Old sutures and young plumes? - New geophysical investigations of the crust and upper mantle in southwestern Scandinavia

Marianne Lanzky Kolstrup

March 16, 2015

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1615

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Contents

Summary

This dissertation investigates the physical structure of the crust and upper mantle in southwestern Scandinavia using different seismological and non-seismologic geophysical methods:

1. The S-wave velocity structure of the crust and crust-mantle transition (the Moho) is derived from joint inversion of P-wave receiver functions and Rayleigh wave phase velocities below 42 seismic stations in southern Norway and Sweden.

2. The P- and S-wave velocity structure of the upper mantle in southwestern Scandinavia is investigated using multiscale, finite-frequency, seismic tomography.

3. Temperature and temperature-dependent density of the lithospheric mantle below southern Norway is estimated from thermal and isostatic modelling using several geophysical data sets as constraints (topography, Moho depth, geoid undulations).

The thesis also deals with several practical and theoretical aspects of regional finite-frequency tomography. We present a robust data processing routine for the measurement of relative finite-frequency travel-time residuals and analyze important features of relative finite-frequency Fréchet kernels in a typical regional seismic network.

The work on the crustal and uppermost mantle structure shows that the crust-mantle transition is not necessarily a simple discontinuity. We find a lateral boundary between an area with a sharp and relatively shallow Moho discontinuity and an area with a deep and gradual crust-mantle transition. This boundary is thought to be Proterozoic in age and to have survived the Caledonian orogeny.

The most important features of the upper mantle tomographic models are: a shallow (50-200 km) channel-like low-velocity anomaly between southern Norway and Denmark; a deep (150-350 km) plume-like anomaly with a diameter of ∼100 km below central southern Norway; and a belt of high velocities below Sweden that delineate the proper Fennoscandian Shield. The wiggly boundary between the low-velocity channel and the high velocities of Sweden follows zones of Carboniferous-Permian rifting and magmatism very closely. The low-velocity channel can be explained by a relatively thin and warm

lithosphere (∼100 km), whereas the high velocities below Sweden indicate both a thick and cold lithosphere, and a depleted composition of the mantle. The deep plume-like anomaly below southern Norway has structural similarities with small-scale upper-mantle plumes below central Europe, and could be caused by a similar upwelling of material with anomalous temperature and/or composition.

We suggest that episodic erosion of an originally thick mantle lithosphere below southern Norway could have weakened the lithospheric plate and triggered uplift of the area in the Cenozoic.

Acknowledgements

First of all I would like to thank my supervisor Valerie Maupin for guidance and support, and for sharing her scientific wisdom with me. This work would not have been possible without Shu-Huei Hung, to whom I am grateful for our collaboration on finite-frequency tomography. I am also indebted to Christophe Pascal for guidance on lithosphere modelling, and to Andreas Köhler and Andy Frasetto for sharing their work and data with me.

I would also like to thank my co-supervisor Christian Weidle and the TopoScandiaDeep working group for discussions and inspiration. I also thank the geologists who helped me to see my geophysical models in a geological light, especially Arild Andresen, Tom Andersen, Roy-Helge Gabrielsen and Mattias Lundmark.

Life is full of small technical problems, and a special thanks goes to Andreas Köhler for helping me with LaTeX and GMT, and to Abishek Rai for help on SAC and BASH.

Last but not least, thanks to Benny and Halfdan for waiting patiently for me to finish this work.

List of publications

This thesis consists of an introduction, three chapters with five reviewed or submitted articles, and concluding remarks.

Chapter 2

1. **Kolstrup, M. L.** and Maupin, V. 2013. A Proterozoic boundary in southern Norway revealed by joint inversion of P-receiver functions and surface waves, Precambrian Research, 238, 186-198.

Chapter 3

1. **Kolstrup, M. L.** & Maupin, V., 2014. A procedure for automatic measurement of relative finite-frequency travel-time residuals - Application to southwestern Scandinavia, Submitted to Journal of Seismology.

2. Maupin, V. & **Kolstrup, M. L.**, 2014. Insights in P- and S-wave relative traveltime tomography from analysing finite-frequency Fréchet kernels, Submitted to Geophysical Journal International.

3. **Kolstrup, M. L.**, Hung, S.-H. & Maupin, V., 2014. Multiscale, finite-frequency P and S tomography of the upper mantle in the southwestern Fennoscandian Shield, Submitted to Geophysical Journal International.

Chapter 4

1. **Kolstrup, M. L.**, Pascal, C., and Maupin, V., 2012. What compensates the topography of southern Norway? Insights from thermo-isostatic modeling, Journal of Geodynamics, 61, 105–119.

Chapter 1

Introduction

1.1 Background

Topography and tectonic processes at plate boundaries are closely linked. We find the highest present-day topography on Earth in the Himalayas where the Indian and Eurasian plates meet in collision. The Andes mountains have an equally impressive relief and originates where the oceanic Nazca plate subducts under the South American plate. Below the sea-surface, we also find impressive mountain ranges associated with sea-floor spreading, for example at the Mid-Atlantic Ridge.

The Wilson Cycle is a paradigmatic concept in plate tectonic theory and describes the opening and closing of an Atlantic-type ocean in three phases: (1) opening and spreading with creation of passive continental margins; (2) spontaneous foundering of the old and cold oceanic plate at the passive margin and initiation of a new subduction zone; (3) closing of the ocean in a continent-continent collision (e.g. Wilson, 1966; Turcotte & Schubert, 2002). The high topography of the Himalayas, the Andes and the Mid-Atlantic Ridge can be well explained by geological processes at the plate boundaries created in the Wilson Cycle.

Inside the continental plates, though, we also find areas with high relief (>1000 m), that cannot be explained by plate boundary processes. Examples of such places are South Africa, the Ural Mountains, the Trans-Antarctic Mountains, Greenland, and the Norwegian Mountains. All of these intraplate mountainous areas seem to have their own distinct characteristics and no undisputed model exists that explains their formation (e.g. Maupin et al., 2013). This dissertation focuses on southern Norway that has the highest topography of the mountains at the eastern North-Atlantic margin (Fig. 1.1).

1.1.1 The Norwegian Mountains and uplift in the North-Atlantic region

The Norwegian Mountains (or the Scandes, Fig. 1.1) follow mainly the trace of the Caledonides, a major mountain range formed about 430 to 410 Ma. After the breakup of the North-Atlantic Ocean, the remnants of the Caledonides can be found in Norway, Scotland, Ireland, Greenland and eastern North America. The Scandes Mountains are often considered as two distinct regions: the Northern and Southern Scandes. The Southern Scandes (in southern Norway) have the highest and more dome-shaped topography, whereas the Northern Scandes have a lesser topographic expression and are restricted to a more narrow belt following the coast (Fig. 1.1). Unlike the Southern Scandes that are thought to be partly supported by low density material in the uppermost mantle (Ebbing et al., 2012), the Northern Scandes seem to be isostatically balanced by their underlying crustal structure.

It has recently been suggested that the topography of the Norwegian Mountains is in fact the remnants of the Caledonides (> 400 My), shaped mostly by erosion and isostatic rebound in response to climatic factors (Nielsen et al., 2009). Contrary to this view, it is widely believed that southern Norway was eroded to sea level before breakup of the North-Atlantic Ocean, and that later uplift happened in several pulses: a regional pulse related to the breakup in the Paleocene-Eocene (66-44 Ma) and a pulse of unknown origin in the Neogene (25-3 Ma) (e.g. Anell et al., 2009; Gabrielsen et al., 2010). The accelerated uplift along the margins of the North Atlantic in Plio-Pleistocene (5-0 Ma) is most likely caused by glacial erosion and isostatic rebound (Anell et al., 2009).

Episodic burial and uplift might be an ongoing process at all passive continental margins on Earth and could be caused by a similar worldwide mechanism, as suggested by Japsen et al. (2012). For the North-Atlantic region, and especially southern Norway, there has been many more or less cite-specific hypotheses about the causes of Cenozoic uplift (e.g. Rohrman & van der Beek, 1996; Nielsen et al., 2002; Doré et al., 2008; Nielsen et al., 2009). A thorough overview is given by Anell et al. (2009), who also discuss the limitations of the different hypotheses. It seems clear, though, that climatic factors and isostatic rebound in response to erosion are very important for shaping the surface topography, but cannot account for a sufficient uplift or the domal shape of southern Norway (Anell et al., 2009, and references therein).

1.1.2 Project TopoScandiaDeep

The project TopoScandiaDeep aimed to image and model the Scandinavian crust and mantle in relation to the topography of the Scandes mountains (Maupin et al., 2013).

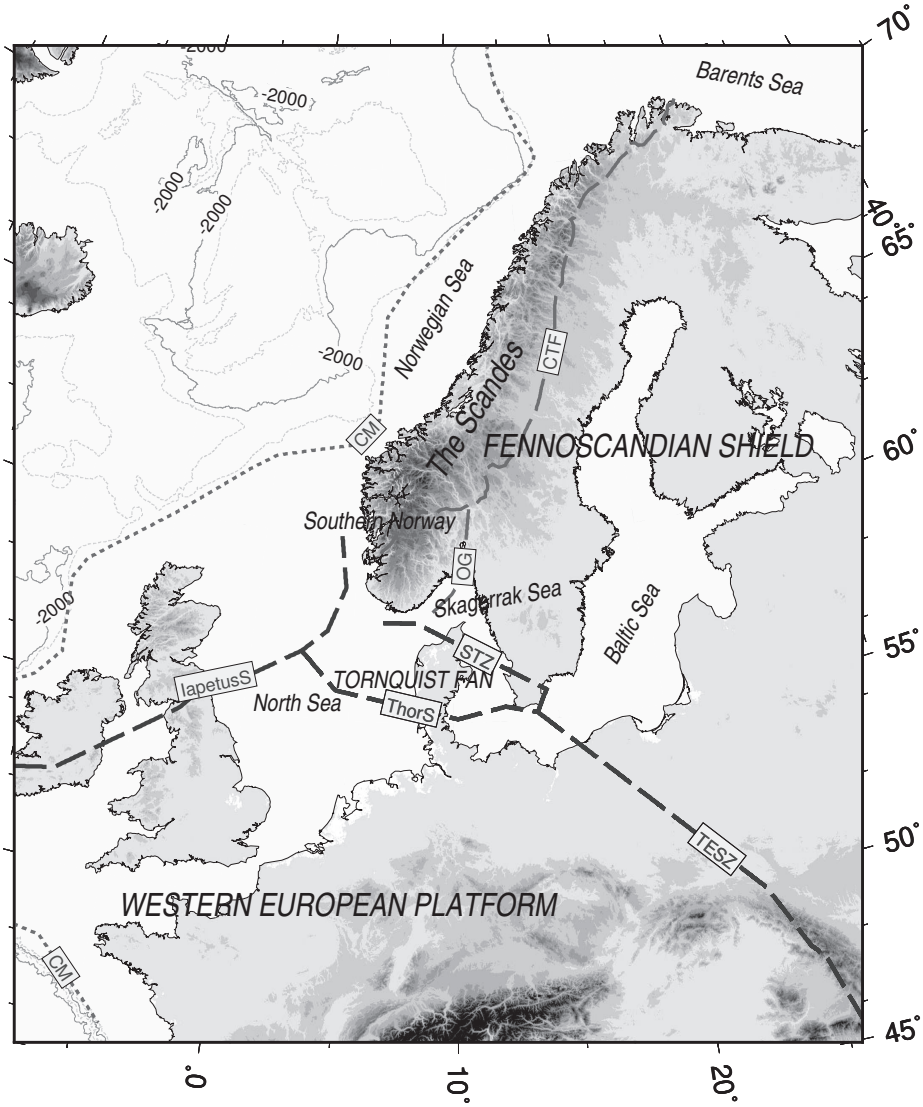


Figure 1.1: Regional topographic map centered around southern Norway. Dashed black lines indicate approximate locations of the Transeuropean Suture Zone (TESZ), the Sorgenfrei-Tornquist Zone (STZ), the Thor Suture (ThorS) and the Iapetus Suture (IapetusS). Grey dashed lines indicate the Oslo-Skagerrak Graben (OG) and the Caledonian Thrust Front (CTF). The grey dotted line depicts the 500 m bathymetry contour that approximates the continental margin (CM). Topography is indicated with shades of gray.

The project was a collaboration between the Universities of Oslo, Aarhus, Copenhagen, Bergen, Amsterdam, Karlsruhe Institute of Technology, NGU, and NORSAR, 2008−2012, and was part of the TOPOEUROPE program. The vision of TOPOEUROPE is to study the relation between the Earth's surface topography and the processes of lithosphere and mantle (Cloetingh et al., 2007). The many seismic studies part of TopoScandiaDeep include active seismics (Stratford & Thybo, 2011a,b), P-wave receiver function imaging (Frasetto & Thybo, 2013; England & Ebbing, 2012), relative and absolute P-wave tomography (Medhus et al., 2009, 2012), S-wave tomography (Wawerzinek et al., 2013), SKS-splitting and S-receiver functions (Wawerzinek, 2012) and surface wave studies (Köhler et al., 2011, 2012; Maupin, 2011). Gravity, magnetics and modelling studies were also part of the project (Ebbing et al., 2012; Gradmann et al., 2013; Maupin et al., 2013).

The seismic studies in TopoScandiaDeep have greatly improved our knowledge about the subsurface structure (a review is given by Maupin et al., 2013), but most data sets were analyzed separately and some of the models have unexplained discrepancies. As different seismic methods have different resolving power, some discrepancy between models is expected, but two major disagreements deserve an investigation:

- The Moho depths estimated in the active seismic study of Stratford & Thybo (2011a) and in the P-wave receiver function imaging of Frasetto & Thybo (2013) differ by up to 10 km in the northeastern part of southern Norway.
- The amplitude of velocity anomalies in the P- and S-wave tomographies of Medhus et al. (2012) and Wawerzinek et al. (2013), respectively, is of the same order of magnitude (approximately $\pm 1.5\%$), but S-wave velocity anomalies are expected to be about twice as large as P-wave anomalies (e.g. Goes et al., 2000). In addition, the spatial distribution of velocity anomalies differ significantly.

It is therefore an integral part of this dissertation to solve some of these discrepancies by performing joint inversions and combined analyses of several seismological data sets from project TopoScandiaDeep. Furthermore, advances in modelling of wave propagation (e.g. Dahlen et al., 2000) and tomographic model parameterizations (e.g. Chiao & Kuo, 2001; Hung et al., 2011) advocate for new tomographic inversions for the P- and S-wave velocity structure of the upper mantle in the area.

1.2 Objectives and thesis structure

The main objective of this work is to improve the seismic velocity models of the crust and upper mantle in southwestern Scandinavia, and especially southern Norway, by combining information from different geophysical data sets and by using complementary methods.

Secondly, to put these models into their geological and tectonophysical context, in order to shed light on the role of the mantle and mantle processes in shaping the topography of southern Norway and the elevated passive continental margins of the North-Atlantic Ocean.

The motivation for the work is described in this introductory chapter. **Chapter 2** is concerned with the structure of the crust and Moho discontinuity in southern Norway and Sweden, and contains a paper on the joint inversion of P-wave receiver functions and surface waves for the S-wave velocity of the crust and uppermost mantle. The characteristics of the crust-mantle transition (or Moho discontinuity) are quantified and analyzed in the geological context of southern Norway.

Chapter 3 comprises a trilogy of papers concerning relative finite-frequency tomography of the upper mantle:

- In the first manuscript we present an automated procedure for the measurement of relative finite-frequency travel-time residuals using a combination of cross-correlation methods. The procedure is used to compute P- and S-wave travel-time residuals from data recorded in a network in southwestern Scandinavia. Travel-time corrections are computed as an integral part of the processing procedure, using a compilation of crustal structure in southwestern Scandinavia that includes the crustal models from Chapter 2.

- The second paper deals with theoretical aspects of relative finite-frequency Fréchet kernels and the computation of finite-frequency and ray-theoretical crustal corrections. These theoretical insights are important for the validation of several assumptions and practical approaches in relative finite-frequency tomography.

- The third paper is devoted to P- and S-wave tomographic imaging of southwestern Scandinavia using multiscale, finite-frequency tomography. The resolution power and effect of different parameterizations is investigated. Ray-theoretical tomographies are performed to ease comparison with previous studies. The variation and covariation of P- and S-wave velocities is analyzed and their tectonophysical implications explored.

Chapter 3 also includes a final section that discusses the deep plume-like structure in the tomographic images and compares it to small-scale upper-mantle plumes in Europe.

Chapter 4 stands apart as the only non-seismological work. The chapter contains a paper about thermal and isostatic modelling of the lithosphere in southern Norway using several data sets (long wavelength topography, Moho depth, geoid undulations), with synthetic prediction of S-wave velocities that takes into account composition and

anelastic attenuation. The synthetic velocities are extensively used in the discussion of the tomographic results in Chapter 4. The method was adapted and refined from work of Pascal (2006) in the master's thesis of Kolstrup (2010) but modelling and extended analysis was done with updated data sets as part of this PhD-dissertation.

Chapter 5 summarizes the main findings in this dissertation and concludes with an outlook for future research and suggestions for dynamic modelling.

Chapter 2

Crust and Moho structure from P-wave receiver functions and surface waves

2.1 Introduction

Seismic waves contain information about both the source that created them, and the structure in the part of the Earth they traveled through. The Earth is a complex medium with a strong layering that creates multiple echoes of the seismic waves, and a seismic record from a far distant earthquake therefore contains information from many places in the Earth. When we want to investigate the structure immediately below the seismic station, i.e. the crust and uppermost mantle, we therefore have to extract the information about the near-receiver structure from a complicated seismic record.

P-wave receiver functions and Rayleigh wave phase velocities are two types of seismic data that provide information about the crust and upper mantle structure below a seismic station. P-wave receiver functions are estimated from a P-wave coda by deconvolving the vertical component of a seismogram from the corresponding radial or transverse component (Langston, 1979). Radial P-wave receiver functions consist of only the direct P-wave pulse and the amplitude of P-to-S conversions of the direct P-wave and its multiples at interfaces beneath the station. The most important conversion is the P-to-S conversion at the Moho discontinuity (or crust-mantle transition). The delay between the direct P-wave and the Moho conversion depends on the depth to the interface and the average S-wave velocity of the crust. The amplitude of the conversion is sensitive to the velocity contrast at the Moho and hence to the finer structure of the crust-mantle transition. P-wave receiver functions are therefore very sensitive to velocity contrasts, but a good average velocity model is needed to properly locate the depth of the interfaces.

Rayleigh waves are surface waves that travel with phase velocities mostly depending on the average S-wave velocities in different depth ranges. Rayleigh wave phase velocities contain the information about the average velocity structure that is lacking in the P-wave receiver functions. The two data sets are therefore strongly complementary and their joint inversion has become widely used (e.g. Özalaybey et al., 1997; Du & Foulger, 1999; Julià et al., 2000, 2003, 2005; Yoo et al., 2007).

In this chapter, we present 42 one-dimensional S-wave velocity profiles of the crust and uppermost mantle in southern Norway and Sweden, and discuss their geological implications. The velocity models are also used as input data in the compilation of crustal structure in southwestern Scandinavia in Chapter 3.2, and for the crustal travel-time corrections in Chapter 3.4. The main findings are summarized in Chapter 5 with an outlook for future research based on ideas from this work.

2.2 Manuscript - Precambrian Research

2.3 Appendix - Supplementary figures

Chapter 3

Relative finite-frequency travel-time residuals, Fréchet kernels, and multiscale tomography of the upper mantle

3.1 Introduction

This chapter contains three papers concerning seismic travel-time tomography. In seismic tomography, measurements of seismic waves from primarily natural sources, such as earthquakes, are used to make models of the deep Earth. One of the major obstacles in tomographic imaging is that earthquakes occur only in specific places around the world and mostly at plate boundaries. Moreover, the placement of receivers is also limited to mainly continental regions and islands. The limited coverage of both sources and receivers results in specific challenges in seismic tomography.

The input data to classic body-wave tomography are travel times of P or S waves. The theoretical travel time in a standard Earth model (e.g. IASP91 or AK135) is subtracted from the observed travel time to give the travel-time residual. The residual is then a measure of how much the travel time differs from the average travel time on Earth, and hence contains information about how much the subsurface structure differs from the average Earth. When many measurements of travel-time residuals are collected at many seismic stations, they can be inverted to models of seismic velocity structure, mostly given as percentage deviations with respect to a standard Earth model.

The theoretical basis of linearized travel-time tomography is the formulation

$$\delta t = \iiint_{\oplus} K(\boldsymbol{x}) \delta s(\boldsymbol{x}) d^3\boldsymbol{x}, \tag{3.1}$$

where δt is the observed travel-time residual, and $K(\boldsymbol{x})$ is the sensitivity (or Fréchet) kernel that relates δt to perturbations in P- or S-wave slowness $s(\boldsymbol{x})$, or velocity $v(\boldsymbol{x}) = 1/s(\boldsymbol{x})$, at every point $\boldsymbol{x}$ in the region $\bigoplus$.

The Earth is a continuous medium, but to link the observations of travel time to the subsurface properties, we need to make a discretization and parameterization of the Earth model. The simplest model parameterization possible is to divide the study volume into blocks of varying slowness or velocity. More often, the model is parameterized in terms of functions that are specified at grid points in the study volume. Discretization, parameterization, and numerical integration of equation 3.1 then leads to a data equation that relates the slowness perturbations (or velocity anomalies) of the discretized study region to the travel-time measurements:

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{m} \tag{3.2}$$

where $\boldsymbol{d}$ is the data vector containing the N travel time measurements δt_i and $\boldsymbol{m}$ is the model vector containing the M slowness or velocity parameters. The elements $\boldsymbol{G}_{ij}$ of the matrix $\boldsymbol{G}$ represents the sensitivity of the ith datum with respect to the slowness perturbation at the jth part of the study region.

As mentioned previously, the distribution of sources and receivers is limited in seismic tomography and this poses specific problems. The main one is that it renders the solution of the inverse problem in equation 3.2 non-unique, i.e. the data coverage is not dense enough to derive a unique model of the subsurface. Hence, a priory information (or regularization) like minimum-norm damping or smoothing is used to constrain the range of possible models, and usually only a preferred model is presented.

The three articles in this chapter are devoted to different parts of equation 3.1, and its discretized version in equation 3.2. The first article is about the measurement of the travel-time residuals, δt_i, in the vector $\boldsymbol{d}$. The second manuscript focuses on some theoretical aspects of the kernels, $K(\boldsymbol{x})$, and their implications on both travel-time measurement and travel-time tomography. The third paper is concerned with the tomographic inversion of the travel-time residuals for the model, $\boldsymbol{m}$, of seismic velocity perturbations.

More specifically, the first paper in section 3.2 presents a processing procedure for the measurement of **relative**, **finite-frequency** travel-time residuals. **Finite-frequency** in this case means that the travel-time residuals are measured in several, narrow frequency bands, making them sensitive to subsurface structures at different spatial scales. The measurement of finite-frequency travel times has become essential due to improvements in wave propagation theory (e.g. Dahlen et al., 2000; Hung et al., 2000) that is now able to take into account the frequency-dependent nature of seismic waves at an affordable com-

putational cost. A better term than **finite-frequency**, is perhaps **multiple-frequency**, which is also used by some authors (e.g. Tian et al., 2009; Sigloch, 2011). The naming of finite-frequency travel times has historical reasons, as older wave propagation theories (ray theory) assume the seismic wave has infinite frequency and follows optical ray trajectories.

Another term that needs an explanation is the use of the word **relative** in seismic tomography. Relative in this case means that the relative travel-time residuals are given with respect to their average in a local or regional seismic network. In practice, this means that the travel times are demeaned for each event (earthquake). This has the advantage that most of the influence of the source and of the structure outside the study area is removed, but also renders the inverted tomographic model relative to an unknown average model. Hence, relative travel-time tomographies are usually regional tomographies, and absolute travel-time tomographies (using absolute travel-time residuals) are usually global tomographies.

The relativization of the data also implies that the kernels $K(\boldsymbol{x})$ should be demeaned in the same manner as the data. The second paper in section 3.3 analyzes relative and absolute kernels in a typical network configuration, and also investigates the effect of the crustal (near-receiver) structure on finite-frequency measurements of travel time.

In the third paper in section 3.4 we use the **multiscale**, finite-frequency tomographic method of Hung et al. (2011) to invert for the P- and S-wave velocity structure of southwestern Scandinavia. The term **multiscale** hints to the type of model parameterization used in the tomographic inversion (Chiao & Kuo, 2001; Chiao & Liang, 2003). The model $\boldsymbol{m}$ is expanded in terms of wavelet basis functions that decomposes the model into a hierarchy of wavelet coefficients at various scales. The coarsest level used here is at a length-scale larger than the entire study region and the finest level is at a length-scale of 25 km. The main advantage of the multiscale parameterization is that it reduces the need for regularization, and allows for a data-adaptive, spatially varying resolution (Hung et al., 2011).

The papers in this chapter have been ordered thematically and not chronologically. Hence, some of the conclusions drawn in the second paper (section 3.3) were not taken into account for the tomography in the third paper (section 3.4). The consequences are briefly discussed in Chapter 5 along with a summary of the results.

3.2 Manuscript - Submitted to Journal of Seismology

Journal of Seismology manuscript No.
(will be inserted by the editor)

A procedure for automatic measurement of relative finite-frequency travel-time residuals

Application to southwestern Scandinavia

Marianne L. Kolstrup · Valérie Maupin

Received: date / Accepted: date

Abstract We present a data processing routine to compute relative finite-frequency travel-time residuals using a combination of the multichannel cross-correlation method (MCCC) and the iterative cross-correlation and stack (ICCS) algorithm. The routine has been tailored for measuring travel times in several frequency bands. Automated data rejection is performed in several steps, ensuring a robust estimate of the array stack, and the ICCS algorithm replaces manual picking of phase onsets before MCCC analysis. The data processing routines are written in Matlab and handle data in several formats. Travel-time corrections for ellipticity, topography and crustal structure are computed alongside the travel-time residuals with options for using global crustal models or local compilations of crustal structure. The routine is applied to data from a network in southwestern Scandinavia.

Body waves, Data processing, Finite-frequency travel times, Travel-time residuals

1 Introduction

Measurement of relative travel-time residuals has formed the backbone of regional teleseismic tomography since the pioneering work of Aki et al (1977). Regional tomographic models have had a considerable impact on our understanding of the upper mantle. The tremendous increase in data quantity and the improvement in inversion techniques have greatly increased tomographic resolution power in the recent years. The quality of the models is however fully dependent on the quality of the travel-time residuals the models are based upon: better measurements will ultimately translate to better models, and good routines for travel-time measurements are therefore crucial.

