-performance computing cluster and using HPC resources from GENCI-IDRIS (grant 2014-047243). Figures were generated using matplotlib library Hunter (2007). The contribution of Alexandre Fournier is IPGP contribution number 3675.

## Résumé du chapitre

Grâce aux avancées réalisées au cours de la dernière décennie dans le domaine de la modélisation de convection mantellique, il est maintenant possible de modéliser au premier ordre la dynamique interne et la tectonique de surface du système croûte-manteau. Nous montrons ici que les données tectoniques (comme la cinématique de surface et la répartition de l'âge du fond marin) et les modèles de convection du manteau avec comportement en plaques peuvent en principe être combinés pour reconstruire la circulation mantellique. Nous présentons une méthode d 'assimilation de données séquentielle, basée sur un schéma sous-optimal dérivé du filtre de Kalman, où les vitesses de surface et les cartes d'âges des fonds océaniques ne sont pas utilisées comme conditions limites pour l'écoulement mais comme données à assimiler. Deux étapes (une prévision suivie d'une analyse) sont répétées séquentiellement pour intégrer des données observées à différents moments. Chaque fois que des observations sont disponibles, l'étape d'analyse détermine l'état le plus probable du manteau à ce moment, en tenant compte d'une estimation préalable (fournie par la prévision) et des nouvelles observations disponibles, en utilisant la meilleure estimation linéaire non biaisée. Entre deux temps d'observation, l'évolution du manteau est régie par le modèle direct de convection mantellique. Pour évaluer son efficacité, cette méthode est appliquée à des cas synthétiques en 2-D (géométrie d'anneau sphérique). Nous comparons les évolutions de référence aux estimations obtenues par assimilation de données. Deux paramètres contrôlent le comportement du schéma: le temps entre deux analyses, et l'amplitude du bruit dans les observations synthétiques. Notre technique se révèle efficace pour reconstruire des évolutions de champ de température, à condition que le temps entre deux analyses soit de 10 Myr. Si l'amplitude de l'erreur a priori sur les observations est importante (30%), notre méthode fournit une meilleure estimation de la tectonique de surface que les observations, en tirant profit de l'information contenue dans la physique de la convection.

# CHAPTER 4

## Ensemble Data Assimilation For Mantle Circulation

To be submitted in:

Bocher, M., Fournier, A., & Coltice, N., in prep. Ensemble data assimilation for mantle circulation, *Nonlinear Processes in Geophysics*

## Abstract

Recent advances in mantle convection modelling led to the release of a new generation of convection codes, able to generate self-consistently plate-like tectonics at their surface. Those models physically link mantle dynamics to surface tectonics. Combined with plate tectonic reconstructions, they have the potential to produce a new generation of mantle circulation models that use data assimilation methods and where uncertainties on plate tectonic reconstructions are taken into account. We recently provided a proof of this concept by applying a suboptimal Kalman Filter to the reconstruction of mantle circulation (Bocher et al., 2016). Here, we propose to go one step further and apply the Ensemble Kalman Filter (EnKF) to this problem. The EnKF is a sequential Monte Carlo method particularly adapted to solve high dimensional data assimilation problems with nonlinear dynamics. We tested the EnKF using synthetic observations consisting of surface velocity and heat flow measurements, on a 2D-spherical annulus model and compared it with the method developed previously. The EnKF performs on average better and is more stable than the former method. Less than 300 ensemble members are sufficient to reconstruct an evolution. We use covariance adaptive inflation and localization to correct for sampling errors. We show that the EnKF results are robust over a wide range of covariance localization parameters. The reconstruction is associated with an estimation of the error, and provides valuable information on where the reconstruction is to be trusted or not.

## 4.1 Introduction

Mantle circulation models are estimates of mantle flow history. They combine two sources of information: observations on the dynamics or 3D structure of the Earth's mantle and a numerical model of mantle convection. In their effort to reconcile both observations and our physical understanding of mantle dynamics, they serve a wide variety of purposes and disciplines. Hager & O'Connell (1979) originally built instantaneous mantle circulation models to understand the effect of plates on large-scale mantle flow. Since then, they have been used, among other applications, to understand the dynamics and evolution of the deep earth mantle structures (Bunge et al., 1998; McNamara & Zhong, 2005; Bower et al., 2013; Davies et al., 2012), to study the evolution of mantle plumes and their relationship to hotspots (Hassan et al., 2016), to infer changes in the Earth's rotation axis (Steinberger & O'Connell, 1997), sea-level (Moucha et al., 2008) or dynamic topography (Flament et al., 2013).

The geodynamics community has developed three alternative approaches to the problem of the reconstruction of mantle circulation. The first approach, backward advection, consists in starting at present by estimating the current density field of the mantle from seismic tomography models (see Conrad & Gurnis, 2003, for a description of this method). This density field is then advected backward in time with plate tectonic reconstructions as imposed boundary condition (Steinberger & O'Connell, 1997). This method has a limited numerical cost and exploits the two most instructive constraints on mantle circulation: plate tectonic reconstructions and seismic tomography. However, this technique neglects thermal diffusion, so it is not able to reconstruct past thermal structures that have completely diffused before present and it is limited to times and regions for which the effect of diffusion is thought to be small. This limits reconstructions to the last $50$ to $75$ Myr (Conrad & Gurnis, 2003) or even to shorter periods if we consider the uncertainties on tomographic models (Bello et al., 2014). The second approach, the semi-empirical sequential method, estimates mantle circulation by integrating plate tectonics reconstructions chronologically into a mantle convection

model. Plate tectonic reconstructions are either introduced as velocity boundary conditions, as first described by Bunge et al. (1998), or with a more sophisticated method, by blending a convection solution with thermal and kinematic models of plates and slabs (Bower et al., 2015). This approach allows the use of models of convection with chemical heterogeneities (McNamara & Zhong, 2005). Also, it is not anymore the reconstruction method that limits the timespan of the reconstruction, but the availability of plate tectonic reconstructions. This led to mantle circulation models integrating up to $450$ Myr of plate reconstruction history (Zhang et al., 2010). However, this method considers plate tectonic reconstructions as perfect estimates of surface tectonics: uncertainties affecting the reconstructions are not taken into account although they are substantial, especially as reconstructions go further in the past (for example, there is almost no information on the state of the ocean floor before $140$ Myr, see e.g. Torsvik et al., 2010). This method also requires the choice of an arbitrary initial temperature field to compute the evolution. The third approach uses data assimilation methods to solve the mantle circulation problem. Data assimilation methods are inverse methods dealing with the specific problem of estimating the evolution of a dynamical system from asynchronous data and a physical model (Evensen, 2009a). The full inverse problem for mantle circulation, as stated by Bunge et al. (2003), would take into account model errors, numerical approximations, errors on plate reconstructions and on the estimation of the current tomography-derived temperature field to provide the best fit given all sources of information. However, solving the full inverse problem of mantle circulation is still a great challenge given the nonlinearities in mantle convection dynamics and the computational power required to compute a realistic forward mantle convection evolution alone (Stadler et al., 2010; Burstedde et al., 2013). So far, variational data assimilation dominates over other methods to estimate mantle circulation (Bunge et al., 2003; Horbach et al., 2014; Ghelichkhan & Bunge, 2016). To simplify the problem, they minimize the misfit between the final temperature field of the mantle circulation model and the one deduced from seismic tomography. These mantle circulation models impose plate tectonic reconstructions as boundary conditions, as in the first two approaches.

Here, we take a different view on data assimilation methods for mantle circulation models by focusing on how to take into account the uncertainties in plate tectonic reconstructions. For almost a decade, 3-D spherical mantle convection models have shown the capability to self-consistently produce plate-like tectonics at their surface (Walzer & Hendel, 2008; Van Heck & Tackley, 2008; Yoshida, 2008; Foley & Becker, 2009). These models physically link surface tectonics comparable to that of the Earth to mantle convection processes (Coltice et al., 2012; Rolf et al., 2014; Mallard et al., 2016). In Bocher et al. (2016), we took advantage of this link to build a sequential data assimilation algorithm able to integrate plate reconstructions into a mantle convection code while taking into account the uncertainties on those plate tectonic reconstructions. This technique assimilates a time series of surface observations chronologically, by repeating two stages, analysis and forecast, until all observations are taken into account. Whenever an observation is available, the analysis evaluates the most likely state of the mantle at this time, considering a prior guess (supplied by the forecast) and the new observations at hand. For this evaluation, we used the classical best linear unbiased estimate (Talagrand, 1997). Then, the forward model of mantle convection forecasts the evolution of the mantle until the next observation time. We tested this algorithm on synthetic experiments. It proved to be efficient in recovering mantle circulation given constraints on the amplitude of errors affecting observations and the timespan between observations. Here we extend this work by applying a more advanced sequential data assimilation method, the Ensemble Kalman Filter (EnKF, described in Evensen, 1994; Burgers et al., 1998). This method is particularly suited for high dimensional nonlinear dynamical models (Evensen, 2009b). Instead of estimating the most likely state of the mantle, the Ensemble Kalman filter provides at each time an approximation of the probability density function of the state of the system in the form of a finite ensemble of states. During the forecast stage, each member of the ensemble evolves independently. For the analysis, we use the second order statistics of the ensemble to correct each ensemble member with the new observations at hand. We evaluate this method with synthetic experiments in 2D-spherical annulus geometry (Hernlund & Tackley, 2008) and compare it to the algorithm developed by Bocher et al. (2016). The EnKF

provides more accurate estimations than the former method, and is even able to reconstruct evolutions that the former method could not. Moreover, the EnKF also estimates locally the error on the reconstruction. The optimal size of the ensemble for our test case is $300$ members. Both covariance inflation and localization eliminate spurious correlations arising from the finite size of the Ensemble that is used to compute them.

This paper is organized as follows. In section 2, we present our simplifications on the general mantle circulation reconstruction problem and the correspondence with the notation in the EnKF algorithm. Then, in section 3, we detail the EnKF method and justify the variants chosen for the application to mantle circulation. Section 4 presents the results obtained on synthetic experiments and compares them to results obtained by the method described in Bocher et al. (2016). Section 5 is a discussion on the choice of the method and the challenges involved in the application of such a method to a realistic setting.

## 4.2 Presentation of the Problem

We aim at reconstructing mantle circulation for the last hundreds of millions of years by combining a mantle convection model with plate tectonic reconstructions, using an Ensemble Kalman Filter. To study the behavior of the Ensemble Kalman filter on such problem, we consider a simplified mantle convection model. This section describes the model used to compute a mantle evolution, the data set assimilated in this evolution, and finally the backbone of Ensemble Kalman Filtering.

### 4.2.1 Mantle Convection Model

At the timescales and lengthscales we are interested in ($\geq$ 10 kyr, $\geq$ 1000 km), the mantle can be modelled as a continuous visco-plastic medium. To compute mantle circulation, we solve the equations of conservation of mass (Eq. 4.1 below), momentum

(Eq. 4.2 below) and energy (Eq. 4.8 below) for an isochemical mantle under the Boussinesq approximation. The system of equations is non-dimensionalized to the thermal diffusion time scale (see Ricard, 2015). Given the high Prandtl number of the mantle (of the order of $10^{24}$), inertia is neglected. With these assumptions, the equations of conservation of mass and momentum become diagnostic equations of the form

$$\nabla \cdot \mathbf{u} = 0, \tag{4.1}$$

$$\nabla \cdot \boldsymbol{\sigma} - \nabla p + \mathrm{Ra} T \boldsymbol{e}_r = 0, \tag{4.2}$$

where $\boldsymbol{\sigma}$, $\mathbf{u}$, $p$, and $T$ are the non-dimensional deviatoric stress, velocity, dynamic pressure, and temperature, respectively. The equations are written in spherical coordinates $(r, \theta, \phi)$, using the physical convention with $r$ the radius, $\theta$ the colatitude and $\phi$ the longitude. The associated unit vectors are $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\phi)$.

$\mathrm{Ra}$ is the Rayleigh number, defined as

$$\mathrm{Ra} = \frac{\rho_0 g_0 \alpha_0 \Delta T a^3}{\mu_0 \kappa_0}, \tag{4.3}$$

with $\rho_0$ the density for $T = 0$, $g_0$ the gravitational acceleration, $\alpha_0$ the thermal expansivity, $\Delta T$ the temperature drop, $a$ the depth of the layer, $\kappa_0$ the thermal diffusivity, $\mu_0$ the dynamic viscosity of the system. The Rayleigh number in our model is $10^6$. It is one or two orders of magnitude lower than that of the Earth, but high enough to ensure chaotic convection. The vertical velocities and shear-stress at the surface and the base of the model are set to zero.

The deformation response of mantle material to stress is implemented as a linear relationship linking the strain rate tensor $\dot{\boldsymbol{\epsilon}}$ to the deviatoric stress tensor $\boldsymbol{\sigma}$ as

$$\boldsymbol{\sigma} = 2\mu_{\mathrm{eff}} \dot{\boldsymbol{\epsilon}} = \mu_{\mathrm{eff}} \left( \nabla \mathbf{u} + (\nabla \mathbf{u})^T \right). \tag{4.4}$$

The effective viscosity $\mu_{\mathrm{eff}}$ takes into account both a viscous Newtonian behavior with

a viscosity $\mu_n$ and a pseudo-plastic behavior with an equivalent "pseudo-plastic viscosity" $\mu_y$,

$$\mu_{\text{eff}} = \min(\mu_n, \mu_y). \tag{4.5}$$

The Newtonian viscosity $\mu_n$ follows an Arrhenius law

$$\mu_n = \mu_0 \exp\left(\frac{E_A}{T + T_1}\right) \tag{4.6}$$

with $\mu_0 = \exp\left(-\frac{E_A}{2T_1}\right)$, $T_1$ the temperature at which $\mu_n = 1$, and $E_A$ the nondimensional activation energy. We implement the decrease of viscosity in the asthenosphere by reducing by a factor of $10$ the viscosity $\mu_n$ when the temperature is above a solidus equation $T_s = T_{s_0} + \nabla_r T_s(r_a - r)$ with $r_a$ the surface value of $r$. The implementation of a weak asthenosphere tends to favor plate-like behavior (Tackley, 2000b; Richards et al., 2001).

The pseudo-plastic part of the effective viscosity $\mu_y$ is defined by

$$\mu_y = \frac{\sigma_{yield}}{2\dot{\epsilon}_{\text{II}}}, \tag{4.7}$$

where $\dot{\epsilon}_{\text{II}}$ is the second invariant of the strain rate tensor and $\sigma_{yield} = \sigma_Y + (r_a - r)\nabla_r \sigma_Y$, with $\sigma_Y$ and $\nabla_r \sigma_Y$ the yield stress at the surface and the depth-dependence of the yield stress, respectively.

The energy conservation equation is the only prognostic equation of the system

$$\frac{\mathrm{D}T}{\mathrm{D}t} = \nabla^2 T + \mathrm{R}_h. \tag{4.8}$$

$\mathrm{R}_h$ is the non-dimensional internal heating rate defined as

$$\mathrm{R}_h = \frac{\rho_0 D^2 H}{k_0 \Delta T} \tag{4.9}$$

with $H$ the dimensional heating rate and $k_0$ the thermal conductivity. We set isothermal top and bottom boundaries with temperatures $T_a$ and $T_b$, respectively. The models presented here have $10\%$ basal heating and $90\%$ internal heating.

These equations are solved using the finite volume, multigrid parallel code STAGYY (Tackley et al., 1993), on a spherical annulus grid. This geometry provides results closer to the spherical grid than cylindrical geometry (Hernlund & Tackley, 2008). In the following, the longitudinal coordinate of a point is $\phi_l$, with $l \in \{1, 2, ..., L\}$ and its radial coordinate is $r_m$ with $m \in \{1, 2, ..., M\}$, $r$ varying from $r_b$ to $r_a$.

Note that this paper focuses on the methodology of ensemble data assimilation for a convecting system similar to that of the Earth's mantle. Hence, we choose a rather simple model that can reproduce plate-like tectonics at the surface. We rely on simplifications such as 2D geometry, incompressible and isochemical mantle and a rheology which does not take into account the history of the material. Although some of the complexities we ignore may play a fundamental role in the reconstruction of the Earth's mantle evolution, we choose to focus in this manuscript on the data assimilation methodology. Moreover, we choose to keep the same parameters as the test case of Bocher et al. (2016) in order to allow direct comparison between the methods. Table 4.1 lists the chosen parameter values.

To ease the comparison with Earth's mantle convection, we rescale the nondimensional time in the evolution, $t$, by the transit time of the convective system. By definition, the transit time of the Earth's mantle is $t_t^E = a^E/v_{rms}^E$, with $a^E$ the thickness of the mantle and $v_{rms}^E$ the root mean square of surface velocities of the Earth, as estimated by plate tectonic reconstructions (Seton et al., 2012). We compute the same value for the model $t_t^m = a/v_{rms}^m$. The scaled time $t^s$ is then $t^s = t\frac{t_t^E}{t_t^m}$.

### 4.2.2 Observations of Mantle Circulation

The state of the Earth's surface is the time integrated expression of mantle circulation. At a global scale, the main source of information for the last $100$ Myr is the database

TABLE 4.1: Values of the parameters of the forward model

| Symbol | Meaning | value |
|---|---|---|
| $\mathrm{Ra}$ | Rayleigh number | $10^6$ |
| $\mathrm{R}_h$ | Non-dimensional internal heating rate | $20.5$ |
| $L$ | number of grid points in longitude | $384$ |
| $M$ | number of grid points in radius | $48$ |
| $r_a$ | Radius of the top of the domain | $2.2$ |
| $r_b$ | Radius of the bottom of the domain | $1.2$ |
| $T_a$ | Temperature at the top of the domain | $0$ |
| $T_b$ | Temperature at the bottom of the domain | $0.9$ |
| $E_A$ | Activation Energy | $23.03$ |
| $T_1$ | Temperature at which $\mu_T = 1$ | $1$ |
| $\beta$ | Factor of viscosity reduction for partial melting | $10$ |
| $T_{s_0}$ | Solidus Temperature at $r = r_a$ | $0.6$ |
| $\nabla_r T_s$ | Radial gradient of the solidus temperature | $2$ |
| $\sigma_Y$ | Yield Stress | $1.10^4$ |
| $\nabla_r \sigma_Y$ | Radial gradient of the yield stress | $2.10^5$ |

of the localization and identification of magnetic anomalies on the seafloor, translated into maps of seafloor ages (Müller et al., 2008a; Seton et al., 2014). This information is complemented with regional geological studies giving constraints on the timing and geometry of tectonic events as well as a synthesis of paleontological, structural geology, stratigraphical, magnetic anomalies, gravity data and seismic studies. In addition, paleomagnetic data provide constraints on the paleolatitude of continental blocks (Besse & Courtillot, 2002).

Plate tectonic reconstructions use the geometric theory of plate tectonics to integrate all these observations. The result is a time series of maps of seafloor ages, plate layout and kinematics. The continuously closed plate algorithm (Gurnis et al., 2012) produces plate tectonic reconstruction maps continuous in space and time (Seton et al., 2012; Müller et al., 2016).

Although we are aware that these plate tectonic reconstruction maps are in themselves models and not direct observations, we propose to develop an assimilation method that use them as data to assimilate in our mantle convection model. This solution is generally chosen in mantle circulation reconstructions (Bunge et al., 2002; Zhang et al., 2010; Bower et al., 2015), because it provides continuous surface boundary conditions in space and time for the period of reconstruction. One advantage of the technique we

develop is that it is possible to consider errors on the data that is assimilated, another is that the reconstructions do not need to be known at all times and at all points on the surface. Hence it is possible, in principle, to design a data assimilation scheme using direct observations. However, this would require further developments both on the database design and on the data assimilation algorithm. Sequential data assimilation methods for mantle circulation are still in their infancy, so we opt for a simpler structure of the data to be assimilated: a time series of maps of surface velocity and seafloor age, as given by plate tectonic reconstructions.

In this study, we limit ourselves to the test of data assimilation in synthetic experiments. In the model described in subsection 4.2.1, the absence of small scale convection at the base of the boundary layer makes the surface heat flux an excellent proxy for the age of the seafloor (Coltice et al., 2012). Consequently, we consider surface heat flux and surface velocity as the data to assimilate.

To our knowledge, the amplitude of the uncertainty on global plate tectonic reconstructions has not yet been assessed. For the synthetic tests we perform in section 4.4, we choose an arbitrary value of $10\%$ of the root mean square value of heat flux and surface velocity, respectively. We further discuss this choice in section 4.5.

### 4.2.3 Ensemble Kalman Filtering Framework: Notations

Our aim is to assimilate a time series of observations (surface velocities and heat fluxes) into a mantle convection model to estimate the evolution of the state of the mantle. We introduce here the general formulation of Ensemble Kalman Filtering and link them to our problem. We use the notation system recommended by Ide et al. (1997).

The time series of data is defined as a set of column vectors $\{\mathbf{y}_1^o, \mathbf{y}_2^o, ..., \mathbf{y}_K^o\}$, where the subscripts $\{1, 2, ..., K\}$ refer to the times at which observations are available. As seen in the previous section, the data used for our experiments are surface velocity and surface

heat flux. The data vector at time $k$ is thus defined as

$$\mathbf{y}_k^o = \left[q_k^o(\phi_1), q_k^o(\phi_2), ..., q_k^o(\phi_L), u_{\phi k}^o(\phi_1), u_{\phi k}^o(\phi_2), ..., u_{\phi k}^o(\phi_L)\right]^T, \tag{4.10}$$

where $q_k^o(\phi_l)$ and $u_{\phi k}^o(\phi_l)$ are the observed values of surface heat flux and surface horizontal velocity at the $k$-th timestep and longitude $\phi_l$, and $(\cdot)^T$ means transpose. We model errors on observations by a random vector of zero mean and covariance matrix $\mathbf{R}_k$ (we suppose unbiased observations). Although $\mathbf{R}_k$ is a diagonal matrix of constant value and size in our experiments, it is not generally the case. Correlations between errors on observations could be specified in $\mathbf{R}_k$.

The evolution of the state of the system is estimated sequentially during the period where observations are available. At each timestep $k \in \{1, 2, ..., K\}$, we define two state vectors: the a priori state, or forecast state $\mathbf{x}_k^f$ and the analysis state $\mathbf{x}_k^a$, which is the state corrected after having assimilated the observations $\mathbf{y}_k^o$. The system of equations developed in section 4.2.1 shows that we can compute velocity, viscosity and pressure values at each grid point from the temperature field. Nevertheless, the relation between surface velocities and the temperature field is nonlinear. We choose to include the whole temperature field and the surface velocities to form an augmented state vector. This simplifies the computations thereafter. The state of the mantle at a timestep $k \in [1, K]$ is defined as

$$\mathbf{x}_k = [T_k(\phi_1, r_1), T_k(\phi_1, r_2), ..., T_k(\phi_L, r_M), u_{\phi k}(\phi_1), u_{\phi k}(\phi_2), ..., u_{\phi k}(\phi_L)]^T, \tag{4.11}$$

where $T_k(\phi_l, r_m)$ and $u_{\phi k}(\phi_l)$ are the values of temperature at the $k$th timestep, longitude $\phi_l$ and radius $r_m$ and surface horizontal velocity at the $k$th timestep and longitude $\phi_l$.

The forecast and analyzed states are uncertain as well. Their uncertainties are represented by two random vectors of zero expectancy and covariance matrices $\mathbf{P}_k^f$ and $\mathbf{P}_k^a$, respectively. We do not compute explicitely these covariance matrices. Instead, we

TABLE 4.2: Notations and dimensions of data assimilation variables

| Symbol | Meaning | size (literal) | size (Value) |
|---|---|---|---|
| $\mathbf{x}$ | state | $LM+L$ | 18 816 |
| $\mathbf{y}$ | data | $L+L$ | 768 |
| $\mathbf{H}$ | observation matrix operator | $(L+L)\times(LM+L)$ | $768\times 18\,816$ |
| $\mathbf{R}$ | observation error covariance matrix | $(L+L)\times(L+L)$ | $768\times 768$ |
| $\mathbf{P}$ | state error covariance matrix | $(LM+L)\times(LM+L)$ | $18\,816\times 18\,816$ |
| $\mathbf{X}$ | ensemble state | $(LM+L)\times N$ | $18\,816\times N$, ($N=96$,288 or 768) |

compute two ensembles of $N$ states $\{\mathbf{x}^f_{kn}\}_{n\in[1,N]}$ and $\{\mathbf{x}^a_{kn}\}_{n\in[1,N]}$, such that their average equals $\mathbf{x}^f_k$ and $\mathbf{x}^a_k$, respectively, and their respective sample covariance matrices approximate $\mathbf{P}^f_k$ and $\mathbf{P}^a_k$. The ensemble of states $\{\mathbf{x}^f_{kn}\}_{n\in[1,N]}$ and $\{\mathbf{x}^a_{kn}\}_{n\in[1,N]}$ are stored in the matrices $\mathbf{X}^f_k$ and $\mathbf{X}^a_k$, where the $n$th column is the state of the $n$th ensemble member $\mathbf{x}^f_{kn}$ and $\mathbf{x}^a_{kn}$, respectively.