Relative travel times are in most cases defined as relative to the mean arrival time of each event, or in some cases relative to a reference station. This removes uncertainties regarding source location and timing, and also most of the influence of the structure outside the region of interest (Aki et al, 1977; Masson and Trampert, 1997). Relative travel times are usually measured using a cross-correlation method. These methods exploit the similarity of seismic waveforms measured at different stations to estimate precise travel times in a seismic network (e.g. Bungum and Husebye, 1971; VanDecar and Crosson, 1990; Sigloch and Nolet, 2006; Pavlis and Vernon, 2010).

The multichannel cross-correlation method (MCCC) of VanDecar and Crosson (1990) is a robust method for relative travel-time measurement that also provides quantitative uncertainty estimates. In the MCCC method, relative delay times Δt_{ij} between all pairs of stations stations i, j are measured using cross correlation of seismograms. These relative delays are then inverted in a least-squares sense for the optimized relative arrival times, t_i, at each station, under the constraint of zero mean, i.e. $\sum t_i = 0$. The equation residuals between observed relative delays and predicted delays are calculated as $\mathrm{res}_{ij} = \Delta t_{ij} - (t_i - t_j)$. The standard deviation of the equation residuals associated with each trace, σ_i, provides an estimate of the timing uncertainty associated with the relative arrival time t_i (VanDecar and Crosson, 1990).

Marianne L. Kolstrup
Department of Geosciences, University of Oslo, Post Box 1047, 0316 Oslo, Norway
Tel.: +47-48032053
Fax: +47-22854215
E-mail: m.l.kolstrup@geo.uio.no

Inherent in all cross-correlation methods is a risk of cycle skipping during analysis (e.g. VanDecar and Crosson, 1990; Chevrot, 2002), meaning that an earlier or later wiggle in the seismogram is interpreted as the arrival of the phase of interest. Initial alignment of traces is therefore necessary such that a narrow window can be used in the search for the maximum of the cross-correlation function. Predictions from theoretical calculations of arrival times (e.g. Crotwell et al, 1999) are usually not sufficient to solve this issue (e.g. Lou et al, 2013), as the travel-time residuals are commonly larger or similar to wave periods in the highest frequency range. Manual picking of approximate phase arrival times is therefore necessary before cross-correlation analysis.

The iterative cross-correlation and stack (ICCS) algorithm (Lou et al, 2013) was recently developed to replace the unavoidable initial phase marking part of the MCCC procedure, and is part of the AIMBAT tool (Automated and Interactive Measurement of Body-wave Arrival Time) (Lou et al, 2013). The ICCS algorithm cross-correlates each trace with the array stack and determines the time lag of the traces with respect to the stack. Each trace is corrected for its time lag and the array stack is updated. This process is repeated iteratively until convergence is reached and all traces are aligned. The MCCC method can then be applied to the aligned traces to slightly refine the measurements of arrival time and most importantly to provide quantitative uncertainty measurements (Lou et al, 2013).

The procedures of both VanDecar and Crosson (1990) and Lou et al (2013), and other recent tools (e.g. Chevrot, 2002; Pavlis and Vernon, 2010), are all tailored for measuring high frequency signals. With the development of finite-frequency theory (Dahlen et al, 2000; Hung et al, 2000) it has become essential to measure travel times in several narrow frequency bands. We find from experience that the measurement of travel times in more narrow high-frequency bands increases the risk of cycle skipping during analysis in both the ICCS and MCCC algorithms, especially for lower quality events from great epicentral distances ($> 75°$). We therefore present a set of processing routines for finite-frequency measurements that remedy this problem. We apply the ICCS and MCCC algorithms to the lowest frequency band first and then to successively higher bands, along with several rounds of automated data rejection that increase the robustness of the initial stack in the ICCS algorithm. In addition, travel-time corrections for ellipticity, topography and crustal structure are computed alongside the relative travel-time residuals.

We apply the procedure to data from a network in southwestern Scandinavia (Weidle et al, 2010), and compile a detailed crustal model for travel-time corrections.

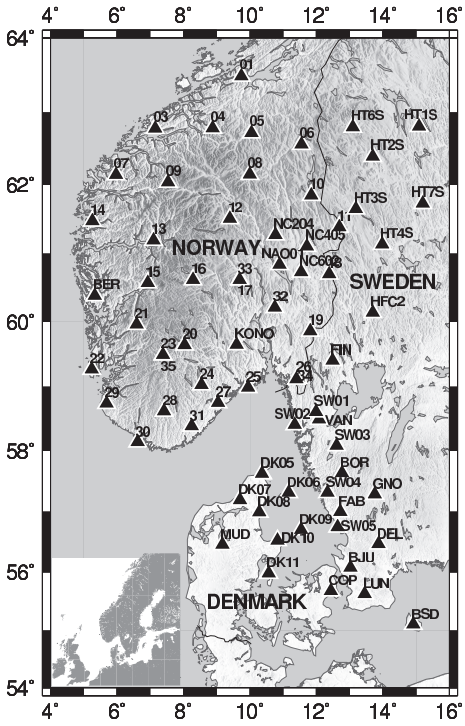


Fig. 1 Topographic map of the study area with location of seismological stations and national borders. Numbers xx are short for stations NWGxx in the MAGNUS network.

2 Measurement of relative travel times

We use seismological data recorded by the temporary MAGNUS network between September 2006 and June 2008 (Weidle et al, 2010), by the temporary DANSEIS network from April 2008 to June 2008, by the temporary CALAS stations (Medhus et al, 2009), and by permanent stations in the study area (Fig. 1).

Direct P and S phases were studied from earthquakes with epicentral distances of $30 - 95°$ (Fig. 2). P-wave phases were studied on the vertical component and S-wave phases on the transverse component.

To have an even azimuthal coverage of the study region it is not possible to only use the best quality events, as events are fewer and of lesser magnitude in the southern and western quadrant (Fig. 2), increasing

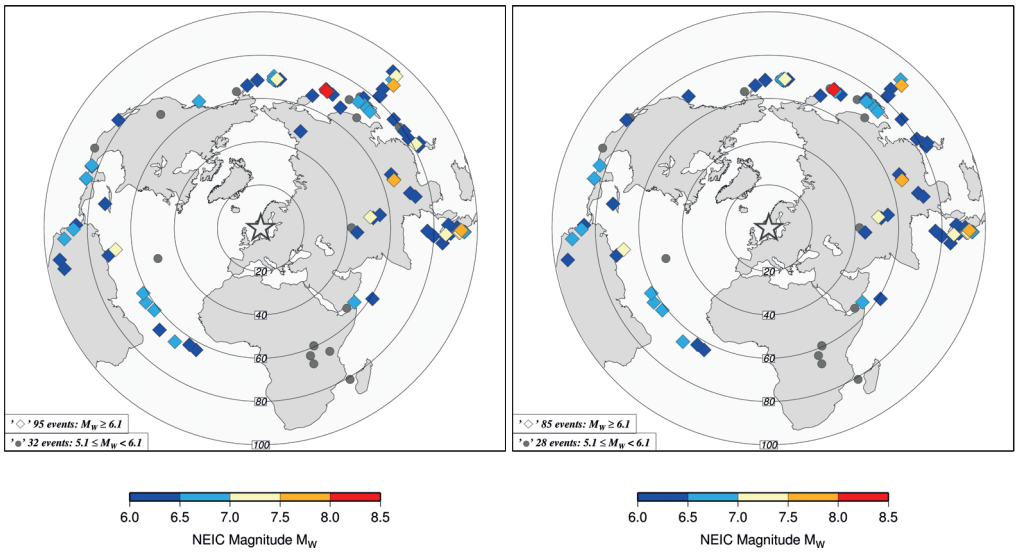


Fig. 2 Maps of earthquake sources for P- and S-wave travel times. The events displayed have given travel-time residuals in at least one frequency band. Left: Events for P-waves. Right: Events for S-waves.

the need for a robust algorithm for travel-time measurement.

In short, the data processing sequence is performed in four steps:

1. Preprocessing: Filtering, calculation of theoretical phase arrival time, and windowing according to the frequency content of the band and the phase in question.
2. A first rejection of data based on quality criteria calculated from the envelope of the data ensures a robust array stack.
3. Alignment of traces by the ICCS algorithm (Lou et al, 2013). Additional rejection of incoherent or noisy data.
4. Relative delay times Δt_{ij} by multichannel cross-correlation and least-squares inversion for zero-mean relative travel times t_i (MCCC, VanDecar and Crosson, 1990). Calculation of uncertainty estimates of the relative travel times.

The steps $1-4$ are performed first on the lowest frequency band, and then on successively higher frequency bands. Each step is described in detail in sections 2.1 to 2.4. Travel-time residuals and corrections for ellipticity and crustal structure are calculated after Step 4 and will be presented in the next section. Various statistical parameters and map views of results are written to a postscript file that can be used for visual quality control. We use data from a 6.4 M event at the Andreanoff Islands ($\Delta = 68°$, back azimuth 5°), 15th April 2008, to illustrate the data processing.

2.1 Preprocessing

Trends and offsets are removed from the recordings, and the individual instrument responses are removed by deconvolution. All data are initially bandpass filtered ($0.0025-5.0$ Hz) using a window of several hundred seconds length and a second-order, zero-phase Butterworth filter, and resampled to a common sampling rate. P-waves are then bandpass filtered in the frequency ranges $0.3-0.125$ Hz ($33-8$ s) and $0.5-2$ Hz ($2-0.5$ s) to isolate the secondary noise peak in southern Norway around 0.2 Hz (5 s). S-waves are bandpass filtered in the ranges $0.03-0.077$ Hz ($33-13$ s) and $0.077-0.125$ Hz ($13-8$ s).

The theoretical arrival time in AK135 (Kennett et al, 1995) of the phase of interest is predicted using taup (Crotwell et al, 1999) in a Matlab implementation (Euler, 2014). A shorter window is chosen around the predicted arrival time such that the window contains at least one full period of the target phase. The upper panels in Fig. 3a and b show P-wave traces in the low- and high-frequency bands, respectively, aligned according to the theoretical arrival time (Crotwell et al, 1999).

The arbitrary choice of time axis has no influence on the final results.

2.2 Data rejection based on the envelope of data

The first round of data rejection excludes traces that have a maximum of their envelope eight times larger than the median of all the envelope maxima for a particular event. This makes sure that strong outliers do not ruin the computation of the mean stack and its envelope (Fig. 3a and b, middle panels). The signal to noise ratio of all the data is estimated from the ratio between the peak and average value of the mean envelope (Fig. 3a and b, middle panels), and events are accepted for analysis based on this ratio. The actual threshold value depends on the frequency band and window length but should generally be above 1.5.

In a new round of data rejection, individual traces are excluded from further analysis based on the difference between the envelope of the trace and the mean envelope of the stack (Fig. 3a and b, lower panels). The threshold for keeping traces is a maximum relative difference of around 1.0 to 2.5, where the actual value depends on the frequency content of the band, with a higher threshold value for higher frequencies. When a trace is rejected from analysis, the mean envelope (Fig. 3, middle panels) and the envelope differences (Fig. 3, lower panels) are recalculated. This procedure is continued iteratively until no more data are rejected, and effectively removes the most noisy traces, ensuring a robust estimate of the array stack before applying the ICCS algorithm (Lou et al, 2013).

2.3 ICCS algorithm tailored for several frequency bands

To avoid cycle skips in the multichannel cross-correlation analysis (MCCC, VanDecar and Crosson, 1990) it is essential that traces are initially aligned, either by manual picking or by an automatic procedure (VanDecar and Crosson, 1990; Lou et al, 2013), and that noisy or incoherent traces are excluded from the analysis. We use the iterative cross-correlation and stack algorithm (ICCS, Lou et al, 2013) to initially align traces, but do not pick an absolute arrival time of the stack as done in the AIMBAT procedure (Lou et al, 2013), since our target is relative arrival time.

In the ICCS algorithm, each individual trace is correlated with the stack and the time lag associated with the maximum of the cross-correlation function is found. Then each individual trace is aligned according to the time lag and relative to the stack in an iterative procedure where the stack is updated for each iteration. Fig. 4 shows the high-frequency P-wave traces before and after alignment by ICCS.

An additional advantage of applying the ICCS algorithm is the computation of quality criteria (Pavlis and Vernon, 2010; Lou et al, 2013) that are used to exclude noisy or incoherent data, which may have escaped the first round of data rejection in Step 2 (section 2.2). After alignment by ICCS, the cross-correlation coefficient and mean spectral coherence between each trace and the stack are calculated. These two parameters have values ranging from 0 to 1, where low values are diagnostic of low quality data. A weighted average of the two is computed and traces with a value lower than a user-defined cut off (usually about 0.5) are excluded. This procedure rejects data with a significantly different shape than the array stack. Fig. 5 shows the low- and high-frequency P-wave stacks after alignment by the ICCS algorithm with a few green-colored incoherent traces that will be rejected from further analysis.

Inherent in both the ICCS and the MCCC algorithms is a risk of cycle skipping in the cross-correlation analysis when the period of the data and the travel-time residuals are of the same order of magnitude. For the low-frequency bands of both the P- and S-waves, the travel-time residuals are much smaller than the period of the wave, and the risk of cycle skipping in the pick of cross-correlation maxima is therefore very low. In contrast, there is a significant risk of cycle skipping in the ICCS algorithm in case of high-frequency low-quality data. This mostly happens in the first iterations where the array stack is less robust, because it is calculated from traces aligned according to the theoretical arrival time. With a poor stack there will be little difference between local and global maxima in the cross-correlation functions, and the algorithm might therefore misplace a few traces, possibly permanently. This will lead to a few outlying estimates of travel-time residuals for the higher frequency bands.

This problem of cycle skipping in the ICCS algorithm is probably enhanced in our data set because most events have epicentral distances $> 70°$, giving relatively strong attenuation of the high-frequency content of the data. For events with shorter epicentral distances, even with magnitudes $M < 6.0$, the problem of cycle skipping in the ICCS algorithm is almost nonexistent.

To remedy the risk of cycle skipping in the first iterations of the ICCS algorithm for the higher frequency bands, we use the low-frequency travel-time residuals from the MCCC analysis in Step 4 (section 2.4) as input time lag to align the traces in the stack. Only reliable travel-time residuals with small standard devia-

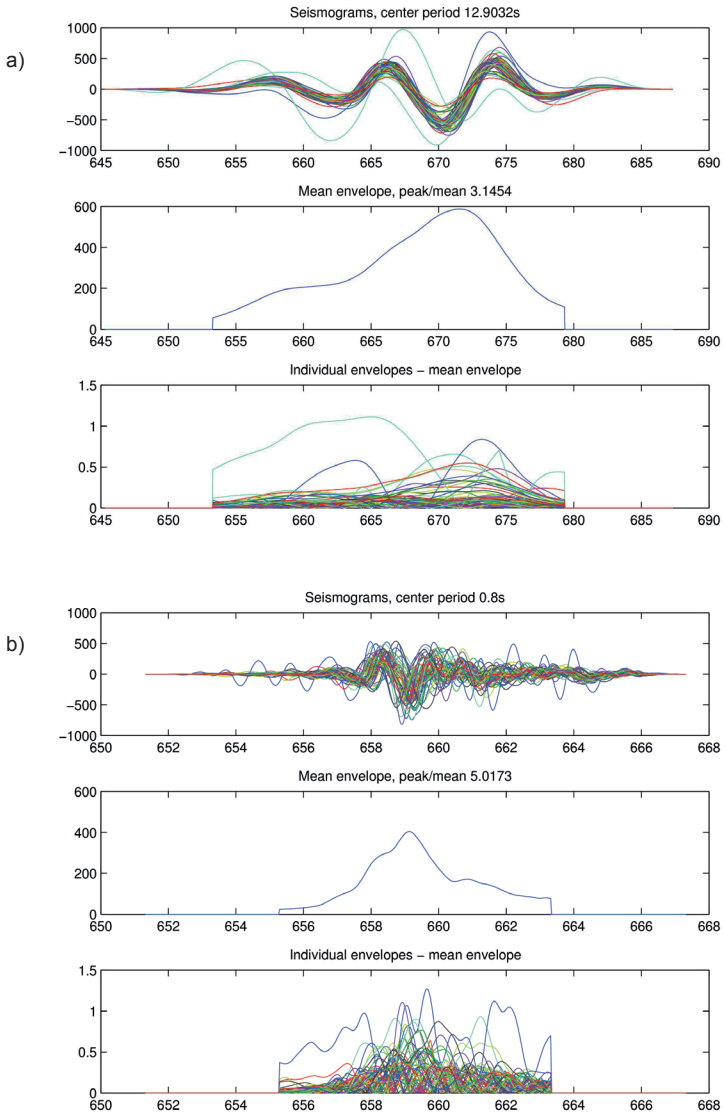


Fig. 3 Elements of the analysis performed in Step 2 of the data processing. a) P-wave low-frequency band. b) P-wave high-frequency band. Upper panel in a) and b) shows traces accepted after data rejection in Step 2. Middle panel in a) and b) shows the mean envelope of the accepted traces. Lower panel in a) and b) shows the difference between individual envelopes and the mean envelope.

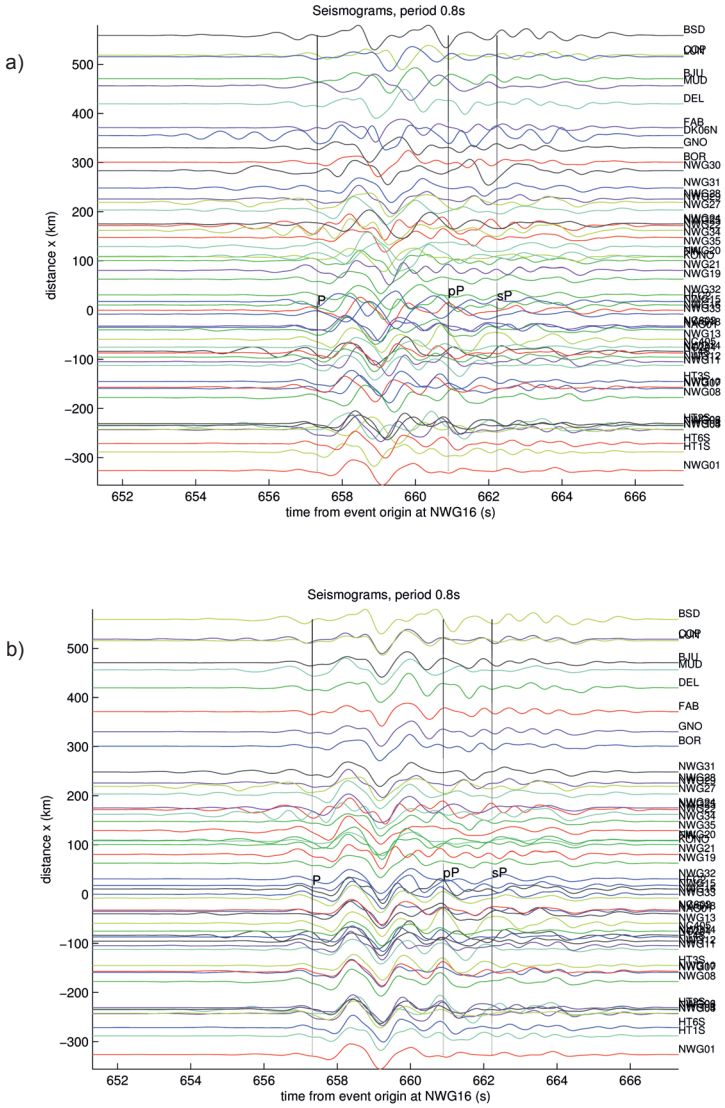


Fig. 4 High-frequency P-wave traces before and after alignment by ICCS. a) Seismograms after alignment by theoretical arrival times from AK135 and data rejection in Step 2 (section 2.2). b) Seismograms after alignment by ICCS and additional rejection of incoherent traces in Step 3 (section 2.3).

tions are used as input time lags. This ensures a better initial stack in the high-frequency bands and reduces the number of iterations in the ICCS algorithm.

The effect of the improved initial stack is especially important for the more band limited S-wave data. Fig. 6 shows the improvement of the initial stack, when the low-frequency S-wave residuals are used to align the traces of the high-frequency band, compared to using the theoretical arrival time from AK135.

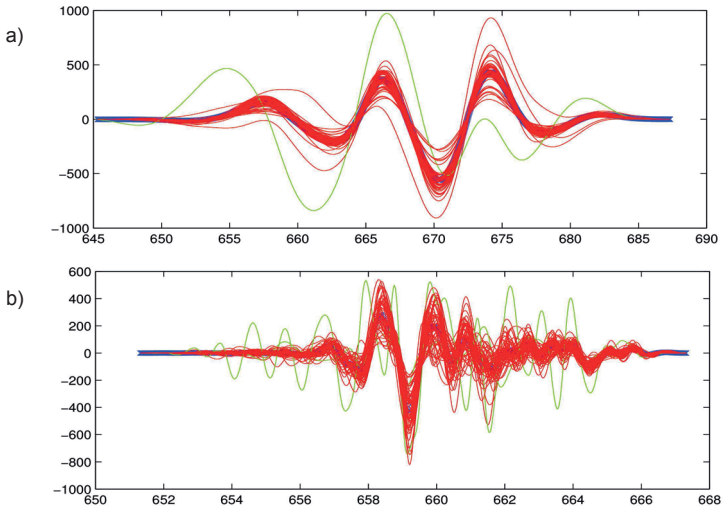


Fig. 5 P-wave stack and individual traces after alignment by ICCS. a) Low-frequency band. b) High-frequency band. The array stack is colored blue, the accepted traces red, and green traces are the ones that will be rejected from analysis before Step 4 (section 2.4).

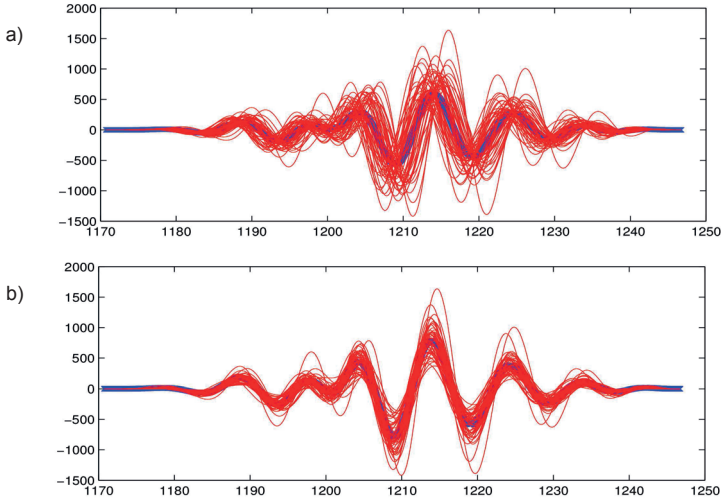


Fig. 6 High-frequency S-wave stack and traces in the first iteration of the ICCS algorithm. The array stack is colored blue and the individual traces red. a) High-frequency traces aligned according to the theoretical arrival time predicted with AK135. b) High-frequency traces aligned according to the low-frequency travel-time residuals from the MCCC analysis in Step 4 (section 2.4).

2.4 Multichannel cross-correlation (MCCC) analysis

In multichannel cross-correlation analysis, the relative delay times Δt_{ij} are measured between all pairs of stations, i, j, using cross correlation. These relative delays are then inverted for the optimized relative arrival times, t_i, as outlined in section 1. If the mean value is subtracted from the final time lags determined by the ICCS algorithm, they are approximately equal to the relative travel-time residuals calculated by the MCCC algorithm (Lou et al, 2013). The main reason for applying MCCC to the traces aligned by ICCS is therefore the calculation of uncertainty estimates, σ_i, from the equation residuals res_{ij}.

The program outputs various statistics from the MCCC analysis, such as σ_i in map plots, and histograms of res_{ij} and of the cross-correlation coefficients. These parameters can be used to determine if an event has a too high noise-level to be included in the data set, and to detect station problems.

The major difference between our implementation and the original Fortran/SAC codes of VanDecar and Crosson (1990) is storage of data. To keep track of which traces to use for each event, we use a set of indexes for all data and a set of indexes for the data accepted for analysis. A mapping between the sets of indexes is then used in the matrix building (eqn. 5-6, VanDecar and Crosson, 1990). An advantage of doing the processing in Matlab and using a mapping between subsets of the data, is that we only store the original data and none of the intermediate processing steps.

Working with high-frequency P waves, VanDecar and Crosson (1990) found that equation residuals $\mathrm{res}_{ij} > 0.5$ s are usually the result of cycle skipping in the cross-correlation analysis. For the low-frequency P-waves we find that a threshold of 0.8 s is appropriate to target the (very few) cycle skips and for the S-waves we use larger thresholds of 2 s for the low frequency band and 1.5 s for the high frequency band. The cross-correlation functions associated with large equation residuals are recalculated in a smaller window near the time lag predicted by the least-squares solution $(t_i - t_j)$, as in VanDecar and Crosson (1990). The data set of relative delay times (Δt_{ij}) is updated with the new measurements before a weighted reinversion using the cross-correlation coefficients between traces as weights. The differences between the unweighted relative travel times and the weighted relative travel times are very small.

3 From relative travel times to travel-time residuals

The relative travel-time residuals are calculated by subtracting the demeaned theoretical arrival times (section 2.1) for each event from the measured relative travel times t_i. Relative travel times and travel-time residuals are plotted for the Andreanoff Islands event in Fig. 7.

The map plots of travel times and travel-time residuals are made automatically for each event and provide a practical way of evaluating the results visually, making the existence of outliers very obvious. At one station, for example, the polarity of the EW componenent was reversed and gave anomalously negative S-wave residuals for events in the eastern quadrants, but not for events in the western quadrants. Had maps only been plotted manually for a few events, this type of station problems might not have been discovered.

The map plots are also useful in evaluating the importance of different phases. In our data set, most events have epicentral distances between 70° and 95°, making the PcP and ScS phases arrive relatively close to the direct arrivals. When isolating the PcP phase and calculating the residuals, the maps of PcP residuals clearly showed that in most cases, the coda of the direct P-wave dominated over the PcP-wave, leaving an imprint of a direct P plane wavefront in the PcP travel-time residuals.

3.1 Ellipticity correction

Relative travel-time corrections for the ellipticity of the Earth (Kennett and Gudmundsson, 1996) are computed using the Matlab implementation of Euler (2014). These corrections are in general very small compared to the residuals (-0.05 to 0.09 s for P-waves and -0.10 to 0.14 s for S-waves).

3.2 Crustal correction

Relative crustal and topographic corrections from 3D models like Crust1.0 (Laske et al, 2013) or detailed local models are also computed using the package of Euler (2014), but we expand the crustal model structure to include a uniform mantle layer from the Moho to 50 km depth in order to have a common reference depth for the crustal corrections.

As the relative travel-time residuals are zero mean, the corrections also have to be zero mean. The amount of data accepted for analysis differ slightly from event to event, and the time span of recording also differ from station to station. Hence, the mean of the corrections

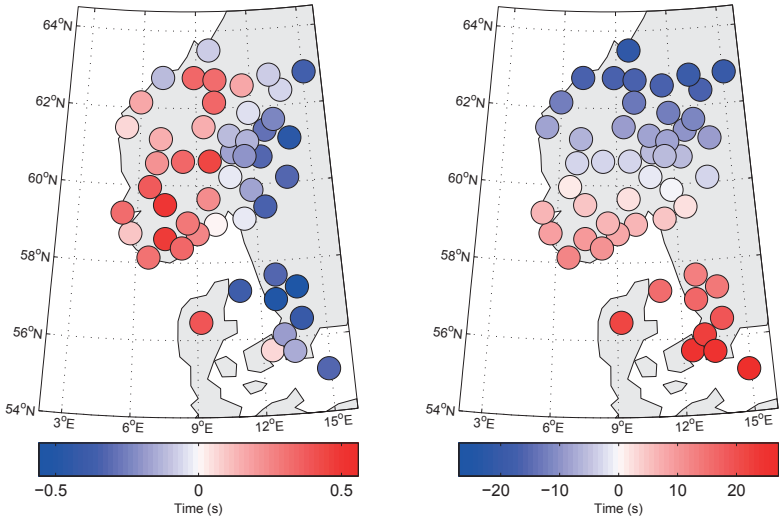


Fig. 7 P-wave low-frequency travel times and travel-time residuals from the 6.4 M event at the Andreanoff Islands ($\Delta = 68^\circ$, back azimuth $= 5^\circ$), 15th April 2008. Left window: relative travel-time residuals. Right window: relative travel times. Note the different colour scales for the travel times and for the residuals.

has to be computed for each event and subtracted from the absolute corrections. It is therefore very convenient to compute all corrections at the same time as computing the travel-time residuals, and this is done routinely in our data processing.

Our study area comprises strongly varying crust in coastal areas, making a global compilation like Crust1.0 (Laske et al, 2013) inappropriate. We therefore use a local crustal model compiled from a range of sources, especially the work of Stratford and Thybo (2011); Lassen and Thybo (2012) and Kolstrup and Maupin (2013). The local crustal model has two sediment layers, three crustal layers and a mantle layer.

The resulting absolute crustal corrections (including topography) can be compared to the relative corrections in Fig. 8. Low-frequency P-wave residuals are shown before and after correction in Fig. 9.

A mantle layer of variable thickness and velocity can easily be added to the model and used to compute corrections for the uppermost mantle structure, if this structure is well known. Information about uppermost mantle velocities are available from noise and surface wave studies in the study area (Köhler et al, 2012, 2014), but due the trade off between Moho depth and subcrustal velocities, we cannot use these uppermost mantle velocities along with a crustal model derived from other studies.

An alternative would be to use only this surface-wave-based model to calculate the corrections for the crust and upper mantle. The lateral resolution of such a model is however lower than what we can get from a compilation of active seismics and receiver function analysis. As variation in Moho depth is clearly the dominant contributor to the corrections, we prefer to apply a procedure that ensures the use of the most accurate estimates of Moho depth below each station.

3.2.1 Details on the compilation of the crustal model

Moho depths in Norway, and in Sweden close to the border, are compiled from receiver function analysis (Svenningsen et al, 2007; Frasetto and Thybo, 2013; Kolstrup and Maupin, 2013) and active seismics (Stratford and Thybo, 2011). Southern Norway has a homogeneous crust (Stratford and Thybo, 2011; Loidl et al, 2014) with lower average velocities than elsewhere in southwestern Scandinavia. Average velocities and layering are taken from the models of Stratford and Thybo (2011) and Kolstrup and Maupin (2013). In northeastern southern Norway, we use the latest results concerning Moho depths (Kolstrup and Maupin, 2013; Loidl et al, 2014) and assign a relatively fast lower crust for this area, with velocities taken from the S-wave velocity profiles of Kolstrup and Maupin (2013) and scaled for V_P. These latest results resolve a former discrepancy of approximately 10 km between the Moho depth inferred from active seismics (Stratford and Thybo, 2011), and

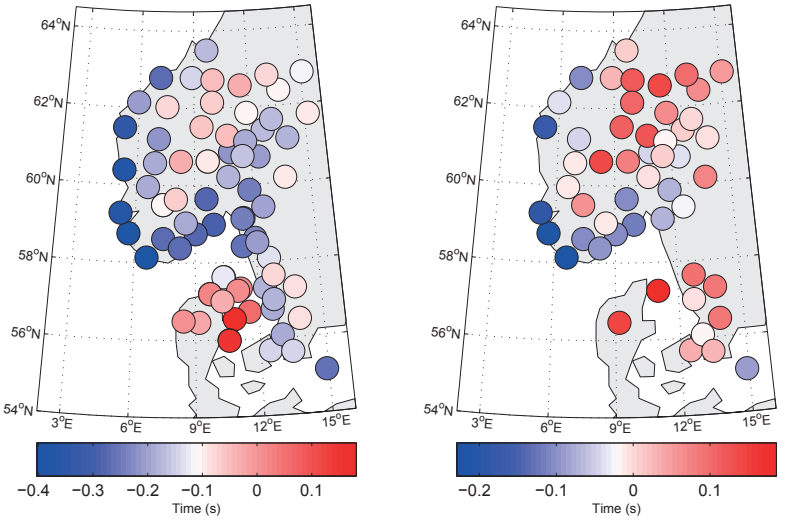


Fig. 8 P-wave crustal corrections. Left window: absolute crustal corrections calculated for all stations in the network. Right window: relative crustal corrections (Andreanoff Islands event).

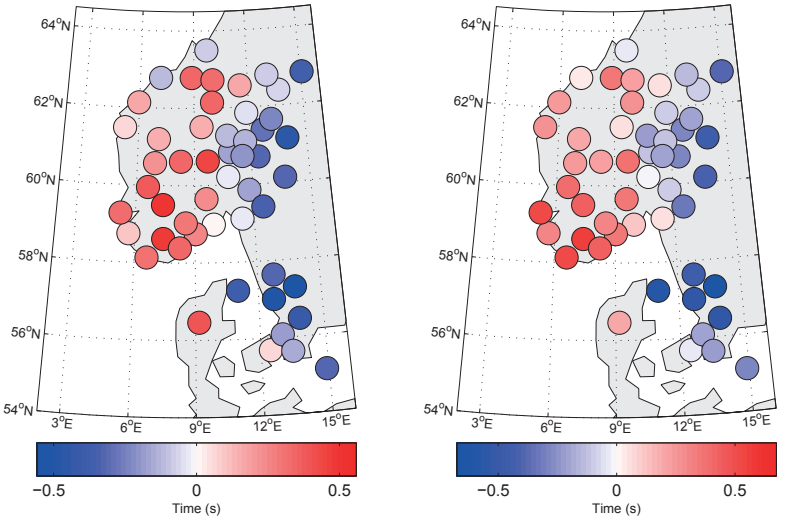


Fig. 9 Travel-time residuals with and without crustal corrections for the P-wave low-frequency band (Andreanoff Islands event). Left window: raw travel-time residuals. Right window: crustal corrected travel-time residuals.

receiver functions (Frasetto and Thybo, 2013; Kolstrup and Maupin, 2013).

Moho depth and thickness of sediment layers in the Danish and Swedish areas are taken from the recent compilation by Lassen and Thybo (2012). Average velocities are calculated for the younger sediments from the profiles in Sandrin and Thybo (2008), with values depending on the thickness of the layers ($V_P = 2.9$ km/s, $V_P = 3.6$ km/s, and $V_P = 4.0$ km/s for layers of thickness 0-2 km, 2-4 km, and > 4 km). The Paleozoic sediments are assigned a P-wave velocity of 5.0 km/s (Medhus et al, 2012).

3.3 Average travel-time residuals

The variation of travel-time residuals with back azimuth provides preliminary information on the spatial distribution of velocity anomalies prior to inversion. Figs 10 and 11 present average P-wave residuals in the high- and low-frequency bands, respectively, and for varying back azimuth. The back-azimuthal ranges are chosen according to natural clusters in the data (Fig. 2). Averages are computed for each station using the inverse of the estimated standard deviations as weights. Residuals are also presented for far distant events (80 to 95°) that arrive very steeply beneath the stations and hence give mostly information about the structure vertically below the station.

Events from the northeast to the southeast (Figs 10a,b,c and 11a,b,c) result in a 2D pattern of early residuals (negative) in the east, and late residuals (positive) in the west. Contrary to this, events from the southwestern and northwestern quadrants result in a more complicated 3D residual pattern. Residuals on the southwestern coast of Norway and Denmark are negative for events from the southwest (Figs 10d and 11d), indicative of a high velocity anomaly in the southwestern part of the study area. Events in the northwestern quadrant (Figs 10e and 11e) give a shorter wavelength pattern of residuals changing from positive to negative values along the western coast of Norway from north to south. The data from far distant events with almost vertical propagation depict a wiggly boundary between east and west (Figs 10f and 11f).

Comparison of the spatial patterns in the low- and high-frequency residuals (Figs 10 and 11), especially in the southern part of the study area, illustrates the different spatial sensitivity of travel times measured in different frequency ranges (Dahlen et al, 2000; Hung et al, 2000), but we do not notice any significant difference in the amplitude of the residuals in the two frequency ranges.

4 Discussion

Our data processing routines are designed for automatic processing with no storage of intermediate processing steps. The processing can first be tested out on a range of representable events from the data set, in order to experiment with different frequency bands, seismic phases, window lengths, data rejection thresholds etc. When suitable parameters are found, the codes can be run automatically, and the results can be inspected visually for a short quality control. As the only human interference is the choice of various parameters and a quality check at the end, there is little risk of drift in the arrival time picking and a high degree of objectiveness in the data rejection procedures.

Medhus et al (2012) and Wawerzinek et al (2013) used basically the same database as we do, but employed different methods for measuring travel times. The P-wave residuals of Medhus et al (2012) were estimated by a combination of cross-correlation measurements and final manual picking of high-frequency P waves ($0.125 - 4$ Hz), whereas the S-wave travel-time residuals of Wawerzinek et al (2013) were found by manual picking of S-waves on magnified waveforms in a lower frequency band ($0.03 - 0.125$ Hz). An advantage of using the same cross-correlation procedure on both the P- and S-wave data set, is that uncertainty estimates are calculated in a consistent manner for both data sets. In a later inversion for seismic velocities these uncertainties can be used to weight the data objectively, suppressing the influence of individual noisy data and low quality events. The two data sets can be inverted either jointly for e.g. V_Φ and V_S, or for V_P and V_S separately. If the same inversion technique and model parameterization is used to invert for V_P and V_S, the resulting models can be compared quantitatively as in e.g. Chou et al (2009) and Hung et al (2011).

Yang and Shen (2006) and Ritsema et al (2009) show that corrections for the crustal structure are frequency-dependent and that the same crustal correction cannot be applied to different frequency bands. The difference is mostly in a different mean level of the corrections at different frequencies. Hence, for the relative crustal corrections, where the mean level of the corrections is removed, there is little difference between the relative crustal corrections in the different frequency bands and the ray-theoretical crustal corrections. We therefore use ray-theoretical crustal corrections as an approximation of the true frequency-dependent crustal corrections.

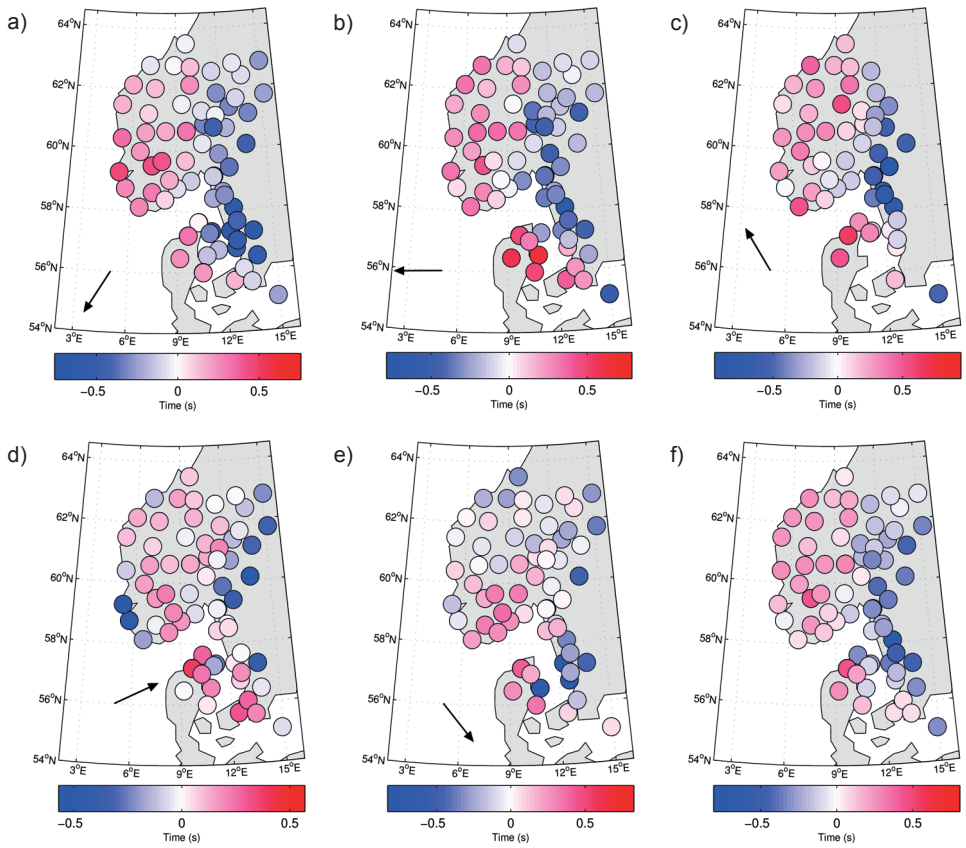


Fig. 10 Average P-wave travel-time residuals in the high-frequency band from different back-azimuthal ranges, corrected for topography, crustal structure and ellipticity. The arrow shows the average direction of the incoming waves.

5 Conclusion

We present an automated processing procedure combining the ICCS and MCCC algorithms and tailored for estimating relative travel-time residuals in several frequency bands. Using the low-frequency travel-time residuals as input to the ICCS algorithm for higher frequency bands increases the reliability of the automated alignment in ICCS and reduces the risk of cycle skipping in the MCCC analysis, especially for distant low quality data.

The data processing routine was applied to data from a network in southwestern Scandinavia and the spatial pattern of P- and S-wave travel-time residuals was analyzed for varying back azimuths in different frequency bands. The application of the same cross-correlation procedure on both the P- and S-wave data set results in a consistent calculation of uncertainty estimates, facilitating a quantitative comparison of the data sets and their joint inversion. Tomographic models from inversion of the P- and S-wave data sets are presented in Kolstrup et al (2014, submitted).

Matlab scripts and example data are available by contacting the authors.

Acknowledgements This work has been done in the framework of the ESF EUROCORES TOPO-EUROPE Program 07-TOPO-EUROPE-FP-014. MAGNUS waveforms were recorded with the mobile KArlsruhe BroadBand Array (KABBA) of the Karlsruhe Institute of Technology, Germany as well as with permanent stations of the NORSAR array and the Norwegian National Seismological Network. We thank Andy Frassetto for providing us with the DANSEIS data (University of Copenhagen). Financial support for the MAGNUS experiment was provided by the Universities of Aarhus, Copenhagen, Karlsruhe and Oslo as well as NORSAR. We are grateful to Garrett Euler (Wash-

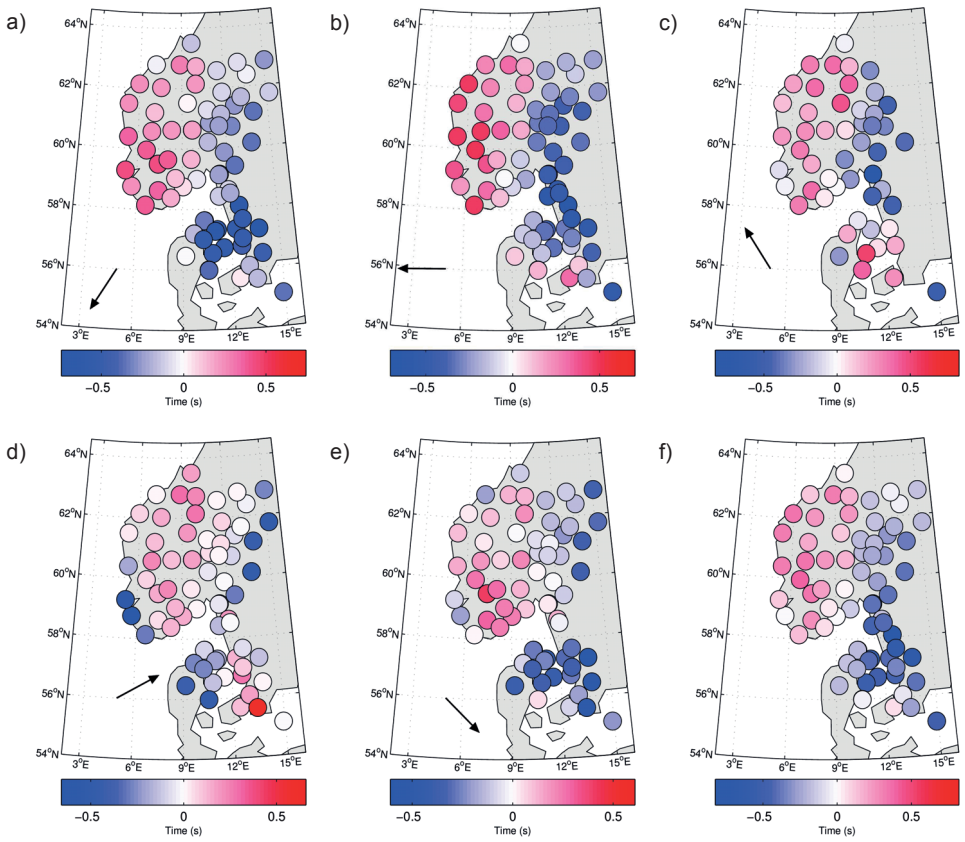


Fig. 11 Same as Fig. 10 but for the low-frequency band.

ington University) for open access to the Project SEIZMO toolbox. Figures have been prepared using the Generic Mapping Tools (Wessel, P. and Smith, W.H.F., 1998) and M_Map (Pawlowicz, 2005).

References

Aki K, Christoffersson A, Husebye ES (1977) Determination of the three-dimensional seismic structure of the lithosphere. Journal of Geophysical Research 82(2):277–296

Bungum H, Husebye E (1971) Errors in time delay measurements. Pure and Applied Geophysics 91(1):56–70

Chevrot S (2002) Optimal measurement of relative and absolute delay times by simulated annealing. Geophysical Journal International 151(1):164–171

Chou HC, Kuo BY, Chiao LY, Zhao D, Hung SH (2009) Tomography of the westernmost Ryukyu subduction zone and the serpentinization of the fore-arc mantle. Journal of Geophysical Research 114:B12,301

Crotwell HP, Owens TJ, Ritsema J (1999) The TauP Toolkit: Flexible seismic travel-time and ray-path utilities. Seismological Research Letters 70(2):154–160

Dahlen F, Hung SH, Nolet G (2000) Fréchet kernels for finite-frequency traveltimes – I. Theory. Geophysical Journal International 141(1):157–174

Euler G (2014) Project SEIZMO. URL http://epsc.wustl.edu/∼ggeuler/codes/m/seizmo/

Frasetto A, Thybo H (2013) Receiver function analysis of the crust and upper mantle in Fennoscandia – isostatic implications. Earth and Planetary Science Letters 381:234–246

Hung SH, Dahlen F, Nolet G (2000) Fréchet kernels for finite-frequency traveltimese – II. Examples. Geophysical Journal International 141(1):175–203

Hung SH, Chen WP, Chiao LY (2011) A data-adaptive, multiscale approach of finite-frequency, traveltime tomography with special reference to P and S wave data from central Tibet. Journal of Geophysical Research 116:B06,307

Kennett B, Gudmundsson O (1996) Ellipticity corrections for seismic phases. Geophysical Journal International 127(1):40–48

Kennett BLN, Engdahl ER, Buland R (1995) Constraints on seismic velocities in the Earth from traveltimes. Geophysical Journal International 122(1):108–124

Köhler A, Weidle C, Maupin V (2012) Crustal and uppermost mantle structure of southern Norway: Results from surface wave analysis of ambient seismic noise and earthquake data. Geophysical Journal International 191(3):1441–1456

Köhler A, Maupin V, Balling N (2014) Surface wave tomography of the Tornquist Fan region, SW Scandinavia, using ambient noise and earthquake data. Work in preparation

Kolstrup ML, Maupin V (2013) A Proterozoic boundary in southern Norway revealed by joint-inversion of P-receiver functions and surface waves. Precambrian Research 238:186–198

Kolstrup ML, Hung SH, Maupin V (2014) Multiscale, finite-frequency P and S tomography of the upper mantle in the southwestern Fennoscandian Shield. Submitted to Geophysical Journal International

Laske G, Masters G, Ma Z, Pasyanos M (2013) Update on CRUST1.0 – A 1-degree global model of Earth's crust. Geophysical Research Abstracts 15:2658

Lassen A, Thybo H (2012) Neoproterozoic and Palaeozoic evolution of SW Scandinavia based on integrated seismic interpretation. Precambrian Research 204:75–104

Loidl B, Behm M, Thybo H, Stratford W (2014) Threedimensional seismic model of crustal structure in Southern Norway. Geophysical Journal International 196(3):1643–1656

Lou X, van der Lee S, Lloyd S (2013) AIMBAT: A Python/Matplotlib tool for measuring teleseismic arrival times. Seismological Research Letters 84(1):85–93

Masson F, Trampert J (1997) On ACH, or how reliable is regional teleseismic delay time tomography? Physics of the Earth and Planetary Interiors 102(1):21–32

Medhus A, Balling N, Jacobsen B, Weidle C, England R, Kind R, Thybo H, Voss P (2012) Upper-mantle structure beneath the Southern Scandes Mountains and the Northern Tornquist Zone revealed by P-wave traveltime tomography. Geophysical Journal International 189(3):1315–1334

Medhus AB, Balling N, Jacobsen BH, Kind R, England RW (2009) Deep-structural differences in southwestern Scandinavia revealed by P-wave travel time residuals. Norwegian Journal of Geology 89:203–214

Pavlis GL, Vernon FL (2010) Array processing of teleseismic body waves with the USArray. Computers & Geosciences 36(7):910–920

Pawlowicz R (2005) M_Map: A mapping package for Matlab. URL http://www.eos.ubc.ca/∼rich/map.html

Ritsema J, Van Heijst H, Woodhouse J, Deuss A (2009) Long-period body wave traveltimes through the crust: implication for crustal corrections and seismic tomography. Geophysical Journal International 179(2):1255–1261

Sandrin A, Thybo H (2008) Deep seismic investigation of crustal extensional structures in the Danish Basin along the ESTRID-2 profile. Geophysical Journal International 173(2):623–641

Sigloch K, Nolet G (2006) Measuring finite-frequency body-wave amplitudes and traveltimes. Geophysical Journal International 167(1):271–287

Stratford W, Thybo H (2011) Seismic structure and composition of the crust beneath the southern Scandes, Norway. Tectonophysics 502(3):364–382

Svenningsen L, Balling N, Jacobsen B, Kind R, Wylegalla K, Schweitzer J (2007) Crustal root beneath the highlands of southern Norway resolved by teleseismic receiver functions. Geophysical Journal International 170(3):1129–1138

VanDecar J, Crosson R (1990) Determination of teleseismic relative phase arrival times using multichannel cross-correlation and least squares. Bulletin of the Seismological Society of America 80(1):150–169

Wawerzinek B, Ritter J, Roy C (2013) New constraints on the 3D shear wave velocity structure of the upper mantle underneath Southern Scandinavia revealed from non-linear tomography. Tectonophysics 602:38–54

Weidle C, Maupin V, Ritter J, Kvarna T, Schweitzer J, Balling N, Thybo H, Faleide JI, Wenzel F (2010) MAGNUS - A Seismological Broadband Experiment to Resolve Crustal and Upper Mantle Structure beneath the Southern Scandes Mountains in Norway. Seismological Research Letters 81(1):76–84

Wessel, P and Smith, WHF (1998) New improved version of the generic mapping tools released. EOS Transactions AGU 79(33):579

Yang T, Shen Y (2006) Frequency-dependent crustal correction for finite-frequency seismic tomography. Bulletin of the Seismological Society of America

96(6):2441–2448

3.3 Manuscript - Submitted to Geophysical Journal International

Geophys. J. Int. () , 1–19

Insights in P- and S-wave relative traveltime tomography from analysing finite-frequency Fréchet kernels

Valérie Maupin and Marianne Lanzky Kolstrup
Department of Geosciences, University of Oslo, Post Box 1047, 0316 *Oslo, Norway*
Tel: +47-22855756
Fax: +47-22854215
email: valerie.maupin@geo.uio.no

Accepted date, Received date, in original form date

SUMMARY
Regional body wave tomography, also called ACH-tomography, is the inversion of relative traveltime residuals of teleseismic body waves measured at regional networks. We analyse the characteristics of the finite-frequency Fréchet kernels for P and S waves for this kind of tomography. Using a simplified geometry enables us to use the complete Green's function in the expression of the Fréchet kernels and analyse terms which are usually neglected, like the importance of the near-field terms and the P-wave traveltime sensitivity to shear-wave velocity variations. By comparing the kernels of the relative residuals and absolute ones, we show that relative residuals have a reduced sensitivity to heterogeneities of large dimensions, and that this reduction is a generalisation of the fact that the average model is not recovered in ACH-tomography. This sensitivity reduction affects equally short- and long-period residuals. We show in addition the presence of a sensitivity reduction at large depth for the long-period waves. Kernels and reflectivity impulse responses of the crust are used to analyse if crustal corrections should be made frequency-dependent in finite-frequency regional tomography. We find that in most cases the frequency dependence due to reverberations is substantial, and that in many realistic network configurations ray theory is unlikely to be well-appropriate to compute crustal corrections for the long-period waves. We also find that the lateral dimensions of the crust affecting the traveltimes is frequency-dependent and reaches, at long periods, 50 km for sedimentary basins and 100 km for Moho depth.