Finally, we introduce the observation operator, which maps a given state vector $\mathbf{x}^e_{kn}$ ($e$ being $f$ or $a$) to the corresponding data $\mathbf{y}^e_{kn}$. If the surface heat flux is approximated by a first order discretization of Fourier's law, then the observation operator is linear, and can be represented by the matrix $\mathbf{H}$ such that

$$\forall k\in\{1,2,...,K\},\forall n\in\{1,2,...,N\},\quad \mathbf{y}^e_{kn}=\mathbf{H}\mathbf{x}^e_{kn}. \tag{4.12}$$

Table 4.2 summarizes the dimensions of the vectors and matrices for our problem.

## 4.3 Ensemble Kalman Filter with Localization and Inflation

The Ensemble Kalman Filter (Evensen, 1994; Burgers et al., 1998) is a sequential data assimilation algorithm using the same equations as the Kalman Filter for the analysis step, but Monte Carlo methods to forecast the error statistics on the state. We explain here how we adapt the Ensemble Kalman Filter to our problem and justify the choice of the starting ensemble.

To implement the EnKF, we used the software environment Parallel Data Assimilation Framework (PDAF, Nerger et al., 2005; Nerger & Hiller, 2013).

### 4.3.1 Initialization: First Analysis and Generation of the starting ensemble

As in Bocher et al. (2016), we compute second order statistics from a series of $400$ decorrelated snapshots of convection simulations. We obtain the first forecast state of average $\mathbf{x}_1^f$ and associated covariance matrix $\mathbf{P}_1^f$. The background covariance matrix $\mathbf{P}_1^f$ is eigendecomposed and rank reduced into $\mathbf{P}_{1r}^f = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$, with $\mathbf{\Lambda}$ a diagonal matrix containing the $n_r = 1928$ largest eigenvalues of $\mathbf{P}_1^f$ (which accounts for $99.98\%$ of its cumulative variance) and $\mathbf{V}$ a matrix of the corresponding eigenvectors.

The first set of observations $\mathbf{y}_1^o$ is assimilated to obtain

$$\mathbf{x}_1^a = \mathbf{x}_1^f + \mathbf{V}\mathbf{A}\mathbf{V}^T\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y}_1^o - \mathbf{H}\mathbf{x}_1^f), \tag{4.13}$$

$$\mathbf{P}_1^a = \mathbf{V}\mathbf{A}\mathbf{V}^T, \tag{4.14}$$

with

$$\mathbf{A} = \left[\mathbf{\Lambda}^{-1} + \mathbf{V}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{V}\right]^{-1}. \tag{4.15}$$

We generate an ensemble of $N$ initial states using the second order exact sampling method (Hoteit, 2001; Pham, 2001). First, A is eigendecomposed

$$\mathbf{A} = \mathbf{V}^a\mathbf{\Lambda}^a\mathbf{V}^{aT}. \tag{4.16}$$

The ensemble members are then computed following

$$\mathbf{X}_1^a = \begin{pmatrix} | & & | \\ \mathbf{x}_{11}^a & \dots & \mathbf{x}_{1N}^a \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \mathbf{x}_1^a & \dots & \mathbf{x}_1^a \\ | & & | \end{pmatrix} + \sqrt{N-1}\mathbf{V}\mathbf{V}^a\mathbf{\Lambda}^{a1/2} \begin{pmatrix} \mathbf{\Omega}_{N\times(N-1)}^T \\ \mathbf{0}_{(n_r-N)\times N} \end{pmatrix}, \quad (4.17)$$

where $\mathbf{\Omega}$ is a random matrix whose columns are vectors forming an orthonormal basis and each of them is orthogonal to $\mathbf{1} = [1, ..., 1]^T$. $\mathbf{\Omega}$ is generated through the algorithm described in the appendix of Nerger et al. (2012). The matrix $\mathbf{\Omega}$ is designed so that the sample mean of the starting ensemble is equal to $\mathbf{x}_1^a$ and its sample covariance matrix is equal to matrix $\mathbf{P}_1^a$ reduced to its $N$ largest eigenvalues.

This method of generating the starting ensemble takes advantage of the extensive knowledge we have on the background statistics of the model. Several other methods have been tested to generate a starting ensemble, such as starting with random decorrelated snapshots of mantle convection obtained from a very long run, second order exact sampling from $\mathbf{x}_1^f$ and $\mathbf{P}_1^f$, and several assimilations of the first observations $\mathbf{y}_1^o$. None of these solutions were as efficient for our problem as the technique used here.

### 4.3.2 Forecast

Between timesteps $k-1$ and $k$, the forward numerical code STAGYY computes independently the evolution of each of the analyzed states $\{\mathbf{x}_{k-1,n}^a\}_{n\in[1,N]}$ to produce a forecast ensemble $\{\mathbf{x}_{k,n}^f\}_{n\in[1,N]}$.

The forecast state is the average of the ensemble

$$\mathbf{x}_k^f = \frac{1}{N}\mathbf{X}_k^f\mathbf{1}, \quad (4.18)$$

and the forecast error covariance matrix is given by the sample covariance matrix of the ensemble of forecast states

$$\mathbf{P}_k^f = \frac{1}{N-1}\mathbf{X}_k^f \left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)\left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)^T \mathbf{X}_k^{fT}. \tag{4.19}$$

After several assimilation cycles, the finite size of the ensemble induces the underestimation of the error variance (van Leeuwen, 1999), and can lead to filter divergence. We observed this behavior in our case, and to stabilize the filter we apply covariance inflation, as suggested in Anderson & Anderson (1999) and Hamill et al. (2001).

We correct the forecast ensemble variance with an inflation factor $\gamma$ according to

$$\mathbf{X}_k^f \leftarrow \frac{1}{N}\mathbf{X}_k^f \mathbf{1}\mathbf{1}^T + \left[\mathbf{X}_k^f \left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)\right]\sqrt{\gamma}, \tag{4.20}$$

where $\leftarrow$ means that we replace the matrix on the left-hand side by the term on the right-hand side. $\gamma$ is computed following the same principles as in the suboptimal Kalman Filter developed in Bocher et al. (2016), i.e. by comparing the error on observations and the standard deviation of the innovation $\mathbf{d}_k$ defined as

$$\mathbf{d}_k = \mathbf{y}_k^o - \frac{1}{N}\mathbf{H}\mathbf{X}_k^f \mathbf{1}. \tag{4.21}$$

The inflation factor is

$$\gamma = \frac{V^d - V^o}{V^f}, \tag{4.22}$$

with

$$V^d = \mathrm{Tr}\left(\mathbf{d}_k \mathbf{d}_k^T\right), \tag{4.23}$$

$$V^o = \mathrm{Tr}(\mathbf{R}_k), \tag{4.24}$$

$$V^f = \mathrm{Tr}\left[\mathbf{H}\mathbf{X}_k^f \left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)\left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)^T \mathbf{X}_k^{fT}\mathbf{H}^T\right], \tag{4.25}$$

where $\mathrm{Tr}(\cdot)$ means the trace. The inflation factor is then truncated between a minimum value of $1$ (to prevent further contraction of the ensemble spread) and a maximum value of $\gamma^+ = 1.25$ (to prevent overspread). Several values of maximum inflation factor have been tested, from $\gamma^+ = 1.1$ to $\gamma^+ = 2$, and showed little impact on the efficiency of the assimilation. A constant inflation factor was also tested, but the results with an adaptive inflation factor were substantially more accurate, especially for the first assimilation times.

### 4.3.3 Analysis

The analyzed state $\mathbf{x}^a_{kn}$ of the $n$th member of the ensemble is

$$\mathbf{x}^a_{kn} = \mathbf{x}^f_{kn} + \mathbf{K}_k \left( \mathbf{y}^o_{kn} - \mathbf{H}\mathbf{x}^f_{kn} \right) \tag{4.26}$$

where $\mathbf{K}_k$ is the Kalman Gain. $\mathbf{y}^o_{kn}$ is the observed data vector $\mathbf{y}^o_k$ to which a random perturbation of zero expectation and covariance matrix $\mathbf{R}_k$ is added, as is recommended in Burgers et al. (1998).

The Kalman Gain is defined as

$$\mathbf{K}_k = (\mathbf{P}^f_k \circ \mathbf{C})\mathbf{H}^T \left[ \mathbf{H}(\mathbf{P}^f_k \circ \mathbf{C})\mathbf{H}^T + \mathbf{R}_k \right]^{-1}, \tag{4.27}$$

where the matrix $\mathbf{P}^f_k$ is the sample covariance matrix of the ensemble of forecast states $\{\mathbf{x}^f_{kn}\}_{n\in[1,N]}$. We use a limited ensemble size (maximum $768$) to estimate $\mathbf{P}^f_k$. Spurious correlations ensue, especially between distant points. To counteract this effect, we implement direct forecast error localization by Schur multiplying (symbol $\circ$) $\mathbf{P}^f_k$ by the localization matrix $\mathbf{C}$, as introduced by Hamill et al. (2001) and Houtekamer & Mitchell (2001). The matrix $\mathbf{C}$ is itself the Schur product of a vertical localization matrix $\mathbf{C}_v$ and a horizontal localization matrix $\mathbf{C}_h$. The value of $\mathbf{C}_v(i,j)$ depends on the absolute radius difference of the $i-th$ and the $j-th$ components of the state vector and on vertical correlation length $\ell_v$. The value of $\mathbf{C}_h(i,j)$ depends on the absolute angle difference of

the $i-th$ and the $j-th$ components of the state vector and on vertical correlation length $\ell_h$. Both values follow a Gaspari-Cohn compactly supported fifth-order piecewise rational function (similar to a Gaussian but with a compact support, Eq. 4.10 of Gaspari & Cohn, 1999).

We also tested the domain localization strategy as described in Janjic et al. (2011), since it is computationally more efficient and already implemented in PDAF. However, it led to a systematic failure of the assimilation. This is due to the nature of our problem: all the observations are located at the surface of the model and we aim at estimating the temperature field over the whole depth of the mantle. A vertical localization is as necessary as a horizontal localization: the localization has to be done in the state space and not only in the data space.

### 4.3.4 Implementation of the Ensemble Kalman Filter

We used the software environment PDAF (Nerger et al., 2005; Nerger & Hiller, 2013) in combination with the mantle convection code STAGYY (Tackley, 2008) to develop an Ensemble Kalman Filter code for mantle convection . PDAF provides a set of core routines computing in parallel the analysis steps for a range of ensemble based data assimilation techniques. It provides as well a set of standard routines to adapt the parallelization of a preexisting parallel forward numerical model and integrate the data assimilation routines. The final product is a highly scalable ensemble data assimilation code running both forecasts and analyses in parallel.

We modified the STAGYY code following the procedure recommended by PDAF (see the online documentation wiki at Nerger, 2016). We also made a few modifications in PDAF routines to allow for direct forecast error localization with the Ensemble Kalman filter. Additionally, we designed a basic observation database so as to load in a single step all the observations used in the data assimilation procedure.

## 4.4 A posteriori Evaluation of the Ensemble Kalman Filter Method

We test the data assimilation scheme on twin experiments using the model described in section 4.2.1. Throughout this section, we compare the results of the Ensemble Kalman Filter for mantle circulation reconstructions to the results computed using the method developed in Bocher et al. (2016), hereafter referred to as method 1.

After describing the setup used for twin experiments, we test the robustness of the EnKF method and compare it to that of method 1. Then, we determine the range of data assimilation parameters which are suitable to conduct an ensemble data assimilation. Finally, we assess the ability of the scheme to actually reconstruct specific geodynamic structures.

### 4.4.1 Twin Experiment Setup

Twin experiments are a way to assess the accuracy of a data assimilation procedure in a controlled environment, where the true evolution is perfectly known.

First, we compute a reference evolution using the forward numerical model, considered as the true evolution, from which we extract the set of true states $\{\mathbf{x}_k^t\}_{k\in[1,K]}$. Here, the timespan of the evolution is $150$ Myr and we sample true states every $10$ Myr. From these states, we compute a time series of surface heat fluxes and surface velocities, following Equation 4.12. We noise these observations with a random Gaussian noise of standard deviation $10\%$ of the root mean square of surface heat flux $q_{rms}$ and surface velocities $v_{rms}$, to obtain the time series of observations to assimilate $\{\mathbf{y}_k^o\}_{k\in[1,K]}$. It follows that the observation error covariance matrix $\mathbf{R}$ is diagonal and does not change with time.

Then, we perform ensemble data assimilation for the data set $\{\mathbf{y}_k^o\}_{k\in[1,K]}$, with the observation error covariance matrix $\mathbf{R}$. We did not consider any model error in the filter

TABLE 4.3: Notations and range of values tested for data assimilation parameters

| Symbol | Meaning | value |
|---|---|---|
| $N$ | number of ensemble members | 96 to 768 |
| $K$ | number of observation times | 16 |
| $\gamma^+$ | maximum inflation factor | 1.25 |
| $\ell_v$ | vertical correlation length | 0.3 to 1 |
| $\ell_h$ | horizontal correlation angle | $\pi/10$ to $\pi/2$ |

we describe, so the parameters of the model used in the data assimilation realizations are the same as those of the reference model.

We present here tests with different assimilation parameters, varying the number of members $N$, the vertical correlation length $\ell_v$ and the horizontal correlation angle $\ell_h$. Table 4.3 details the range of parameters tested.

We compute four different evolutions to test the accuracy of the ensemble Kalman filter for different cases. Figure 4.3 shows the initial and final states of these evolutions, together with the result of global error evolution, and will be discussed in the next section.

### 4.4.2 Robustness of the Assimilation Algorithm

The evolution of the global error on the estimated temperature field over the time period $\{1, ..., K\}$ is

$$\left[\epsilon_T^f(1), \epsilon_T^a(1), \epsilon_T^f(2), ..., \epsilon_T^f(K), \epsilon_T^a(K)\right] \tag{4.28}$$

where $\epsilon_T^e(k)$, $e$ standing for $a$ (analysis) or $f$ (forecast) is

$$\epsilon_T^e(k) = \sqrt{\frac{\sum_{l=1}^{L}\sum_{m=1}^{M}\left(\overline{T}_k^e(\phi_l, r_m) - T_k^t(\phi_l, r_m)\right)^2 \mathcal{V}(\phi_l, r_m)}{\sum_{l=1}^{L}\sum_{m=1}^{M} T_k^t(\phi_l, r_m)^2 \mathcal{V}(\phi_l, r_m)}} \tag{4.29}$$

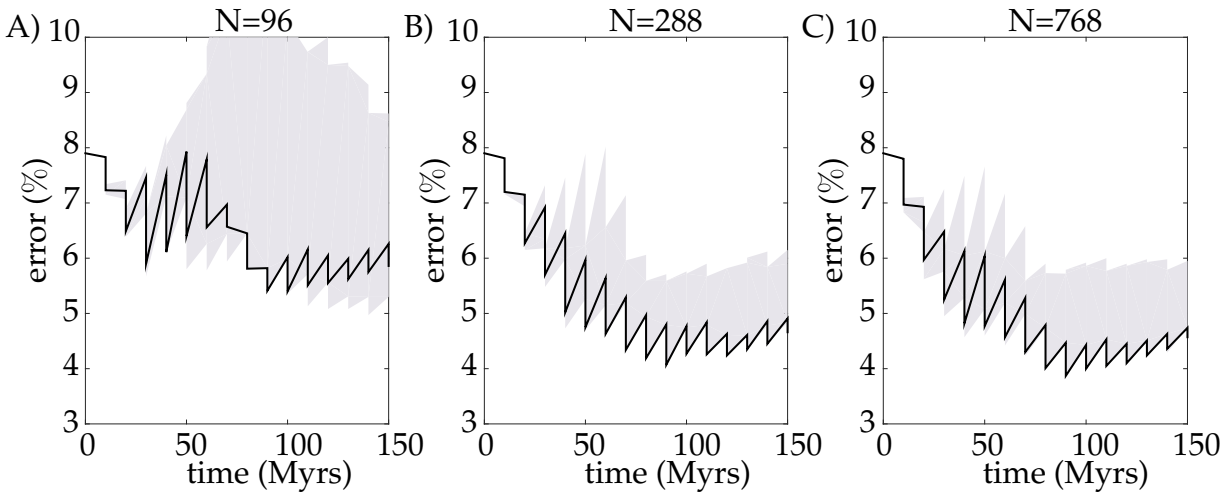


FIGURE 4.1: Time evolution of the errors on the estimated temperature field obtained from data asssimilations with the same 150 Myr dataset, but different assimilation parameters. The size of the ensemble is A) $N = 96$, B) $N = 288$ and C) $N = 768$. The assimilations are computed for any combination of data assimilation parameters: $\gamma^+ = 1.25$, $\ell_v = 0.3$, $0.5$, $0.7$ and $1$ and $\ell_h = \pi/10$, $\pi/8$, $\pi/6$, $\pi/4$ and $\pi/2$. The black line represent the evolution of the error for the best assimilation A)$N = 96$, $\ell_z = 0.5$, $\ell_h = \pi/6$ and $\gamma^+ = 1.25$ B)$N = 288$, $\ell_z = 0.7$, $\ell_h = \pi/10$ and $\gamma^+ = 1.25$ and C)$N = 768$, $\ell_z = 0.5$, $\ell_h = \pi/4$ and $\gamma^+ = 1.25$. The grey area is delimited by the maximum and minimum values of errors at each time, for all data assimilations.

with $\mathcal{V}(\phi_l, r_m)$ the volume of the grid cell at longitude $\phi_l$ and radius $r_m$, and $\overline{T}^e_k(\phi_l, r_m)$ the average temperature of the estimated ensemble (either forecast or analysis) at longitude $\phi_l$ and radius $r_m$.

We test the EnKF on one evolution, with sizes of the ensemble $N = 96$, $288$ and $768$ and for each combination of the following data assimilation parameters values: vertical correlation length $\ell_v = 0.3$, $0.5$, $0.7$ and $1$ and horizontal correlation angle $\ell_h = \pi/10$, $\pi/8$, $\pi/6$, $\pi/4$ and $\pi/2$ . We show in Figure 4.1, for each ensemble size, the maximum and minimum values of errors obtained for all these parameters, as a function of time.

To determine the best assimilation, we compute the average error after analysis

$$\overline{\epsilon}^a_T = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{\epsilon}^a_T(k). \tag{4.30}$$

For each ensemble size, the error evolution of the best assimilation (in the sense of minimum $\overline{\boldsymbol{\epsilon}}^a_T$) is also shown in Figure 4.1.

For any size of the ensemble, it is possible to find a set of parameters leading to a stabilization of the global error on the temperature field after a few analyses. The time after which the solution is stabilized corresponds to the transit time of the physical model (70 Myr). We can see that for $N = 288$ and $N = 768$, any combination of vertical and horizontal correlation lengths leads to errors lower than the first analysis. Although the error is decreasing through time for any combination of data assimilation parameters, the difference between the maximum and the minimum errors obtained is greater than $1\%$, which is large given that the first analysis error is already below $8\%$. The best error evolutions for $N = 288$ and $N = 768$ are very similar, with a minimum error of $4.07\%$ and $3.87\%$ after $90$ Myr, and an average global error after analysis of $5.01\%$ and $4.85\%$, respectively. During the assimilation of a dataset, most of the computational time is dedicated to the forecast step, so the data assimilation with $768$ members is $2.7$ times longer than the assimilation with $288$ members, on the account of the embarrassingly parallel nature of the forecast phase. Since we obtain very similar results for $N = 288$ and $N = 768$, we favor the assimilation with $288$ members.

We compute the error on the estimated temperature from the true temperature field. However, in a realistic case, the true temperature is not known, and the evaluation of the data assimilation algorithm is based on the study of the statistics of the innovation vector $\mathbf{d}_k$ at analysis $k$

$$\mathbf{d}_k = \mathbf{y}_k^o - \mathbf{H}\mathbf{x}_k^f. \tag{4.31}$$

At each analysis time, we compute the Euclidean norm of the instantaneous innovation $d_k$ and the Euclidean norm of the cumulative mean innovation $\overline{\mathbf{d}}_k$

$$d_k = \| \mathbf{d}_k \| \qquad \text{and} \qquad \overline{\mathbf{d}}_k = \left\| \frac{1}{k} \sum_{i=1}^{k} \mathbf{d}_i \right\| \tag{4.32}$$

Before computing these norms, we normalize the part of the innovation corresponding to surface heat flux and velocities by their respective root mean square values.

The norm of the instantaneous innovation $d_k$ measures the distance between the forecast data and the observation, and therefore gives indications on the success or failure of the assimilation. The evolution of the cumulative mean of the innovation $\overline{\mathbf{d}}_k$ allows us to check the consistency of the data assimilation algorithm at first order. Indeed, the derivation of the EnKF equations assumes that the error on observations $\mathbf{y}^o$ and the error on the forecast data $\mathbf{H}\mathbf{x}^f$ are unbiased. Such hypotheses imply that the statistically expected value of $\mathbf{d}$ is zero, which means that the norm of the cumulative innovation should converge to zero as the number of analyses increases.

Figure 4.2 shows the evolution of $d_k$ and $\overline{\mathbf{d}}_k$ as a function of the number of analyses for data assimilations with different sizes of ensemble and their respective optimum vertical and horizontal correlation lengths. The norm of the instantaneous innovation (Figure 4.2B) first decreases during the first $8$ analyses, i.e. $70$ My, and then oscillates for the rest of the assimilation. We observe the same behavior for the evolution of the error on temperature $\epsilon_T^e(k)$: the norm of the instantaneous innovation for ensemble sizes of $N = 288$ and $768$ are very similar, and lower than that obtained with $N = 96$. Figure 4.2A shows the cumulative innovation constantly decreasing throughout the assimilation, with comparable values for $N = 288$ and $N = 768$, and slightly higher values for $N = 96$.

We also tested the assimilation algorithm for 4 different evolutions, with the optimal parameters for an ensemble size of $N = 288$ members ($\ell_v = 0.7$ and $\ell_h = \pi/10$). Figure 4.3 shows the initial and final temperature fields of the evolutions, together with the evolution of the global error, the spread of the ensemble, and the error evolution for the method 1.