Key words: Seismic tomography, Body waves, Theoretical seismology

1 INTRODUCTION

Since the pioneering work of Aki et al. (1977), body wave tomography using relative arrival times of teleseismic mantle phases at regional networks, often called ACH tomography after the three authors of this first article, has provided an enormous amount of information about upper mantle structure. Techniques have been developed to improve traveltime measurement and inversion, a major step forward being the development of finite-frequency Fréchet kernels (Dahlen et al. 2000; Hung et al. 2000) and the associated measurement of frequency-dependent traveltimes and amplitudes (Sigloch & Nolet 2006). Images of finite-frequency kernels have greatly contributed to a better understanding of how traveltimes and amplitudes are sensitive to variations in seismic wave velocities in the Earth. The kernels have been shown for a variety of waves and wave combinations. Except for an example with two stations in Hung et al. (2000), they have however not been shown and analysed in the framework of regional body-wave tomography, despite that they have been used in several such tomographies (Hung et al. 2004; Yang et al. 2006; Tian et al. 2009; Hung et al. 2011). The particular geometry of regional body wave tomographies, with many receivers close to each other, an inversion box located right beneath the network at relatively close proximity compared to the source-station distance, and the usage of relative traveltimes, does provide a number of particularities to the associated finite-frequency Fréchet kernels that are worth studying in order to shed some light on the limitations but also on the resolution power of these tomographies.

First of all, a central processing procedure in regional body wave tomography consists of measuring not the waves absolute arrival times, but their relative times at the stations of the network. The residuals thus obtained are with respect to an unknown mean traveltime (VanDecar & Crosson

1990), having the well-known consequence of rendering tomographic models relative to an unknown average model (Aki et al. 1977; Lévêque & Masson 1999; Bastow 2012). This is valid for the original methodology, based on onset times, and for the more recent methodlogy based on finite-frequency correlation times. This also implies that kernels of relative traveltimes and not kernels of absolute traveltimes should be used in the tomography. We analyse here how relative kernels differ from absolute ones. This difference highlights the resolution power of relative residuals to heterogeneities of large lateral dimensions. We also analyse the sensitivity of relative traveltimes to deep heterogeneities in order to assess the influence of velocity variations below the inversion model.

The finite-frequency Fréchet kernels developed by Dahlen et al. (2000) and Hung et al. (2000) are computed assuming several assumptions, in particular that only far-field forward scattering contributes significantly to traveltime variations. The far-field forward-scattering approximation is less valid for heterogeneities close to the source or close to the station. The longest periods used in finite-frequency tomography are typically 20 to 30 s, implying wavelengths that are not small compared to the distance from the inverted model to the stations, and one may question the validity of the far-field approximated kernels in such a case. In order to evaluate the error done by using the far-field approximation, we analyse therefore the difference between the far-field approximated kernels and full-wave kernels.

Another consequence of the far-field forward-scattering approximation is to restrict the P-wave traveltimes dependence to the one on compressional-wave speed, and the S-wave traveltimes dependence to the one on shear-wave speed. In the general case, P-wave traveltimes are also influenced by shear-wave speed variations. Zhang & Shen (2008) have analysed this influence and found it substantial, in particular close to the source and receiver. Their study is however geared towards global tomography and we complement it here with a similar analysis in the framework of regional studies.

Due to the sparcity of the ray-path crossings just below the stations, regional body-wave tomography cannot usually resolve the upper part of the model, in particular the crustal thickness and seismic velocity. The crustal influence on traveltimes is however substantial and needs to be corrected for (e.g., Waldhauser et al. 2002). Crustal corrections are usually computed using ray theory (e.g., Tian et al. 2007), providing frequency-independent results. Finite-frequency tomography plays on the different sensitivities of the different frequencies to the structure. It is well-established that the reverberations in the crust affect the traveltimes at different frequencies in different ways, and that the crustal corrections should therefore be frequency-dependent not to bias absolute traveltime finite-frequency tomographies (Ritsema et al. 2009; Yang & Shen 2006; Obayashi et al. 2004). A proper correction of the traveltimes based on the reflectivity of the crust assuming no lateral variation has therefore been proposed by these authors. We analyse here if the demeaning of the data done in finite-frequency relative bodywave tomography reduces the need for the frequency-dependent correction in regional tomographies. The amount of frequency dependence for relative traveltime crustal corrections due to reverberations in crustal models with varying Moho depth and sedimentary thickness is therefore briefly analyzed. In addition, by studying the sensitivity of the frequency-dependent traveltimes to velocity variations at Moho and crustal depths we analyze how 1-D models appropriate to compute crustal corrections should be derived.

After a presentation of the methodology adopted to compute the Fréchet kernels and of the very simplified geometry that we have used, we study all aspects described above by analysing SH- and P-wave traveltime kernels at short and long periods from mantle to upper crustal depth.

2 METHODOLOGY FOR THE COMPUTATION OF THE KERNELS

2.1 Sensitivity kernels

Traveltime anomalies at a given station, δt, are commonly measured by finding the maximum of the cross-correlation of the observed wavefield with a reference one. For small wavefield variations, δt may be calculated using the following integral (Dahlen et al. 2000; Zhao et al. 2005):

$$\delta t = \frac{-\int_{t_1}^{t_2} \dot{u}_l(\vec{s},t)\,\delta u_l(\vec{s},t)\,\mathrm{d}t}{P_l} \tag{1}$$

where $u_l(\vec{s},t)$ is the reference wavefield and $\delta u_l(\vec{s},t)$ is the difference between the observed and reference wavefields, both recorded along the e_l unit vector at the location of the station $\vec{s}$ in the time interval $[t_1,t_2]$. The time derivative is represented by $\dot{}$ and P_l is a normalisation factor equal to

$$P_l = \int_{t_1}^{t_2} |\dot{u}_l(\vec{s},t)|^2 \mathrm{d}t \tag{2}$$

How this traveltime anomaly δt depends on structural variations can be expressed at first order using sensitivity kernels. Restricting the structural variations to those in compressional wave velocity $\delta\alpha$ and shear wave velocity $\delta\beta$, the traveltime anomalies can be expressed as an integral over the Earth volume:

$$\delta t = \int (K_\alpha(\vec{x})\delta\alpha(\vec{x}) \; + K_\beta(\vec{x})\delta\beta(\vec{x}))\,\mathrm{d}x^3 \tag{3}$$

where the kernels $K_\alpha(\vec{x})$ and $K_\beta(\vec{x})$ are (Zhang & Shen (2008), after Zhao et al. (2005)):

$$K_\alpha(\vec{x}) = \frac{2\,\rho(\vec{x})\,\alpha(\vec{x})}{P_l} \int_{t_1}^{t_2} \dot{u}_l(\vec{s},t) \int_{-\infty}^{+\infty} e_l \cdot \{(\nabla \cdot G^T(\vec{s},t;\vec{x},\tau)(\nabla\cdot\vec{u}(\vec{x},\tau)\}\,\mathrm{d}\tau\mathrm{d}t \tag{4}$$

$$K_\beta(\vec{x}) = \frac{2\,\rho(\vec{x})\,\beta(\vec{x})}{P_l} \int_{t_1}^{t_2} \dot{u}_l(\vec{s},t) \int_{-\infty}^{+\infty} e_l \cdot \{\nabla G^{213}(\vec{s},t;\vec{x},\tau) : [(\nabla\vec{u})(\vec{x},\tau) + (\nabla\vec{u})^T(\vec{x},\tau)] - 2(\nabla\cdot G^T(\vec{s},t;\vec{x},\tau))(\nabla\cdot\vec{u}(\vec{x},\tau))\}\,\mathrm{d}\tau\mathrm{d}t \tag{5}$$

where ρ is the density in the reference space and $\vec{u}$ and G are respectively the displacement of the reference field and the Green's function in the reference space. $()_{213}$ represents the transposition of the first and second indices in a third-order

tensor. These equations express how traveltime variations can be seen as the result of the interference of a reference field with the scattered waves created by the interaction of the same reference field with point scatterers distributed in the volume of the Earth.

The kernels defined above are for variations in α and β, assuming no variation in density. It is important to realize that the variations in α and β need to be independent of each other, meaning that a variation in α implies a variation in the Lamé parameter λ only, while a variation in β implies a variation in both λ and rigidity μ such that $\delta\lambda + 2\,\delta\mu = 0$.

The reference wavefield may be a P-wave, an S-wave, or any combination of them, both in the K_α and K_β kernels. In addition, G may includes far-field P waves, far-field S waves and the mid and near-field terms. For an S-wave reference field, as $\nabla \cdot \vec{u}$ is zero, K_α is identically zero, expressing that the traveltimes of S waves are independent of variations in α (or, more precisely, of λ only), as already noted by Zhang & Shen (2008). For a P-wave reference wavefield, both K_α and K_β are non-zero, expressing that the traveltime of an incoming P wave is affected by variations in both α and β, although the influence of variations in β goes to zero in the limit of ray theory. Zhang & Shen (2008) have shown that the kernels K_β for P waves may not be negligible, and we will reanalyse this element in the context of relative tomography. For P waves we will therefore analyse two kernels, and denote them K_α^P and K_β^P. The non-zero kernel for S waves will be denoted K_β^S.

2.2 Sensitivity kernels for relative traveltimes

Images of finite-frequency kernels for traveltimes have been shown for many source-receiver geometries and for simple as well as complex wavepaths (Hung et al. 2000). Sensitivity kernels for relative traveltimes have not been shown as much attention. The kernel for the traveltime difference between two stations can simply be calculated by taking the difference between the kernels of the absolute traveltimes at the two stations (Dahlen et al. 2000). Relative phase arrivals used in body wave tomography are usually not the simple difference between traveltimes at two stations, but the result of an optimization procedure based on finding the traveltimes that satisfy in a least-squares sense the arrival time differences measured between all pairs of available stations, under a constraint of zero mean (VanDecar & Crosson 1990). In case of error-free data, this procedure is equivalent to removing, for each event, the average arrival time at all stations from the arrival time at each station. The kernels to be used for the inversion of relative traveltimes are therefore the kernels at each individual station minus the average of all kernels for a particular event.

$$K'_c(\vec{x}) = K_c(\vec{x}) - \overline{K_c}(\vec{x}) \tag{6}$$

where c is α or β, K'_c are the kernels for the relative traveltimes and $\overline{K_c}$ is the mean kernel:

$$\overline{K_c}(\vec{x}) = \frac{1}{N} \sum_{stations} K_c(\vec{x}) \tag{7}$$

where N is the number of stations. In the following, we denote K_c as absolute kernels and K'_c as relative kernels.

Differential kernels for a wave recorded at two neighboring stations have been shown for body waves (Hung et al. 2000) and for surface waves (Chevrot & Zhao 2007; de Vos et al. 2013), showing clearly how the sensitivity of differential measurements concentrate close to the stations with alternative positive and negative lobes. As far as we know, relative kernels have not been studied. We expect them to concentrate the sensitivity close to the network, but not to show the same alternation of positive and negative lobes as the differential kernels.

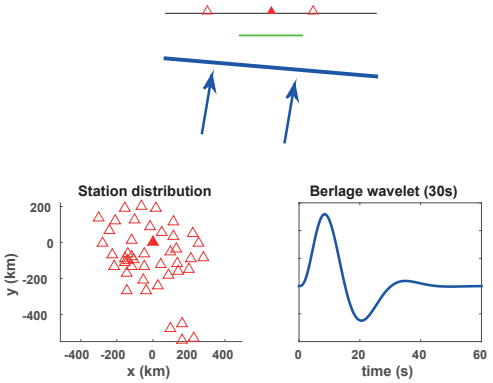


Figure 1. Vertical slice showing the geometry of the incoming plane wave (in blue), of the stations (red triangles, filled for the central station), and of a penny-shaped heterogeneity (in green) below the central station; map showing the location of the stations with the central station shown as a filled triangle at (0,0); inset showing the Berlage wavelet at 30 s central period.

2.3 A plane wave as reference field

Since we target regional tomographic studies, we adopt here a very simple strategy to calculate the kernels: we assume that the reference field is a plane wave propagating upwards towards a series of receivers located on a horizontal plane embedded in an infinite homogeneous 3-D space (Fig. 1). The plane wave simulates the wave coming from a distant source. Using an infinite space enables us to use the analytical form of the complete Green's function (Aki & Richards 1980), making possible the analysis of the importance of different components of the field, being P-wave far-field, S-wave far-field or near-field terms. Since the Green's function has a simple analytical form in the time domain, this also provides a very efficient way of computing the kernels (eqs. 4 and 5) directly in the time domain. The main drawback is that we do not account for the effect of the free surface.

In our case, the reference field is a plane P or SH wave coming with prescribed incidence and azimuth. The time dependence of the incoming pulse is a Berlage wavelet of order 2 (Aldridge 1990), with a prescribed central frequency f and an attenuation $Q = 6f$. The frequency content of this causal wavelet (Fig. 1) resembles the one commonly achieved in frequency-filtered body waves, an important element to simulate finite-frequency kernels that reproduce properly the sensitivity of the measurements to the structure.

The kernels will be calculated using the vertical component of the wavefield for the P waves and the horizontal transversal component for the SH waves. We use mainly

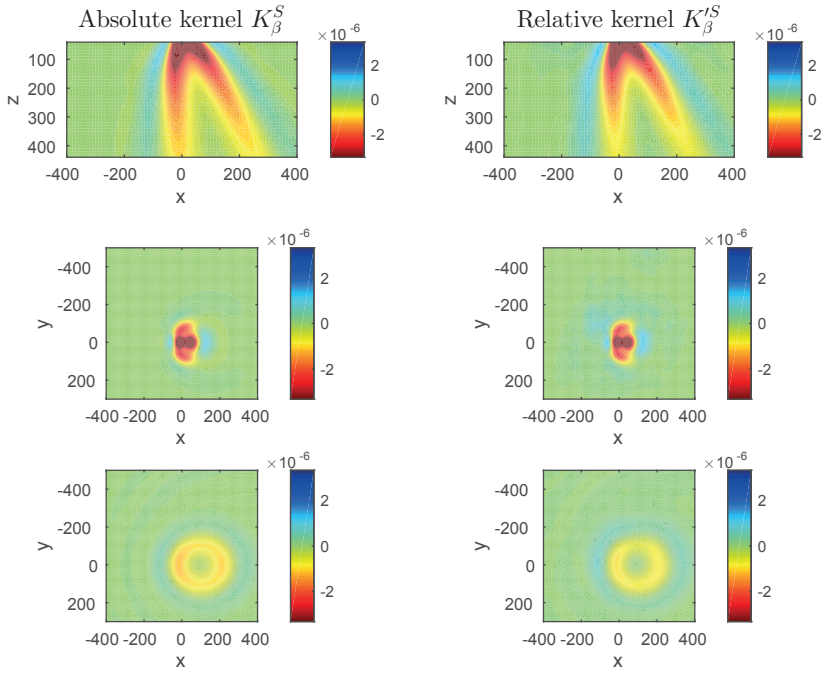


Figure 2. Absolute (left panels) and relative (right panels) kernels for a plane SH wave with period 30 s and incidence 15° at the central station shown in Fig. 1. The upper panels show a vertical slice at the location of the station and the lower panels show horizontal slices at 50 and 350 km depth.

a station distribution inpired by a recent body-wave tomography (Kolstrup et al. submitted 2014) based mostly on permanent stations and a temporary network in southern Norway (Weidle et al. 2010). This geometry reproduces the characteristics of many broadband deployments, with 45 stations deployed in a 500 km × 500 km area with a typical interstation distance of 50 km (Fig. 1).

3 MANTLE KERNELS FOR SH TRAVELTIMES

3.1 Absolute and relative kernels K_β^S and $K_\beta^{'S}$ at 30 and 3 s periods

Fig. 2 displays absolute and relative kernels (eqs. 5 and 6) at a central station (see Fig 1) for an SH wave of central period 30 s, propagating with incidence 15° in a model with compressional wave velocity of 8.2 km/s and shear wave velocity of 4.5 km/s. The kernels are shown only for mantle depths in this example, as we will focus on near-station effects in a later section. The kernels show maximum values close to the station and the typical widening and doughnut shape at depth. A secondary lobe of positive values can also be seen. The amplitude of this secondary lobe, and of eventual further lobes not seen here, is strongly dependent on the frequency content of the wave, increasing in amplitude for more monochromatic waves. The amplitude spectrum of the Berlage wavelet used here has a maximum at the central frequency, but is not very monochromatic and should therefore simulate quite well the frequency content of real data, giving realistic values for the outer lobes.

The relative kernel (right panels) is the absolute kernel (left panels) minus the mean of the absolute kernels at all stations (eq. 7). Its exact shape depends therefore on the geometry of the network and on the location of the analysed station in the network. We have chosen here to show the relative kernel at a central station (Fig. 1) since the absolute kernels overlap more in the middle of the network, giving a larger mean kernel and therefore a larger difference between absolute and relative kernels. The vertical slices show that the absolute and relative kernels are similar in the upper part of the model and that the difference increases with depth, as expected, since the partial overlap of absolute kernels at different stations increases with depth. This larger difference at depth can also be seen in the horizontal slices at 50 and 350 km depth. Compared to the absolute kernels, the relative kernels have a weakened main negative ring and a strengthened positive ring further away from the station.

At shorter periods (Fig. 3), the kernels are more narrow, as expected, and absolute and relative kernels look very similar. This is of course related to the fact that the narrow absolute kernels overlap little, resulting in a small mean kernel. There is however a shift away from zero in the mean value of the relative kernels which is not well rendered by the color scale of the plots. This has important consequences for large heterogeneites, as we shall see in the next section.

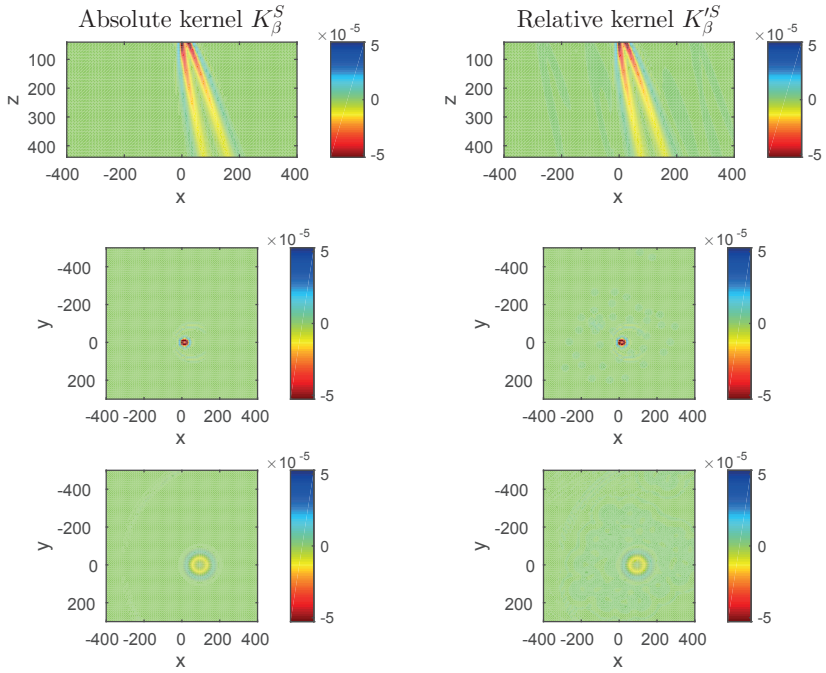


Figure 3. The same as Fig. 2 for an SH wave of period 3 s.

3.2 Delay times

In order to visualize the kernels in a different way, we have calculated the traveltime anomalies produced by penny-shaped 1 km thick heterogeneities of increasing diameter located horizontally at depths of 50 to 800 km below a central receiver. An example of such a penny-shaped heterogeneity is shown in Fig. 1. The travel time anomaly for a heterogeneity of radius R located at depths z_1 to $z_2 = z_1 + 1$ is simply:

$$\delta t = \delta\beta \int_{z_1}^{z_2} \int_0^R K_\beta^S \, 2\pi \, r \, \mathrm{d}r \, \mathrm{d}z \qquad (8)$$

The amplitude of the constant heterogeneity $\delta\beta$ is chosen to be -1 km/s in order to produce positive delays. We use vertically propagating incident plane waves in the examples shown. The delays are plotted as a function of increasing diameter of the heterogeneity for different depths, with a thicker line for the deepest heterogeneity.

3.2.1 Absolute delays

The delays for SH waves of 30 and 3 s periods measured at the central receiver are shown in Fig. 4 for depths from 50 to 400 km. The slightly hacked shape of the curves is due to limitations in the discretization of the model and is not significant. We can nicely see that the absolute delays (upper panels) increase in amplitude with the diameter of the heterogeneity and converge for large heterogeneities towards a value of 0.05 s, corresponding to the value expected for a plane wave traversing an infinite horizontal slab of 1 km thickness or, equivalently, to the value obtained using ray theory for such a slab. The diameter required to reach convergence increases with the square root of the period, as expected, and is about 200 km for 3 s period and 600 km for 30 s period. The increase with diameter is not monotonic: the curve has a maximum larger than the delay predicted by ray theory, especially at the largest depths, and some oscillations related to the contributions of the outer rings of the kernels. The small oscillations visible in the short-period delays at large diameters are all on different curves and are due to P waves scattered far away from the ray path, but arriving in the right time window to interfere with the reference SH wave.

3.2.2 Relative delays

The relative delays (middle panels) are similar to the absolute ones for small diameters, but very different for increasing diameter: the delays decrease to zero for large heterogeneities. This is in accordance with the well-known fact that relative traveltimes are not sensitive to changes in the average structure. Here we can precise this statement and quantify the sensitivity of the data to structures of different sizes.

The delays are almost zero for heterogeneities larger than 600 km, basically independently of the depth of the heterogeneity. This demonstrates simply that the traveltimes are not affected by heterogeneities with lateral dimensions equal or larger than the dimension of the network itself. But

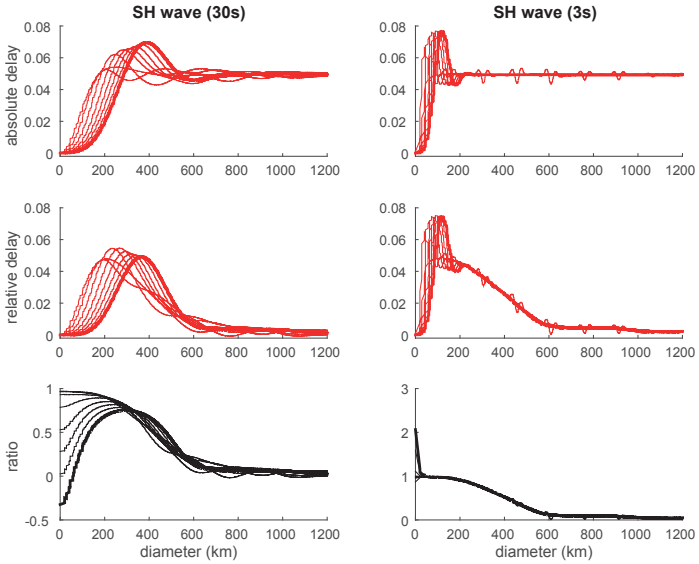


Figure 4. Absolute and relative SH-wave delays observed at the central station (see Fig. 1) due to a penny-shaped heterogeneity of thickness 1 km and shear-wave velocity variation of -1 km/s as a function of the diameter of the heterogeneity. The depth of the heterogeneity varies from 50 to 400 km in steps of 50 km. The delays for the deepest heterogeneity are shown with thicker lines. The incoming waves are vertically propagating SH waves with periods of 30 s for the left panels and 3 s for the right panels. The upper panels show the absolute delays; the middle panels show the relative delays; the lower panels show the ratio between absolute and relative delays.

we also see a decrease in delays starting at heterogeneities about half-the-size of the network. Simulations done with other networks and less centrally located stations (Fig. 5) show that this is a general feature and that the decrease in sensitivity depends directly and rather linearly on the dimension of the network and very little on the number of stations used, as long of course as this number does not get very small. The conclusions drawn here for a 500 km by 500 km network can therefore easily be adapted to other networks.

3.2.3 Sensitivity of relative tomography

In order to visualize better the difference between the absolute and relative delays, we have plotted their ratios in the lower panels of Fig. 4. The ratios are a direct measure of the reduction in sensitivity of the relative residuals compared to the absolute ones. We can note two main features in these plots.

Firstly, the decrease to zero for large heterogeneities is gradual, as seen in the previous section, starting at heterogeneties of about half-the-network. In addition to being rather independent of the depth of the heterogeneity, as also seen in the previous section, we see here that this decrease is also quite frequency-independent. The frequency independence confirms that this sensitivity reduction is a pure geometric consequence of the dimension of the network.

Secondly, we can notice that the ratios are almost 1 at short periods for heterogeneities of diameters of up to 200 km, showing that the relative residuals have the same sensitivity as the absolute ones in this parameter range. At long periods, the same is observed for shallow heterogeneities, but at larger depths, the ratio does not reach 1. For example, at the largest depth of 400 km shown in Fig. 4, we see that the maximum ratio is 0.75, meaning that the relative residual is at the maximum about 75% of the corresponding absolute residual. The smaller relative residual is due to the fact that rather large heterogeneities at depth affect the absolute traveltimes at many different stations, thus contributing significantly to the average traveltime, and thus giving smaller relative residuals.

Since the sensitivity kernels increase in diameter with depth, we expect this sensitivity reduction of the relative residuals to get stronger with depth. If there is a depth for which relative residuals are small enough to be negligible, this could be used to constrain the depth at which to cut the model when performing a tomographic inversion. Although the maximum delay ratio continues to decrease with increasing depth, this decrease is not very fast. We get for example a value of 60% at 800 km depth at the central station. Although our conclusions would need verification in a spherical Earth, this indicates that the decrease in sensitivity is not such that it can alone justify cutting the model at a given depth in relative body wave tomography. Arguments concerning the level of heterogeneity are still needed in order to justify the model vertical dimension, although in a less strong formulation than could be anticipated from ray theory. In addition, the reduction in sensitivity occurs at long periods, but not a short periods in the range of models than we have investigated (lower right panel of Fig 4).