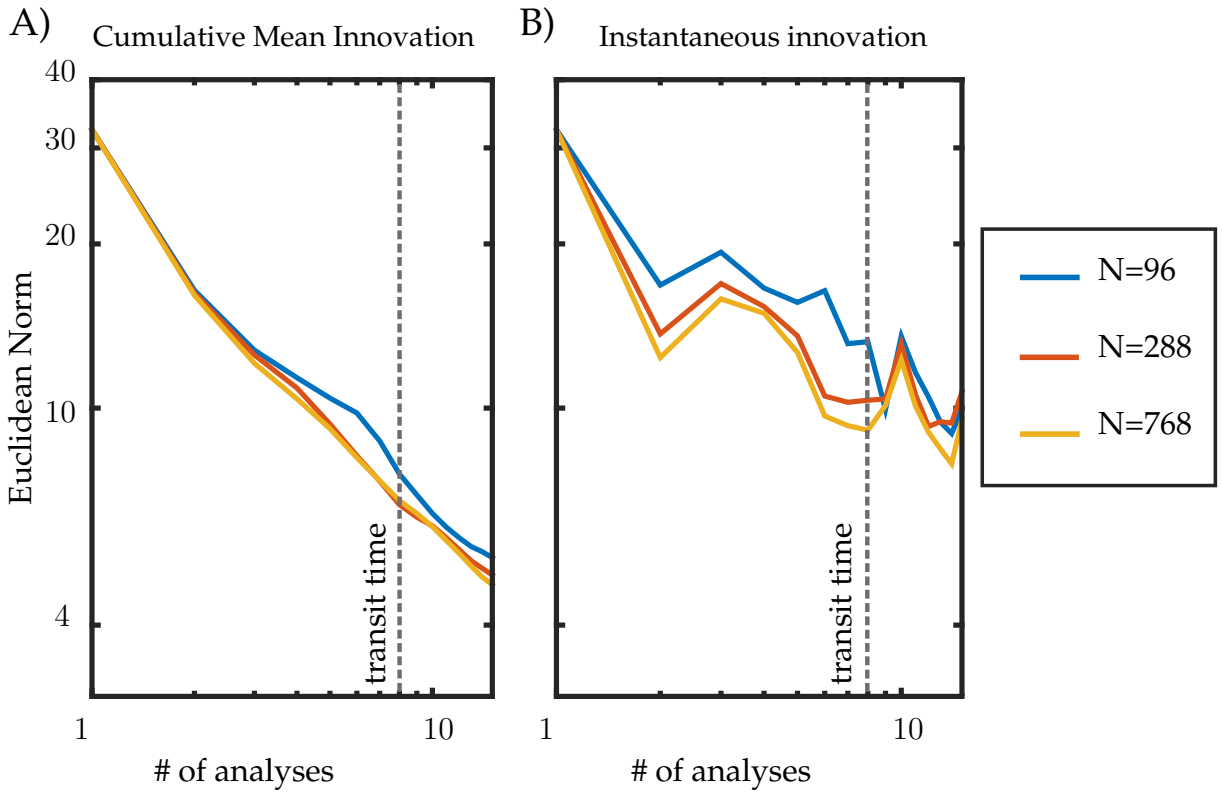


FIGURE 4.2: Evolution of A) the cumulative mean innovation and B) the norm of the instantaneous innovation, as a function of number of analyses and for different ensemble sizes. For each size of the ensemble, the evolutions correspond to the best combinations of correlation length parameters: $N = 96$, $\ell_z = 0.5$, $\ell_h = \pi/6$ and $\gamma^+ = 1.25$ ; $N = 288$, $\ell_z = 0.7$, $\ell_h = \pi/10$ and $\gamma^+ = 1.25$ and $N = 768$, $\ell_z = 0.5$, $\ell_h = \pi/4$ and $\gamma^+ = 1.25$.

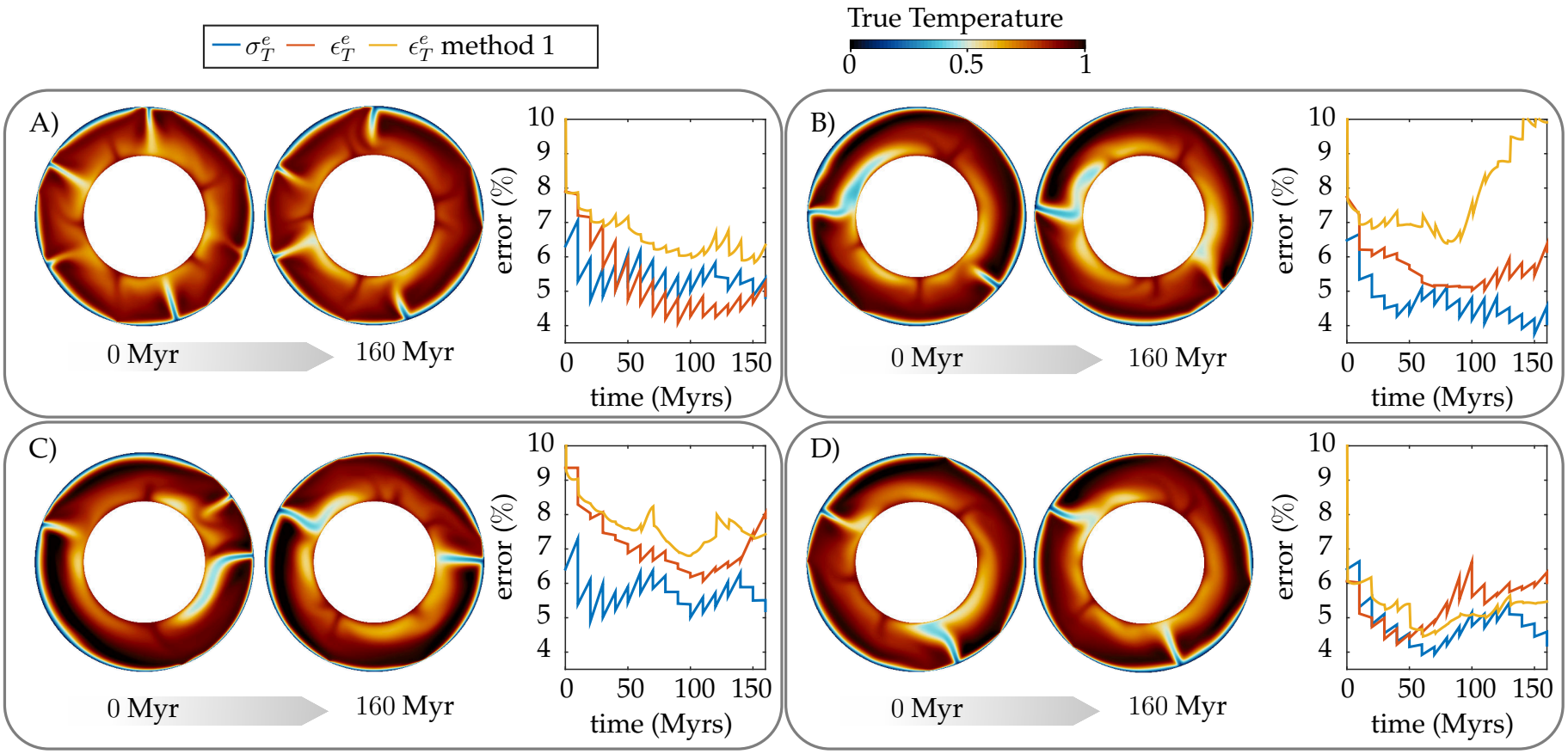


FIGURE 4.3: Evolution of the error ($\epsilon_T^e$, red) as a function of time for 4 different evolutions with $N = 288$, $\gamma^+ = 1.25$, $\ell_v = 0.7$ and $\ell_h = \pi/10$, compared to the evolution of the spread of the ensemble ($\sigma_T^e$, blue) and the evolution of the error with the technique of Bocher et al. (2016) ($\epsilon_T^e$ method 1, yellow). The initial and final states of the true evolutions are represented on the left of each corresponding graph.

The spread of the ensemble is an estimation of the error on the state. We compare the evolution of $\epsilon_T^e$ to the global standard deviation of the temperature field of the ensemble:

$$\left[\sigma_T^f(1), \sigma_T^a(1), \sigma_T^f(2), ..., \sigma_T^f(K), \sigma_T^a(K)\right] \tag{4.33}$$

with $\sigma_T^e(k)$ defined as

$$\sigma_T^e(k) = \sqrt{\frac{\sum_{n=1}^{N}\sum_{l=1}^{L}\sum_{m=1}^{M}\left(T_{kn}^e(\phi_l, r_m) - \overline{T}_k^e(\phi_l, r_m)\right)^2 \mathcal{V}(\phi_l, r_m)}{(N-1)\sum_{l=1}^{L}\sum_{m=1}^{M}\overline{T}_k^e(\phi_l, r_m)^2 \mathcal{V}(\phi_l, r_m)}}. \tag{4.34}$$

We compute the error for an estimated evolution with the method 1 using Equation 4.29.

Although we ran the four evolutions with the same values of physical parameters (as described in Table 4.1), they show different configurations: Evolution A has a shorter wavelength of convection, with the persistence of 4 subductions, 3 ridges and 5 upwellings, the death of one ridge and creation of two. Evolutions B, C and D have longer wavelengths of convection, with two major downwellings, stable throughout the evolutions. In evolution B, one of these downwellings has a very large negative temperature anomaly at the bottom of the domain. In evolution C, the remnant of a subduction merges with a larger subduction into a single downwelling.

In the 4 cases, the errors on the estimated temperature field systematically decrease during the analysis step for the EnKF algorithm. The errors stay below the first analysis error for evolutions A, B and C, while they reach slightly higher values for evolution D. The error of the EnKF is always lower than that obtained with method 1 for the first $50$ My. The average error is lower for the EnKF than for method 1 in 3 out of 4 cases. The average standard deviation of the ensemble (ensemble spread) is of the same order of magnitude as the true error. However, its evolution is not the same as the true error, with differences between both of more than $2\%$ for some part of evolution C, for example. For evolutions C and D, the results of the two methods are comparable

whereas the assimilation with EnKF performs better than method 1 for evolutions A and B.

For evolution B, method 1 fails to reconstruct accurately the evolution, with the error reaching values greater than $10\%$ at the end of the assimilation. This case is further investigated on Figures 4.4 and 4.5. Figure 4.4 compares the true temperature field evolution with the analyzed temperature field of method 1 and of the Ensemble Kalman filter with $N = 288$, $\ell_v = 0.7$ and $\ell_h = \pi/10$. The sudden increase in the error of the estimated temperature field for method 1 seen on Figure 4.3B happens after around $80$ Myr of assimilation, when the direction of bending at the bottom of the domain changes for the downwelling on the left side (see Figure 4.4, second row). The analyzed temperature field of Method 1 does not predict this change of direction (see Figure 4.4, first row), while the analyzed temperature field of the ENKF predicts it (see Figure 4.4, third row). Method 1 computes only the evolution of the best estimate of the system. The computation of only one estimate ignores that, in this case, a slight perturbation of the estimated state could lead to a totally different dynamics. On the contrary, the EnKF method computes the evolution of an ensemble of perturbed solutions and thus takes into account the nonlinearity of the solution, at least for the forecast stage. Figure 4.5 shows examples of the analyzed temperature fields of different ensemble members for evolution B, after $80$ Myr of assimilation. Although the average temperature fields displays a downwelling bending to the right, the ensemble members show a wide variety of downwelling geometries.

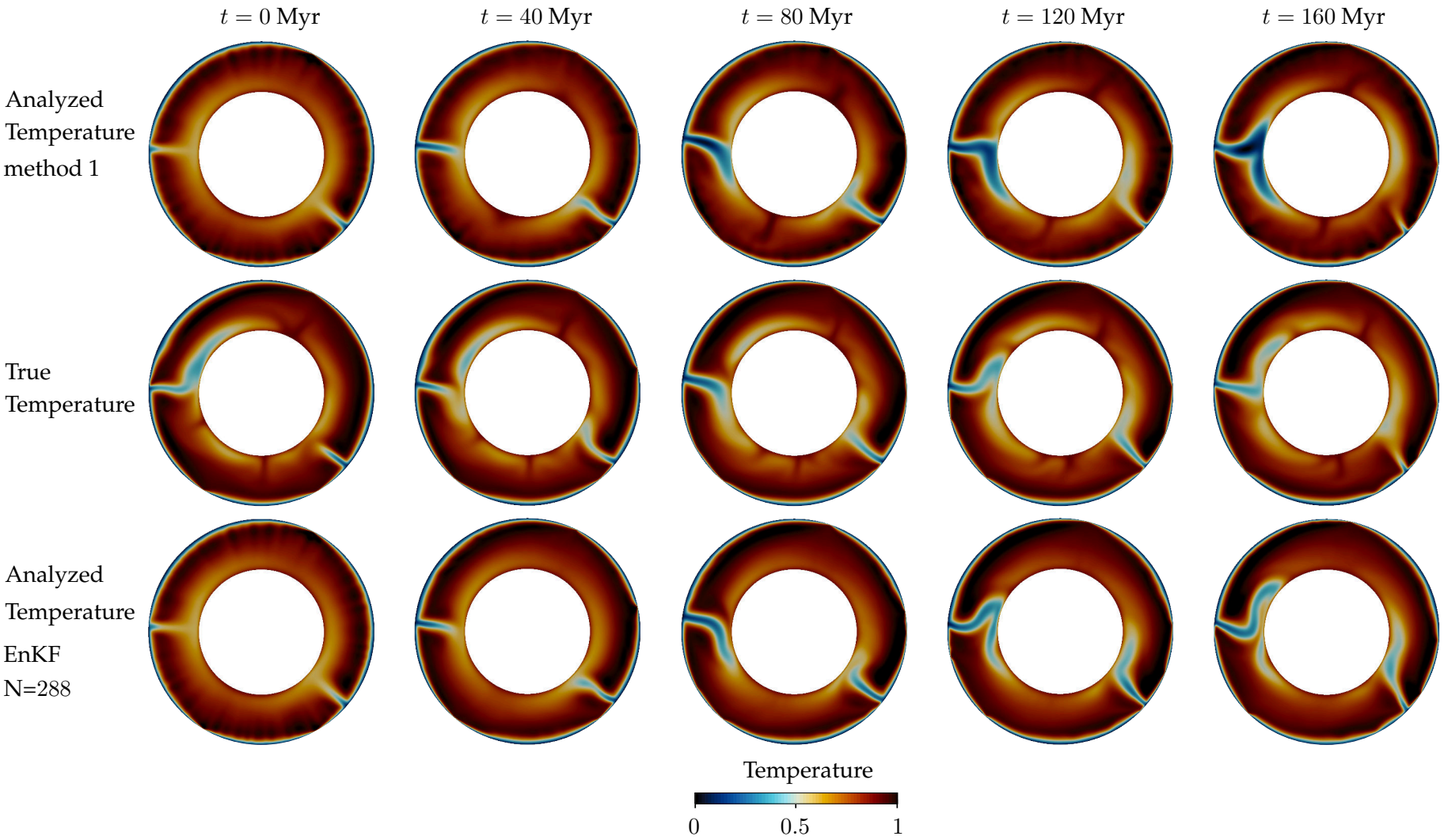


FIGURE 4.4: Comparison of temperature field evolutions for evolution B. The first row depicts the evolution of the analyzed temperature field with method 1. The second row is the true evolution of the temperature field. The third row is the evolution of the analyzed temperature field with ensemble Kalman filter, $N = 288$, $\ell_v = 0.7$ and $\ell_h = \pi/10$.

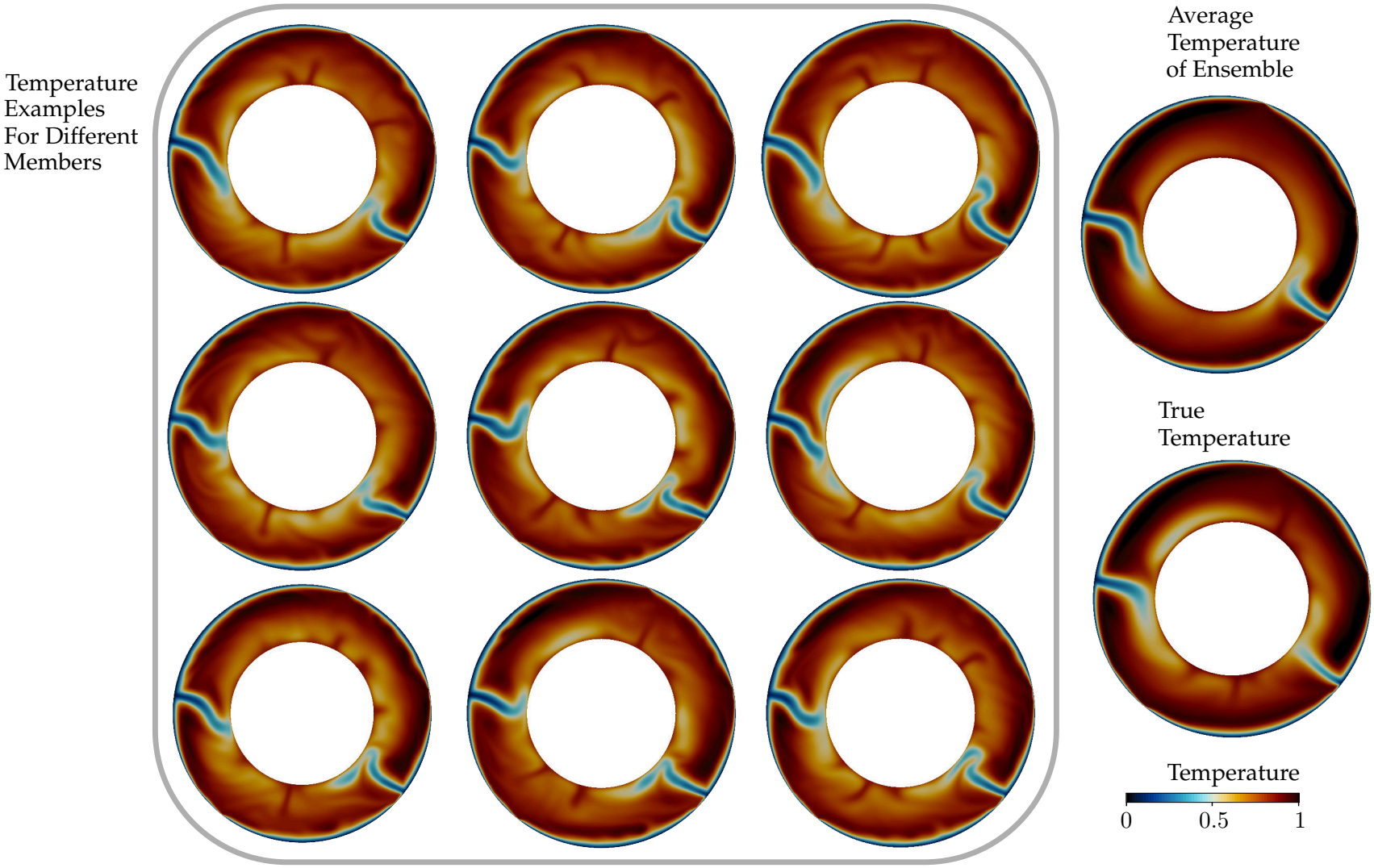


FIGURE 4.5: Example of temperature fields of the members of the ensemble. This example is taken after 80 Myr, for the assimilation of evolution B, with ENKF $N = 288$, $\ell_v = 0.7$ and $\ell_h = \pi/10$.

### 4.4.3 Effect of the data assimilation parameters on the quality of the estimation

As shown in Figure 4.1, the choice of $N$, $\ell_v$ and $\ell_h$ is critical to minimize errors in the assimilation, with errors on the estimated temperature field varying from $4\%$ to more than $10\%$ according to the choice of parameters. We investigate further the effect of these parameters by comparing the average global errors after analyses, $\overline{\epsilon}^a_T$, for different combinations of $N$, $\ell_v$ and $\ell_h$. Figure 4.6 displays the values of $\overline{\epsilon}^a_T$ for sizes of ensemble $N = 96$, $288$ and $768$ (Figures 4.6A, B and C respectively) with $\ell_v$ varying between $0.3$ and $1$, and $\ell_h$ between $\pi/10$ and $\pi/2$. As in Figure 4.1, we observe a dichotomy between assimilations with $N = 96$ members, with higher errors, and assimilations with $N = 288$ and $768$, with lower errors.

For each size of ensemble $N$ we identify the pair $(\ell_v, \ell_h)$ that leads to the assimilation with the lowest error $\overline{\epsilon}^a_{Tmin}(N)$. From this minimum value $\overline{\epsilon}^a_{Tmin}(N)$, we select all the pairs $(\ell_v, \ell_h)$ that lead to data assimilation with global errors less than $\overline{\epsilon}^a_{Tmin}(N) + 0.2\%$. As the size of the ensemble increases, the optimal lengths of correlations $(\ell_v, \ell_h)$ tend to increase. This is a classical effect (Houtekamer & Mitchell, 1998), observed in ensemble Kalman filters for various dynamical systems. As $N$ increases, the amplitude of noise in the sample correlation matrix $\mathbf{P}^f$ decreases, and small, yet real, correlations between distant points can be taken into account (Hamill et al., 2001). Between ensemble sizes of $N = 96$ and $N = 288$ the zone of optimal correlations is displaced towards the greater vertical correlation lengths. When we increase the size of the ensemble from $N = 288$ to $N = 768$, the zone of optimal correlations is displaced towards greater horizontal correlation angles. So the accurate estimation of correlations between points on the same vertical level needs less samples than between points on the same horizontal level. This is due to the specifics of mantle convection dynamics. The highly nonlinear rheology produces plates at the surface with values of velocity and temperature that may vary substantially (by one or two orders of magnitude) on short distances in the horizontal direction, especially because of pseudoplasticity. On the contrary, highly viscous cold downwellings establish a strong continuity in the vertical direction.

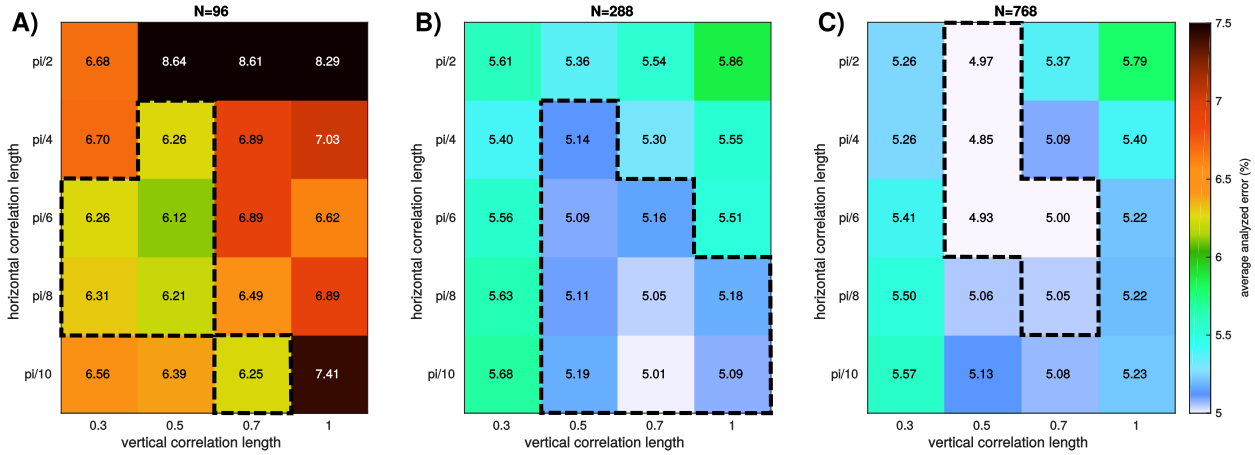


FIGURE 4.6: Value of the mean analyzed error for the same evolution and different vertical and horizontal correlation lengths. A) for 96 ensemble members, B) for 288 ensemble members, C) 768 ensemble members. The dashed lines delimit the zones for which errors are less than $\overline{\epsilon}^a_{Tmin}(N) + 0.2\%$.

For the ensemble size $N = 288$ and all the values of $(\ell_v, \ell_h)$, we additionally evaluate the average global ensemble spread

$$\overline{\sigma}^a_T = \frac{1}{K}\sum_{k=1}^{K} \sigma^a_T(k), \tag{4.35}$$

the average forecast error on data

$$\overline{\boldsymbol{\epsilon}}^f_y = \frac{1}{K}\sum_{k=1}^{K} \frac{d_k}{\parallel \mathbf{y}^o_k \parallel} \tag{4.36}$$

and the norm of the cumulative innovation for $K$ analyses:

$$\overline{\mathbf{d}}_K = \left\| \frac{1}{K}\sum_{k=1}^{K} \left(\mathbf{y}^o_k - \mathbf{H}\mathbf{x}^f_k\right) \right\|. \tag{4.37}$$

These three values are indicators of the accuracy of the assimilation and can be computed in the case of an assimilation with Earth data, unlike $\overline{\epsilon}^a_T$.

Figure 4.7 represents these results along with the true error $\overline{\epsilon}^a_T$. The ensemble of optimal data assimilation parameters is also outlined ($\overline{\epsilon}^a_T < \overline{\epsilon}^a_{Tmin}(N) + 0.2\%$).

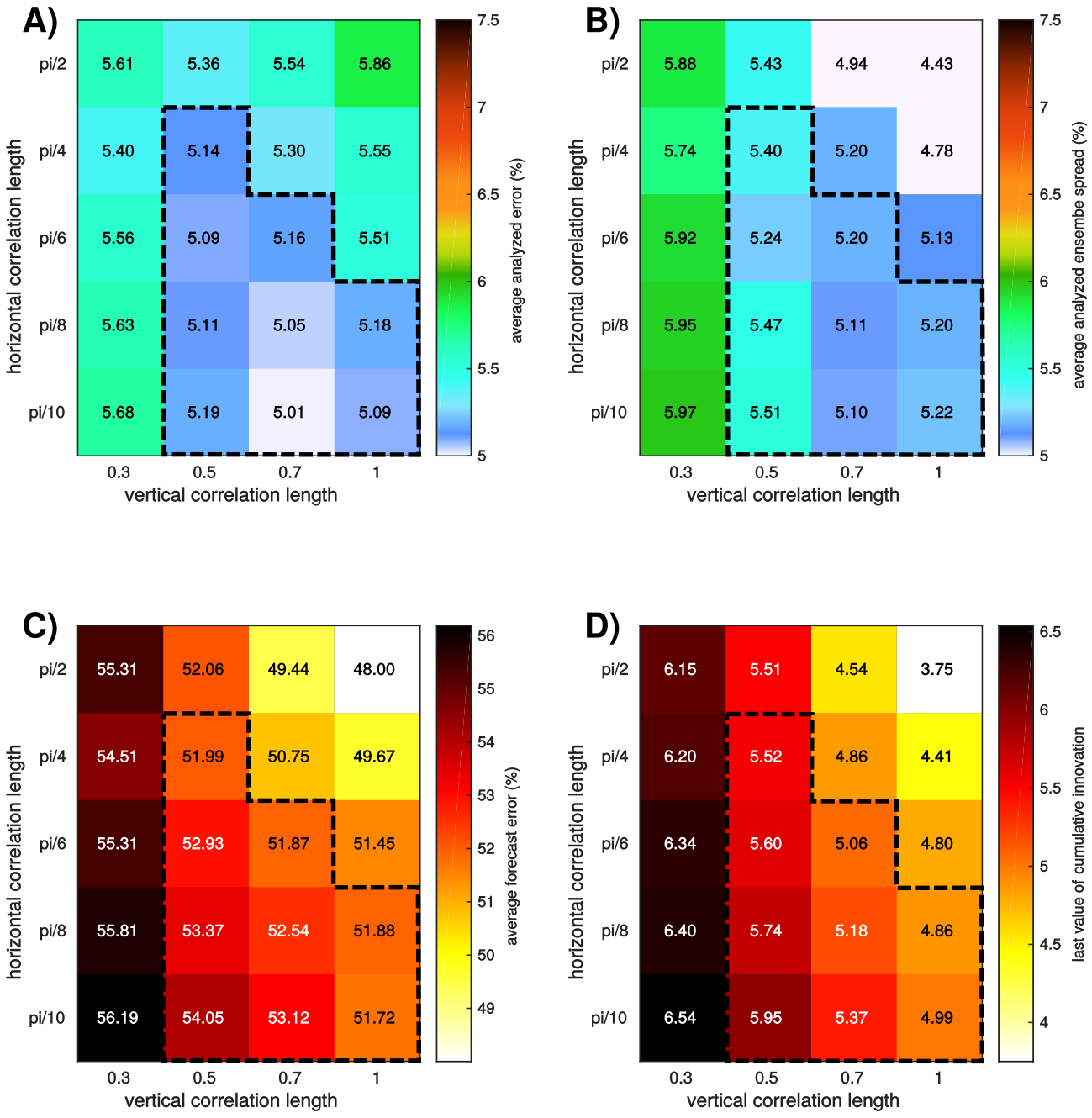


FIGURE 4.7: Value of A) mean analyzed error, B) mean ensemble spread, C) average forecast error on data, D) norm of cumulative innovation after $K$ analyses for $N = 288$, and different vertical and horizontal correlation lengths. The dashed line delimits the zone for which errors are less than $\bar{\epsilon}^a_{Tmin}(288) + 0.2\%$.

Overall, the average ensemble spread $\overline{\sigma}_T^a$ (Figure 4.7B) decreases when $\ell_h$ and $\ell_v$ increase, with a minimum for $\ell_h = \pi/2$ and $\ell_v = 1$. The higher the correlation lengths, the more covariances will be taken into account in the analysis, and the analyzed members will be closer to each others and $\overline{\sigma}_T^a$ lower. The average ensemble spread $\overline{\sigma}_T^a$ is of the same order of magnitude as the true error $\overline{\epsilon}_T^a$. Moreover, there is a local minimum of $\overline{\sigma}_T^a$ at $\ell_v = 0.7$ and $\ell_h = \pi/10$. These parameters correspond to the minimum true error $\overline{\epsilon}_T^a$.

The average forecast errors and the norm of the cumulative innovations display the same behavior: they decrease with increasing vertical and horizontal correlation lengths. The longer the correlations lengths, the closer the forecast data are to the observations, and the less biased the assimilation. This means that a better fit to the observations does not necessarily imply a better fit to the true temperature field. In a realistic context, the result of the assimilation should be checked against independent data to evaluate its accuracy. In the case of the Earth's mantle, independent data could be for example the geoid or tomographic models.

### 4.4.4 Accuracy of the Reconstruction of Geodynamic Structures

We focus on three key flow structures: 1) downwelling slabs (subduction) 2) ridges, i.e. shallow structures resulting from divergent plates at the surface, 3) plumes, hot upwellings raising from the base of the model.

Figure 4.8 shows the final state of the assimilation after $150$ My for the evolution A of Figure 4.3. We selected 3 assimilations: EnKF96, an Ensemble Kalman Filter with $N = 96$, $\ell_v = 0.5$ and $\ell_h = \pi/6$ (first row), EnKF288 an EnKF with $N = 288$, $\ell_v = 0.7$ and $\ell_h = \pi/10$ (second row) and the assimilation with method 1 (third row). We do not show the Ensemble Kalman Filter with $768$ members since the resulting temperature field is almost indistinguishable from that of EnKF288. The first column represents the true temperature field, which is the same for all assimilations. The second column is the analyzed temperature field, i.e. the average of the temperature fields of the analyzed

ensemble members. The third column is the absolute temperature error, and the fourth column is the standard deviation of the ensemble spread, which is an estimate of the error on the analyzed temperature field.

Globally, the EnKF288 and EnKF96 solutions for the temperature field are smoother than the solution of method 1. We observe this difference especially in the asthenosphere, the part of the mantle below the top boundary layer. For method 1, the asthenosphere shows short wavelength temperature variations. These variations are absent from the true temperature field and are inconsistent with convection solutions with the chosen parameters. They stem from the amplification of the noise in the observations during the analysis. Moreover, the asthenosphere of the analyzed temperature field of method 1 is hotter than the true temperature.

Both EnKF96 and EnKF288 reconstruct successfully the ridges locations and structures, as testified by their error fields. On the contrary, method 1 fails to reconstruct the ridge on the top right of the domain. It also predicts a ridge that does not exist in the true state (in the top left quadrant). On the right of the domain, another ridge is associated with a vertical positive temperature anomaly underneath. This pattern is found regularly under ridges when applying method 1. This is due to the use of a constant forecast error covariance matrix, $\mathbf{P}_0^f$ for the analysis. This constant matrix does not take into account the specifics of the dynamics under a ridge, where the positive anomaly is generally shallow. We do not observe this detrimental effect in the EnKF assimilations, where we compute the forecast error covariance matrix $\mathbf{P}_k^f$ at each analysis time from the forecast ensemble.

All three assimilations reconstruct the subductions and predict accurately the bending direction of slabs at the base of the model. Method 1 tends to underestimate the amplitude of the negative temperature anomalies whereas both EnKF assimilations overestimate them. This is especially noteworthy for the bottom left subduction. Moreover, the estimated slabs are wider than the true slabs. However, we note two arguments in favor of the EnKF: first, the estimation of the slab improves when the size of the ensemble increases and second, the local standard deviations of the ensemble indicates

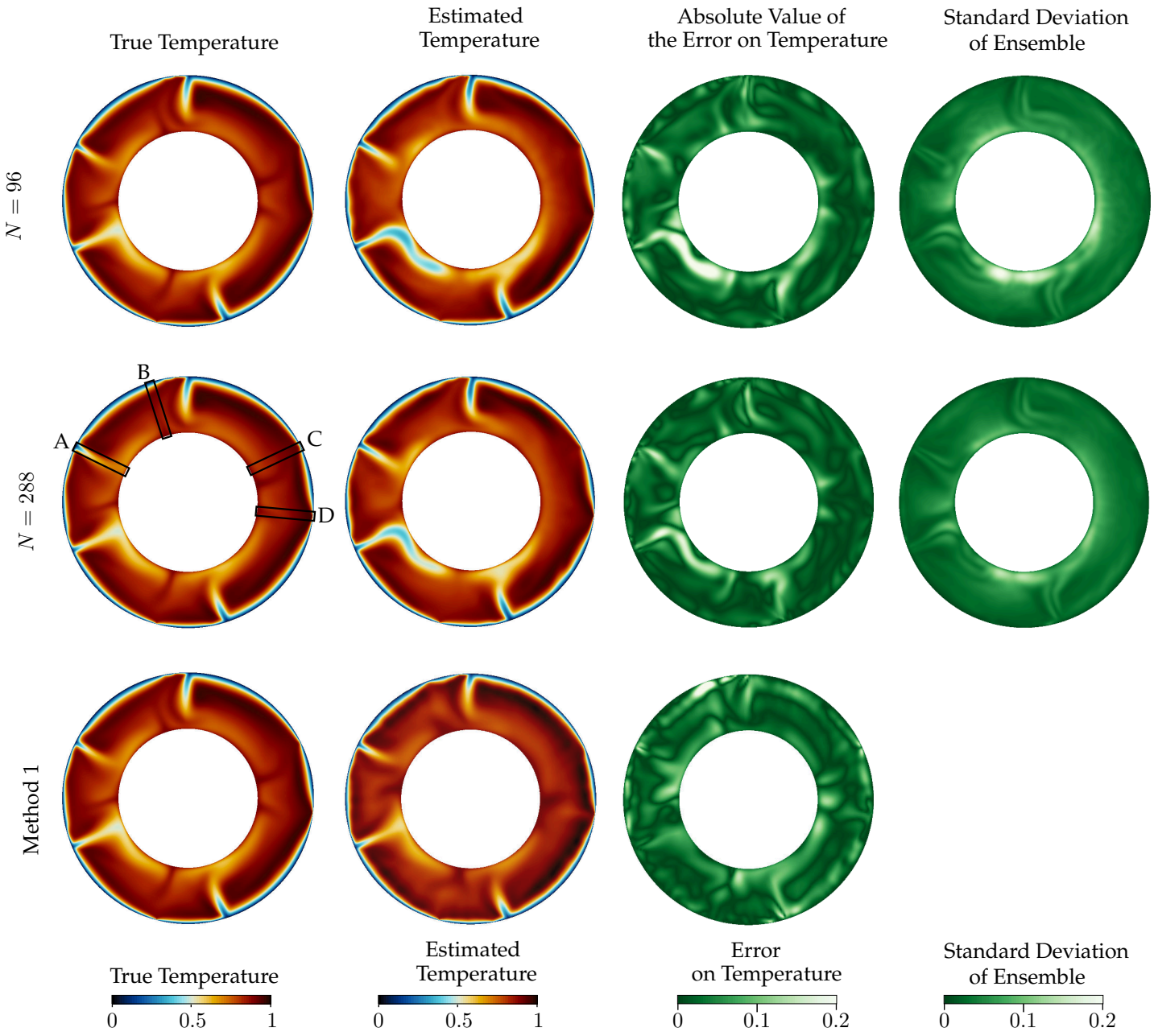


FIGURE 4.8: Comparison of estimated states after 150 Myr for the evolution A of Figure 4.3. First row: Ensemble Kalman Filter with $N = 96$, $\ell_v = 0.5$ and $\ell_h = \pi/6$; second row: $N = 288$, $\ell_v = 0.7$ and $\ell_h = \pi/10$; third row for method 1 (third row). The first column represents the true temperature field at 150 Myr, the second column is the analyzed temperature field, the third is the absolute error on temperature value and the fourth is the estimated error on the analyzed field (spread of the ensemble). On the true temperature field of EnKF288 we framed the location of the subduction (A), plume (B), ridge initiation (C) and stable ridge (D) studied in Figure 4.9.

that the estimation in this part of the domain is less accurate.

Both EnKF288 and EnKF96 solutions do not show any plume at the base of the mantle. However, the ensemble spread shows a greater uncertainty on the places where plumes occur. Method 1 predicts the approximate location of all plumes, but their geometry is not accurate. Method 1 provides only one estimate of the temperature field. In this evolution, the plumes are allowed to develop. EnKF96 and EnKF288 provide an ensemble of states. Each state develops plumes at different locations and their averages show only a slightly hotter anomaly over a wide area of possible location for the plumes, as we showed earlier in Figure 4.5 for another assimilation.

To illustrate how different flow structures are reconstructed, we represent on Figure 4.9 the evolution of temperature through time of the ensemble members of EnKF288 at points on the same vertical for a subduction, a plume, a ridge initiation and a stable ridge. Figure 4.8 shows the location of these geodynamical features on the true temperature field. We plot the temperature evolutions at the surface, mid-mantle and at the bottom of the domain. Note that the surface and bottom values of temperature actually correspond to the values of the first points below the surface and above the bottom of the domain, respectively.

At the surface, the temperature is corrected accurately at each analysis, with a difference between the true temperature and the analyzed temperature of less than 0.01. The correction associated with the analysis gradually decreases with depth due to both covariance localization and the dynamics of the system. After $70$ My (i.e. one transit time), the true value of temperature falls within the range of values predicted by the ensemble for all geodynamical contexts and all depths.

For the subduction, the correction is first done on the surface, and then propagates gradually in depth. The reconstruction of mid-mantle temperature becomes accurate after $40$ My, and at the bottom of the model after $70$ My, which is the value of the transit time. At the surface, the spread of the ensemble decreases as more data are assimilated. On the contrary, the spread of the ensemble remains steady for mid-mantle depths and at the bottom of the domain. For these depths, only the average temperature varies.

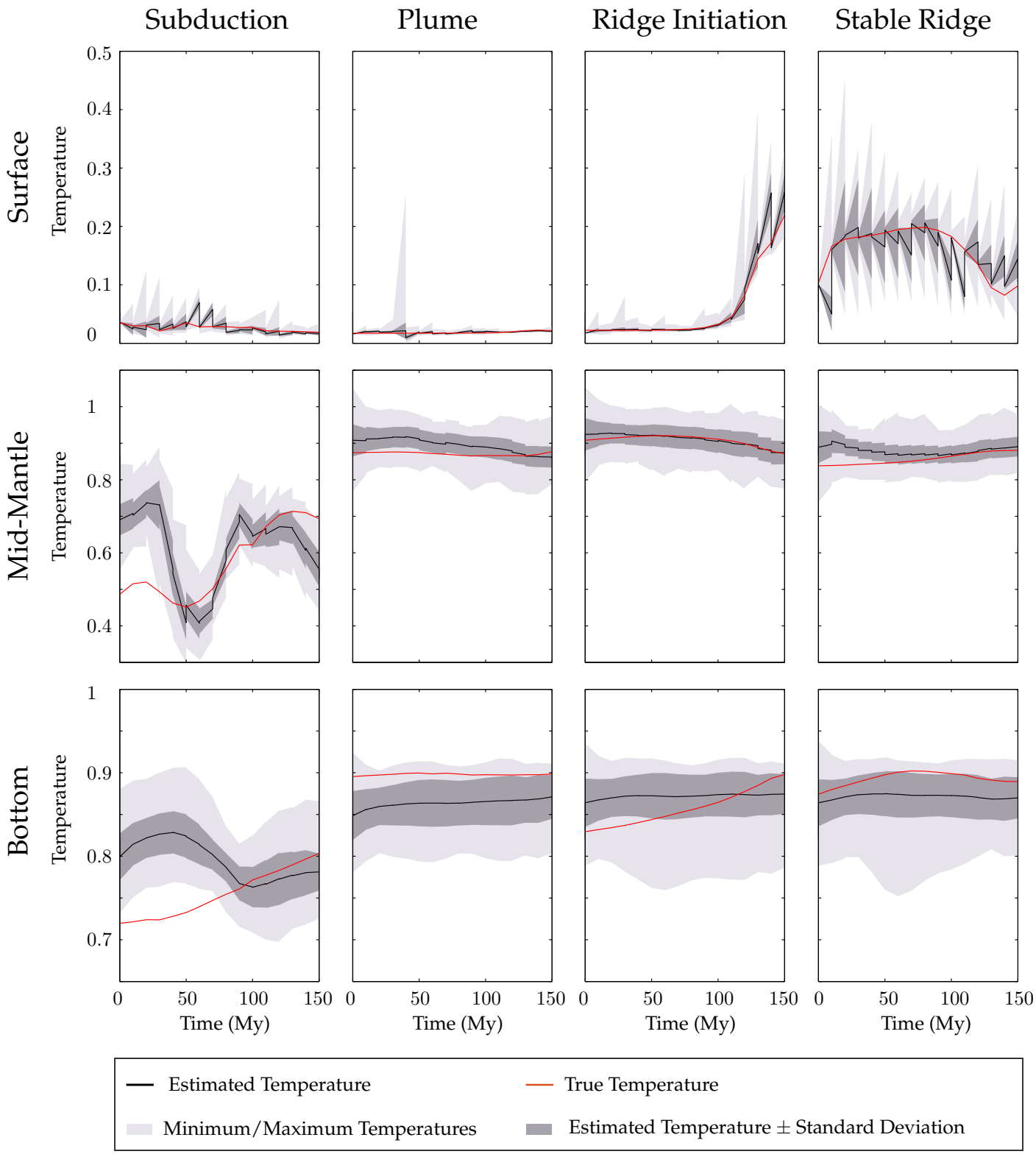


FIGURE 4.9: Evolution of the temperature value during assimilation at the surface (first row), in the mid mantle (second row) and at the bottom of the domain (third row), for a ridge initiation (first column), a subduction (second column), an upwelling (third column) and a stable ridge (fourth column).

At the surface for the plume, the spread of the ensemble is very low except for a peak at $40$ My, which corresponds to an instability, corrected after one analysis. We note that this instability affects greatly method 1 since it leads to the false prediction of the ridge seen in Figure 4.8. At mid-mantle, the ensemble average is slowly converging to the true temperature. At the bottom, the estimated temperature is lower than the true temperature, although it slightly increases throughout the assimilation.

The ridge initiation shows how new observations affect the spread of the ensemble. At the surface, the spread of the ensemble remains low until $100$ My, the time of initiation of the ridge. From then on, the estimated temperature increases and the ensemble members follow the cycle of increasing spread during forecast and dramatic decrease of spread during analysis. The temperature in the mid-mantle is estimated with a very good accuracy after $50$ My. On the contrary, the assimilation does not predict the evolution of the temperature at the bottom of the domain, although the true temperature falls within the zone defined by the standard deviation of the ensemble after $50$ My.

For the stable ridge, the spread of the ensemble at the surface is increasing during forecast and decreasing dramatically during the analysis. At mid-mantle, the estimated temperature becomes accurate after $100$ My. At the bottom of the domain the temperature is underestimated although it follows the variations of the true temperature: increase of temperature at the beginning of the assimilation and slight decrease at the end of the assimilation.

## 4.5 Discussion

We chose the Ensemble Kalman Filter method for its ease of implementation and flexibility to adapt to different forward numerical models. Indeed, as long as the nature of the state and observations does not change, the computation of the analysis step remains the same regardless of the convection code used. On the contrary, the alternative method, variational data assimilation, requires the development of an adjoint

code that needs further development for each additional complexity added to the forward model (see Kalnay et al., 2007, for a comparison of EnKF and 4D variational methods). For the mantle circulation problem, this results in a series of derivation of the adjoint considering different approximations (Ismail-Zadeh et al., 2003a; Bunge et al., 2003; Ghelichkhan & Bunge, 2016; Worthen et al., 2014). The ability of a data assimilation scheme to adapt to different numerical codes is a particularly important issue for mantle convection since models are in constant evolution, with current developments including the implementation of chemistry, nonlinear rheologies, elasticity, phase transition and compressibility (see e.g. Zhong et al., 2015, for a review of recent developments of mantle convection codes). In particular, this ease of implementation allows us to work on models producing self-consistently plate-like tectonics at their surface, and hence to obtain forecasts whose data can be ultimately compared with plate reconstructions.

The application of the Ensemble Kalman Filter to the mantle circulation problem is the continuation of the simpler sequential filter that we developed in an earlier work (Bocher et al., 2016). The main difference between the two filters is that the EnKF evaluates the state covariance matrix with an ensemble of members. This ensemble approach allows the nonlinear evolution of errors during the forecast stage. This leads to a higher precision in the reconstruction, but also to a more robust scheme, able to reconstruct evolutions which could not be reconstructed with the former method. Moreover, the ensemble assimilation provides an estimate of the errors on the reconstruction at each point of the domain. The estimation of errors could be a valuable information for plate tectonic reconstructions for regions and times where data are scarce, by showing the possible alternative scenarios supported by the ensemble.

This gain in information and quality for reconstructions comes with a price. While we could perform the former assimilation method in one core hour, the method developed here requires several hundreds to several thousands of core hours. However, an efficient parallelization using the PDAF software (Nerger & Hiller, 2013) in combination with the parallel code STAGYY produces a highly parallel ensemble filter, able to

perform the assimilations on $768$ cores in $20$ min for an ensemble of $96$ members and $3$ hours for an ensemble of $768$ members.

The important computational cost of the EnKF limited us in the number of assimilations we could test. After checking the stability of the assimilation results on four different evolutions, we chose to focus on studying the effect of the parameters of the Ensemble data assimilation: the size of the ensemble and the vertical and horizontal correlation lengths. The optimum size of the ensemble for our problem is of the order of $288$ members. Indeed, almost tripling the number of members leads to a decrease of the average error of less than $0.2\%$, and on the contrary, dividing the size of the ensemble by $3$ leads to an increase of the average error of more than $1\%$. Although these differences in errors appear to be small, they affect the quality of the reconstruction of thermal structures, as is illustrated in figure 4.8. The average errors on the temperature field for the estimates shown in figure 4.8 range between $4.8$ and $5.8\%$, so the difference in errors are less than $1\%$. Locally, this translates into the presence (or absence) of geodynamic structures (like ridges and upwellings) which are artefacts. Covariance localization proved to be important to minimize the error in the reconstruction of mantle structure: for $288$ members, the difference in the average error is of $0.85\%$ between the optimal correlation length and the least favorable one. We also investigated the statistics of the cumulative innovation and of the average forecast error for different ensemble sizes and correlation lengths. In a realistic case, these are the only variables available to evaluate the quality of the assimilation. The variation of both cumulative innovation and average forecast error as a function of ensemble size show the same tendency as the average error on the temperature field. On the contrary, the correlation lengths minimizing the norm of the cumulative innovation and the average forecast error were different from the ones minimizing the error on the temperature field. This shows the limits of these indicators to determine the optimal parameters for the assimilation. In a realistic case, rigorous a posteriori evaluation of a data assimilation result would require comparison of the prediction made with independent observations (Talagrand, 2014). For mantle circulation, seismic tomography, topography, true polar wander or the geoid could play this role.

By construction, sequential data assimilation methods do not propagate new information back in time. In the case of the reconstruction of mantle circulation, this is a clear disadvantage since the information on the Earth's surface tectonics tends to become more reliable as we get closer to present-day. Consequently, a natural extension of the present work would be to implement an Ensemble Kalman Smoother (Evensen & Van Leeuwen, 2000; Van Leeuwen, 2001). In the same way as the EnKF uses sample spatial correlations of the ensemble to update the state of the system with new observations, the Ensemble Kalman smoother uses sample time and space correlations of successive ensembles to update former states with the new observations. Evensen (2003) shows how an Ensemble Kalman Smoother can be implemented with a minimal computational cost alongside a preexisting EnKF. Moreover, Nerger et al. (2014) shows that such algorithm is efficient for nonlinear models, and that in their test case, optimal localization parameters for the Ensemble Kalman Smoother coincide with optimal localization parameters for the EnKF.