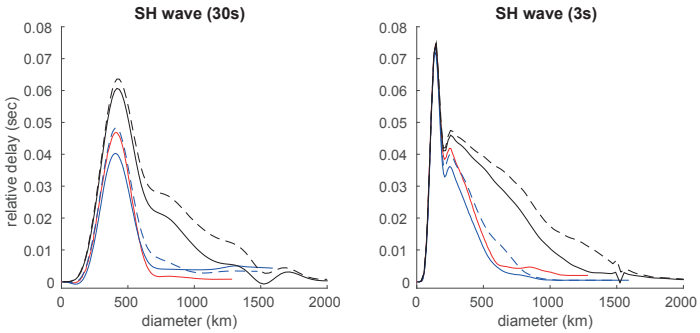


Figure 5. Relative SH-wave delays as in Fig. 4 for different network sizes and station locations. The depth of the heterogeneities is 300 km. Full lines show the relative delays for a centrally located station in the network of Fig. 1 (red), a network with the same aperture but four times more stations (blue) and a network with four times more stations over a four times larger area (twice larger in x and in y)(black). Dashed lines show the relative delays at stations located about 80 (blue) and 160 km (black), respectively, from the outskirts of the second and third networks.

3.2.4 Importance of P-wave and near-field contributions

Scattered S waves, P waves and near-field terms are generated when a plane wave, P or S, hits a heterogeneity. The kernels and delays shown above include all scattered terms. The kernels developed by Dahlen et al. (2000) that are used in most finite-frequency tomographies do not, as they are based on the paraxial ray approximation, assuming that only far-field and close-to-forward-scattering terms are significant. The scattering amplitude is then replaced in all directions by the value valid in the forward direction. This results in only like-type (P-P and S-S) scattering in the expressions.

These assumptions have been found appropriate for global tomography (Hung et al. 2000), but may be less valid for regional body wave tomography as the inverted model is close to the receivers, where both non-forward scattering and near field may be present. We cannot test here the forward-scattering approximation but we can test the importance of S-P and P-S scattering and of the near field by calculating kernels with only part of the total Green's function in eqs. 4 and 5.

For an incident SH wave at 30 s period, we find that the S-far-field terms contribute about 80% of the delays from heterogeneities at 100 km depth and above 90% at 200 km depth. The remaining is mostly due to near-field terms, with a small contribution of about 4% from the P-far-field term. At short period (3 s), the P-far-field term has about the same amplitude, but the near-field term is negligible.

At relevant mantle depths and periods, we find therefore that the forward-scattering far-field approximation is appropriate to compute the kernels for regional S-wave finite-frequency tomography.

4 MANTLE KERNELS FOR P TRAVELTIMES

As opposed to S waves, the traveltimes of the P waves are sensitive to variations in both the compressional wave velocity and the shear wave velocity. The sensitivity of P-wave traveltimes to shear-wave velocity β is neglected in most studies as it is zero both in ray theory and in the assumption of dominant forward scattering used to calculate the finite-frequency kernels of Dahlen et al. (2000). It is however pointed out by Zhang & Shen (2008) that the amplitude of the shear wave kernels for P-wave traveltimes, K_β^P, may be substantial, in particular close to the source and receiver where the forward assumption is not well appropriate. The kernels $K_\beta^{'P}$ may therefore be of particular importance in regional tomography.

4.1 Relative kernels $K_\alpha^{'P}$ and $K_\beta^{'P}$ at 20 and 2 s periods

Figure 6 shows the relative kernels $K_\alpha^{'P}$ (left plots) and $K_\beta^{'P}$ (right plots) at 2 s period. The kernels are shown in the mantle only and with the same the colour scale in order to visualize their relative importance. We see clearly that $K_\alpha^{'P}$

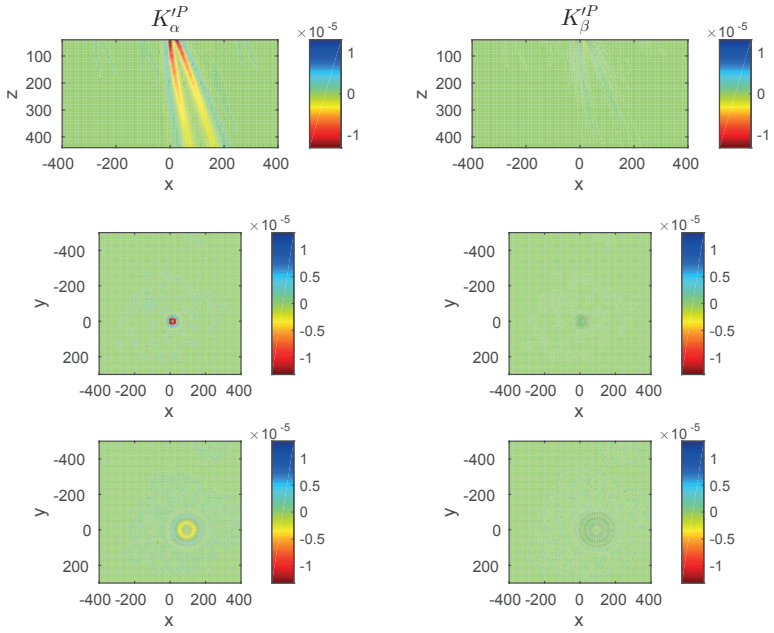


Figure 6. Relative kernels for a plane P wave with period 2 s and incidence 30°. The left and right panels show respectively the kernels of the relative P-wave traveltimes with respect to the compressional wave velocity, K^P_α, and to the shear wave velocity, K^P_β. The station position and the position of the slices is the same as in Fig. 2.

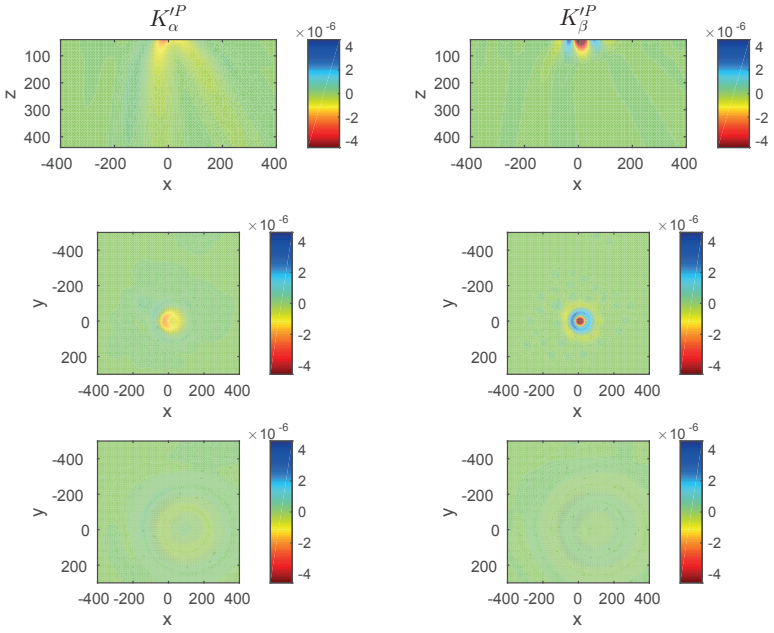


Figure 7. The same as Fig. 6 for a P wave of period 20 s.

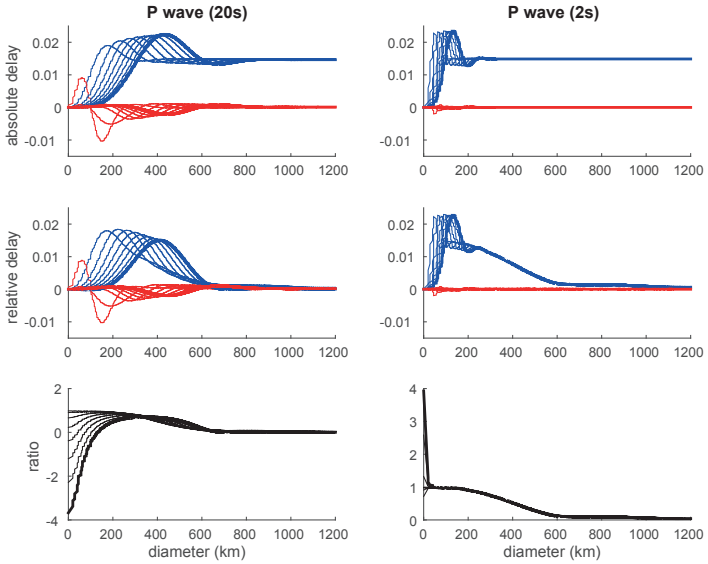


Figure 8. The same as Fig. 4 for P waves of period 20 s (left panels) and 2 s (right panels). The delays produced by heterogeneities with a compressional wave velocity variation of -1 km/s are shown in blue, while those produced by heterogeneities with a shear wave velocity variation of -1 km/s are shown in red. The lower panels show the ratio of relative to absolute delays for a variation of compressional wave velocity only. The depth of the heterogeneities varies from 50 to 400 km in steps of 50 km.

is dominating and very similar to the kernel $K_\beta^{'S}$ for short-period SH waves shown in the previous section (Fig. 3). $K_\beta^{'P}$ is very weak and barely visible, showing that mantle S-wave velocity variations do not influence notably the traveltimes of short-period P waves. We will analyse later the influence of crustal heterogeneities.

The situation is different at the longer period of 20 s (Fig. 7): $K_\beta^{'P}$ (right panels) is now larger than $K_\alpha^{'P}$ (left panels) and is controling the colour scale. It is therefore not easy to see that $K_\alpha^{'P}$ is still very similar to $K_\beta^{'S}$ at long periods (Fig. 2). The $K_\beta^{'P}$ kernels are large, but are very different in shape from the $K_\alpha^{'P}$. They have a negative value in a rather small area close to the ray surrounded by a positive value, as also shown by Zhang & Shen (2008) for the absolute kernels. The central negative value is related to scattering of S waves whereas the outer ring is related to scattering of P waves. Both contributions decrease with increasing depth for two different reasons: S waves generated at great depth arrive too late to interfere with the reference P wave and therefore do not contribute to its traveltime anomaly; P waves scattered at large depth are scattered with a small scattering angle and have therefore a small amplitude since P-to-P scattering by a point heterogeneity in shear wave velocity has a node in the forward direction (see Aki & Richards (1980) vol. 2). Consequently, the kernels $K_\beta^{'P}$ are confined to shallower depth than the $K_\alpha^{'P}$ ones. In the present case, $K_\beta^{'P}$ is very small below depths of 100 km.

We do not show the absolute kernels here. As the main kernels $K_\alpha^{'P}$ are very similar to the $K_\beta^{'S}$ kernels for the SH waves shown in the previous section, they are affected by the same sensitivity reduction at depth and for large heterogeneities. The analysis of this reduction in the previous section therefore also applies to P waves and will not be repeated. In addition, some of these effects will be seen in the delays presented in the next section. Concerning K_β^P, there is no interesting difference between the relative and absolute kernels, as their amplitude is small at depths larger than 100 km, where the difference between relative and absolute kernels starts to appear.

At 100 km depth and 20 s period, more than 90% of K_α^P is contributed by the far-field P waves. The remaining is due to near-field terms since far-field S waves do not contribute at all to K_α^P. We conclude that for P-wave regional tomography also, usage of the far-field kernels of Dahlen et al. (2000) is well-appropriate.

4.2 Delay times

Fig. 8 presents the delays produced by penny-shaped heterogeneities of 1 km thickness and α or β variations of -1 km/s. Concerning delays related to α variations (blue curves), we see that the absolute delays (upper panels) converge nicely for large heterogeneities to the theoretical infinite slab value of 0.015 s, while the relative ones (middle panels) go to zero, as in the SH case. The relative-to-absolute delays ratios for α variations (lower panels) is very similar to the one seen in the previous section for SH waves (Fig. 4).

The delays due to variations in β (red curves) are very different. First, the relative delays are not noteworthy different from the absolute ones. They are negligible at 2 s period at all mantle depths. At 20 s period, they are not negligible compared to those in α at the shallowest depth of 50 km shown here. Small heterogeneities in β with a diam-

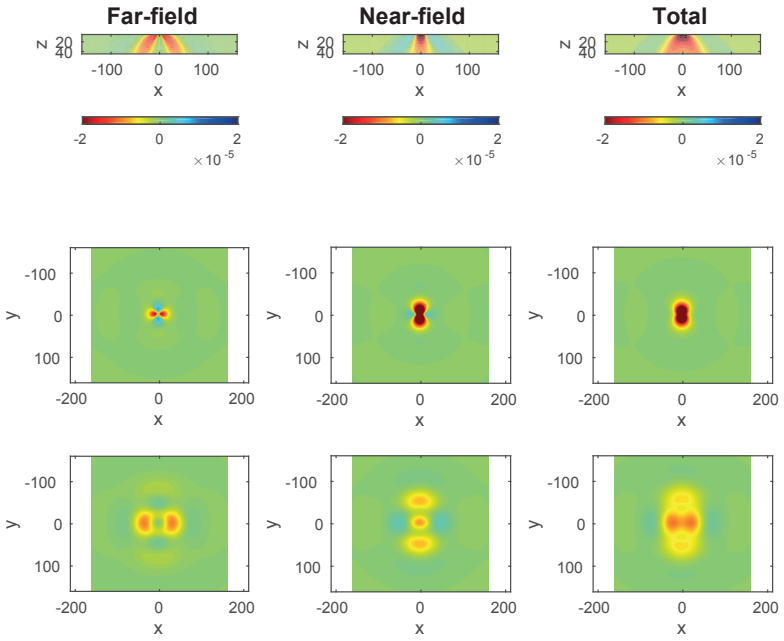


Figure 9. Absolute kernels with respect to shear wave velocity for an S wave of period 30 s. The left panels show the far-field component of the kernels, the middle panels show the near-field and the right panels show the total kernels. The maximum value of the near-field term is out of scale by a factor of 10. From top to bottom: vertical slices through receiver location and horizontal slices at 5 and 40 km depth. Incidence angle is zero and the reference model has crustal velocities.

eter of 50 km produce a positive delay which is larger than the corresponding one for a variation in α. This delay becomes negative and reaches a minimum value for a diameter of about 150 km, for then to converge to zero for larger heterogeneities, in good agreement with the results of Zhang & Shen (2008) and with the theoretical value of zero delay in P-wave traveltime for a horizontal infinite slab with only shear-wave velocity variation. At larger depths, the amplitude of the delays due to β variations decreases quickly. At 100 km depth, the maximum amplitude is about -0.005 s for a heterogeneity of 200 km diameter.

At 20 s period, the delay of up to -0.01 s due to β variations at 50 km depth is not negligible at all compared with the reference ray value of 0.015 s for an infinite horizontal slab in α. In a real case, the total P-wave delay will be the sum of the delays due to the α and β variations. In the realistic case of a shear wave velocity variation of the same sign and magnitude than the compressional velocity one, we observe that the negative delay due to β variations cancel the positive delay due to α variations that overshoot the ray-theory value. This is valid not only at 50 km depth, but at larger depths also, and results in the ray-theory value being close to the actual delay in a larger area of the parameter space than we would expect from inspection of $K_\alpha^{'P}$ and $K_\beta^{'P}$ separately. This favorable effect is lost at shorter periods.

5 KERNELS AND DELAYS IN THE CRUST

In ACH-type body wave tomography, the structure of the crust is usually not inverted for, but its influence on the traveltimes is taken into account by a crustal correction applied to each station individually (e.g., Allen et al. 2002; Martin et al. 2005). By analysing the kernels in the crust, and most importantly at Moho depth where the major contribution to crustal corrections usually originates, we can identify the dimensions of the crustal heterogeneities that affect the traveltimes. Since the difference between absolute and relative kernels decreases with depth and is not significant in the crust for the usual station spacing, we show here only absolute kernels. In addition we use vertically propagating plane waves. As our simplified geometry does not allow for taking into account the effect of the free surface, we complement this analysis of the 3-D effects by a brief analysis of the effects of free surface and crustal reverberations in 1-D models.

5.1 SH waves in the middle to lower crust

Fig. 9 shows the kernels for an SH wave of 30 s period from 5 to 45 km depth. The total kernel is shown to the right and its decomposition into near field terms (mid-field terms and near-field proper) in the middle panels and far field in the left panels. The reference model has crustal velocities of 6.5 km/s for P and 3.85 km/s for S waves.

We see that the contribution of the near field is impor-

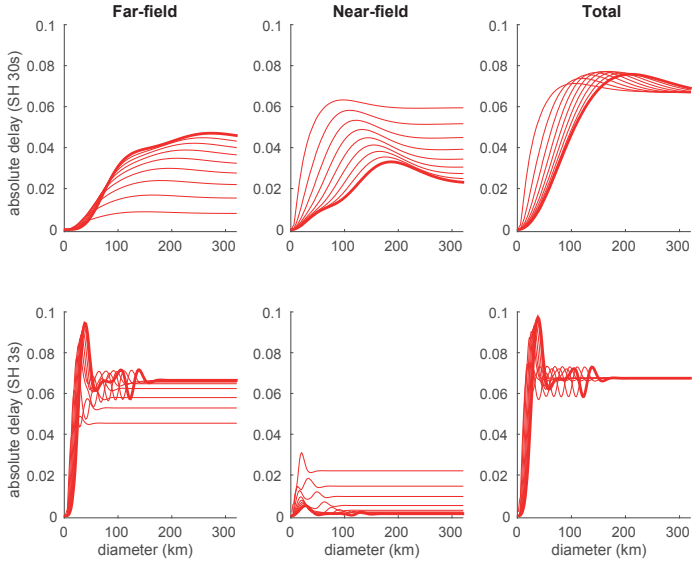


Figure 10. Delays due to heterogeneities in β of -1 km/s at crustal depths for S waves at periods of 30 s (top panels) and 3 s (bottom panels). The left panels show the contribution to the delays by the far-fiels terms; the middle panels show the contribution by the near-field terms; the right panels show the sum of the two contributions. Depth varies from 5 km to 45 km by increments of 5 km. The delays at 45 km depth are shown with a thicker line. Conventions otherwise as in Fig. 4.

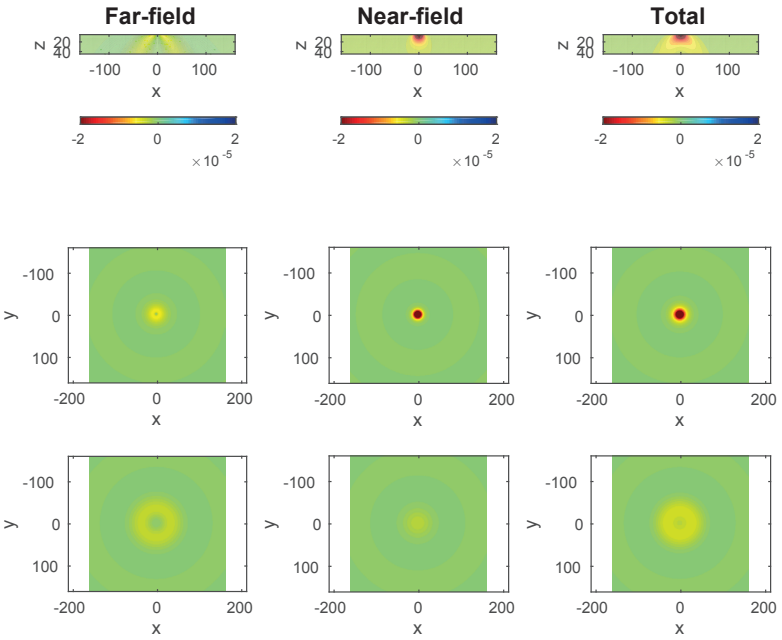


Figure 11. The same as Fig. 9 but for a P wave of period 20 s and for kernels with respect to compressional wave velocity. The maximum value of the near-field term is out of scale by a factor of 7.5.

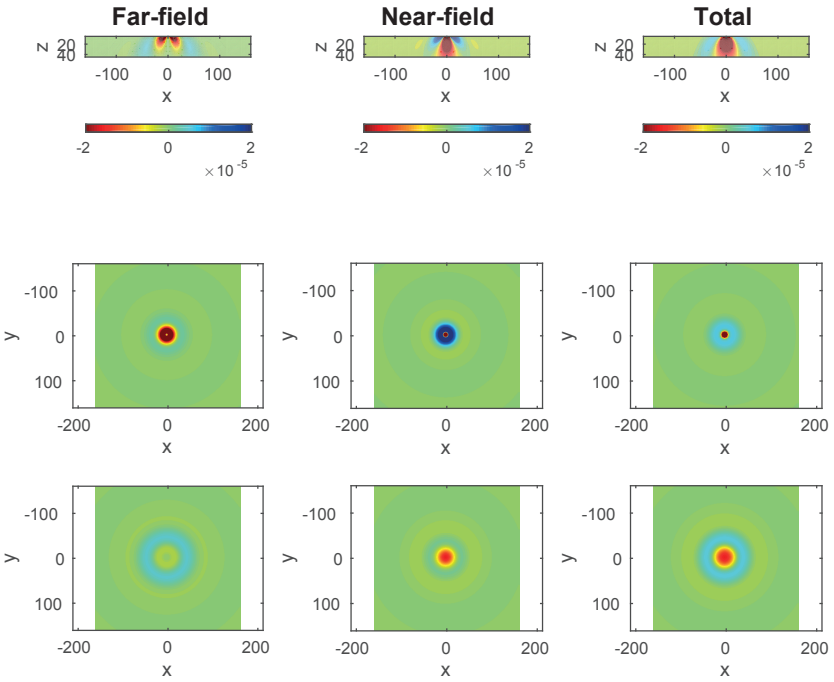


Figure 12. The same as Fig. 11 but for the kernels with respect to shear wave velocity. The maximum value of the far-field term is out of scale by a factor of 2 and the maximum value of the near-field term is out of scale by a factor of almost 20.

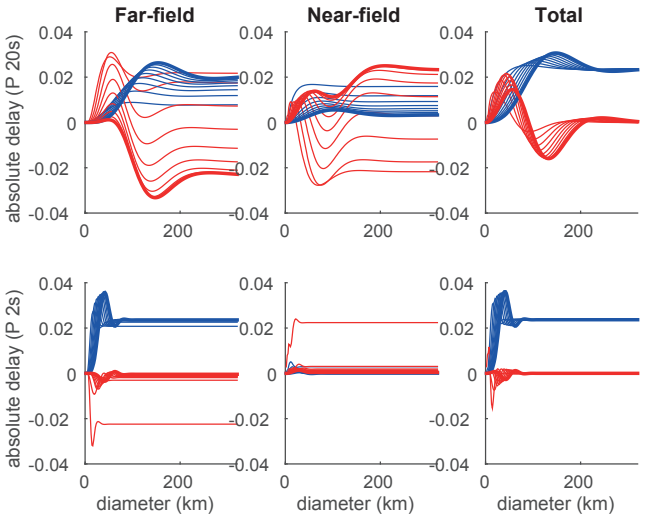


Figure 13. The same as Fig. 10 for P waves of period 20 s (top) and 2 s (bottom). Delays for shear wave velocity variations of -1 km/s are shown in red and delays for compressional wave velocity variations of -1 km/s also are shown in blue.

tant in the whole crustal range, and even clearly dominant at 5 km depth. The sensitivity of the traveltime to the structure directly beneath the station is greatly enhanced by the near-field terms compared to the sensitivity induced by the far-field terms only. Thus, the "doughnut-hole" is absent when near-field terms are included. If the Moho discontinuity is at about 40 km depth, the long period SH waves are sensitive to its variation in an area of about 100 km diameter.

This can also be inferred from Fig. 10 where the delays due to penny-shaped heterogeneities are displayed. We see that at long periods (upper panels) the delays reach the ray-theory value of 0.067 s for diameters of about 100 km. The contribution of the near field is significant and delays calculated using only the far-field terms would not be correct. At shorter periods (lower panels), the far-field terms dominate at depths larger than 10 km, but it is interesting to note that, as at long periods, the near field contributes significantly to convergence of the delay towards the ray-theory value at the shallowest depths. The traveltimes depend on the crustal structure in an area of about 20 to 30 km in diameter for SH waves of 3 s period.

5.2 P waves in the middle to lower crust

The sensitivity of long-period P waves to crustal structure is illustrated in Figs 11 to 13 with kernels shown with the same color scale as for SH waves in Fig. 9. The near field dominates the sensitivity with respect to α (Fig. 11) at shallow depths and contributes significantly to the sensitivity with respect to β (Fig. 12). The importance of the near field can also be seen in its contribution to delays by penny-shaped heterogeneities (middle panels in Fig. 13). At long periods, heterogeneities in α with a diameter of 50 km in the shallow crust and 100 km at Moho depth produce delays close to 0.024 s as predicted by ray theory. Heterogeneities in β may contribute to the total delay almost as much as heterogeneities in α, but in a more complicated way, with positive, negative or no delay depending on the lateral extent of the heterogeneity. At shorter periods (Fig. 13, lower panels), the sensitivity to α is more concentrated in a narrow region below the station and the sensitivity to β can be neglected.

5.3 Delays due to Moho topography in an example model

In order to illustrate the difference between crustal delays calculated with finite-frequency kernels and with ray theory, we have calculated both of them in an synthetic example of Moho depth. The model is build by interpolation of crustal information beneath a number of stations in southern Scandinavia, compiled in Kolstrup (2015) (Fig. 14). We get here a rough model with Moho varying from 28 to 46 km depth and with significant small scale features. The model does not necessarily give a realistic view of the precision we usually have on Moho depth and on its lateral variations around seismological stations, but is used to test the influence of a rough Moho on the associated traveltime corrections. A flat Moho case at 32.5 km depth is also used for reference. We have singled out 5 stations located in regions with different Moho characteristics: little Moho variation (station 3), gradients (stations 1 and 4), a rather sharp crustal root (station 5), and a locally thin crust (station 2).

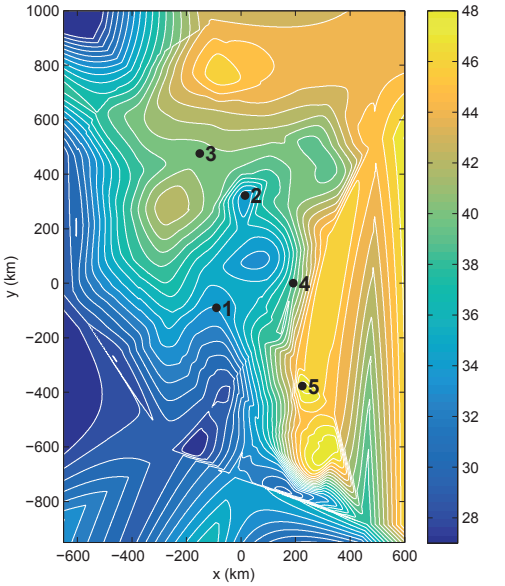


Figure 14. Map of Moho depth and location of the 5 stations for which delays have been calculated. The stations are numbered by increasing Moho depth. The Moho depth is an interpolation of Moho information compiled in Kolstrup (2015). Contour lines are drawn every 1 km.