As a first approach to test the EnKF for mantle circulation reconstructions, we chose a fairly simple convection model. As already discussed in Section 4.2.1, a more realistic mantle model would have, among other things, a 3D-spherical shell geometry and a higher Rayleigh number. This would substantially increase the size of the data assimilation problem. However, we followed the procedure as described in Nerger & Hiller (2013) to implement the EnKF. This results in a highly scalable filter, enabling the computation of the EnKF assimilation in a reasonable time. An increase in the Rayleigh number also implies thiner boundary layers, slabs and plumes. This could translate into lower optimum correlation lengths for the EnKF. A more realistic model would additionally include a viscosity increase in the lower mantle (Ricard et al., 1993), and the presence of continents. This would tend to lengthen the wavelength of convection in the lower mantle and therefore might ease the mantle circulation reconstruction (see for example Ricard, 2015, sections 7.02.6.3.2 and 7.02.6.7 for a discussion of both effects on mantle convection).

In the synthetic experiments of section 4.4, the convection model used to produce the

series of data is the same as the forward model used during the assimilation. For an application with Earth data, this will not be the case. The equations solved in models of mantle convection still hold some shortcomings (Ricard, 2015). Moreover, theories, observations and experiments do not yet fully constrain parameters, especially rheological ones (King, 2016), and variations in rheology affect the reconstructions of mantle circulation (Bello et al., 2015). Hence it could be fundamental to take into account model errors. A first order solution is to increase the inflation parameter $\gamma$ in Equation 4.20: this would overall increase the a priori uncertainty on the mantle estimation. Performing experiments where the model used to compute the observation is different from the model used for the assimilation would provide us with more information on how to implement model errors. Another solution would be to consider the joint assimilation of the state and model parameters. Although it is in principle possible for the EnKF (Evensen, 2009b), it could be computationally not tractable. Indeed, the response of mantle dynamics to different rheological parametrization is highly nonlinear, and their inversion calls for the development of techniques focusing on rheology, such as adjoint based inversions of rheological parameters (Worthen et al., 2014; Ratnaswamy et al., 2015) or further applications of the recently developed pattern recognition techniques for mantle convection (Atkins et al., 2016).

The choice of the synthetic experiments assimilation window of $150$ Myr is a compromise between having the possibility to compute assimilations for various cases and having an assimilation window covering most part of the timespan of plate tectonic reconstructions (Seton et al., 2012; Müller et al., 2016; Torsvik et al., 2010). A real assimilation could take into account a longer timespan and therefore improve the assimilation results. However, the structure of the dataset used for the synthetic experiments is a very idealized version of the actual plate reconstruction models. We already discuss this issue in Bocher et al. (2016). In the following, we supplement and update this discussion in the light of research that has recently come to the fore.

First, we set a time series of data covering the whole surface of the domain and regularly available, every $10$ Myr. Plate tectonic reconstructions data are more complex.

They are based on the estimation of finite relative rotations between individual plates, structured into a hierarchy describing global relative motions and anchored in an absolute reference frame. The span of each finite relative rotation is determined depending on the amount and quality of information available for a specific context and therefore varies depending on plate pairs and times. The average span of finite rotations of recent plate models is of the order of $10$ Myr (Torsvik et al., 2010) to $5$ Myr (Müller et al., 2016), but varies over time with for example $1$ Myr resolution for the last $20$ Myr in some regions (Merkouriev & DeMets, 2014), or some gaps in the data such as during the cretaceous superchron from $121$ to $83$ Myr ago (Granot et al., 2012). The continuously closed plate algorithm (Gurnis et al., 2012) produces plate tectonic reconstruction maps continuous in space and time which allows the creation of a series of global plate reconstructions at regular intervals. Nonetheless, creating such a regularized time series of reconstruction might miss tectonic events. Instead, we could adapt the frequency of analyses to the varying plate reconstruction resolution. Additional synthetic experiments with a time-series of data whose frequency evolves through time are necessary to explore the limits of such method.

Second, the observations were perturbed independently with a Gaussian noise of $10\%$ of the respective root mean square value of surface heat flux and surface velocities. The estimation of uncertainties on absolute plate motion models involves estimation of both uncertainties in relative plate motion and on the absolute reference frame (Müller & Wessel, 2015). The main source of information on the motion of plates comes from the map of seafloor magnetic anomalies. Hellinger (1981) developed a method to compute relative motion of plates and associated uncertainties inferred from magnetic anomaly identifications. Recently, Seton et al. (2014) built an open source community database. It gathers seafloor magnetic anomaly identifications, and estimation with Hellinger (1981) method of plates relative motion and associated uncertainties. This database could be used in the future as a basis to automatically produce global plate motion histories and assess their uncertainties. To our knowledge, this has not been done so far at a global scale. On a regional scale and for recent time ($5$ to $20$ Myr), Iaffaldano et al. (2012) applied the trans-dimensional hierarchical Bayesian method to

reduce noise in finite rotation data and produce time series of high resolution plate relative motions. More recently, Iaffaldano & Bunge (2015) applied this technique to the relative motion of the pacific plate with North America for the last $75$ Myr. The uncertainties on relative plates velocities ranges from $5$ to $40\%$ of the root mean square surface velocity. As we go further back in time, the quantification of relative plate motion uncertainties becomes hazardous: most of the seafloor created before $150$ Myr has been destroyed by subduction. These plate tectonic reconstructions involve interpretation of different types of data, with a limited spatial coverage and relies heavily on human expertise. For these epochs, maintaining very high uncertainties on the regions where few data supports the reconstructions would be a solution.

## 4.6 Conclusion

We applied the Ensemble Kalman Filter algorithm to the reconstruction of mantle circulation through time. We chose a formulation with covariance inflation and localization to minimize the effect of sampling errors in the estimation of the forecast error covariance matrix. Synthetic "twin" experiments with different evolutions and for different parameters allowed us to assess the efficiency of the algorithm and to determine the optimal parameters for the assimilation.

This work builds on the developments of a first approach to sequential data assimilation for mantle circulation made in Bocher et al. (2016). The EnKF is more robust and on average more accurate than the former method. Additionally, the Ensemble Kalman Filter provides not only an estimate of mantle circulation, but also detailed maps of uncertainties on this estimation.

We evaluate the accuracy of the EnKF as a function of three main parameters: the size of the ensemble, and two covariance localization parameters, namely the vertical correlation length and horizontal correlation angle. We find that the optimal size of the ensemble is of the order of $300$ members. For this ensemble size, the optimal vertical

correlation length corresponds to two thirds of the domain thickness, and the optimal horizontal correlation angle is of $\pi/10$ (around $2000$ km). These values should be reevaluated as the dynamical model becomes more realistic.

The EnKF was implemented using the parallel data assimilation framework PDAF in a preexisting mantle convection code, STAGYY. The resulting code is highly scalable, which means that the application of the EnKF to realistic data assimilation with plate reconstructions and a 3D spherical mantle model is within reach in a foreseeable future.

## acknowledgements

The research leading to these results has received funding from the European Research Council within the framework of the SP2-Ideas Program ERC-2013-CoG, under ERC grant agreement 617588. Calculations were performed using HPC resources from GENCI-IDRIS (grant 2014-047243). The contribution of Alexandre Fournier is IPGP contribution number XXX.

## Résumé du chapitre

Les récentes avancées en modélisation de la convection mantellique ont conduit à l'apparition d'une nouvelle génération de codes de convection, capables de générer de façon auto-cohérente des tectoniques de type plaque à leur surface. Ces modèles relient physiquement la dynamique du manteau à la tectonique de surface. Associés à des reconstructions de la tectonique des plaques, ils ont le potentiel de produire une nouvelle génération de modèles de circulation mantellique qui utilisent des méthodes d'assimilation de données et tiennent compte des incertitudes sur les reconstructions de la tectonique des plaques. Nous avons récemment fourni une preuve de ce concept en appliquant un filtre de Kalman sous-optimal à la reconstruction de la circulation mantellique (Bocher et al., 2016). Ici, nous proposons d'aller un plus loin et d'appliquer le filtre de Kalman d'ensemble (EnKF) à ce problème. L'EnKF est une méthode de Monte Carlo séquentielle particulièrement adaptée pour résoudre des problèmes d'assimilation de données de grande dimension avec une dynamique non linéaire. Nous avons testé l'EnKF en utilisant des observations synthétiques comprenant des données de vitesse de surface et de flux de chaleur, sur un modèle d'anneau sphérique 2D et l'avons comparée avec la méthode développée précédemment. L'EnKF fonctionne en moyenne mieux et est plus stable que l'ancienne méthode. Moins de 300 membres d'ensemble sont suffisants pour reconstituer une évolution. Nous utilisons l'inflation et la localisation adaptative sur la matrice de covariance pour corriger les erreurs d'échantillonnage. Nous montrons que les résultats de l'EnKF sont robustes sur une large gamme de paramètres de localisation de la matrice de covariance. La reconstruction est associée à une estimation de l'erreur, et fournit des informations précieuses sur les endroits où la reconstruction est fiable.

# CHAPTER 5

# Discussion and Conclusion

During this thesis, we successively proposed two algorithms to reconstruct mantle circulation by combining observations on surface tectonics with a mantle convection model, using sequential data assimilation methods.

We first developed a suboptimal sequential data assimilation scheme based on the Kalman filter (Chapter 3). This scheme is efficient in recovering the temperature field of a convective system with plate-like tectonics at its surface over several $100$ Myr. We tested the robustness of the method by conducting synthetic experiments in 2D spherical annulus geometry. We obtained accurate results, even for the location of plumes at the base of the models, for periods of at least $300$ Myr, provided the time between analyses is shorter than $15$ Myr and the noise in observation is lower than $30\%$.

These encouraging results motivated us to further explore the data assimilation problem of reconstructing mantle circulation from surface tectonics (Chapter 4). The Ensemble Kalman Filter seemed to be a method of choice given its ability to deal with high dimensional problems and nonlinear models. It indeed proved to be more robust

and on average more accurate than the former method, while additionally providing detailed maps of uncertainties on the estimated evolution.

The synthetic experiments we performed in both studies are by no means exhaustive, and many more tests remain to be done to thoroughly evaluate data assimilation methods in the context of mantle circulation reconstruction. Among them, we identify the need for:

- further twin experiments using a more realistic setting:
  - from the dynamical model side, this means, among other things, tests in 3D spherical geometry, the introduction of continents, a higher Rayleigh number, a viscosity jump at mid-mantle, but also taking into account the chemical aspects of mantle convection (see for example Ricard, 2015, section 6 for an overview of the specificities of mantle convection).
  - from the observed data side, this means using an observation error covariance matrix which takes into account the correlations between errors on adjacent points, and the correlation between errors on observed surface velocities and observed surface heat fluxes. For models developing small-scale convection, It will also be necessary to use directly seafloor age maps instead of surface heat fluxes as observed data.
- the design of "sister experiments", to assess the effect of errors in mantle convection models. The principle of these sister experiments would be to compute a reference evolution with a given set of model parameters, and then to apply the data assimilation algorithm using a different set of parameters. This type of experiment will inform us on the impact an error in the forward model will have in the estimation of mantle circulation. This information could be of prior importance to succeed in the application of these data assimilation algorithms to Earth data.

Of course, the ultimate test for a data assimilation scheme is its application to Earth

data. In chapter 4, we developed a data assimilation code for mantle circulation reconstruction by combining a parallel ensemble data assimilation framework (PDAF Nerger et al., 2005; Nerger & Hiller, 2013) with the state-of-the-art parallel mantle convection code STAGYY (see Tackley, 2008). The high scalability of the resulting code, and the modelling capabilities of STAGYY (see Section 1.5) make the application of data assimilation to real data technically possible. The computation of a mantle circulation reconstruction will further require:

- a time series of the layout of plates and associated plate motions. This time series is easily accessible thanks to the open-access policy of teams working on plate tectonics reconstructions (e.g. Müller et al., 2016).

- a comprehensive assessment of the errors on plate tectonics reconstructions, taking into account the errors in both relative and absolute plate motions, and consistent with the formalism used in data assimilation. This task can only be achieved by a tight collaboration with teams working on plate tectonics reconstructions. As highlighted in section 2.2.2, tools have been developed to estimate the uncertainties on relative and absolute plate motions (see e.g. Wessel & Müller, 2015 for a recent review on these techniques). Hence, a first order estimation of uncertainties in plate tectonics reconstructions is possible, even though, as we go back in time, this estimation might not reflect true uncertainties on plate tectonics reconstructions (see discussion at the end of Section 2.2.2).

- choices in the parametrization for the mantle convection model. This choice of parameters will be a compromise between obtaining an internal dynamics and a surface tectonics as close to the Earth as possible, and the limits in computational power. In section 1.4, we argued for the use of a composite Newtonian and pseudoplastic rheology to obtain a mantle convection model self-consistently producing plate-like tectonics at its surface. This type of models has been investigated in pioneering studies using 3D spherical geometry (e.g. Richards et al., 2001; Walzer & Hendel, 2008; Van Heck & Tackley, 2008; Yoshida, 2008; Foley & Becker, 2009). Further work has also involved the statistical comparison of

the surface dynamics of this type of models with plate tectonics reconstructions (Coltice et al., 2012, 2013; Rolf et al., 2014; Mallard et al., 2016). Finally, they have recently been used to reconstruct mantle circulation using the semi-empirical sequential method (Bello et al., 2015; Coltice, in prep). These studies provide a solid basis to design a mantle convection model for data assimilation applications.

To evaluate the result of such mantle circulation reconstructions, we will need to test it against independent data. We propose two types of tests:

1. comparison of mantle circulation reconstructions against data independent from plate tectonics reconstructions. At present, these include seismic tomography (see Section 2.2.1), dynamic topography and the geoid (see e.g. Steinberger, 2016), and measurement of heat fluxes at the surface (Jaupart & Mareschal, 2015). In the past, this could include sea level changes (Müller et al., 2008b) and evidences of vertical motions (Flament et al., 2013).

2. Another possible test is to perform data assimilation over a restricted timespan to obtain an estimation of the state of the mantle at some time in the past. Then, we can compute a free evolution (i.e. unconstrained by data) of this estimated state up to the present day. By comparing the free evolution with observations, we can assess the predicting power of mantle circulation reconstructions.

Section 2.3 gives a partial overview of the possible applications of mantle circulation reconstructions: studies on the effect of surface tectonics on the deep mantle structure (e.g. McNamara & Zhong, 2005; Schuberth et al., 2009b,a; Zhang & Zhong, 2011; Davies et al., 2012; Bower et al., 2013; Bull et al., 2014; Hassan et al., 2015), and investigation of alternative scenarios of absolute plate motions (e.g. O'Neill et al., 2005; Doubrovine et al., 2012; Shephard et al., 2012; Nerlich et al., 2016; Zahirovic et al., in press). Other applications include studies of changes in the Earth's rotation axis (Steinberger & O'Connell, 1997), the sea-level (Moucha et al., 2008), dynamic topography (Flament et al., 2013), and the effect of mantle circulation on core dynamics (Olson et al., 2015).

Using data assimilation allows the integration of tectonic data into convection models in a more consistent way, since the convection model we use generates naturally plate-like tectonics at its surface, and plate tectonics reconstructions are not enforced through kinematic boundary conditions, but taken into account in a statistically consistent fashion.

Compared to other disciplines, data assimilation for mantle convection is still in its infancy. Knowledge, experience and technical advances of data assimilation in other fields, especially atmospheric sciences and oceanography, provide an extensive amount of resources to design algorithms for the mantle circulation reconstruction problem. A direct extension of our work would be for example to develop an ensemble Kalman smoother algorithm (Evensen & Van Leeuwen, 2000; Van Leeuwen, 2001) to better estimate past mantle circulation using data on more recent states of the mantle. New methods are also being developed to take better into account the nonlinearities in dynamical models and the resulting nongaussianity in the statistics of errors (Bocquet et al., 2010). Such solutions should be considered to build and improve data assimilation algorithms for mantle circulation.

We started this dissertation by remarking the coexistence of two different theories that set the framework for understanding the dynamics of the Earth's mantle and crust system: plate tectonics and mantle convection. The successive data assimilation developments that we proposed are an attempt to combine the products of these two theories: plate tectonics reconstructions on one hand and mantle convection codes on the other hand. The continuous progress of mantle convection codes in modelling Earth's inner dynamics and surface tectonics suggest a more ambitious, long-term goal: the raw observation on surface tectonics (the mapping of magnetic anomalies on the ocean floor, the observation and dating of tectonic features, paleomagnetic data...) could be integrated directly into mantle convection models, without using the theoretical framework of plate tectonics.

The revolution of Earth sciences which took place in the 60's (see Section 2.2.2), with the emergence of plate tectonics theory, was not merely a change in paradigm. It required

the Earth science community to gather and connect observations coming from various disciplines to build a unifying global theory of Earth's surface motions. In doing so, it set a new standard for the functioning of solid Earth Science. It brought the Earth science community to consider geological processes as a multidisciplinary and multi-scale problem: to understand the global history of the solid Earth, we need to navigate between micro, macro, local, regional and global scales and reconcile observations of different nature, coming from various disciplines, in a consistent framework. Plate tectonics theory provided such global framework and proved to be accurate at first order in its description of surface dynamics. However, plate tectonics is, in essence, a kinematic theory: it describes the motions at the surface of the Earth, but not their cause. Moreover, it is an idealized model of surface motions: plates are not perfectly rigid (e.g. Gordon, 1998), and plate boundaries are not infinitely small but can be on the contrary diffuse, in the oceans (Gordon, 2000; Wiens et al., 1985) as well as in the continents, as the presence of massive mountain ranges attest. Mantle convection theory provides, on the contrary, a global physical framework which takes into account those diffuse and continuous deformations, and describes both the solid Earth inner dynamics and surface tectonics. Many challenges remain before we can directly integrate raw observations in mantle convection models to produce a physically consistent reconstruction of both the Earth's mantle and surface tectonics. Undoubtedly, data assimilation methods will be one of the pieces of this puzzle.

# Discussion et Conclusion

Au cours de cette thèse, nous avons successivement proposé deux algorithmes pour reconstruire la circulation mantellique en combinant les observations sur la tectonique de surface avec un modèle de convection mantellique, en utilisant des méthodes d'assimilation de données séquentielle.

Nous avons d'abord développé un schéma d'assimilation séquentielle sous-optimal basé sur le filtre de Kalman (Chapitre 3). Ce schéma est efficace pour reconstituer le champ de température d'un système convectif avec une tectonique de type plaque à sa surface sur plusieurs $100$ Myr. Nous avons testé la robustesse de la méthode en effectuant des expériences synthétiques en géométrie sphérique 2D. Nous avons obtenu des reconstitutions fidèles, même pour l'emplacement des panaches à la base des modèles, après des périodes d'assimilation d'au moins $300$ Myr, à condition que le temps entre les analyses soit inférieur à $15$ Myr et que l'amplitude du bruit dans les observations soit inférieure à $30\%$.

Ces résultats encourageants nous ont poussés à explorer davantage le problème de l'assimilation des données pour la reconstitution de la circulation mantellique à partir de la tectonique de surface (Chapitre 4). Le filtre de Kalman d'Ensemble semblait être

une méthode de choix compte tenu de sa capacité à traiter des problèmes aux dimensions élevées et des modèles non linéaires. Il s'est en effet révélé plus robuste et en moyenne plus précis que l'ancienne méthode, tout en fournissant en plus des cartes détaillées des incertitudes sur l'évolution estimée.

Les expériences de synthèse que nous avons réalisées dans les deux études ne sont nullement exhaustives et de nombreux autres tests restent à faire pour évaluer soigneusement les méthodes d'assimilation des données dans le contexte de la reconstitution de la circulation mantellique. Parmi eux, nous identifions la nécessité:

- d'expériences jumelles supplémentaires utilisant des paramètres plus réalistes:
  - du côté du modèle dynamique, cela signifie, entre autres, des tests en géométrie sphérique 3D, l'introduction de continents, un nombre de Rayleigh plus élevé, un saut de viscosité au milieu du manteau, mais également prendre en compte les aspects chimiques de la convection mantellique (voir par exemple Ricard, 2015, section 6 pour un aperçu des spécificités de la convection mantellique).
  - du côté des données observées, cela signifie utiliser une matrice de covariance des erreurs sur les données observées qui prend en compte les corrélations entre les erreurs sur des points adjacents et la corrélation entre les erreurs sur les vitesses de surface observées et les flux de chaleur de surface observés. Pour les modèles développant une convection à petite échelle, Il sera également nécessaire d'utiliser directement les cartes de l'âge des fonds marins au lieu des flux de chaleur de surface comme données observées.
- la conception d'«expériences sœurs», pour évaluer l'effet des erreurs dans les modèles de convection mantellique. Le principe de ces expériences sœurs serait de calculer une évolution de référence avec un ensemble de paramètres physiques donnés, puis d'appliquer l'algorithme d'assimilation de données en utilisant un ensemble différent de paramètres.

Bien sûr, le test ultime pour un schéma d'assimilation de données est son application

aux données réelles. Dans le chapitre 4, nous avons développé un code d'assimilation de données pour la reconstitution de la circulation mantellique en combinant un code d'assimilation de données d'ensemble parallélisé (PDAF Nerger et al., 2005; Nerger & Hiller, 2013) avec le code parallèle de convection STAGYY (voir Tackley, 2008). La grande scalabilité du code obtenu et les capacités de modélisation de STAGYY (voir section 1.5) rendent techniquement possible l'application de l'assimilation de données à des données terrestres. Le calcul d'une reconstitution de la circulation mantellique exige en outre:

- une série chronologique de la disposition des plaques et de leurs mouvements. Cette série chronologique est facilement accessible grâce à la politique d'accès libre des équipes travaillant sur les reconstitutions de tectonique des plaques (par exemple Müller et al., 2016).

- une évaluation complète des erreurs sur les reconstitutions de la tectonique des plaques, en tenant compte des erreurs dans les mouvements de plaques relatifs et absolus, et en cohérence avec le formalisme utilisé dans l'assimilation des données. Cette tâche ne peut être réalisée que par une étroite collaboration avec les équipes travaillant sur les reconstitutions de la tectonique des plaques. Comme on l'a souligné dans la section 2.2.2, des outils ont été développés pour estimer les incertitudes sur les mouvements de plaques relatifs et absolus (voir, par exemple, Wessel & Müller, 2015 pour une revue récente de ces techniques). Par conséquent, une estimation au premier ordre des erreurs sur les reconstitutions de la tectonique des plaques est possible, même si cette estimation pourrait ne pas refléter les incertitudes véritables sur les reconstitutions, en particulier pour les temps très anciens (voir discussion à la fin de la section 2.2.2).

- des choix dans la paramétrisation pour le modèle de convection mantellique. Ce choix de paramètres constituera un compromis entre d'une part l'obtention d'une dynamique interne et d'une tectonique de surface aussi proche que possible de la Terre et d'autre part les limites en termes de puissance de calcul disponible. Dans la section 1.4, nous avons plaidé en faveur de l'utilisation d'une rhéologie

composite newtonienne et pseudoplastique pour obtenir un modèle de convection mantellique produisant de manière auto-cohérente une tectonique de type plaque à sa surface. Ce type de modèles a été étudié dans des études pionnières utilisant la géométrie sphérique 3D (par exemple Richards et al., 2001; Walzer & Hendel, 2008; Van Heck & Tackley, 2008; Yoshida, 2008; Foley & Becker, 2009). Des travaux ultérieurs ont également porté sur la comparaison statistique de la dynamique de surface de ce type de modèles avec des reconstitutions de tectonique des plaques (Coltice et al., 2012, 2013; Rolf et al., 2014; Mallard et al., 2016). Enfin, ils ont été récemment utilisés pour reconstruire la circulation mantellique en utilisant la méthode séquentielle semi-empirique (Bello et al., 2015; Coltice, in prep). Ces études fournissent une base solide pour concevoir un modèle de convection mantellique adapté à l'application d'algorithmes d'assimilation de données.

Pour évaluer les reconstitutions de la circulation mantellique obtenues par assimilation de données, nous devrons les comparer avec des données indépendantes. Nous proposons deux types de tests:

1. la comparaison des reconstitutions de la circulation mantellique avec des données indépendantes des reconstitutions de la tectonique des plaques. Pour la circulation mantellique actuelle, on peut envisager par exemple les modèles de tomographie sismique (voir section 2.2.1), la topographie dynamique et le géoïde (voir, par exemple Steinberger, 2016), et les mesures de flux de chaleur à la surface (Jaupart & Mareschal, 2015). Dans le passé, les données indépendantes pourraient inclure les changements du niveau marin (Müller et al., 2008b) et les indices de mouvements verticaux (Flament et al., 2013).

2. Un autre test possible consiste à effectuer l'assimilation de données sur un intervalle de temps restreint pour obtenir une estimation de l'état du manteau à un moment donné dans le passé. Ensuite, nous pouvons calculer une évolution libre (c'est-à-dire non contrainte par les données) de cet état estimé jusqu'à ce jour. En comparant l'évolution libre avec les observations, nous pourrons évaluer le

pouvoir prédictif des reconstitutions de la circulation mantellique.

La section 2.3 donne un aperçu partiel des applications possibles des reconstitutions de la circulation mantellique: étude de l'effet de la tectonique de surface sur la structure du manteau inférieur (e.g. McNamara & Zhong, 2005; Schuberth et al., 2009b,a; Zhang & Zhong, 2011; Davies et al., 2012; Bower et al., 2013; Bull et al., 2014; Hassan et al., 2015), et étude de scénarios alternatifs de mouvements absolus des plaques (par exemple O'Neill et al., 2005; Doubrovine et al., 2012; Shephard et al., 2012; Nerlich et al., 2016; Zahirovic et al., in press). D'autres applications incluent les études des changements dans l'axe de rotation de la terre (Steinberger & O'Connell, 1997), le niveau marin (Moucha et al., 2008), la topographie dynamique (Flament et al., 2013) et l'effet de la circulation mantellique sur la dynamique du noyau (Olson et al., 2015).

L'utilisation de l'assimilation de données permet l'intégration de données tectoniques dans les modèles de convection d'une manière plus cohérente, puisque le modèle de convection que nous utilisons génère naturellement une tectonique de type plaque à sa surface et que les reconstitutions de la tectonique des plaques ne sont pas imposées par des conditions limites cinématiques, mais intégrées d'une manière statistiquement cohérente.

Comparée à d'autres disciplines, l'assimilation de données pour la convection mantellique est encore à ses débuts. Les connaissances, l'expérience et les progrès techniques de l'assimilation de données dans d'autres domaines, en particulier les sciences de l'atmosphère et l'océanographie, fournissent une quantité importante de ressources pour concevoir des algorithmes adaptés au problème de la reconstitution de la circulation mantellique. Une extension directe de notre travail serait par exemple de développer un lisseur de Kalman (Evensen & Van Leeuwen, 2000; Van Leeuwen, 2001) pour mieux estimer la circulation mantellique passée en utilisant des données sur les états les plus récents du manteau. De nouvelles méthodes sont également développées pour mieux prendre en compte les non-linéarités dans les modèles dynamiques et la nongaussianité qui en résulte dans la statistique des erreurs (Bocquet et al., 2010). Ces solutions devraient être envisagées pour construire et améliorer les algorithmes

d'assimilation de données pour la circulation mantellique.

Nous avons commencé cette thèse en soulignant la coexistence de deux théories différentes qui sont utilisées pour comprendre la dynamique du système croûte-manteau terrestre: la tectonique des plaques et la convection mantellique. Les développements successifs d'algorithmes d'assimilation de données que nous avons proposés sont une tentative de combiner les produits de ces deux théories: les reconstitutions de la tectonique des plaques d'une part et les codes de convection mantellique d'autre part. Le développement continu des codes de convection mantellique dans la modélisation de la dynamique interne de la Terre et de la tectonique de surface suggère un objectif à long terme plus ambitieux: les observations brutes sur la tectonique de surface (cartes des anomalies magnétiques des fonds océaniques, données paléomagnétiques...) pourraient être intégrées directement dans les modèles de convection mantellique, sans utiliser le cadre théorique de la tectonique des plaques.

La révolution des sciences de la Terre qui a eu lieu dans les années 60 (voir la section 2.2.2), avec l'émergence de la théorie de la tectonique des plaques, n'était pas simplement un changement de paradigme. Il a fallu que la communauté des sciences de la Terre rassemble et relie les observations provenant de diverses disciplines pour construire une théorie globale expliquant les mouvements de la surface terrestre. Ce faisant, elle a établi une nouvelle norme pour le fonctionnement de sciences de la Terre Solide. Elle a amené la communauté des sciences de la terre à considérer les processus géologiques comme un problème multidisciplinaire et multi-échelles: pour comprendre l'histoire globale de la Terre solide, nous devons naviguer entre les échelles microscopique, macroscopique, locale, régionale et globale et concilier des observations de nature différente, provenant de diverses disciplines, dans un cadre cohérent. La théorie de la tectonique des plaques a fourni un tel cadre global et s'est révélée juste dans sa description de la dynamique de surface, au premier ordre. Cependant, la tectonique des plaques est une théorie cinématique: elle décrit les mouvements à la surface de la Terre, mais pas leur cause. En outre, il s'agit d'un modèle idéalisé des mouvements de surface: les plaques ne sont pas rigoureusement rigides, et les limites des plaques

ne sont pas infiniment petites, mais peuvent être au contraire diffuses, dans les océans (Gordon, 2000; Wiens et al., 1985) ainsi que sur les continents, comme en témoignent la présence de massifs montagneux. La théorie de la convection mantellique fournit au contraire un cadre physique global qui tient compte de ces déformations diffuses et continues et décrit à la fois la dynamique interne de la Terre solide et la tectonique de surface. De nombreux défis restent avant que nous puissions intégrer directement les observations brutes dans les modèles de convection mantellique pour produire une reconstitution physiquement cohérente de la circulation dans le manteau terrestre et de la tectonique de surface. Sans aucun doute, les méthodes d'assimilation de données seront l'une des pièces de ce puzzle.

# Bibliography

Alboussière, T. & Ricard, Y., 2013. Reflections on dissipation associated with thermal convection, *Journal of Fluid Mechanics*, **725**, R1.

Alisic, L., Gurnis, M., Stadler, G., Burstedde, C., & Ghattas, O., 2012. Multi-scale dynamics and rheology of mantle flow with plates, *Journal of Geophysical Research: Solid Earth (1978–2012)*, **117**(B10).

Anderson, J. L. & Anderson, S. L., 1999. A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Monthly Weather Review*, **127**(12), 2741–2758.

Antolik, M., Gu, Y. J., Ekström, G., & Dziewonski, A. M., 2003. J362d28: a new joint model of compressional and shear velocity in the earth's mantle, *Geophysical Journal International*, **153**(2), 443–466.

Argus, D. F., Gordon, R. G., Heflin, M. B., Ma, C., Eanes, R. J., Willis, P., Peltier, W. R., & Owen, S. E., 2010. The angular velocities of the plates and the velocity of earth's centre from space geodesy, *Geophysical Journal International*, **180**(3), 913–960.

Atkins, S., Valentine, A. P., Tackley, P. J., & Trampert, J., 2016. Using pattern recognition to infer parameters governing mantle convection, *Physics of the Earth and Planetary*

*Interiors*, **257**, 171–186.

Aubert, J. & Fournier, A., 2011. Inferring internal properties of earth's core dynamics and their evolution from surface observations and a numerical geodynamo model, *Nonlin. Processes Geophys*, **18**, 657–674.

Auer, L., Boschi, L., Becker, T., Nissen-Meyer, T., & Giardini, D., 2014. Savani: A variable resolution whole-mantle model of anisotropic shear velocity variations based on multiple data sets, *Journal of Geophysical Research: Solid Earth*, **119**(4), 3006–3034.

Balachandar, S., 1998. Eigendecomposition of the two-point correlation tensor for optimal characterization of mantle convection, *Geophysical journal international*, **132**(1), 111–127.

Balachandar, S., Yuen, D., Reuteler, D., & Lauer, G., 1995. Viscous dissipation in three-dimensional convection with temperature-dependent viscosity, *Science*, **267**(5201), 1150.

Barnett-Moore, N., Hosseinpour, M., & Maus, S., 2016. Assessing discrepancies between previous plate kinematic models of mesozoic iberia and their constraints, *Tectonics*, **35**, 1843–1862.

Batchelor, G. K., 2000. *An introduction to fluid dynamics*, Cambridge university press.

Baumgardner, J. R., 1985. Three-dimensional treatment of convective flow in the earth's mantle, *Journal of Statistical Physics*, **39**(5-6), 501–511.

Becker, T. & Boschi, L., 2016. *SMEAN 2*, `http://www-udc.ig.utexas.edu/external/becker/tdata.html` [Accessed: 2016-09-01].

Becker, T. W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochemistry, Geophysics, Geosystems*, **3**(1).

Bello, L., Coltice, N., Rolf, T., & Tackley, P. J., 2014. On the predictability limit of convection models of the earth's mantle, *Geochemistry, Geophysics, Geosystems*, **15**, 2319–2328.

Bello, L., Coltice, N., Tackley, P. J., Müller, R. D., & Cannon, J., 2015. Assessing the role of slab rheology in coupled plate-mantle convection models, *Earth and Planetary Science Letters*, **430**, 191–201.

Bengtsson, L., Ghil, M., & Källén, E., 1981. *Dynamic meteorology: data assimilation methods*, vol. 36, Springer.

Benioff, H., 1949. Seismic evidence for the fault origin of oceanic deeps, *Geological Society of America Bulletin*, **60**(12), 1837–1856.

Bercovici, D., 1993. A simple model of plate generation from mantle flow, *Geophysical Journal International*, **114**(3), 635–650.

Bercovici, D., 1995. A source-sink model of the generation of plate tectonics, *Journal of geophysical research*, **100**(B2), 2013–2030.

Bercovici, D., 2003. The generation of plate tectonics from mantle convection, *Earth and Planetary Science Letters*, **205**(3), 107–121.

Bercovici, D., 2015. Mantle dynamics: An introduction and overview, *Treatise on Geophysics*, **7**, 1–22.

Bercovici, D. & Ricard, Y., 2014. Plate tectonics, damage and inheritance, *Nature*, **508**(7497), 513–516.

Bercovici, D., Schubert, G., & Glatzmaier, G. A., 1992. Three-dimensional convection of an infinite-prandtl-number compressible fluid in a basally heated spherical shell, *Journal of Fluid Mechanics*, **239**, 683–719.

Bercovici, D., Tackley, P., & Ricard, Y., 2015. The generation of plate tectonics from mantle dynamics, *Treatise on Geophysics*, **7**, 271–318.

Besse, J. & Courtillot, V., 2002. Apparent and true polar wander and the geometry of the geomagnetic field over the last 200 myr, *Journal of Geophysical Research: Solid Earth*, **107**(B11).

Bird, P., 2003. An updated digital model of plate boundaries, *Geochemistry, Geophysics, Geosystems*, **4**(3).

Blankenbach, B., Busse, F., Christensen, U., Cserepes, L., Gunkel, D., Hansen, U., Harder, H., Jarvis, G., Koch, M., Marquart, G., et al., 1989. A benchmark comparison for mantle convection codes, *Geophysical Journal International*, **98**(1), 23–38.

Blayo, E., Bocquet, M., Cosme, E., & Cugliandolo, L. F., 2014. *Advanced data assimilation for geosciences*, Oxford University Press.

Blewitt, G., 2015. 3.11 - gps and space-based geodetic methods, in *Treatise on Geophysics*, pp. 307 – 338, ed. Schubert, G., Elsevier, second edition edn.

Bocher, M., Coltice, N., Fournier, A., & Tackley, P., 2016. A sequential data assimilation approach for the joint reconstruction of mantle convection and surface tectonics, *Geophysical Journal International*, **204**(1), 200–214.

Bocher, M., Fournier, A., & Coltice, N., in prep. Ensemble data assimilation for mantle circulation, *Nonlinear Processes in Geophysics*.

Bocquet, M., Pires, C. A., & Wu, L., 2010. Beyond gaussian statistical modeling in geophysical data assimilation, *Monthly Weather Review*, **138**(8), 2997–3023.

Bonan, B., Nodet, M., Ozenda, O., & Ritz, C., 2014. Data assimilation in glaciology, in *Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics: Special Issue, June 2012*, Oxford University Press.

Boschi, L. & Dziewonski, A. M., 2000. Whole earth tomography from delay times of p, pcp, and pkp phases: Lateral heterogeneities in the outer core or radial anisotropy in the mantle?, *Journal of Geophysical Research: Solid Earth*, **105**(B6), 13675–13696.

Bower, D. J., Gurnis, M., & Seton, M., 2013. Lower mantle structure from paleogeographically constrained dynamic earth models, *Geochemistry, Geophysics, Geosystems*, **14**(1), 44–63.

Bower, D. J., Gurnis, M., & Flament, N., 2015. Assimilating lithosphere and slab history in 4-d earth models, *Physics of the Earth and Planetary Interiors*, **238**, 8–22.

Bull, A. L., Domeier, M., & Torsvik, T. H., 2014. The effect of plate motion history on the longevity of deep mantle heterogeneities, *Earth and Planetary Science Letters*, **401**,

172–182.

Bunge, H.-P. & Grand, S. P., 2000. Mesozoic plate-motion history below the northeast pacific ocean from seismic images of the subducted farallon slab, *Nature*, **405**(6784), 337–340.

Bunge, H.-P., Richards, M. A., Baumgardner, J. R., et al., 1996. Effect of depth-dependent viscosity on the planform of mantle convection, *Nature*, **379**(6564), 436–438.

Bunge, H.-P., Richards, M. A., Lithgow-Bertelloni, C., Baumgardner, J. R., Grand, S. P., & Romanowicz, B. A., 1998. Time scales and heterogeneous structure in geodynamic earth models, *Science*, **280**(5360), 91–95.

Bunge, H.-P., Richards, M., & Baumgardner, J., 2002. Mantle–circulation models with sequential data assimilation: inferring present–day mantle structure from plate–motion histories, *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **360**(1800), 2545–2567.

Bunge, H.-P., Hagelberg, C., & Travis, B., 2003. Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography, *Geophysical Journal International*, **152**(2), 280–301.

Burgers, G., Jan van Leeuwen, P., & Evensen, G., 1998. Analysis scheme in the ensemble kalman filter, *Monthly Weather Review*, **126**(6), 1719–1724.

Burstedde, C., Stadler, G., Alisic, L., Wilcox, L. C., Tan, E., Gurnis, M., & Ghattas, O., 2013. Large-scale adaptive mantle convection simulation, *Geophysical Journal International*, **192**(3), 889–906.

Busse, F., Christensen, U., Clever, R., Cserepes, L., Gable, C., Giannandrea, E., Guillou, L., Houseman, G., Nataf, H., Ogawa, M., et al., 1994. 3d convection at infinite prandtl number in cartesian geometry—a benchmark comparison, *Geophysical & Astrophysical Fluid Dynamics*, **75**(1), 39–59.

Čadek, O., Ricard, Y., Martinec, Z., & Matyska, C., 1993. Comparison between newtonian and non-newtonian flow driven by internal loads, *Geophysical Journal International*, **112**(1), 103–114.

Cammarano, F., Goes, S., Vacher, P., & Giardini, D., 2003. Inferring upper-mantle temperatures from seismic velocities, *Physics of the Earth and Planetary Interiors*, **138**(3), 197–222.

Cane, M. A., Kaplan, A., Miller, R. N., Tang, B., Hackert, E. C., & Busalacchi, A. J., 1996. Mapping tropical pacific sea level: Data assimilation via a reduced state space kalman filter, *Journal of Geophysical Research: Oceans (1978–2012)*, **101**(C10), 22599–22617.

Chang, S.-J., Ferreira, A. M., Ritsema, J., Heijst, H. J., & Woodhouse, J. H., 2015. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations, *Journal of Geophysical Research: Solid Earth*, **120**(6), 4278–4300.

Charney, J., Halem, M., & Jastrow, R., 1969. Use of incomplete historical data to infer the present state of the atmosphere, *Journal of the Atmospheric Sciences*, **26**(5), 1160–1163.

Chassignet, E. P. & Verron, J., 2006. *Ocean weather forecasting: an integrated view of oceanography*, Springer Science & Business Media.

Chorin, A. J., Marsden, J. E., & Marsden, J. E., 1990. *A mathematical introduction to fluid mechanics*, vol. 3, Springer.

Christensen, U., 1984. Convection with pressure-and temperature-dependent non-newtonian rheology, *Geophysical Journal International*, **77**(2), 343–384.

Christensen, U. & Harder, H., 1991. 3-d convection with variable viscosity, *Geophysical Journal International*, **104**(1), 213–226.

Christensen, U. R. & Yuen, D. A., 1984. The interaction of a subducting lithospheric slab with a chemical or phase boundary, *Journal of Geophysical Research: Solid Earth*,

**89**(B6), 4389–4402.

Coffin, M., Gahagan, L., & Lawver, L., 1997. Present-day plate boundary digital data compilation, *University of Texas Institute for Geophysics Technical Report*.

Cohn, S. E., Sivakumaran, N., & Todling, R., 1994. A fixed-lag kalman smoother for retrospective data assimilation, *Monthly Weather Review*, **122**(12), 2838–2867.

Coltice, N., in prep. Tentative forecasts of tectonics with mantle convection models, *PLOS One*.

Coltice, N., Rolf, T., Tackley, P., & Labrosse, S., 2012. Dynamic causes of the relation between area and age of the ocean floor, *Science*, **336**(6079), 335–338.

Coltice, N., Seton, M., Rolf, T., Müller, R., & Tackley, P. J., 2013. Convergence of tectonic reconstructions and mantle convection models for significant fluctuations in seafloor spreading, *Earth and Planetary Science Letters*, **383**, 92–100.

Conrad, C. P. & Gurnis, M., 2003. Seismic tomography, surface uplift, and the breakup of gondwanaland: Integrating mantle convection backwards in time, *Geochemistry, Geophysics, Geosystems*, **4**(3).

Cosme, E., Brankart, J.-M., Verron, J., Brasseur, P., & Krysta, M., 2010. Implementation of a reduced rank square-root smoother for high resolution ocean data assimilation, *Ocean Modelling*, **33**(1), 87–100.

Cox, A. & Doell, R. R., 1963. Geomagnetic polarity epochs and pleistocene geochronometry, *Nature*, **198**, 1049–1051.

Crameri, F., 2013. *The interaction between subduction-related mantle currents and surface topography*, Ph.D. thesis, Diss., Eidgenössische Technische Hochschule ETH Zürich.

Crameri, F. & Tackley, P. J., 2014. Spontaneous development of arcuate single-sided subduction in global 3-d mantle convection models with a free surface, *Journal of Geophysical Research: Solid Earth*, **119**(7), 5921–5942.

Crameri, F. & Tackley, P. J., 2015. Parameters controlling dynamically self-consistent plate tectonics and single-sided subduction in global models of mantle convection, *Journal of Geophysical Research: Solid Earth*, **120**(5), 3680–3706.

Crameri, F., Schmeling, H., Golabek, G., Duretz, T., Orendt, R., Buiter, S., May, D., Kaus, B., Gerya, T., & Tackley, P., 2012. A comparison of numerical surface topography calculations in geodynamic modelling: an evaluation of the 'sticky air'method, *Geophysical Journal International*, **189**(1), 38–54.

Daley, R., 1993. *Atmospheric data analysis*, no. 2, Cambridge university press.

Davies, D. R., Goes, S., Davies, J. H., Schuberth, B., Bunge, H.-P., & Ritsema, J., 2012. Reconciling dynamic and seismic models of earth's lower mantle: The dominant role of thermal heterogeneity, *Earth and Planetary Science Letters*, **353**, 253–269.

Davies, G. F. & Richards, M. A., 1992. Mantle convection, *The Journal of Geology*, **100**, 151–206.

DeMets, C., Gordon, R. G., & Argus, D. F., 2010. Geologically current plate motions, *Geophysical Journal International*, **181**(1), 1–80.

Deparis, V., Legros, H., & Ricard, Y., 1995. Mass anomalies due to subducted slabs and simulations of plate motion since 200 my, *Physics of the Earth and Planetary Interiors*, **89**(3), 271–280.

Dietz, R. S., 1961. Continent and ocean basin evolution by spreading of the sea floor, *Nature*, **190**(4779), 854–857.

DiVenere, V. & Kent, D. V., 1999. Are the pacific and indo–atlantic hotspots fixed? testing the plate circuit through antarctica, *Earth and Planetary Science Letters*, **170**(1), 105–117.

Doubrovine, P. V., Steinberger, B., & Torsvik, T. H., 2012. Absolute plate motions in a reference frame defined by moving hot spots in the pacific, atlantic, and indian oceans, *Journal of Geophysical Research: Solid Earth*, **117**(B9).

Du Toit, A. L., 1937. *Our wandering continents: an hypothesis of continental drifting*, Oliver and Boyd.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference earth model, *Physics of the earth and planetary interiors*, **25**(4), 297–356.

Elsasser, W. M., 1946a. Induction effects in terrestrial magnetism part i. theory, *Physical Review*, **69**(3-4), 106.

Elsasser, W. M., 1946b. Induction effects in terrestrial magnetism part ii. the secular variation, *Physical Review*, **70**(3-4), 202.

Engdahl, E. R., van der Hilst, R., & Buland, R., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bulletin of the Seismological Society of America*, **88**(3), 722–743.

Evans, B. & Kohlstedt, D. L., 1995. *Rheology of Rocks*, pp. 148–165, American Geophysical Union.

Evans, D. A., 2003. True polar wander and supercontinents, *Tectonophysics*, **362**(1), 303–320.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics, *Journal of Geophysical Research: Oceans*, **99**(C5), 10143–10162.

Evensen, G., 2003. The ensemble kalman filter: Theoretical formulation and practical implementation, *Ocean dynamics*, **53**(4), 343–367.

Evensen, G., 2007. *Data assimilation*, Springer.

Evensen, G., 2009a. *Data assimilation: the ensemble Kalman filter*, Springer Science & Business Media.

Evensen, G., 2009b. The ensemble kalman filter for combined state and parameter estimation, *IEEE Control Systems*, **29**(3), 83–104.

Evensen, G. & Van Leeuwen, P. J., 2000. An ensemble kalman smoother for nonlinear dynamics, *Monthly Weather Review*, **128**(6), 1852–1867.

Ewing, M. & Heezen, B. C., 1956. Some problems of antarctic submarine geology, in *Antarctica in the International Geophysical Year: Based on a Symposium on the Antarctic*, pp. 75–81, Wiley Online Library.

Flament, N., Gurnis, M., & Müller, R. D., 2013. A review of observations and models of dynamic topography, *Lithosphere*, **5**(2), 189–210.

Flament, N., Gurnis, M., Williams, S., Seton, M., Skogseid, J., Heine, C., & Dietmar Müller, R., 2014. Topographic asymmetry of the south atlantic from global models of mantle flow and lithospheric stretching, *Earth and Planetary Science Letters*, **387**, 107–119.

Foley, B. J. & Becker, T. W., 2009. Generation of plate-like behavior and mantle heterogeneity from a spherical, viscoplastic convection model, *Geochemistry, Geophysics, Geosystems*, **10**(8).

Forte, A., Simmons, N., & Grand, S., 2015. Constraints on seismic models from other disciplines—constraints on 3-d seismic models from global geodynamic observables: Implications for the global mantle convective flow, *Treatise of Geophysics*, **1**, 853–907.

Foulger, G. R., Panza, G. F., Artemieva, I. M., Bastow, I. D., Cammarano, F., Evans, J. R., Hamilton, W. B., Julian, B. R., Lustrino, M., Thybo, H., et al., 2013. Caveats on tomographic images, *Terra Nova*, **25**(4), 259–281.

Fournier, A., Hulot, G., Jault, D., Kuang, W., Tangborn, A., Gillet, N., Canet, E., Aubert, J., & Lhuillier, F., 2010. An introduction to data assimilation and predictability in geomagnetism, *Space Science Reviews*, **155**(1-4), 247–291.

Gaspari, G. & Cohn, S. E., 1999. Construction of correlation functions in two and three dimensions, *Quarterly Journal of the Royal Meteorological Society*, **125**(554), 723–757.

Gerya, T., 2011. Future directions in subduction modeling, *Journal of Geodynamics*, **52**(5), 344–378.

Gerya, T. V., Connolly, J. A., & Yuen, D. A., 2008. Why is terrestrial subduction one-sided?, *Geology*, **36**(1), 43–46.