The delays are calculated for vertically propagating waves by integration of the finite-frequency kernels in the region located between 26 and 52 km depth. In addition to the kernels calculated with the complete Green's function, we also use far-field-only kernels to approximate those defined in Dahlen et al. (2000). In the far-field case, P-wave delays are calculated using only K_α^P and a Green's function containing only the far-field P wave, and S-wave delays are calculated using K_β^S and a Green's function containing only the far-field S wave.

Due to the large seismic velocity contrast between crust and mantle, non-linearity effects are significant. In order to minimize them, the reference structure is chosen to have velocities α_0 and β_0 intermediate between crustal (α_c and β_c) and mantle ones (α_m and β_m). We use an expression for the delay in terms of slowness instead of velocity, as this reduces the non-linearity and the associated biases significantly.

$$\delta t = \int \left(-\alpha_0^2 K_\alpha(\vec{x})\left(\frac{1}{\alpha(\vec{x})}-\frac{1}{\alpha_0}\right)-\beta_0^2 K_\beta(\vec{x})\left(\frac{1}{\beta(\vec{x})}-\frac{1}{\beta_0}\right)\right) \mathrm{d}x^3 \tag{9}$$

These delays are compared to the ones using ray theory:

$$\delta t_R = h_c\left(\frac{1}{c_c}-\frac{1}{c_0}\right)+h_m\left(\frac{1}{c_m}-\frac{1}{c_0}\right) \tag{10}$$

where c is α for P waves and β for S waves, and h_c and h_m are the thicknesses directly below the station of the crustal layer (starting at 26 km depth) and mantle layers (ending at 52 km depth) respectively.

The delays calculated by the different methods are compared in Fig. 15. For the flat Moho case at 32.5 km depth (F),

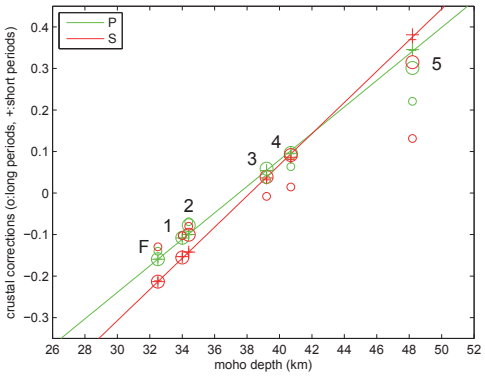


Figure 15. Comparison of the delays calculated with ray theory (full lines) and integration of finite-frequency kernels at short periods (+) and long periods (o) using the total Green's functions (large symbols) or the far-field approximation of the Green's function (small symbols). Due to similar values, the short-period far-field symbols (small +) are often hidden in the symbols for short-period total-field values (large +). Curves and symbols are shown in green for P waves and in red for S waves. The finite-frequency delays have been calculated at the 5 stations shown in Fig. 14 and for a flat Moho at 32.5 km depth (F).

the delays for the long-period waves (30 s for S waves and 20 s for P waves) cannot be distinguished from those for the short-period waves (3 s for S waves and 2 s for P waves) and fit very well with the ray-theory values, except for the far-field-only delays at long periods. The situation is very much the same for stations 1, 3 and 4: the Moho variations at these stations are such that short-period delays, long-period delays and ray theory are very similar, at least provided the kernels are computed using the complete Green's function at long periods. At station 5, where we have a crustal root, the long-period delays using the complete Green's function are smaller than the short-period and ray-theory ones by about 0.06 s for S waves and 0.04 s for P waves, corresponding to an apparent Moho about 2 km shallower than right below the station for S waves and about 1.5 km for P waves. This is in agreement with the fact that the long periods are sensitive to a larger area around the station and therefore "see" a more shallow Moho. The opposite happens at station 2, located in a locally thin crust: the long periods see a deeper Moho by about 1 km for S waves. The bias when using far-field-only kernels is significant.

We conclude that for long periods P and SH waves, a Moho depth map averaged over distances of about 100 km can be used. For shorter period waves that are sensitive to Moho variations at smaller scale, we may use a more precise Moho map if available. Except in cases of sharp roots or thinnings, the difference between the delays in the different frequency ranges is not significant. Considering the precision we usually have on Moho depth, there is no advantage in using full-wave finite-frequency kernels in the calculation of the crustal corrections and the less compationally expensive far-field-only kernels are not appropriate at these depths.

5.4 Frequency dependence of traveltimes due to reverberations in the crust

The approach used above cannot take into account the phase delay related to the reverberations in the crustal waveguide (Yang & Shen 2006; Ritsema et al. 2009). As far as we know, this delay has been neglected in all regional bodywave tomographies. This is of course fine for tomographies based on onset times and for the high-frequency data of the finite-frequency tomographies. Concerning the low-frequency data, it has been shown that the interference of the primary wave with the crustal reverberations affect the measured absolute traveltimes and that high-frequency and low-frequency traveltimes are significantly different (Obayashi et al. 2004; Yang & Shen 2006; Ritsema et al. 2009). Obayashi et al. (2004) shows in particular that the analysis of PP waves reflecting at oceanic locations with thin crust requires a frequency-dependent correction. In the case of regional bodywave tomographies, the absolute difference between the traveltimes of the short and long period waves is not relevant since the average traveltime, calculated in each frequency band separately, is substracted. The relevant information concerning crustal correction is only the variation across the network of the difference relative to ray theory. Even if this difference is large, it may be neglected in a relative tomography if it varies little across the network.

In order to quantify this effect, we have computed the delays of long-period and short-period waves across simple crustal models based on modifications of model ak135f (Kennett et al. 1995). Crustal thickness has been varied from 20 to 50 km by simply adjusting the thickness of the upper crust. The effect of sedimentary layers has been evaluated by replacing part of the upper crust by a sedimentary layer with P- and S-wave velocities of 3.8 and 2.0 km/s. The response of the crustal layer has been calculated for P and SH waves using the reflectivity method, as in Obayashi et al. (2004), Yang & Shen (2006) and Ritsema et al. (2009), using software from Levin & Park (1997). In order to mimic measurement procedures, the frequency-dependent crustal traveltimes are taken as the time of the maximum of the impulse response zero-phase bandpassed around the appropriate central frequency (using here the same periods of 2, 3, 20 and 30 s as in the other figures).

Except for structures with 1 to 2 km of sediments, where a delay of up to 0.1 s is observed for SH waves, the traveltimes of the high frequency waves do not differ from those calculated with ray theory by more than the sampling rate of 0.015 s. The traveltime difference between low frequencies and ray theory is on the other hand substantial (Fig. 16). For P waves and Moho depths larger than 35 km, correponding to about half the P wavelength in the crust, the bias when using ray theory is smaller than 0.1 s. For a thinner crust, we have a negative bias, showing that the long periods arrive earlier than predicted by ray theory by up to 0.4 s, in agreement with Yang & Shen (2006). The variation is larger for SH waves, with an overall amplitude of almost 1 s that persists for thicker crusts. The reverberations in the sedimentary layer introduce significant negative delays varying over a range of 0.5 and 0.9 s for P and SH waves, respectively. These delays actually suppress much of the corresponding sedimentary layer ray correction, simply expressing that long period waves are not sensitive to the

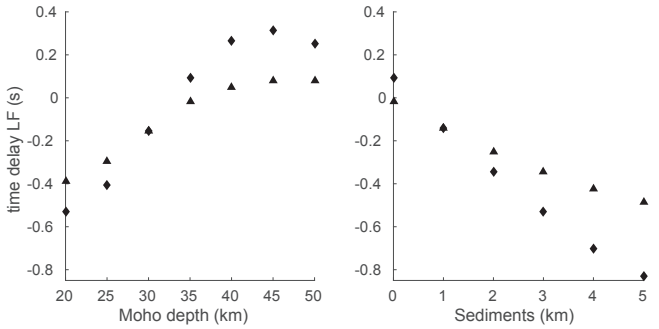


Figure 16. Traveltimes of P (triangles) and SH (diamonds) long-period waves minus ray-theory traveltimes as a function of Moho depth (left) and thickness of an upper sedimentary layer (right). Used frequency ranges are [0.03, 0.07] Hz for P waves and [0.023, 0.043] Hz for SH waves.

presence of a layer much thinner than their wavelength, and behave as if this layer had been replaced by the material below (or around).

At long periods, the influence of reverberations in the crust, either due to Moho variations or sedimentary layers, is not small compared to the ray-theory based corrections. For P waves, it might be reasonable to neglect this effect if the crustal thickness is at least 35 km across the network. It is also possible to deal with the sedimentary layers correction by simply omitting them when calculating the long-period corrections (but not the short-period ones).

A complete analysis of the effect of reverberations is out of the scope of the present paper, especially because it is strongly dependent on the frequency range used to filter the waves. The short analysis we present is used to compare with 3-D effects studied here and can be considered, in view of our frequency range, as giving an upper limit of the expected effects in realistic data cases. The importance of the reverberations has to be evaluated in each study depending on the used frequency range and on the lateral variations in the crust. We see however that they may in general be of the same order of magnitude as the ray-theory corrections. If the lateral crustal variations are large enough to require some crustal corrections, it is therefore likely that the effect of the reverberations should also be taken into account at long periods, either by a complete analysis using the impulse responses at the individual stations or at least by omitting thin low-velocity layers at low frequencies.

5.5 Delays from sedimentary basins

Besides Moho variations, sedimentary basins are the main contributors to lateral variations of seismic velocities at crustal depth and to crustal corrections (e.g. Waldhauser et al. (2002) for the Alps). We have just seen that the long-period waves traveltimes are in practice not affected by the sedimentary layers. Still we discuss these delays together with those of the high frequencies in order to identify the lateral dimensions over which models should be averaged to compute the proper traveltime corrections if one wishes to do it with the reflectivity method.

Figure 17 shows the delays due to sedimentary layers of 1 km thickness at several depths in the upper 5 km of the crust for SH as well as for P waves, both at long and short periods. The sedimentary rocks have velocities of 3.5 and 2.0 km/s while the reference surrounding upper crust has crystalline velocities of 5.8 and 3.46 km/s. The delays are calculated using slowness variations (eq. 9) and can be compared to those predicted by ray theory (0.211 s for SH waves and 0.113 s for P waves).

The short periods (left panels) show a delay of the order of the one predicted by ray theory for a heterogeneity with a diameter of about 10 km. A heterogeneity of this size produces a smaller delay at long periods and a larger heterogeneity of about 50 km is needed to reach the ray theoretical value at a depth of 5 km. At shallower depth, the ray value is however reached from a smaller heterogeneity.

An interesting feature concerning the P-wave delays (bottom plots in Fig. 17) is that the contribution from the S-wave variation (red lines) is not negligible compared to the one from the P-wave variation (blues lines) and that it greatly contributes to the fact that the total delay (black

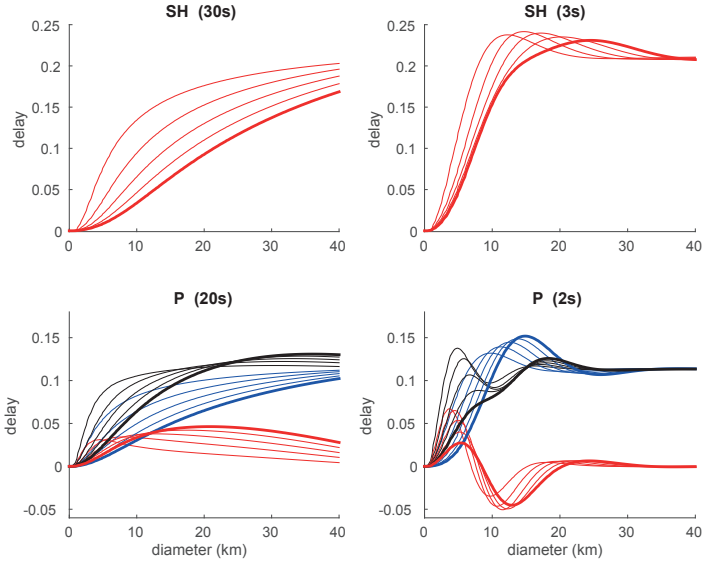


Figure 17. Delays due to sedimentary layers of 1 km thickness at 1 to 5 km depth for SH waves at periods of 30 s (top left) and 3 s (top right) and for P waves at periods of 20 s (bottom left) and 2 s (bottom right). The delays at 5 km depth are shown with thicker lines. The reference model har upper crustal P- and S-wave velocities of 5.8 and 3.46 km/s and the sedimentary rocks have velocities of 3.5 and 2.0 km/s. The black lines show the sum of the delays due the P- and S-wave velocity variations. Conventions otherwise as in Fig. 8.

lines) converges to the ray value for much smaller heterogeneities than the delay due to P-wave variations alone, especially at long periods.

In conclusion, we can say that basin corrections using ray theory are reasonably valid at short periods for all relevant depths provided the basin is at least 10 km in diameter. For sedimentary rocks at depths of around 5 km, the dimensions of the basin should be larger for the ray corrections to be valid. As basin depth and lateral dimensions are not independent of each other, we do not expect very small basins to be very deep. Pull-apart basins, which have small lateral dimensions compared to depth, still show on average a factor 10 between horizontal dimensions and depth (Gurbuz 2010). This corresponds to dimensions of 50 km for a 5 km depth basin. The deep basins around the Alps have similar dimensions (Waldhauser et al. 2002). The geometry of sedimentary basins is thus such that ray theory can normally be used to calculate their contributions to crustal corrections at short periods. Structures averaged laterally over 50 km can be used to evaluate the crustal corrections at longer periods, assuming no major 3-D effect occurs.

6 DISCUSSION AND CONCLUSIONS

Using a simple approach of an incident plane wave in a homogeneous space, we have simulated the geometry of a regional body wave tomography and calculated the associated Fréchet finite-frequency kernels for P and S waves at short and long periods. Our major objective has been to analyse two elements relating to relative traveltime tomography: how the demeaning of the traveltime data influence the finite-frequency kernels and the model resolution, and if crustal traveltime corrections should be done with a better approximation than the frequency-independent ray theory. In addition, we have analysed the influence of S-wave velocity variations on P-wave traveltimes and the validity of the forward-scattering far-field approximation for the computation of the kernels.

6.1 Demeaned kernels and the unknown average model in ACH tomography

The kernels of the relative traveltime residuals are the kernels of the absolute traveltime residuals minus their average at all stations. As far as possible, relative kernels should be used in relative tomography in order to be consistent with the inverted dataset. The additional computational effort needed to compute relative kernels is small and does not usually justify using absolute ones, at least as long as storage of all kernels is not a problem.

It is well-known that the demeaning of the residuals implies that models obtained by relative traveltime tomography are relative to an unknown average model (Aki et al. 1977; Lévêque & Masson 1999; Bastow 2012). What does this "average" mean in terms of lateral dimensions has however not been addressed. It is usually assumed that the average over the whole network is unknown and that the heterogeneities with smaller dimensions are recovered depending on the "resolution" of the dataset (Lévêque & Masson 1999), without questionning if the demeaning has an influence on this resolution. Analysing relative finite-frequency kernels, it becomes possible to quantify more precisely the influence of the demeaning on the lateral resolution.

We find quite simply that the sensitivity of the relative traveltimes compared to the sensitivity of the absolute ones decreases gradually with increasing lateral dimension of the heterogeneity and reaches zero for heterogeneites of the dimension of the station network. This decrease is not very frequency-dependent and almost not depth-dependent for the high-frequency waves. The fact that we have a gradual reduction in sensitivity highlights the fact that not only the lateral average is unknown in relative tomography, but also that the data are less sensitive to heterogeneities of dimensions of for example two-thirds the size of the network, and that this sensitivity decrease can be easily quantified. Inverting with the relative kernels instead of the absolute ones can of course compensate for this decrease in sensitivity, but the noise level in the data will anyway limit how well these heterogeneites can be resolved. This also highlights that, although adding a good regional reference model to the tomographic one, as proposed by Lévêque & Masson (1999) and Bastow (2012), will help to resolve many issues concerning vertical images and geodynamic interpretation, it cannot completely compensate for the information that is lost during the demeaning of the data.

6.2 Demeaned kernels and depth extend of regional tomographies

Another important issue in regional tomography is to which extend heterogeneities located outside the inversion box may affect the inverted model and to which depth the inversion should be carried out. The demeaning of the data does of course remove the influence of large-scale variations between the source and the inversion region, but the possible presence of medium-scale variations with a significant influence on relative traveltimes has been pointed out (Masson & Trampert 1997; Zhao et al. 2013), and methodologies have been developed to reduce this problem by adding absolute traveltime information to the relative ones (Weidle et al. 2005). To our knowledge, this issue has only been discussed in the framework of ray theory. As finite-frequency kernels measure better than rays the traveltime sensitivity to heterogeneities, it could be useful to use them to revisit this issue.

It is out of the scope of the present paper to answer this question fully, but we can contribute to a first quantification by analysing the amplitude of the relative kernels at large depth. As the absolute kernels increase in diameter with depth, they overlap more at depth, leading to a smaller amplitude of the relative kernels at large depths. One could imagine that, at some depth, the relative kernels become small enough to become negligible.

We have analysed if this decrease in sensitivity can be used to identify at which depth heterogeneities do not influence the relative traveltimes anymore, and thereby pinpoint at which depth the models can be cut when performing a regional tomography. With a 500 km × 500 km network, the long period waves (30 s for S and 20 s for P) have a sensitivity reduction of 25% for the best recoverable heterogeneities at 400 km depth. This sensitivity reduction increases to 40% at 800 km depth. There is no such reduction in sensitivity at short periods (3 s for S and 2 s for P). We conclude that the sensitivity reduction cannot be used as a strong argument for defining the maximum depth of the model as it acts only on long-period waves and in a moderate way. It will however contribute to strengthen other arguments based on the level of heterogeneity expected in the mantle.

6.3 Crustal corrections

It is well-established that reverberations in the crust cause a significant delay (usually negative) of the long-period waves compared to the short-period ones, and that this has to be taken into account in global absolute traveltime tomographies, especially when an oceanic crust is involved (Obayashi et al. 2004; Yang & Shen 2006; Ritsema et al. 2009). We have analysed here if the relative traveltimes, since they are demeaned by their frequency-dependent average and are usually recorded on land, are less affected by this phenomenon. The answer to this question depends of course on the degree of lateral crustal variation below the network and used frequency range. Except for P waves in purely crystalline crusts thicker than 35 km, we find that this additional delay may have an order of magnitude similar to the ray-theory crustal corrections themselves, and cannot therefore be neglected for realistic network configurations. In the absence of sediments, if Moho is varying from 20 to 50 km below the network, the difference between the required corrections compared to what is predicted by ray theory reaches 0.5 s for 20s P waves and 1 s for 30s SH waves in the examples we have computed. The delays are particularly significant in the presence of sedimentary layers: the negative delays in practice compensate for the ray corrections, expressing that the long-period waves traveltimes are not affected by layers of a few km and should not be corrected for them. In that case, ray theory can still be used by simply omitting the sediments for the long-period corrections.

Neglecting the reverberations, we have also analysed the lateral dimensions of the crust and sedimentary layers beneath the station that affect the traveltimes. Using a Moho map containing a realistic level of details and variation, we find that at longer periods, and in case of a sharp crustal root or locally thin crust, P and S waves at 20 to 30 s period are sensitive to a smooth version of the Moho topography with a length scale of about 100 km. Should the Moho map include significant topography at scale smaller than that, as for example in the Alps (Waldhauser et al. 2002), corrections should rather be done using a filtered version of the Moho map.

Similar conclusions are drawn for crustal corrections related to sedimentary basins. Our analysis provides the tools necessary to evaluate with which level of lateral details the basins should be modelled for computing the crustal corrections at different periods.

6.4 P-wave traveltimes sensitivity to β

The S-wave traveltimes are sensitive only to variations in β. The P-wave traveltimes are on the other hand sensitive to variations in both α and β, as pointed out by Zhang & Shen (2008). We show that the β-sensitivity is negligible at short periods, but that variations in β down to 100 km depth may produce significant delays for long-period P-waves, depending on their lateral dimensions. This should be kept in mind when computing corrections related to crustal and shallow upper mantle variations. Variations in β deeper in

the mantle, within the inversion model if inverting from 100 km depth and down, have little influence on the P-wave traveltimes and do not need to be inverted for.

6.5 Validity of the far-field forward-scattering approximation

Also from 100 km depth downwards, we find that the far-field forward-scattering approximation, which implies than only like-type scattering (P-P or S-S) contributes to the traveltime variation, approximates the kernels at long periods with a precision of 80% for S waves and above 90% for P waves. At short periods, this precision is even better. We conclude therefore that it is legitimate to use the finite-frequency kernels of Dahlen et al. (2000) in regional body wave tomography and to benefit from their advantages in terms of computing times compared to more exact expressions.

ACKNOWLEDGEMENTS

VM would like to acknowledge the hospitality of the Department of Geology of the University of Leicester where most of this work was done while she was on sabbatical leave. We would also like to thank two anomynous reviewers for rapid reviews which pointed out that we had severely underestimated the effect of the crustal reverberations.

REFERENCES

Aki, K. & Richards, P., 1980. *Quantitative seismology: theory and methods*, W.H. Freeman and Cie, San Fransisco.

Aki, K., Christoffersson, A., & Husebye, E., 1977. Determination of the three-dimensional seismic structure of the lithosphere, *Journal of Geophysical Research*, **82**, 277–296.

Aldridge, D., 1990. The Berlage wavelet, *Geophysics*, **55**(11), 1508–1511.

Allen, R., Nolet, G., Morgan, W., Vogfjord, K., Bergsson, B., Erlendsson, P., Foulger, G., Jakobsdottir, S., Julian, B., Pritchard, M., Ragnarsson, S., & Stefansson, R., 2002. Imaging the mantle beneath Iceland using integrated seismological techniques, *Journal of Geophysical Research*, **107**(B12 - 2325).

Bastow, I. D., 2012. Relative arrival-time upper-mantle tomography and the elusive background mean, *Geophysical Journal International*, **190**(2), 1271–1278.

Chevrot, S. & Zhao, L., 2007. Multiscale finite-frequency Rayleigh wave tomography of the Kaapvaal craton, *Geophysical Journal International*, **169**, 201–215.

Dahlen, F., Hung, S.-H., & Nolet, G., 2000. Frechet kernels for finite-frequency traveltimes - I. Theory, *Geophysical Journal International*, **141**, 157–174.

de Vos, D., Paulssen, H., & Fichtner, A., 2013. Finite-frequency sensitivity kernels for two-station surface wave measurements, *Geophysical Journal International*, **194**, 1042–1049.

Gurbuz, A., 2010. Geometric characteristics of pull-apart basins, *Lithosphere*, **2**(3), 199–206.

Hung, S.-H., Dahlen, F., & Nolet, G., 2000. Frechet kernels for finite-frequency traveltimes - II. Examples, *Geophysical Journal International*, **141**, 175–203.

Hung, S.-H., Shen, Y., & Chiao, L., 2004. Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach, *Journal of Geophysical Research*, **109**(B08305), 1–16.

Hung, S.-H., Shen, Y., & Chiao, L., 2011. A data-adaptive, multiscale approach of finite-frequency, traveltime tomography with special reference to P and S wave data from central Tibet, *Journal of Geophysical Research*, **116**(B06307), 1–16.

Kennett, B. L. N., Engdahl, E. R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, **122**(1), 108–124.

Kolstrup, M., 2015. *Old sutures and young plumes? - New geophysical investigations of the crust and upper mantle in southwestern Scandinavia*, Ph.D. thesis, University of Oslo.

Kolstrup, M., Hung, S.-H., & Maupin, V., submitted 2014. Multiscale, finite-frequency P and S tomography of the upper mantle in the southwestern Fennoscandian Shield, *Geophysical Journal International*.

Lévêque, J.-J. & Masson, F., 1999. From ACH tomographic models to absolute velocity models, *Geophysical Journal International*, **137**(3), 621–629.

Levin, V. & Park, J., 1997. P-SH conversions in a flat-layered medium with anisotropy of arbitrary orientation, *Geophysical Journal International*, **131**(2), 253–266.

Martin, M., Ritter, J., & CALIXTO Working Grp, 2005. High-resolution teleseismic body-wave tomography beneath SE Romania - I. Implications for three-dimensional versus one-dimensional crustal correction strategies with a new crustal velocity model, *Geophysical Journal International*, **162**(2), 448–460.

Masson, F. & Trampert, J., 1997. On ACH, or how reliable is regional teleseismic delay time tomography?, *Physics of the Earth and Planetary Interiors*, **102**(1-2), 21 – 32.

Obayashi, M., Suetsugu, D., & Fukao, Y., 2004. PP-P differential traveltime measurement with crustal correction, *Geophysical Journal International*, **157**(3), 1152–1162.

Ritsema, J., van Heijst, H., Woodhouse, J., & Deuss, A., 2009. Long-period body wave traveltimes through the crust: implication for crustal corrections and seismic tomography, *Geophysical Journal International*, **179**, 1255–1261.

Sigloch, K. & Nolet, G., 2006. Measuring finite-frequency body-wave amplitudes and traveltimes, *Geophysical Journal International*, **167**(1), 271–287.

Tian, Y., Hung, S.-H., Nolet, G., Montelli, R., & Dahlen, F. A., 2007. Dynamic ray tracing and traveltime corrections for global seismic tomography, *Journal of Computational Physics*, **226**(1), 672–687.

Tian, Y., Sigloch, K., & Nolet, G., 2009. Multiple-frequency SH-wave tomography of the western US upper mantle, *Geophysical Journal International*, **178**(3), 1384–1402.

VanDecar, J. & Crosson, R., 1990. Determination of teleseismic relative phase arrivals using multi-channel cross-correlation and least squares, *Bulletin of the Seismological Society of America*, **80**, 150–169.

Waldhauser, F., Lippitsch, R., Kissling, E., & Ansorge, J., 2002. High-resolution teleseismic tomography of upper-mantle structure using an a priori three-dimensional crustal model, *Geophysical Journal International*, **150**(2), 403–414.

Weidle, C., Maupin, V., Ritter, J., Kvaerna, T., Schweitzer, J., Balling, N., Thybo, H., Faleide, J., & Wenzel, F., 2010. MAGNUS–A seismological broadband experiment to resolve crustal and upper mantle structure beneath the southern Scandes mountains in Norway, *Seismological Research Letters*, **81**(1), 76–84.

Weidle, C., Widiyantoro, S., & CALIXTO Working Grp, 2005. Improving depth resolution of teleseismic tomography by simultaneous inversion of teleseismic and global P-wave traveltime data - application to the Vrancea region in Southeastern Europe, *Geophysical Journal International*, **162**(3), 811–823.

Yang, T. & Shen, Y., 2006. Frequency-dependent crustal correction for finite-frequency seismic tomography, *Bulletin of the Seismological Society of America*, **96**(6), 2441–2448.