Ghelichkhan, S. & Bunge, H.-P., 2016. The compressible adjoint equations in geodynamics: derivation and numerical assessment, *GEM-International Journal on Geomathematics*, **7**(1), 1–30.

Ghil, M., 1989. Meteorological data assimilation for oceanographers. part i: Description and theoretical framework, *Dynamics of Atmospheres and Oceans*, **13**(3-4), 171–218.

Ghil, M. & Malanotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography, *Advances in geophysics*, **33**, 141–266.

Ghil, M., Cohn, S., Tavantzis, J., Bube, K., & Isaacson, E., 1981. Applications of estimation theory to numerical weather prediction, in *Dynamic meteorology: Data assimilation methods*, pp. 139–224, Springer.

Ghosh, A. & Holt, W. E., 2012. Plate motions and stresses from global dynamic models, *Science*, **335**(6070), 838–843.

Ghosh, A., Becker, T., & Zhong, S., 2010. Effects of lateral viscosity variations on the geoid, *Geophysical Research Letters*, **37**(1).

Glatzmaier, G. A., 1988. Numerical simulations of mantle convection: Time-dependent, three-dimensional, compressible, spherical shell, *Geophysical & Astrophysical Fluid Dynamics*, **43**(2), 223–264.

Glišović, P. & Forte, A., 2016. A new back-and-forth iterative method for time-reversed convection modeling: Implications for the cenozoic evolution of 3-d structure and dynamics of the mantle, *Journal of Geophysical Research: Solid Earth*, **121**, 4067–4084.

Glišović, P. & Forte, A. M., 2014. Reconstructing the cenozoic evolution of the mantle: Implications for mantle plume dynamics under the pacific and indian plates, *Earth and Planetary Science Letters*, **390**, 146–156.

Goldreich, P. & Toomre, A., 1969. Some remarks on polar wandering, *Journal of Geophysical Research*, **74**(10), 2555–2567.

Gordon, R. B., 1967. Thermally activated processes in the earth: Creep and seismic attenuation, *Geophysical Journal International*, **14**(1-4), 33–43.

Gordon, R. G., 1998. The plate tectonic approximation: Plate nonrigidity, diffuse plate boundaries, and global plate reconstructions, *Annual Review of Earth and Planetary Sciences*, **26**(1), 615–642.

Gordon, R. G., 2000. Diffuse oceanic plate boundaries: strain rates, vertically averaged rheology, and comparisons with narrow plate boundaries and stable plate interiors, *The history and dynamics of global plate motions*, pp. 143–159.

Grand, S. P., van der Hilst, R. D., & Widiyantoro, S., 1997. High resolution global tomography: a snapshot of convection in the earth, *Geological Society of America Today*, **7**(4).

Granot, R., Dyment, J., & Gallet, Y., 2012. Geomagnetic field variability during the cretaceous normal superchron, *Nature Geoscience*, **5**(3), 220–223.

Griggs, D., 1939a. Creep of rocks, *The Journal of Geology*, **47**(3), 225–251.

Griggs, D. T., 1939b. A theory of mountain-building, *American Journal of Science*, **237**(9), 611–650.

Gu, Y. J., Dziewonski, A. M., Su, W., & Ekström, G., 2001. Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities, *Journal of Geophysical Research: Solid Earth*, **106**(B6), 11169–11199.

Gurnis, M., 1986a. Quantitative bounds on the size spectrum of isotopic heterogeneity within the mantle, *Nature*, **323**, 317–320.

Gurnis, M., 1986b. Stirring and mixing in the mantle by plate-scale flow: Large persistent blobs and long tendrils coexist, *Geophysical Research Letters*, **13**(13), 1474–1477.

Gurnis, M. & Hager, B. H., 1988. Controls of the structure of subducted slabs, *Nature*, **335**(6188), 317–321.

Gurnis, M., Turner, M., Zahirovic, S., DiCaprio, L., Spasojevic, S., Müller, R. D., Boyden, J., Seton, M., Manea, V. C., & Bower, D. J., 2012. Plate tectonic reconstructions with continuously closing plates, *Computers & Geosciences*, **38**(1), 35–42.

Hadamard, J., 1902. Sur les problèmes aux dérivées partielles et leur signification physique, *Princeton university bulletin*, **13**(49-52), 28.

Hager, B. & Richards, M., 1989. Long-wavelength variations in earth's geoid: physical models and dynamical implications, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **328**(1599), 309–327.

Hager, B., Clayton, R., Richards, M., Comer, R., & Dziewonski, A., 1985. Lower mantle heterogeneity, dynamic topography and the geoid, *Nature*, **313**(6003), 541–545.

Hager, B. H. & O'Connell, R. J., 1979. Kinematic models of large-scale flow in the earth's mantle, *Journal of Geophysical Research: Solid Earth (1978–2012)*, **84**(B3), 1031–1048.

Hales, A., 1936. Convection currents in the earth., *Geophysical Journal International*, **3**(9), 372–379.

Hamill, T. M., Whitaker, J. S., & Snyder, C., 2001. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter, *Monthly Weather Review*, **129**(11), 2776–2790.

Harder, H. & Hansen, U., 2005. A finite-volume solution method for thermal convection and dynamo problems in spherical shells, *Geophysical Journal International*, **161**(2), 522–532.

Haskell, N. A., 1937. The viscosity of the asthenosphere, *American Journal of Science*, (193), 22–28.

Hassan, R., Flament, N., Gurnis, M., Bower, D. J., & Müller, D., 2015. Provenance of plumes in global convection models, *Geochemistry, Geophysics, Geosystems*, **16**(5), 1465–1489.

Hassan, R., Müller, R. D., Gurnis, M., Williams, S. E., & Flament, N., 2016. A rapid burst in hotspot motion through the interaction of tectonics and deep mantle flow,

*Nature*, **533**(7602), 239–242.

Heezen, B. C. & Tharp, M., 1977. World ocean floor panorama, *U.S. Navy*.

Heezen, B. C., Tharp, M., & Ewing, M., 1959. The floors of the oceans i. the north atlantic, *Geological Society of America Special Papers*, **65**, 1–126.

Heirtzler, J., Dickson, G., Herron, E., Pitman, W., & Le Pichon, X., 1968. Marine magnetic anomalies, geomagnetic field reversals, and motions of the ocean floor and continents, *Journal of Geophysical Research*, **73**(6), 2119–2136.

Heirtzler, J. R. & Le Pichon, X., 1965. Crustal structure of the mid-ocean ridges: 3. magnetic anomalies over the mid-atlantic ridge, *Journal of Geophysical Research*, **70**(16), 4013–4033.

Heirtzler, J. R., Le Pichon, X., & Baron, J. G., 1966. Magnetic anomalies over the reykjanes ridge, in *Deep Sea Research and Oceanographic Abstracts*, vol. 13, pp. 427–443, Elsevier.

Hellinger, S., 1981. The uncertainties of finite rotations in plate tectonics, *Journal of Geophysical Research: Solid Earth*, **86**(B10), 9312–9318.

Hernlund, J. W. & Tackley, P. J., 2008. Modeling mantle convection in the spherical annulus, *Physics of the Earth and Planetary Interiors*, **171**(1), 48–54.

Hess, H. H., 1962. History of ocean basins, *Petrologic studies*, **4**, 599–620.

Holmes, A., 1931. Radioactivity and earth movements., *Nature*, **128**, 496.

Holmes, A., 1944. *Principles of physical geology*, Thomas Nelson and Sons.

Holt, W. E., Shen-Tu, B., Haines, J., & Jackson, J., 2000. On the determination of self-consistent strain rate fields within zones of distributed continental deformation, in *The History and Dynamics of Global Plate Motions*, pp. 113–141, American Geophysical Union.

Horbach, A., Bunge, H.-P., & Oeser, J., 2014. The adjoint method in geodynamics: derivation from a general operator formulation and application to the initial condition problem in a high resolution mantle circulation model, *GEM-International Journal on Geomathematics*, **5**(2), 163–194.

Hosseinpour, M., Williams, S., Seton, M., Barnett-Moore, N., & Müller, R. D., 2016. Tectonic evolution of western tethys from jurassic to present day: coupling geological and geophysical data with seismic tomography models, *International Geology Review*, **58**, 1–30.

Hoteit, I., 2001. *Filtres de Kalman réduits et efficaces pour l'assimilation de données en océanographie*, Ph.D. thesis, Université Joseph-Fourier-Grenoble I.

Hoteit, I. & Pham, D.-T., 2004. An adaptively reduced-order extended kalman filter for data assimilation in the tropical pacific, *Journal of marine systems*, **45**(3), 173–188.

Houser, C., Masters, G., Shearer, P., & Laske, G., 2008. Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms, *Geophysical Journal International*, **174**(1), 195–212.

Houtekamer, P. L. & Mitchell, H. L., 1998. Data assimilation using an ensemble kalman filter technique, *Monthly Weather Review*, **126**(3), 796–811.

Houtekamer, P. L. & Mitchell, H. L., 2001. A sequential ensemble kalman filter for atmospheric data assimilation, *Monthly Weather Review*, **129**(1), 123–137.

Hunter, J. D., 2007. Matplotlib: A 2d graphics environment, *Computing In Science & Engineering*, **9**(3), 90–95.

Iaffaldano, G. & Bunge, H.-P., 2015. Rapid plate motion variations through geological time: Observations serving geodynamic interpretation, *Annual Review of Earth and Planetary Sciences*, **43**, 571–592.

Iaffaldano, G., Bodin, T., & Sambridge, M., 2012. Reconstructing plate-motion changes in the presence of finite-rotations noise, *Nature communications*, **3**, 1048.

Iaffaldano, G., Hawkins, R., Bodin, T., & Sambridge, M., 2014. Redback: Open-source software for efficient noise-reduction in plate kinematic reconstructions, *Geochemistry, Geophysics, Geosystems*, **15**(4), 1663–1670.

Ide, K., Courtier, P., Ghil, M., & Lorenc, A., 1997. Unified notation for data assimilation: operational, sequential and variational, *Practice*, **75**(1B), 181–189.

Isacks, B., Oliver, J., & Sykes, L. R., 1968. Seismology and the new global tectonics, *Journal of Geophysical Research*, **73**(18), 5855–5899.

Ismail-Zadeh, A. & Tackley, P., 2010. *Computational methods for geodynamics*, Cambridge University Press.

Ismail-Zadeh, A., Korotkii, A., Naimark, B., & Tsepelev, I., 2003a. Three-dimensional numerical simulation of the inverse problem of thermal convection, *Computational Mathematics and Mathematical Physics*, **43**(4), 581–599.

Ismail-Zadeh, A., Korotkii, A., & Tsepelev, I., 2003b. Numerical approach to solving problems of slow viscous flow backwards in time, in *Computational fluid and solid mechanics*, pp. 938–941, Elsevier Science Amsterdam.

Ismail-Zadeh, A., Schubert, G., Tsepelev, I., & Korotkii, A., 2004. Inverse problem of thermal convection: numerical approach and application to mantle plume restoration, *Physics of the Earth and Planetary Interiors*, **145**(1), 99–114.

Ismail-Zadeh, A., Korotkii, A., Schubert, G., & Tsepelev, I., 2007. Quasi-reversibility method for data assimilation in models of mantle dynamics, *Geophysical Journal International*, **170**(3), 1381–1398.

Ismail-Zadeh, A., Korotkii, A., & Tsepelev, I., 2016. *Data-driven Numerical Modelling in Geodynamics: Methods and Applications*, Springer.

Janjic, T., Nerger, L., Albertella, A., Schröter, J., & Skachko, S., 2011. On domain localization in ensemble-based kalman filter algorithms, *Monthly Weather Review*, **139**(7), 2046–2060.

Jarvis, G. T. & McKenzie, D. P., 1980. Convection in a compressible fluid with infinite prandtl number, *Journal of Fluid Mechanics*, **96**(03), 515–583.

Jaupart, C. & Mareschal, J., 2015. Heat flow and thermal structure of the lithosphere, in *Treatise on Geophysics (Second Edition)*, vol. 6, pp. 217 – 252, ed. Schubert, G., Elsevier, Oxford.

Kageyama, A. & Sato, T., 2004. "yin-yang grid": An overset grid in spherical geometry, *Geochemistry Geophysics Geosystems*, **5**(9), Q09005.

Kalnay, E., 2003. *Atmospheric modeling, data assimilation, and predictability*, Cambridge university press.

Kalnay, E., Li, H., Miyoshi, T., YANG, S.-C., & BALLABRERA-POY, J., 2007. 4-d-var or ensemble kalman filter?, *Tellus A*, **59**(5), 758–773.

Karato, S.-i., 1998. Seismic anisotropy in the deep mantle, boundary layers and the geometry of mantle convection, in *Geodynamics of Lithosphere & Earth's Mantle*, pp. 565–587, Springer.

Karato, S.-I., 2013. Rheological properties of minerals and rocks, pp. 94–144, Wiley Online Library.

Karato, S.-i. & Karki, B. B., 2001. Origin of lateral variation of seismic wave velocities and density in the deep mantle, *Journal of Geophysical Research: Solid Earth*, **106**(B10), 21771–21783.

King, S. D., 2016. Reconciling laboratory and observational models of mantle rheology in geodynamic modelling, *Journal of Geodynamics*, **100**, 33–50.

Koelemeijer, P., Ritsema, J., Deuss, A., & van Heijst, H.-J., 2016. Sp12rts: a degree-12 model of shear-and compressional-wave velocity for earth's mantle, *Geophysical Journal International*, **204**(2), 1024–1039.

Kogan, M. G. & Steblov, G. M., 2008. Current global plate kinematics from gps (1995–2007) with the plate-consistent reference frame, *Journal of Geophysical Research: Solid Earth*, **113**(B4).

Kohlstedt, D., 2015. Properties of rocks and minerals-constitutive equations, rheological behavior, and viscosity of rocks, *Treatise on Geophysics*, **2**, 441–472.

Kohlstedt, D., Evans, B., & Mackwell, S., 1995. Strength of the lithosphere: Constraints imposed by laboratory experiments, *Journal of Geophysical Research: Solid Earth*, **100**(B9), 17587–17602.

Kreemer, C., Blewitt, G., & Klein, E. C., 2014. A geodetic plate motion and global strain rate model, *Geochemistry, Geophysics, Geosystems*, **15**(10), 3849–3889.

Kronbichler, M., Heister, T., & Bangerth, W., 2012. High accuracy mantle convection simulation through modern numerical methods, *Geophysical Journal International*, **191**(1), 12–29.

Kustowski, B., Ekström, G., & Dziewoński, A., 2008. Anisotropic shear-wave velocity structure of the earth's mantle: A global model, *Journal of Geophysical Research: Solid Earth*, **113**(B6).

Lattés, R. & Lions, J. L., 1969. *The method of quasi-reversibility: applications to partial differential equations*, Elsevier.

Le Dimet, F.-X. & Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus A*, **38**(2), 97–110.

Le Pichon, X., 1968. Sea-floor spreading and continental drift, *Journal of Geophysical Research*, **73**(12), 3661–3697.

Lekic, V., Cottaar, S., Dziewonski, A., & Romanowicz, B., 2012. Cluster analysis of global lower mantle tomography: A new class of structure and implications for chemical heterogeneity, *Earth and Planetary Science Letters*, **357**, 68–77.

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, *Geochemistry, Geophysics, Geosystems*, **9**(5).

Li, X.-D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *Journal of Geophysical Research: Solid Earth*, **101**(B10), 22245–22272.

Lithgow-Bertelloni, C. & Richards, M. A., 1998. The dynamics of cenozoic and mesozoic plate motions, *Reviews of Geophysics*, **36**(1), 27–78.

Lithgow-Bertelloni, C., Richards, M. A., Ricard, Y., O'Connell, R. J., & Engebretson, D. C., 1993. Toroidal-poloidal partitioning of plate motions since 120 ma, *Geophysical Research Letters*, **20**(5), 375–378.

Liu, L. & Gurnis, M., 2008. Simultaneous inversion of mantle properties and initial conditions using an adjoint of mantle convection, *Journal of Geophysical Research: Solid Earth (1978–2012)*, **113**(B8).

Liu, Q. & Gu, Y., 2012. Seismic imaging: from classical to adjoint tomography, *Tectonophysics*, **566**, 31–66.

Lourenço, D. L., Rozel, A., & Tackley, P. J., 2016. Melting-induced crustal production helps plate tectonics on earth-like planets, *Earth and Planetary Science Letters*, **439**, 18–28.

Lowman, J. P., 2011. Mantle convection models featuring plate tectonic behavior: An overview of methods and progress, *Tectonophysics*, **510**(1), 1–16.

Mahalanobis, P., 1936. On the generalized distance in statistic, *Proc. Nat. Inst. Sci. India (Calcutta)*, **2**, 49–55.

Mallard, C., Coltice, N., Seton, M., Müller, R. D., & Tackley, P. J., 2016. Subduction controls the distribution and fragmentation of earth's tectonic plates, *Nature*, **535**(7610), 140–143.

Maloof, A. C., Halverson, G. P., Kirschvink, J. L., Schrag, D. P., Weiss, B. P., & Hoffman, P. F., 2006. Combined paleomagnetic, isotopic, and stratigraphic evidence for true polar wander from the neoproterozoic akademikerbreen group, svalbard, norway, *Geological Society of America Bulletin*, **118**(9-10), 1099–1124.

Mason, R. G. & Draff, A., 1961. Magnetic survey off the west coast of north america, 32 n. latitude to 42 n. latitude, *Geological Society of America Bulletin*, **72**(8), 1259–1265.

Masters, G., Johnson, S., Laske, G., Bolton, H., & Davies, J., 1996. A shear-velocity model of the mantle [and discussion], *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **354**(1711), 1385–1411.

Masters, G., Laske, G., Bolton, H., & Dziewonski, A., 2000. The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure, in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale*, pp. 63–87, American Geophysical Union.

Mattern, E., Matas, J., Ricard, Y., & Bass, J., 2005. Lower mantle composition and temperature from mineral physics and thermodynamic modelling, *Geophysical Journal International*, **160**(3), 973–990.

Matthews, K. J., Müller, R. D., Wessel, P., & Whittaker, J. M., 2011. The tectonic fabric of the ocean basins, *Journal of Geophysical Research: Solid Earth*, **116**(B12).

McKenzie, D. & Morgan, W., 1969. Evolution of triple junctions, *Nature*, **224**, 125–133.

McKenzie, D. & Parker, R., 1967. The north pacific: an example of tectonics on a sphere, *Nature*, **216**, 1276–1280.

McKenzie, D. P., Roberts, J. M., & Weiss, N. O., 1974. Convection in the earth's mantle: towards a numerical simulation, *Journal of Fluid Mechanics*, **62**(03), 465–538.

McNamara, A. K. & Zhong, S., 2005. Thermochemical structures beneath africa and the pacific ocean, *Nature*, **437**(7062), 1136–1139.

Meibom, A. & Anderson, D. L., 2004. The statistical upper mantle assemblage, *Earth and Planetary Science Letters*, **217**(1), 123–139.

Merkouriev, S. & DeMets, C., 2006. Constraints on indian plate motion since 20 ma from dense russian magnetic data: Implications for indian plate dynamics, *Geochemistry, Geophysics, Geosystems*, **7**(2).

Merkouriev, S. & DeMets, C., 2014. High-resolution neogene reconstructions of eurasia-north america plate motion, *Geophysical Journal International*, **198**(1), 366–384.

Molnar, P. & Stock, J., 1987. Relative motions of hotspots in the pacific, atlantic and indian oceans since late cretaceous time, *Nature*, **327**(6123), 587–591.

Molnar, P. & Stock, J. M., 1985. A method for bounding uncertainties in combined plate reconstructions, *Journal of Geophysical Research: Solid Earth (1978–2012)*, **90**(B14), 12537–12544.

Moresi, L. & Solomatov, V., 1998. Mantle convection with a brittle lithosphere: thoughts on the global tectonic styles of the earth and venus, *Geophysical Journal International*, **133**(3), 669–682.

Moresi, L., Gurnis, M., & Zhong, S., 2000. Plate tectonics and convection in the earth's mantle: Toward a numerical simulation, *Computing in Science & Engineering*, **2**(3), 22–33.

Moresi, L., Dufour, F., & Mühlhaus, H.-B., 2003. A lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials, *Journal of Computational Physics*, **184**(2), 476–497.

Morgan, W. J., 1968. Rises, trenches, great faults, and crustal blocks, *Journal of Geophysical Research*, **73**(6), 1959–1982.

Morgan, W. J., 1971. Convection plumes in the lower mantle, *Nature*, (230), 42–43.

Morley, L., 1963. An explanation of magnetic banding in ocean basins, *The Oceanic Lithosphere, Sea*, **7**, 1717–1719.

Moucha, R., Forte, A. M., Mitrovica, J. X., Rowley, D. B., Quéré, S., Simmons, N. A., & Grand, S. P., 2008. Dynamic topography and long-term sea-level variations: There is no such thing as a stable continental platform, *Earth and Planetary Science Letters*, **271**(1), 101–108.

Moulik, R., Lekic, V., Romanowicz, B., & Dziewonski, A., 2016. *3-dimensional Reference Earth Model Webpage*, `https://www.geol.umd.edu/facilities/seismology/rem-3d/` [Accessed: 2016-09-01].

Mühlhaus, H., Moresi, H., & Cada, M., 2004. Emergent anisotropy and flow alignment in viscous rock, in *Computational Earthquake Science Part II*, pp. 2451–2463, Springer.

Müller, D. & Wessel, P., 2015. Pate tectonics, *Treatise on Geophysics*, **6**, 45–93.

Müller, R. D., Royer, J.-Y., & Lawver, L. A., 1993. Revised plate motions relative to the hotspots from combined atlantic and indian ocean hotspot tracks, *Geology*, **21**(3), 275–278.

Müller, R. D., Sdrolias, M., Gaina, C., & Roest, W. R., 2008a. Age, spreading rates, and spreading asymmetry of the world's ocean crust, *Geochemistry, Geophysics, Geosystems*, **9**(4).

Müller, R. D., Sdrolias, M., Gaina, C., Steinberger, B., & Heine, C., 2008b. Long-term sea-level fluctuations driven by ocean basin dynamics, *Science*, **319**(5868), 1357–1362.

Müller, R. D., Seton, M., Zahirovic, S., Williams, S. E., Matthews, K. J., Wright, N. M., Shephard, G. E., Maloney, K., Barnett-Moore, N., Hosseinpour, M., et al., 2016. Ocean basin evolution and global-scale plate reorganization events since pangea breakup, *Annual Review of Earth and Planetary Sciences*, **44**(1), 107–138.

Nansen, F., 1928. *The earth's crust : its surface-forms, and isostatic adjustment*, Avhandlinger utgitt av Det norske Videnskapsakademi i Oslo, 1. Mat. Naturv. Klasse; 1927, Dybwad [in Komm.].

Nerger, L., 2016. *PDAF: the Parallel Data Assimilation Framework*, `http://pdaf.awi.de/` [Accessed: 2016-09-01].

Nerger, L. & Hiller, W., 2013. Software for ensemble-based data assimilation systems—implementation strategies and scalability, *Computers & Geosciences*, **55**, 110–118.

Nerger, L., Hiller, W., & Schröter, J., 2005. The parallel data assimilation framework: experiences with kalman filtering, in *Use of High Performance in Meteorology, Proceedings of the 11th ECMWF Workshop. World Scientific, Singapore*, pp. 63–86.

Nerger, L., Janjic, T., Schröter, J., & Hiller, W., 2012. A unification of ensemble square root kalman filters, *Monthly Weather Review*, **140**(7), 2335–2345.

Nerger, L., Schulte, S., & Bunse-Gerstner, A., 2014. On the influence of model nonlinearity and localization on ensemble kalman smoothing, *Quarterly Journal of the Royal Meteorological Society*, **140**(684), 2249–2259.

Nerlich, R., Colli, L., Ghelichkhan, S., Schuberth, B., & Bunge, H.-P., 2016. Mantle convection models constrain central neo-tethys ocean reconstructions, *Geophysical Research Letters*, p. Advance online publication.

Ogg, J., 2012. Chapter 5 - geomagnetic polarity time scale, in *The Geologic Time Scale*, pp. 85 – 113, Elsevier.

Olson, P. & Bercovici, D., 1991. On the equipartition of kinetic energy in plate tectonics, *Geophysical research letters*, **18**(9), 1751–1754.