Yang, T., Shen, Y., van der Lee, S., Solomon, S. C., & Hung, S.-H., 2006. Upper mantle structure beneath the Azores hotspot from finite-frequency seismic tomography, *Earth and Planetary science letters*, **250**(1-2), 11–26.

Zhang, Z. & Shen, Y., 2008. Cross-dependence of finite-frequency compressional waveforms to shear seismic wave speeds, *Geophysical Journal International*, **174**, 941–948.

Zhao, D., Yamamoto, Y., & Yanada, T., 2013. Global mantle heterogeneity and its influence on teleseismic regional tomography, *Gondwana Research*, **23**(2), 595 – 616.

Zhao, L., Jordan, T., Olsen, K., & Chen, P., 2005. Frechet kernels for imaging regional Earth structure based on three-dimensional reference models, *Bulletin of the Seismological Society of America*, **95**(6), 2066–2080.

3.4 Manuscript - Submitted to Geophysical Journal International

Geophys. J. Int. () , 1–30

Multiscale, finite-frequency P and S tomography of the upper mantle in the southwestern Fennoscandian Shield

Marianne L. Kolstrup[1], Shu-Huei Hung[2], Valerie Maupin[1]
[1] *Department of Geosciences, University of Oslo, N-0371 Oslo, Norway*
[2] *Department of Geosciences, National Taiwan University, Taipei, Taiwan*
Tel: +47-22856678
Fax: +47-22854215
email: m.l.kolstrup@geo.uio.no

Accepted date, Received date, in original form date

SUMMARY
We image the P- and S-wave structure of the upper mantle in southwestern Scandinavia using a wavelet-based, multiscale parameterisation and finite-frequency theory to model wave propagation. Relative travel time residuals of direct P- and S-waves are measured in a high and low frequency band and are corrected for crustal structure using a detailed model for the study area. A range of resolution tests are used to find optimal damping values not only for variations in V_P and V_S separately, but also for perturbations in their ratio V_P/V_S. The tests show that features down to a size of 100 (150) km can be well resolved in the P (S) tomography. To ease comparison with previous studies we also perform ray-theoretical multiscale tomographies, and to test the degree of vertical smearing we evaluate different parameterisations in the vertical direction (wavelet-based multiscale and convolutional quelling). Our finite-frequency, multiscale images of variations in V_P and V_S confirm the existence of low velocities below southern Norway and Denmark and high velocities beneath the shield proper in Sweden, as seen in previous studies, but add more details to this simplified picture. The low velocities below southern Norway and Denmark are confined to a channel-like structure at about 100 to 200 km depth, and the lateral transition from low to high velocities follows zones of Carboniferous-Permian extension and magmatism very closely. A deeper low velocity anomaly below central southern Norway emerges from the channel at 150 km depth and extends to a depth of 350 km. In the Swedish area we infer high velocity anomalies in V_P and V_S, and negative anomalies in V_P/V_S that indicate a strongly depleted mantle. We propose that the episodic erosion and convective removal of an originally thick mantle lithosphere below southern Norway to its current thickness of about 100 km could have been a trigger for episodic uplift in the Mesozoic and Cenozoic.

Key words: Body waves, Finite-frequency tomography, Cratons, Fennoscandian Shield, Dynamics of lithosphere and mantle

1 INTRODUCTION

Recent geophysical studies in the western Fennoscandian Shield have revealed a region of anomalously low seismic velocities in the uppermost mantle below southern Norway compared to the otherwise high velocities of the Fennoscandian Shield mantle (Weidle & Maupin 2008; Maupin 2011; Medhus et al. 2012; Wawerzinek et al. 2013). The general picture of a low-velocity anomaly and a transition to the Fennoscandian Shield is present in all studies, but different seismic models show differences in amplitude and location of the anomaly (Medhus et al. 2012; Wawerzinek et al. 2013) that have not been explained.

The low velocity anomaly grossly coincides with the high topography in southern Norway, which has an unknown and highly debated origin (e.g. Nielsen et al. 2009b; Lidmar-Bergstrom & Bonow 2009; Nielsen et al. 2009a; Gabrielsen et al. 2010). It is in particular important to ascertain whether the low velocity upper mantle contributes to the isostatic compensation of the high topography (e.g. Pascal & Olesen 2009; Medhus et al. 2012; Ebbing et al. 2012; Gradmann et al. 2013). A robust estimate of the magnitude and location of the low velocity anomaly and the transition to the Fennoscandian Shield proper is therefore crucial.

The Fennoscandian Shield contains an Archean core

(3.7-2.6 Ga) in the northeast and progressively younger domains towards the southwest (Gaál & Gorbatschev 1987) that accreted to the Archean core in the Paleoproterozoic (e.g., Gaál & Gorbatschev 1987; Nironen 1997; Högdahl et al. 2004). The study region (Fig.1) covers most of the youngest part of the Fennoscandian Shield in the southwest (Gaál & Gorbatschev 1987), which is mainly the product of Mesoproterozoic continental growth (Gaál & Gorbatschev 1987; Bingen et al. 2005) and late Mesoproterozoic deformation in the Sveconorwegian Orogeny (1.25-0.90 Ga, Gaál & Gorbatschev 1987). The southern part of the study area was created by the docking of the microcontinent Avalonia (440 Ma), and the later Caledonian Orogeny (420-400 Ma) (Roberts 2003; Cocks & Torsvik 2006) that also heavily deformed the northwestern part. The last major tectonic events in the study region were the late Carboniferous-Permian rifting with extension and magmatism in the Oslo and Skagerrak Grabens and along the Sorgenfrei-Tornquist Zone (305-220 Ma, Neumann et al. 2004; Larsen et al. 2008) and formation of the Danish Basin (Sørensen 1986; Frederiksen et al. 2001) before the North-Atlantic breakup ∼55 Ma.

Regional tomographies based on Rayleigh wave dispersion or full-waveform inversion covering our study area (Weidle & Maupin 2008; Rickers et al. 2013) infer V_{SV} and V_{SH} velocities 6 % lower than in AK135 (Kennett et al. 1995) below southern Norway and high velocities below neighbouring Sweden. Local studies drawing on surface waves (Maupin 2011; Köhler et al. 2012) infer V_{SV} velocities 2-3 % lower than AK135 (Kennett et al. 1995). New tomographies of Europe using full waveform inversion (Zhu et al. 2012, 2013; Fichtner et al. 2013) invert jointly for V_Φ, V_{SH} and V_{SV} and cover our study area, but do not provide enough details on a local scale to accurately image the lateral variations in the western Fennoscandian lithosphere.

The regional P- and S-wave tomographies in the area (Medhus et al. 2012; Wawerzinek et al. 2013) have better resolution than the surface waves studies but generally lower amplitudes. The relative V_S anomalies in the ray-based tomography of Wawerzinek et al. (2013) are up to ±1.5 %, an amplitude similar to the one found for V_P velocities in Medhus et al. (2012). It is not clear whether this similarity in magnitude stems from an actual similarity in nature, or from a higher degree of smoothing in the S-wave tomography.

The preservation of amplitudes of velocity anomalies is critical when it comes to the tectonothermal interpretation of tomographies. As seismic velocities are more sensitive to temperature than to composition (e.g. Goes et al. 2000; Cammarano et al. 2003), velocity anomalies are usually interpreted as temperature anomalies. Below southern Norway the velocity anomaly measured with surface waves has been interpreted as a temperature anomaly of about 200°C (Maupin 2011; Maupin et al. 2013), but integrated petrological modelling suggest that a difference in composition is also needed to explain the transition from low to high seismic velocities (Gradmann et al. 2013). The V_P/V_S ratio is often used as a diagnostic of compositional anomalies (e.g. Lee 2003; Artemieva 2007) but we cannot access this ratio from tomographies of V_P and V_S derived using different parameterisations and regularizations, as tomographic images are non-unique and strongly dependent on the choice of regularization (Trampert & Snieder 1996; Chiao & Liang 2003; Chiao et al. 2010).

We wish to close the gap between local relative travel time tomography and regional full-waveform inversion and present in this study tomographic models of V_P and V_{SH} that can be compared quantitatively. The models are derived using a wavelet-based multiscale parameterisation combined with finite-frequency theory to model wave propagation (Hung et al. 2010, 2011) and a finite-frequency data set of travel time residuals from Kolstrup (2015).

Travel times measured in multiple frequency bands are sensitive to heterogeneities at different spatial scales and hence increase the amount of information available for each source-receiver pair in the data set, giving better recovery of the magnitude of velocity anomalies in tomographic models (Hung et al. 2004; Sigloch 2008). Multiscale parameterisations based on wavelet transforms (Chiao & Kuo 2001; Chiao & Liang 2003; Chevrot & Zhao 2007; Loris et al. 2007; Charléty et al. 2013) are data adaptive and preserve the long-wavelength amplitude spectra in sparsely sampled regions without losing resolution in densely covered regions. This is an advantage compared to using a detailed local parameterisation, e.g. blocks as often used in regional body wave tomography, which emphasizes spatial resolution but reduces the magnitude of long-wavelength anomalies through the high need of smoothing regularization.

We outline the main features of multiscale, finite-frequency tomography in section 3, after a summary of the data sets (section 2). Preferred tomographic models are presented in section 4.1 and followed by extensive tests of the results (sections 4.2 to 4.5). We discuss our findings in relation to thermal and dynamic modelling of the area and the current debate about the origin of the Scandinavian Mountains (section 5).

2 DATA

The data used in this study was recorded by the temporary MAGNUS network between September 2006 and June 2008 (Weidle et al. 2010), by the temporary DANSEIS network from April 2008 to June 2008, by the temporary CALAS stations (Medhus et al. 2009), and by permanent stations in the study area (Fig. 1).

The data examined come from earthquakes with magnitude $M > 5.0$ that occurred at epicentral distances of 30 to 91.5° (Fig. 2). Most events occurred at distances greater than 70°, making the separation of phases like P (S) and PcP (ScS) difficult and giving a weak PcP (ScS) signal (Kolstrup 2015). Therefore, only the direct P and S phases are included in the data set for the tomography.

P-waves are measured on the vertical component of the seismograms and S-waves are measured on the transverse component. The data are bandpass filtered using different second-order, zero-phase Butterworth filters. For P-waves, the high and low frequency bands are 0.3-0.125 Hz (33-8 s) and 0.5-2 Hz (2-0.5 s), respectively. This isolates the secondary noise peak at around 0.2 Hz (5 s) in southern Norway. S-waves are bandpass filtered in the ranges 0.03-0.077 Hz (33-13 s) and 0.077-0.125 Hz (13-8 s). Due to the relatively high noise level above 0.125 Hz (8 s), it was not possible to measure higher frequency S-waves, as also found by Wawerzinek et al. (2013). Travel times are measured using an automated processing procedure tailored for measuring travel

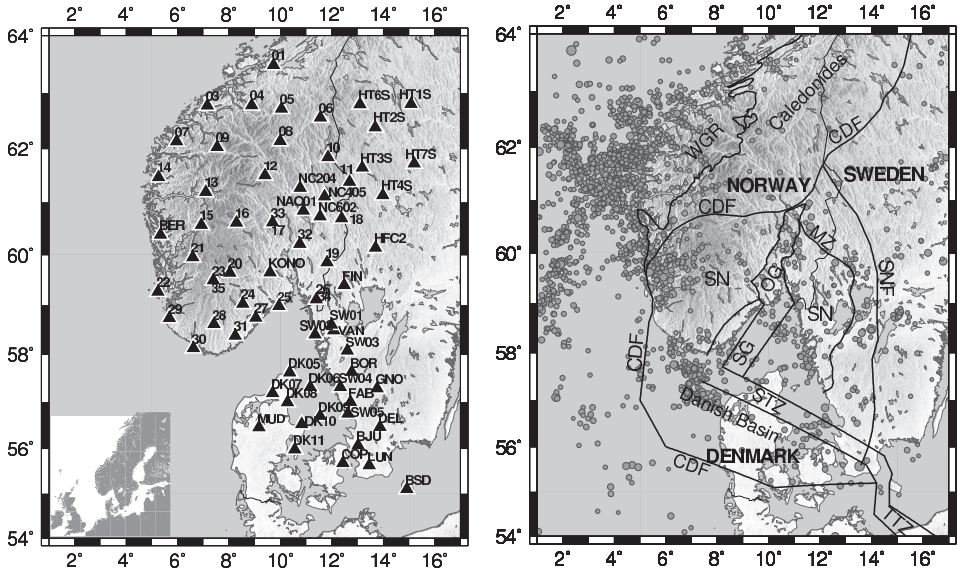


Figure 1. Station network and tectonic overview of the study region. Left: Seismological stations with topography and a regional map in the lower left corner. Numbers xx are abbreviations of the MAGNUS station names NWGxx. Right: Major tectonic features with topography and seismic events from 1980 – 2011 with $M \geq 2.0$ (FENCAT 2011). CDF, Caledonian front; MZ, Mylonite Zone; OG, Oslo Graben; SG, Skagerrak Graben; SN, Sveconorwegian; SNF, Sveconorwegian front; STZ, Sorgenfrei-Tornquist Zone; TTZ,

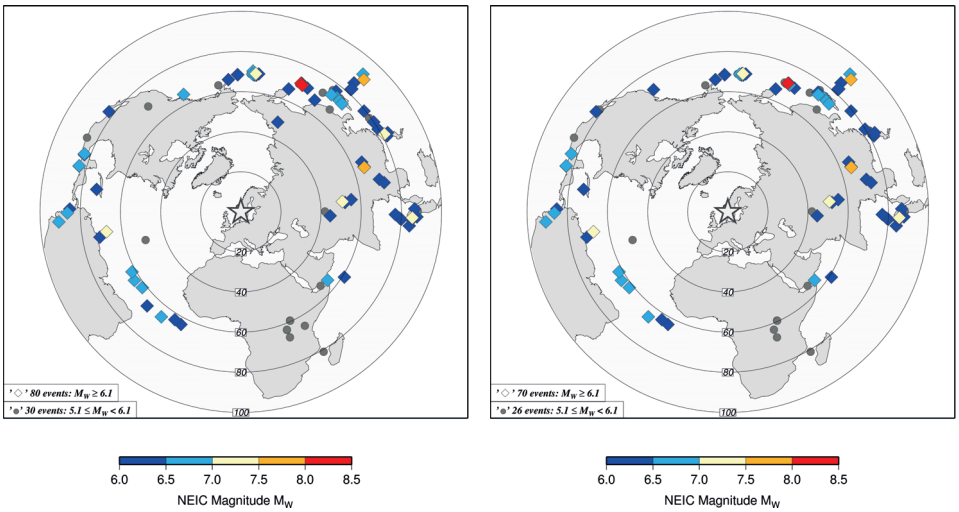


Figure 2. Map of earthquake sources used in the tomographies. The events displayed have given travel time data in at least one frequency band. Left: Events used in the P-wave tomography. Right: Events used in the S-wave tomography.

times in several frequency bands (Kolstrup 2015). The multichannel cross-correlation method (MCCC) of VanDecar & Crosson (1990) is preceded by several steps of automatic data rejection and by a preliminary picking of arrival times using the iterative cross correlation and stack algorithm (ICCS) of Lou et al. (2013).

The multichannel cross-correlation method measures all relative delays between stations Δt_{ij} and solves in a least-squares sense for the optimized arrival times t_i at each station under a constraint of $\sum t_i = 0$. This implies that the measured travel times are relative to the mean travel time for each event, removing uncertainty regarding source location and timing (VanDecar & Crosson 1990) but also removing information about absolute velocities in the study region.

The relative arrival times are then transformed to travel time residuals by subtracting demeaned theoretical arrival times calculated in the spherically-symmetric, continental Earth model AK135 (Kennett et al. 1995). The travel time residuals are corrected for the ellipticity of the Earth and for variations in topography and crustal structure (Kennett & Gudmundsson 1996; Euler 2014; Kolstrup 2015) using a detailed crustal model based on information from a wide range of sources and compiled in Kolstrup (2015). Especially useful for the crustal corrections are the Moho depths and S-wave velocity models from joint inversion of P-receiver functions and surface waves (Kolstrup & Maupin 2013) beneath the temporary Magnus stations and beneath permanent stations in Norway and Sweden, as these models give information directly beneath the stations with no need for interpolation.

The crustal corrections are performed using ray theory and are therefore equal for both high and low frequency bands. It is well-eastablished that crustal reverberations cause a significant difference between travel times of long- and short-period waves that must be taken into account in global absolute travel-time tomography (Obayashi et al. 2004; Yang & Shen 2006; Ritsema et al. 2009). For regional relative tomography, though, the data are demeaned in each frequency band and it is therefore only the relative variation of crustal travel times in the study region that is important (Maupin & Kolstrup 2014, in revision). Ray-theoretical corrections in some cases overestimate the influence of the crust on low-frequency travel times (Maupin & Kolstrup 2014, in revision), and we therefore also invert for a tomographic model without crustal corrections (section 4.3).

The resulting data set consists of 4205 and 3927 direct P arrivals in the high and low frequency bands, respectively, from 110 events (Fig. 2). For S-waves, the corresponding numbers are 3766 and 3187, from 96 events (Fig. 2).

Figure 3 shows map views of P-wave travel time residuals from opposite back azimuths, corrected for crustal structure and ellipticity of the Earth. The residuals are calculated for each station as weighted averages of all residuals from the northeastern and southwestern quadrants, respectively, using the inverse of the estimated standard deviations in measurement error (VanDecar & Crosson 1990) as weights.

The high frequency residuals from the northeastern quadrant (Fig. 3a) show a simple, almost bimodal, picture of early residuals in the east and late residuals in the west, whereas the high frequency residuals from the southwestern quadrant (Fig. 3c) are early in both the southwestern and easternmost areas. This indicates a more complicated 3D subsurface structure than suggested by the northeastern residuals, with the presence of high velocities off coast southwestern Norway.

The P-wave residuals from the lower frequency bands do not differ in magnitude from the higher frequency residuals, but differ in the spatial distribution of positive and negative residuals (Figs 3b and d), especially in the southern part of the of the study area where the long-period waves are less sensitive to the presence of sedimentary layers (Maupin & Kolstrup 2014, in revision). For events from the southwestern quadrants the changed pattern in negative residuals in the southwestern areas suggest the presence of a high velocities off coast western Denmark that are sensed by the lower frequency residuals due to their broader kernels. Hence, providing data from several frequency bands clearly increases the amount of information about subsurface structure and calls for finite-frequency tomography.

3 THEORY AND METHOD

Multiscale, finite-frequency tomography was described in detail by Hung et al. (2011) and we only review the main points and features important for our results.

The basis of linearised travel time tomography is the formulation

$$\delta t = \iiint_{\bigoplus} K(\boldsymbol{x}) \delta s(\boldsymbol{x}) d^3\boldsymbol{x}, \tag{1}$$

where δt is the observed travel time residual with respect to given spherically-symmetric, standard Earth model, and $K(\boldsymbol{x})$ is the sensitivity kernel that relates δt to perturbations in P- or S-wave slowness $s(\boldsymbol{x})$ at every point $\boldsymbol{x}$ in the region $\bigoplus$. To construct the kernels $K(\boldsymbol{x})$ we use the finite-frequency theory developed by Dahlen et al. (2000) in which the Born approximation for forward scattering is combined with ray theory for body waves and the paraxial approximation for wavefronts away from the central ray.

As mentioned in section 2, the travel time residuals δt are demeaned for each event. This implies that the appropriate kernel for a relative travel-time measurement is the one associated with this particular source-station path minus the average kernel for all the data from this particular source (e.g. Chevrot & Zhao 2007). For most network configurations though, the difference between the usual kernels and the relative ones is small enough not to bias the tomography (Maupin & Kolstrup 2014, in revision). Another effect of the demeaning is that the model inverted from the relative travel time residuals will be relative to the unknown 1D average model of the study region (Aki et al. 1977).

The seismic network used in this study covers an area of approximately 500 km by 800 km and is discretised into a grid of 65 ($2^6 + 1$) by 65 ($2^6 + 1$) by 33 ($2^5 + 1$) (=139,425) nodes with a grid spacing of 25 km and until a depth of 800 km. Even though we correct the travel time residuals for crustal structure, we include two layers in the crust and upper mantle at 0-25 km and 25-50 km, as crustal corrections are never perfect (Martin et al. 2005; Maupin & Kolstrup 2014, in revision) and the model most be free to account for those imperfections in the uppermost layers.

Numerical integration of the kernels $K(\boldsymbol{x})$ around each

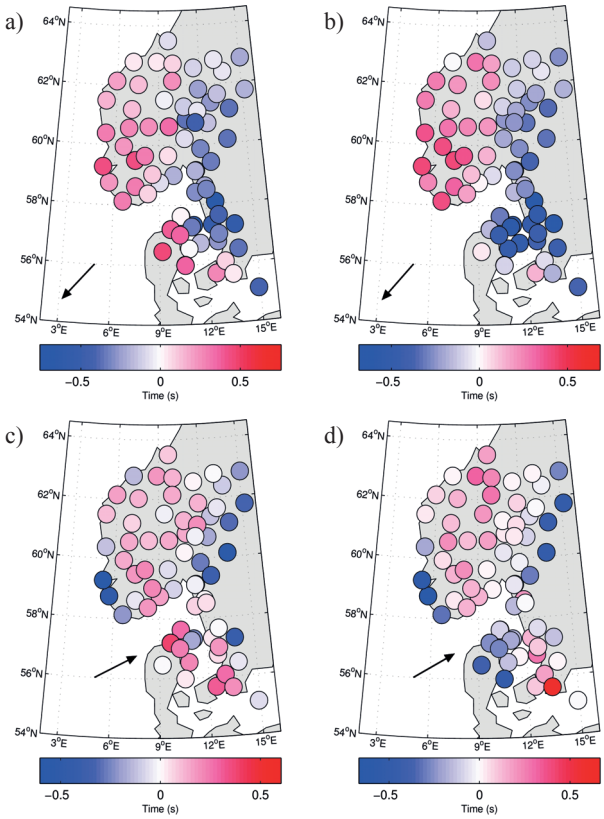


Figure 3. Average P-wave travel time residuals from the northeastern and southwestern quadrants. The arrow shows the direction of the incoming wave. a) High frequency residuals from the northeastern quadrant. b) Low frequency residuals from the northeastern quadrant. c) High frequency residuals from the southwestern quadrant. d) Low frequency residuals from the southwestern quadrant.

node leads to a data equation that relates the slowness perturbation at each node to the travel time measurements:

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{m} \tag{2}$$

where $\boldsymbol{d}$ is the data vector containing the N travel time measurements δt_i and $\boldsymbol{m}$ is the model vector containing the M (139,425) slowness parameters. The elements $\boldsymbol{G}_{ij}$ of the matrix $\boldsymbol{G}$ represents the sensitivity of the ith datum with respect to the slowness perturbation at the jth node.

Instead of solving eqn 2 directly for the model parameters $\boldsymbol{m}$, using regularizations like norm damping and smoothing, the multiscale parameterisation expands the model parameters $\boldsymbol{m}$ in terms of wavelet basis functions in 3D (e.g. Chiao & Kuo 2001; Chevrot & Zhao 2007; Hung et al. 2011), giving an equation similar to eqn 2, but with both $\boldsymbol{G}$ and $\boldsymbol{m}$ transformed into wavelet space, utilizing the CDF (2,2) wavelet (Cohen et al. 1992) and the lifting scheme (Sweldens 1996). This decomposes the model into a hierarchy of wavelet coefficients at various scales, with up to 6 scale-levels of successive refinements. The coarsest level (= 1) is at a length-scale of the entire domain and the finest level (= 6) is at a length-scale of the grid spacing (25 km).

Hence, the model parameters that are inverted for, using weighted, damped least-squares solution (DLS) (e.g., Menke 1984), are wavelet coefficients associated with each grid node. Minimum norm damping (ℓ_2 norm) in the DLS solution therefore acts on the wavelet coefficients and at several scale levels at a time, instead of uniformly on the entire model. This has the desirable effect that the finest scales being resolved are spatially varying, minimizing the need for regularization through a priori information, and preserving the amplitudes of long wavelength structures to a larger degree than conventional grid-based parameterisations (Chiao & Kuo 2001; Hung et al. 2011).

The conventional ℓ_2 norm is used in the DLS-solution as it allows for convenient linearisation of the inversion problem, but recent work suggests that the use of the ℓ_1 norm on wavelet coefficients gives a sparser and hence more robust solution (Loris et al. 2007; Charléty et al. 2013) that better recovers sharp discontinuities in the model. In a later work,

Loris et al. (2010) performed more realistic synthetic experiments where the classic ℓ_2 norm penalty on the Laplacian of the model (smoothing) worked almost as well or even better than some of the ℓ_0- and ℓ_1-wavelet methods. The proper choice of wavelet transformation and regularization strategy therefore seems to be dependent on the target of the study (Loris et al. 2010), being for example recovery of amplitudes or sharpness of boundaries, and is an ongoing field of research (Simons et al. 2011; Chevrot et al. 2012; Charléty et al. 2013; Yuan & Simons 2014).

4 RESULTS

4.1 Images of δlnV_P, δlnV_S and $\delta ln(V_P/V_S)$

Preferred models of δlnV_P ($= \delta V_P/V_P$) and δlnV_S ($= \delta V_S/V_S$) are shown in mapwiews in Figs 4 and 5, respectively, and in four profiles across the study region in Figs 6 and 7.

To visualize regions that are well sampled by the data, we plot the square root of the diagonal elements of the matrix $\boldsymbol{G}^T\boldsymbol{G}$, where $\boldsymbol{G}^T$ is the transpose of the matrix $\boldsymbol{G}$ in eqn 2. $\boldsymbol{G}^T\boldsymbol{G}$ is used as a proxy for the resolution matrix, as each of its diagonal elements is associated with a grid node and is the squared sum of the sensitivity kernels that contribute to this node. Comparison of the $\boldsymbol{G}^T\boldsymbol{G}$ images (Fig. 8) with resolution tests (section 4.4) indicates that robustly constrained regions have $\boldsymbol{G}^T\boldsymbol{G}$ values exceeding 0.5 % of the overall maximum at all nodes, and that the data provide reliable images down to a depth of approximately 500 km for P and 400-500 km for S. In all images we do not display the model at nodes with a value below the threshold of 0.5 % (Figs 4 to 7.).