Olson, P., Deguen, R., Rudolph, M. L., & Zhong, S., 2015. Core evolution driven by mantle global circulation, *Physics of the Earth and Planetary Interiors*, **243**, 44–55.

O'Neill, C., Müller, D., & Steinberger, B., 2005. On the uncertainties in hot spot reconstructions and the significance of moving hot spot reference frames, *Geochemistry, Geophysics, Geosystems*, **6**(4).

Oreskes, N., 1999. *The rejection of continental drift: Theory and method in American earth science*, Oxford University Press.

Oreskes, N., 2003. *Plate tectonics: An insider's history of the modern theory of the Earth*, Westview Press.

Panning, M. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophysical Journal International*, **167**(1), 361–379.

Parmentier, E., Turcotte, D., & Torrance, K., 1976. Studies of finite amplitude non-newtonian thermal convection with application to convection in the earth's mantle, *Journal of Geophysical Research*, **81**(11), 1839–1846.

Pekeris, C. L., 1935. Thermal convection in the interior of the earth, *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, **3**(8), 343–367.

Pham, D. T., 2001. Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Monthly Weather Review*, **129**(5), 1194–1207.

Pouilloux, L., Kaminski, E., & Labrosse, S., 2007. Anisotropic rheology of a cubic medium and implications for geological materials, *Geophysical Journal International*, **170**(2), 876–885.

Puster, P., Hager, B. H., & Jordan, T. H., 1995. Mantle convection experiments with evolving plates, *Geophysical research letters*, **22**(16), 2223–2226.

Ratcliff, J. T., Tackley, P. J., Schubert, G., & Zebib, A., 1997. Transitions in thermal convection with strongly variable viscosity, *Physics of the earth and planetary interiors*, **102**(3), 201–212.

Ratnaswamy, V., Stadler, G., & Gurnis, M., 2015. Adjoint-based estimation of plate coupling in a non-linear mantle flow model: theory and examples, *Geophysical Journal International*, **202**(2), 768–786.

Rawlinson, N., Pozgay, S., & Fishwick, S., 2010. Seismic tomography: a window into deep earth, *Physics of the Earth and Planetary Interiors*, **178**(3), 101–135.

Rawlinson, N., Fichtner, A., Sambridge, M., & Young, M. K., 2014. Chapter one-seismic tomography and the assessment of uncertainty, *Advances in Geophysics*, **55**, 1–76.

Ricard, Y., 2007. Physics of mantle convection, *Treatise on Geophysics*, **7**, 31–87.

Ricard, Y., 2015. Physics of mantle convection, *Treatise on Geophysics*, **7**, 23–71.

Ricard, Y., Fleitout, L., & Froidevaux, C., 1984. Geoid heights and lithospheric stresses for a dynamic earth, in *Annales Geophysicae*, vol. 2, pp. 267–286.

Ricard, Y., Vigny, C., & Froidevaux, C., 1989. Mantle heterogeneities, geoid, and plate motion: A monte carlo inversion, *Journal of Geophysical Research: Solid Earth (1978–2012)*, **94**(B10), 13739–13754.

Ricard, Y., Richards, M., Lithgow-Bertelloni, C., & Le Stunff, Y., 1993. A geodynamic model of mantle density heterogeneity, *Journal of Geophysical Research: Solid Earth*, **98**(B12), 21895–21909.

Ricard, Y., Durand, S., Montagner, J.-P., & Chambat, F., 2014. Is there seismic attenuation in the mantle?, *Earth and Planetary Science Letters*, **388**, 257–264.

Richards, M. A. & Hager, B. H., 1984. Geoid anomalies in a dynamic earth, *Journal of Geophysical Research: Solid Earth*, **89**(B7), 5987–6002.

Richards, M. A., Yang, W.-S., Baumgardner, J. R., & Bunge, H.-P., 2001. Role of a low-viscosity zone in stabilizing plate tectonics: Implications for comparative terrestrial planetology, *Geochemistry, Geophysics, Geosystems*, **2**(8).

Ritsema, J., Deuss, A., Van Heijst, H., & Woodhouse, J., 2011. S40rts: a degree-40 shear-velocity model for the mantle from new rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophysical Journal International*, **184**(3), 1223–1236.

Rolf, T. & Tackley, P., 2011. Focussing of stress by continents in 3d spherical mantle convection with self-consistent plate tectonics, *Geophysical Research Letters*, **38**(18).

Rolf, T., Coltice, N., & Tackley, P., 2014. Statistical cyclicity of the supercontinent cycle, *Geophysical Research Letters*, **41**(7), 2351–2358.

Rowley, D. B., 2008. Extrapolating oceanic age distributions: lessons from the pacific region, *The Journal of Geology*, **116**(6), 587–598.

Runcorn, S., 1961. Climatic change through geological time in the light of the palaeomagnetic evidence for polar wandering and continental drift, *Quarterly Journal of the Royal Meteorological Society*, **87**(373), 282–313.

Runcorn, S., 1962. Convection currents in the earth's mantle, *Nature*, **195**, 1248–1249.

Rutherford, I. D., 1972. Data assimilation by statistical interpolation of forecast error fields, *Journal of the Atmospheric Sciences*, **29**(5), 809–815.

Sandwell, D. T. & Smith, W. H., 2009. Global marine gravity from retracked geosat and ers-1 altimetry: Ridge segmentation versus spreading rate, *Journal of Geophysical Research: Solid Earth*, **114**(B1).

Sasaki, Y., 1970. Some basic formalisms in numerical variational analysis, *Monthly Weather Review*, **98**, 875–883.

Schubert, G., Turcotte, D. L., & Olson, P., 2001. *Mantle convection in the Earth and planets*, Cambridge University Press.

Schuberth, B., Bunge, H.-P., & Ritsema, J., 2009a. Tomographic filtering of high-resolution mantle circulation models: Can seismic heterogeneity be explained by temperature alone?, *Geochemistry, Geophysics, Geosystems*, **10**(5).

Schuberth, B., Bunge, H.-P., Steinle-Neumann, G., Moder, C., & Oeser, J., 2009b. Thermal versus elastic heterogeneity in high-resolution mantle circulation models with pyrolite composition: high plume excess temperatures in the lowermost mantle, *Geochemistry, Geophysics, Geosystems*, **10**(1).

Seton, M., Müller, R., Zahirovic, S., Gaina, C., Torsvik, T., Shephard, G., Talsma, A., Gurnis, M., Turner, M., Maus, S., et al., 2012. Global continental and ocean basin reconstructions since 200ma, *Earth-Science Reviews*, **113**(3), 212–270.

Seton, M., Whittaker, J. M., Wessel, P., Müller, R. D., DeMets, C., Merkouriev, S., Cande, S., Gaina, C., Eagles, G., Granot, R., et al., 2014. Community infrastructure and repository for marine magnetic identifications, *Geochemistry, Geophysics, Geosystems*, **15**(4), 1629–1641.

Shephard, G. E., Bunge, H.-P., Schuberth, B. S., Müller, R., Talsma, A., Moder, C., & Landgrebe, T., 2012. Testing absolute plate reference frames and the implications for the generation of geodynamic mantle heterogeneity structure, *Earth and Planetary Science Letters*, **317**, 204–217.

Shephard, G. E., Müller, R. D., & Seton, M., 2013. The tectonic evolution of the arctic since pangea breakup: Integrating constraints from surface geology and geophysics

with mantle structure, *Earth-Science Reviews*, **124**, 148–183.

Simmons, N. A., Forte, A. M., & Grand, S. P., 2006. Constraining mantle flow with seismic and geodynamic data: a joint approach, *Earth and Planetary Science Letters*, **246**(1), 109–124.

Simmons, N. A., 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 heterogeneity, *Geophysical Journal International*, **177**(3), 1284–1304.

Simmons, N. A., Forte, A. M., Boschi, L., & Grand, S. P., 2010. Gypsum: A joint tomographic model of mantle density and seismic wave speeds, *Journal of Geophysical Research: Solid Earth*, **115**(B12).

Smagorinsky, J., Miyakoda, K., & Strickler, R., 1970. The relative importance of variables in initial conditions for dynamical weather prediction1, *Tellus*, **22**(2), 141–157.

Solomatov, V., 1995. Scaling of temperature-and stress-dependent viscosity convection, *Physics of Fluids (1994-present)*, **7**(2), 266–274.

Spasojevic, S., Liu, L., & Gurnis, M., 2009. Adjoint models of mantle convection with seismic, plate motion, and stratigraphic constraints: North america since the late cretaceous, *Geochemistry Geophysics Geosystems*, **10**(5), Q05W02.

Stadler, G., Gurnis, M., Burstedde, C., Wilcox, L. C., Alisic, L., & Ghattas, O., 2010. The dynamics of plate tectonics and mantle flow: From local to global scales, *Science*, **329**(5995), 1033–1038.

Stein, C., Schmalzl, J., & Hansen, U., 2004. The effect of rheological parameters on plate behaviour in a self-consistent model of mantle convection, *Physics of the Earth and Planetary Interiors*, **142**(3), 225–255.

Steinberger, B., 2016. Topography caused by mantle density variations: observation-based estimates and models derived from tomography and lithosphere thickness, *Geophysical Journal International*, **205**(1), 604–621.

Steinberger, B. & O'Connell, R. J., 1997. Changes of the earth's rotation axis owing to advection of mantle density heterogeneities, *Nature*, **387**, 169–173.

Steinberger, B. & Torsvik, T. H., 2008. Absolute plate motions and true polar wander in the absence of hotspot tracks, *Nature*, **452**(7187), 620–623.

Steinberger, B., Sutherland, R., & O'connell, R. J., 2004. Prediction of emperor-hawaii seamount locations from a revised model of global plate motion and mantle flow, *Nature*, **430**(6996), 167–173.

Stemmer, K., Harder, H., & Hansen, U., 2006. A new method to simulate convection with strongly temperature-and pressure-dependent viscosity in a spherical shell: Applications to the earth's mantle, *Physics of the Earth and Planetary Interiors*, **157**(3), 223–249.

Stixrude, L. & Lithgow-Bertelloni, C., 2005. Thermodynamics of mantle minerals—i. physical properties, *Geophysical Journal International*, **162**(2), 610–632.

Stixrude, L. & Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals-ii. phase equilibria, *Geophysical Journal International*, **184**(3), 1180–1213.

Su, W.-j. & Dziewonski, A. M., 1997. Simultaneous inversion for 3-d variations in shear and bulk velocity in the mantle, *Physics of the Earth and Planetary Interiors*, **100**(1), 135–156.

Swinbank, R., Shutyaev, V., & Lahoz, W. A., 2012. *Data assimilation for the earth system*, vol. 26, Springer Science & Business Media.

Sykes, L. R., 1967. Mechanism of earthquakes and nature of faulting on the mid-oceanic ridges, *Journal of Geophysical Research*, **72**(8), 2131–2153.

Tackley, P. & Xie, S., 2003. Stag3d: A code for modeling thermo-chemical multiphase convection in earth's mantle, in *Second MIT Conference on Computational Fluid and Solid Mechanics*, pp. 1524–1527.

Tackley, P. J., 1996. Effects of strongly variable viscosity on three-dimensional compressible convection in planetary mantles, *Journal of Geophysical Research: Solid Earth*

*(1978–2012)*, **101**(B2), 3311–3332.

Tackley, P. J., 1998a. Self-consistent generation of tectonic plates in three-dimensional mantle convection, *Earth and Planetary Science Letters*, **157**(1), 9–22.

Tackley, P. J., 1998b. Three-dimensional simulations of mantle convection with a thermo-chemical basal boundary layer: D"?, *The core-mantle boundary region*, pp. 231–253.

Tackley, P. J., 2000a. The quest for self-consistent generation of plate tectonics in mantle convection models, in *The history and dynamics of global plate motions*, pp. 47–72, Wiley Online Library.

Tackley, P. J., 2000b. Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations, *Geochemistry, Geophysics, Geosystems*, **1**(8).

Tackley, P. J., 2000c. Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations 2. strain weakening and asthenosphere, *Geochemistry, Geophysics, Geosystems*, **1**(8).

Tackley, P. J., 2002. Strong heterogeneity caused by deep mantle layering, *Geochemistry, Geophysics, Geosystems*, **3**(4), 1–22.

Tackley, P. J., 2008. Modelling compressible mantle convection with large viscosity contrasts in a three-dimensional spherical shell using the yin-yang grid, *Physics of the Earth and Planetary Interiors*, **171**(1), 7–18.

Tackley, P. J., 2015. Mantle geochemical geodynamics, in *Treatise on Geophysics*, vol. 7, pp. 521–584, Elsevier.

Tackley, P. J. & King, S. D., 2003. Testing the tracer ratio method for modeling active compositional fields in mantle convection simulations, *Geochemistry, Geophysics, Geosystems*, **4**(4).

Tackley, P. J., Stevenson, D. J., Glatzmaier, G. A., & Schubert, G., 1993. Effects of an endothermic phase transition at 670 km depth in a spherical model of convection in

the earth's mantle, *Nature*, **361**(6414), 699–704.

Talagrand, O., 1997. Assimilation of observations, an introduction, *Journal-Meteorological Society of Japan Series 2*, **75**, 81–99.

Talagrand, O., 2003. A posteriori validation of assimilation algorithms, in *Data assimilation for the earth system*, pp. 85–95, Springer.

Talagrand, O., 2014. Errors. a posteriori diagnostics, in *Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics: Special Issue, June 2012*, p. 229, Oxford University Press.

Tarduno, J. A. & Cottrell, R. D., 1997. Paleomagnetic evidence for motion of the hawaiian hotspot during formation of the emperor seamounts, *Earth and Planetary Science Letters*, **153**(3), 171–180.

Tarduno, J. A. & Gee, J., 1995. Large-scale motion between pacific and atlantic hotspots, *Nature*, **378**(6556), 477–480.

Tarduno, J. A., Duncan, R. A., Scholl, D. W., Cottrell, R. D., Steinberger, B., Thordarson, T., Kerr, B. C., Neal, C. R., Frey, F. A., Torii, M., et al., 2003. The emperor seamounts: Southward motion of the hawaiian hotspot plume in earth's mantle, *Science*, **301**(5636), 1064–1069.

Thielmann, M., Kaus, B., & Popov, A., 2015. Lithospheric stresses in rayleigh–bénard convection: effects of a free surface and a viscoelastic maxwell rheology, *Geophysical Journal International*, **203**(3), 2200–2219.

Thurber, C. & Ritsema, J., 2015. 1.10 - theory and observations - seismic tomography and inverse methods, in *Treatise on Geophysics (Second Edition)*, pp. 307 – 337, ed. Schubert, G., Elsevier, Oxford, second edition edn.

Todling, R. & Cohn, S. E., 1994. Suboptimal schemes for atmospheric data assimilation based on the kalman filter, *Monthly Weather Review*, **122**(11), 2530–2557.

Torsvik, T. H., Müller, R. D., Van der Voo, R., Steinberger, B., & Gaina, C., 2008. Global plate motion frames: toward a unified model, *Reviews of Geophysics*, **46**(3).

Torsvik, T. H., Rousse, S., Labails, C., & Smethurst, M. A., 2009. A new scheme for the opening of the south atlantic ocean and the dissection of an aptian salt basin, *Geophysical Journal International*, **177**(3), 1315–1333.

Torsvik, T. H., Steinberger, B., Gurnis, M., & Gaina, C., 2010. Plate tectonics and net lithosphere rotation over the past 150my, *Earth and Planetary Science Letters*, **291**(1), 106–112.

Torsvik, T. H., van der Voo, R., Doubrovine, P. V., Burke, K., Steinberger, B., Ashwal, L. D., Trønnes, R. G., Webb, S. J., & Bull, A. L., 2014. Deep mantle structure as a reference frame for movements in and on the earth, *Proceedings of the National Academy of Sciences*, **111**(24), 8735–8740.

Tosi, N., Stein, C., Noack, L., Hüttig, C., Maierova, P., Samuel, H., Davies, D., Wilson, C., Kramer, S., Thieulot, C., et al., 2015. A community benchmark for viscoplastic thermal convection in a 2-d square box, *Geochemistry, Geophysics, Geosystems*, **16**(7), 2175–2196.

Trampert, J. & Fichtner, A., 2013. Global imaging of the earth's deep interior: Seismic constraints on (an)isotropy, density and attenuation, in *Physics and Chemistry of the Deep Earth*, pp. 324–350, John Wiley and Sons, Ltd.

Trompert, R. & Hansen, U., 1998. Mantle convection simulations with rheologies that generate plate-like behaviour, *Nature*, **395**(6703), 686–689.

Turcotte, D. L. & Oxburgh, E., 1972. Mantle convection and the new global tectonics, *Annual Review of Fluid Mechanics*, **4**(1), 33–66.

Van der Hilst, R., Widiyantoro, S., & Engdahl, E., 1997. Evidence for deep mantle circulation from global tomography, *Nature*, **386**, 578–584.

Van Der Meer, D. G., Spakman, W., van Hinsbergen, D. J., Amaru, M. L., & Torsvik, T. H., 2010. Towards absolute plate motions constrained by lower-mantle slab remnants, *Nature Geoscience*, **3**(1), 36–40.

Van Heck, H. & Tackley, P., 2008. Planforms of self-consistently generated plates in 3d spherical geometry, *Geophysical Research Letters*, **35**(19).

Van Keken, P., King, S., Schmeling, H., Christensen, U., Neumeister, D., & Doin, M.-P., 1997. A comparison of methods for the modeling of thermochemical convection, *Journal of Geophysical Research: Solid Earth*, **102**(B10), 22477–22495.

van Leeuwen, P. J., 1999. Comment on "data assimilation using an ensemble kalman filter technique", *Monthly Weather Review*, **127**(6), 1374–1377.

Van Leeuwen, P. J., 2001. An ensemble smoother with error estimates, *Monthly weather review*, **129**(4), 709–728.

Vening-Meinesz, F., 1948. *Gravity expeditions at sea 1923-1938. Vol. IV. Complete results with isostatic reduction, interpretation on the results*, Nederlandse Commissie voor Geodesie.

Vine, F. J. & Matthews, D. H., 1963. Magnetic anomalies over oceanic ridges, *Nature*, **199**(4897), 947–949.

Vynnytska, L. & Bunge, H.-P., 2015. Restoring past mantle convection structure through fluid dynamic inverse theory: regularisation through surface velocity boundary conditions, *GEM-International Journal on Geomathematics*, **6**(1), 83–100.

Walzer, U. & Hendel, R., 2008. Mantle convection and evolution with growing continents, *Journal of Geophysical Research: Solid Earth*, **113**(B9).

Wang, X., Holt, W. E., & Ghosh, A., 2015. Joint modeling of lithosphere and mantle dynamics: Evaluation of constraints from global tomography models, *Journal of Geophysical Research: Solid Earth*, **120**(12), 8633–8655.

Weertman, J. & Weertman, J. R., 1975. High temperature creep of rock and mantle viscosity, *Annual Review of Earth and Planetary Sciences*, **3**, 293.

Wegener, A., 1924. The origin of oceans and continents, *Trans. JGA Skerl. Originally published as Die Entstehung der Kontinente und Ozeane, 3rd ed.(Braunschweig: Friedrich Vieweg und Sohn, 1922). New York: Dutton*.

Weinstein, S. A. & Olson, P. L., 1992. Thermal convection with non-newtonian plates, *Geophysical Journal International*, **111**(3), 515–530.

Wessel, P., 2001. Global distribution of seamounts inferred from gridded geosat/ers-1 altimetry, *Journal of Geophysical Research: Solid Earth*, **106**(B9), 19431–19441.

Wessel, P. & Müller, R., 2015. Plate tectonics, in *Treatise on Geophysics (Second Edition)*, pp. 45 – 93, ed. Schubert, G., Elsevier, Oxford, second edition edn.

Wessel, P., Harada, Y., & Kroenke, L. W., 2006. Toward a self-consistent, high-resolution absolute plate motion model for the pacific, *Geochemistry, Geophysics, Geosystems*, **7**(3).

Wessel, P., Matthews, K. J., Müller, R. D., Mazzoni, A., Whittaker, J. M., Myhill, R., & Chandler, M. T., 2015. Semiautomatic fracture zone tracking, *Geochemistry, Geophysics, Geosystems*, **16**(7), 2462–2472.

Wiens, D. A., DeMets, C., Gordon, R. G., Stein, S., Argus, D., Engeln, J. F., Lundgren, P., Quible, D., Stein, C., Weinstein, S., et al., 1985. A diffuse plate boundary model for indian ocean tectonics, *Geophysical Research Letters*, **12**(7), 429–432.

Williams, S., Flament, N., Müller, R. D., & Butterworth, N., 2015. Absolute plate motions since 130 ma constrained by subduction zone kinematics, *Earth and Planetary Science Letters*, **418**, 66–77.

Wilson, J. T., 1963. A possible origin of the hawaiian islands, *Canadian Journal of Physics*, **41**(6), 863–870.

Wilson, J. T., 1965a. Evidence from ocean islands suggesting movement in the earth, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **258**(1088), 145–167.

Wilson, J. T., 1965b. A new class of faults and their bearing on continental drift, *Nature*, **207**, 343–347.

Worthen, J., Stadler, G., Petra, N., Gurnis, M., & Ghattas, O., 2014. Towards adjoint-based inversion for rheological parameters in nonlinear viscous mantle flow, *Physics*

*of the Earth and Planetary Interiors*, **234**, 23–34.

Wunsch, C., 2006. *Discrete inverse and state estimation problems: with geophysical fluid applications*, vol. 2, Cambridge University Press Cambridge.

Yoshida, M., 2008. Mantle convection with longest-wavelength thermal heterogeneity in a 3-d spherical model: Degree one or two?, *Geophysical Research Letters*, **35**(23).

Yoshida, M., 2010. Preliminary three-dimensional model of mantle convection with deformable, mobile continental lithosphere, *Earth and Planetary Science Letters*, **295**(1), 205–218.

Yoshida, M. & Kageyama, A., 2004. Application of the yin-yang grid to a thermal convection of a boussinesq fluid with infinite prandtl number in a three-dimensional spherical shell, *Geophysical research letters*, **31**(12).

Zahirovic, S., Matthews, K. J., Flament, N., Müller, R. D., Hill, K. C., Seton, M., & Gurnis, M., in press. Tectonic evolution and deep mantle structure of the eastern tethys since the latest jurassic, *Earth-Science Reviews*.

Zhang, N. & Zhong, S., 2011. Heat fluxes at the earth's surface and core–mantle boundary since pangea formation and their implications for the geomagnetic superchrons, *Earth and Planetary Science Letters*, **306**(3), 205–216.

Zhang, N., Zhong, S., Leng, W., & Li, Z.-X., 2010. A model for the evolution of the earth's mantle structure since the early paleozoic, *Journal of Geophysical Research: Solid Earth*, **115**(B6).

Zhang, S. & Christensen, U., 1993. Geoid anomalies from cenozoic subduction in semi-dynamical flow models including a phase boundary, *Geophysical research letters*, **20**(21), 2383–2386.

Zhang, S. & Yuen, D., 1996. Intense local toroidal motion generated by variable viscosity compressible convection in 3-d spherical-shell, *Geophysical research letters*, **23**(22), 3135–3138.

Zhong, S., 2001. Role of ocean-continent contrast and continental keels on plate motion, net rotation of lithosphere, and the geoid, *Journal of Geophysical Research: Solid Earth (1978–2012)*, **106**(B1), 703–712.

Zhong, S., Gurnis, M., & Moresi, L., 1998. Role of faults, nonlinear rheology, and viscosity structure in generating plates from instantaneous mantle flow models, *Journal of Geophysical Research: Solid Earth*, **103**(B7), 15255–15268.

Zhong, S., Zuber, M. T., Moresi, L., & Gurnis, M., 2000. Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection, *Journal of Geophysical Research: Solid Earth*, **105**(B5), 11063–11082.

Zhong, S., McNamara, A., Tan, E., Moresi, L., & Gurnis, M., 2008. A benchmark study on mantle convection in a 3-d spherical shell using citcoms, *Geochemistry, Geophysics, Geosystems*, **9**(10).

Zhong, S., Yuen, D. A., & Moresi, L. N., 2015. Numerical methods for mantle convection, *Treatise on Geophysics*, **7**, 197–222.