The most striking structure of the models is a low-velocity region extending from western Norway to Denmark, strongest in the uppermost layers at depths between 100 and 200 km. Higher velocities are found both west and east of this region, and the low-velocity anomaly therefore has a horizontal, channel-like appearance. The channel seems to be adjacent to, and distinct from, a more circular anomaly below the northwestern end of the Oslo Graben, which extends deeper down to 350 km. The high velocity area inferred off coast southwestern Norway and Denmark is present in both the P- and S-wave images and was predicted already from the raw data (section 2). In the eastern part of the study area, high velocities extend from north to south and until a depth of around 300 to 350 km. The highest V_P anomalies of +2.5% are found east of the Sveconorwegian Mylonite Zone, while the high V_S anomalies of up to +4% are distributed more uniformly in the Sveconorwegian and Svecofennian areas (Fig. 1).

The wiggly boundary between the high velocities in the east and low velocities in the west has a steep, almost vertical angle in the southern part between Denmark and Sweden, and is more gently dipping between Norway and Sweden in the northern part (Figs 6 and 7). The boundary follows the Sorgenfrei-Tornquist Zone very closely in southern Sweden and Denmark and continues along the west side of the Skagerrak and Oslo Grabens before it terminates in the northern part of the study area with no association to major tectonic structures (Fig. 1)

The P- and S-wave images are very similar in all features but the magnitude of the anomalies, which is larger for S as expected (e.g. Kennett et al. 1998; Goes et al. 2000). The amplitude variations are around -2.0 and $+2.5$ % for P-waves and around -3.5 and $+4.0$ % for S-waves. The S-wave models also appear slightly more smooth, possibly due to the longer wavelength of the S-wave data.

In order to compare the images of δlnV_P and δlnV_S quantitatively we calculate the perturbations in the ratio V_P/V_S, i.e. $\delta ln(V_P/V_S) = \delta lnV_P - \delta lnV_S$ (Fig. 9) (e.g. Chou et al. 2009). The ratio V_P/V_S is often used as a diagnostic of compositional anomalies (Lee 2003; Artemieva 2007; Afonso et al. 2010), and it is therefore more straightforward to compare $\delta ln(V_P/V_S)$ to petrological results than the ratio $\delta lnV_P/\delta lnV_S$. In addition, $\delta ln(V_P/V_S)$ is more stable in numerical computations (Chou et al. 2009).

A consistent data set of common source-receiver pairs is usually preferred when computing V_P/V_S-ratios from tomographies (Chou et al. 2009; Hung et al. 2011) or when jointly inverting P- and S-wave data sets (Kennett et al. 1998). In our study we keep all available data to compute $\delta ln(V_P/V_S)$, as the event coverage is almost equal for the two data sets (Fig. 2) and evaluation of the resolution matrix proxy G^TG shows comparable data coverage for the P- and S-wave tomographies (Fig. 8). In addition, equal source-receiver combinations do not guarantee compatibility of the P and S models as the wavelengths and kernels of the P- and S-wave data sets are anyway different for the two wave types and result in different resolutions. The perturbations in V_P/V_S are quite sensitive to small scale variations in the models and it is important not to interpret underdamped solutions or regions of the images with low resolution (Hung et al. 2011). We use therefore a stronger cut off criteria of 1 % for the square root values of the diagonal elements of G^TG in the plots of $\delta ln(V_P/V_S)$ (Fig. 9).

The values of $\delta ln(V_P/V_S)$ reflect the covariation of V_P and V_S and point consistently to high values of V_P/V_S under southern Norway and low values below Sweden down to a depth of 500 km (Fig. 9). The difference is about 3 % between east and west at 150-200 km depth, and even larger in the shallowest part of the mantle.

All anomalies in V_P and V_S, except for the circular low velocity anomaly in the middle of the study area, have smaller magnitudes below depths of 200 km. Whether this is due to strong heterogeneity in the lithospheric mantle and/or loss of resolution at greater depths is an important question to answer to properly interpret the results. It is also interesting to investigate if the circular low velocity anomaly could be connected to the deeper mantle but appear disconnected due to decreased resolution at depth, and to estimate the amount of vertical smearing upwards from such a deep anomaly. Equally important is the quantification of horizontal smearing between separate anomalies in the uppermost mantle, before we interpret the channel-like low velocity region as a hot finger reaching out from Norway to Denmark, especially as the data coverage is best in the east-west direction. As we calculate anomalies in the V_P/V_S ratio it is also important to quantify uncertainties in $\delta ln(V_P/V_S)$. Before going to a more detailed interpretation of the models in section 5, we will investigate their robustness in sections 4.2 to 4.4

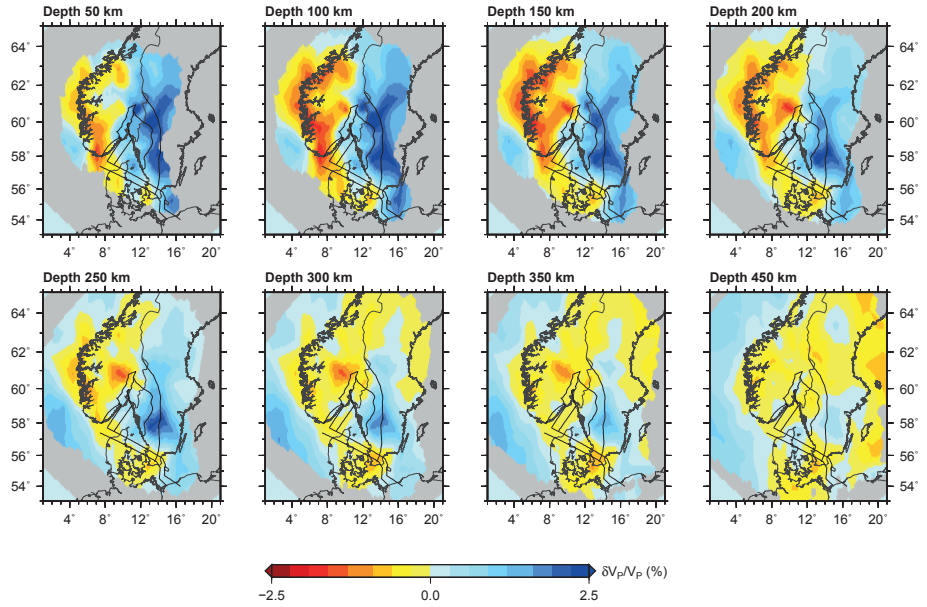


Figure 4. Map views of δlnV_P from at depths from 50 to 450 km derived using the full 3D multiscale parameterisation. Deep crustal structures from Fig. 1 are drawn with solid black lines (Mylonite Zone, Oslo and Skagerrak Grabens, Sorgenfrei-Tornquist Zone, Sveconorwegian Front, and Teisseyre-Tornquist Zone). Regions with poor resolution have been masked out.

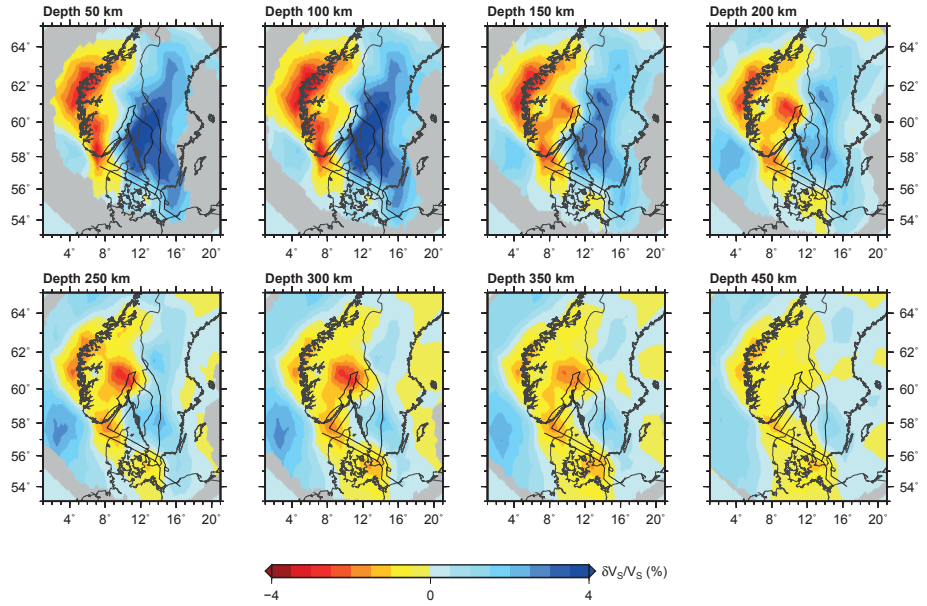


Figure 5. The same as Fig. 4 for δlnV_S.

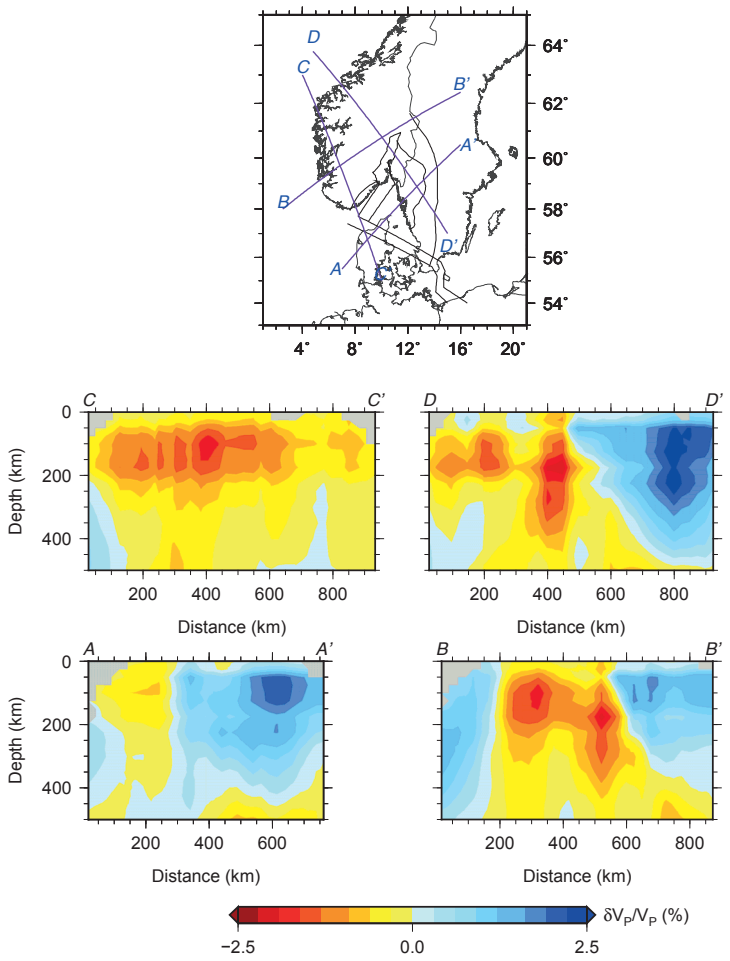


Figure 6. Profiles of δlnV_P derived using the 3D multiscale parameterisation. Regions with poor resolution have been masked out.

4.2 Vertical multiscale vs vertical convolutional quelling

The inverted multiscale model is built from synthesizing the largest scale-level first, and hence there is a risk that the resulting structure will be vertically smoothed if the incoming waves are steeply incident with few ray crossings at greater depths. In our data set, epicentral distances are greater than 70° for most events and this might increase the risk of vertical smearing from the multiscale parameterisation.

The degree of vertical smearing can be tested by switching off the multiscale parameterisation in the vertical direction and use instead convolutional quelling (Meyerholtz et al. 1989; Chiao et al. 2010) for vertical regularization. Convolutional quelling penalizes model roughness by assigning an a priori correlation length to the model (Hung et al. 2004; Chiao et al. 2010).

We therefore invert for subsurface structure using both the full 3D multiscale parameterisation (hereafter called multiscale) and using a hybrid parameterisation with multiscale in the horizontal directions and convolutional quelling with a correlation length of 25 km in the vertical direction (hereafter called hybrid).

Figure 10 shows tradeoff curves comparing the multiscale and the hybrid schemes for models inverted using the same range of damping values. Model variance, defined as in Chiao et al. (2006), is a measure of model uncertainty resulting from data error. This quantity is more appropriate than the usual model norm to calculate L-curves in tomographies with wavelet parameterisation, where focus is not on damping the small scales but removing the non-constrained features. It is therefore the model variance which is shown as a function of data misfit, calculated as the squared dif-

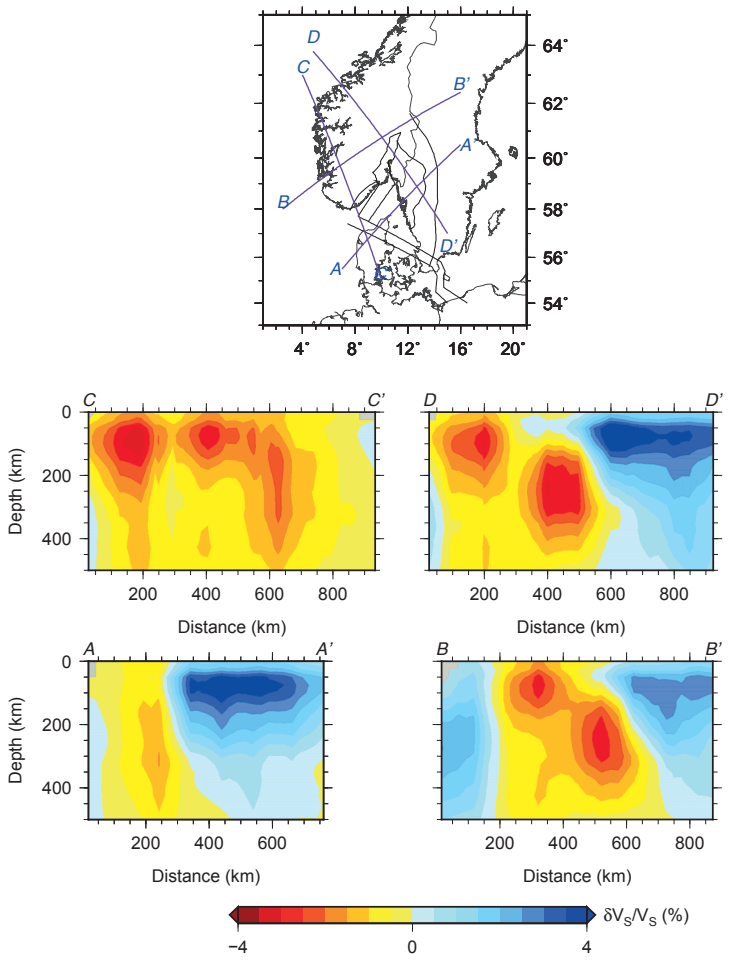


Figure 7. The same as Fig. 6 for δlnV_S.

ference between observed and predicted data normalized by the sum of the squares of the observed data.

Preferred models are chosen near the point of maximum curvature and shown with solid symbols in Fig. 10. The difference between models with damping values around the point of maximum curvature is quite small, with higher damping values suppressing the amplitudes of resulting anomalies slightly. As we perform both P- and S-wave tomographies, we also perform tests like in Chou et al. (2009) to find optimal damping values not only for δlnV_P and δlnV_S separately but also for $\delta ln(V_P/V_S)$. In these tests we calculate synthetic data for a range of synthetic models with similar magnitude for δlnV_P and δlnV_S (section 4.4). Hence the anomalies in V_P/V_S should be zero and we search for the damping values giving the smallest artificial anomalies in V_P/V_S. The chosen damping values in Fig. 10 represent a compromise between model variance, data misfit and stability of the perturbations in $\delta ln(V_P/V_S)$.

The full multiscale inversion clearly has a higher data fit than the hybrid scheme and is the best inversion scheme in that sense. To compare the models derived from the two parameterisations we present in Figs 11 and 12 map views of the hybrid solutions for P and S, corresponding to the multiscale map views in Figs 4 and 5. The greatest difference between the models is that the multiscale ones have larger anomalies and more small-scale perturbations than the smooth-looking hybrid models.

A wide range of resolution tests of the hybrid and multiscale schemes, whereof some are presented in section 4.4, show that the full multiscale inversion is better at recovering the amplitude and location of the anomalies. Overdamped multiscale inversions are very similar to underdamped hy-

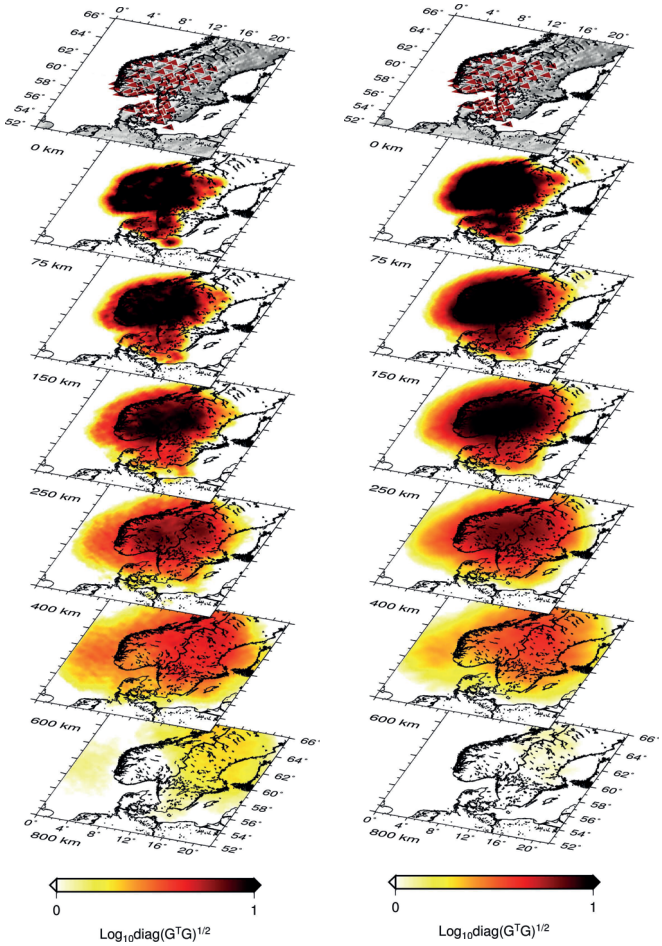


Figure 8. Images of the square root of the diagonal elements of the matrix G^TG. Left: For the P-wave tomography. Right: For the S-wave tomography.

brid inversions, and damping values can be chosen such that the results are close to be indistinguishable. As smoothing regularizations are very common, smooth images are typically preferred over more rough pictures of the subsurface by geophysicists, either consciously or unconsciously. The hybrid scheme produces the more smooth and typical tomographic pictures, but all resolution tests (section 4.4), as well as the data misfit, point out the 3D multiscale as the superior parameterisation. Hence, our preferred model is the one obtained with the full 3D multiscale inversion. When testing our results in section 4.4 we pay special attention to vertical resolution and perform all tests using both parameterisations.

4.3 Influence of crustal travel-time corrections

The vertical smearing is most severe in the uppermost parts where the overlap between kernels is smallest, making the influence of crustal corrections quite important. Figure 13 shows P-wave velocities in the uppermost layers inverted from the data set without crustal corrections (Kolstrup 2015), and Fig. 14 shows the corresponding velocities inverted from the corrected data set. The models are grossly similar, but the corrected images have significantly weaker low velocity anomalies below southern Norway and Denmark in the depths from 0 to 150 km and stronger high velocity anomalies in the eastern part of the study area. Similar conclusions are drawn for the S-wave models (supplementary Figs. 4 and 5).

The increased magnitude of the channel-like low velocity region in the uncorrected tomography (Fig. 13) is due to

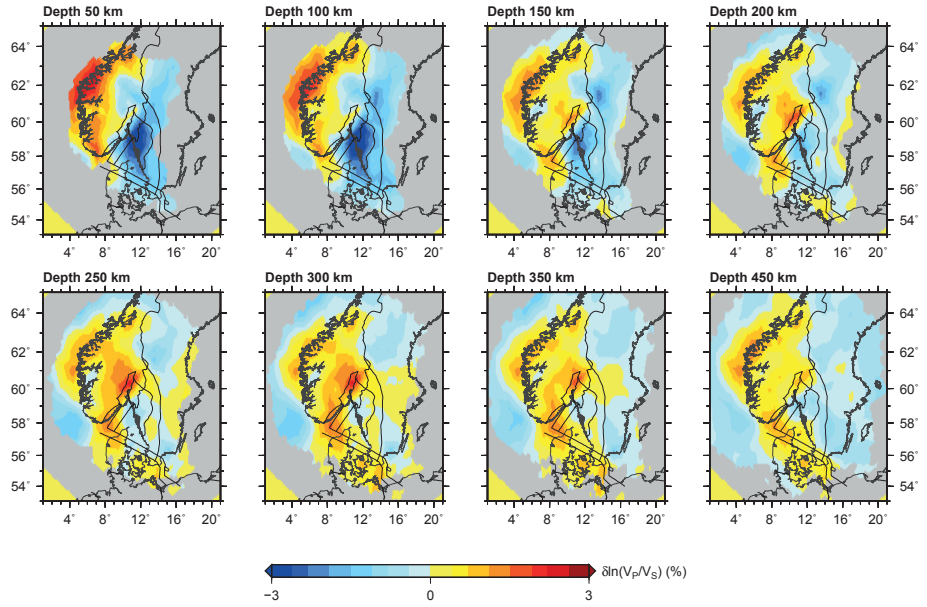


Figure 9. The same as Figs 4 and 5 for $\delta ln(V_P/V_S)$.

the crust in southern Norway being thicker inland than at the coast, and due to the much thicker sedimentary layers in Denmark compared to the rest of the study area. This gives positive (late) crustal travel time corrections in central southern Norway and Denmark. In addition, the thinner crust at the Norwegian coast gives negative (early) travel time corrections that, when not corrected for, result in an overestimation of the high velocities off coast southern Norway. These effects combine to give the expression of a more narrow low velocity channel with greater magnitude of associated anomalies than in our corrected images (Fig. 14).

It is important to note that the ray-theoretical crustal corrections applied here are not entirely consistent with the finite-frequency measurements, as crustal reverberations have a significant frequency-dependent effect on absolute travel times derived with waveform cross-correlation methods (Obayashi et al. 2004; Yang & Shen 2006; Ritsema et al. 2009). Frequency-dependent crustal corrections (Obayashi et al. 2004; Yang & Shen 2006; Ritsema et al. 2009) have heretofore been based on the reflectivity method that does not take into account 2D and 3D effects (Yang & Shen 2006), making them difficult to apply in a region with a high level of 3-dimensional variation. Recent work shows that, compared to reflectivity-based corrections, ray-theoretical corrections in general overestimate the influence of sedimentary layers and of a thin crust (< 30 km) on arrival times measured in the lowest frequency bands (Maupin & Kolstrup 2014, in revision). In our case, ray-theory and reflectivity-based corrections do not differ by more than 0.05 s in Norway and Sweden, but differ by up to 0.6 s for the low-frequency P-wave band and for both S-wave bands in Denmark due to the overestimation by ray-theory of the influence of the sedimentary layers. The ray-theoretical corrections can therefore be viewed as a maximum bound on the crustal corrections, and the true model in the southwestern part therefore lies somewhere between the ray-theoretically corrected model (Fig. 14), and the model with no corrections (Fig. 13) in the upper 150 km. Similarly for the S-wave model (supplementary Figs. 4 and 5).

Our results show that a robust estimate of the crustal corrections below each station is important to avoid artefacts in the upper 150 km in the tomographic images. In addition, it is important not to lock the crust during inversion, such that the crustal layers can allow for additional perturbations (Martin et al. 2005). We also add station and event correction terms in the inversion but assign them a quite low weight, giving them little influence on the results.

4.4 Resolution tests

We calculate synthetic travel time data using the G matrices built from the actual P- and S-wave data sets with 3D finite-frequency kernels, hence using exactly the same source-receiver configuration as in the real tomographies. Random errors with a mean amplitude of about 5% of the average size of the synthetic data is added prior to inversion.

For each synthetic model we test velocity perturbations of 3 % in V_P, and both 3 % and 5 % in V_S. Equal anomalies in V_P and V_S of 3 % are used to determine the damping values that give the smallest fluctuations in the $\delta ln(V_P/V_S)$ as in Chou et al. (2009).

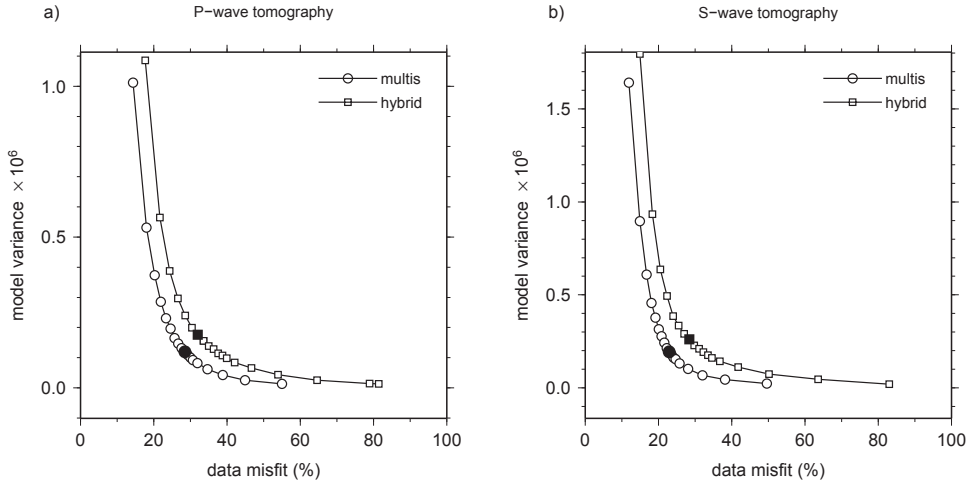


Figure 10. Tradeoff between model variance and data misfit for the 3D multiscale parameterisation and the hybrid parameterisation with vertical convolutional quelling. a) P-wave models. b) S-wave models. Solid symbols denote the preferred solutions that are shown in Figs 4-7 (multiscale) and Figs 11-12 (hybrid).

4.4.1 Horizontal and vertical resolution

We use a range of classical chequerboard tests to estimate the size of objects that can be resolved horizontally and vertically. The starting point is chequerboards of size 200x200x200 km^3, and Fig. 15 shows the synthetic input for V_P. The full 3D multiscale inversion results for P- and S-waves are shown in Figs 16 and 17, and the corresponding hybrid results for P-waves are shown in Fig. 18.

It is obvious from Figs 16 and 18 that both the hybrid and multiscale parameterisations recover the input model quite well, but the full multiscale inversion is significantly better at recovering the amplitude and location of anomalies. This result is consistent for all resolution tests, and we show only results for the 3D multiscale after this first example.

Comparing P- and S-wave results (Figs 16 and 17), we see that the P-wave tomography has better resolution, especially below 200 km depth, but objects of size 200x200x200 km^3 are quite well resolved in both tomographies. It is also seen that the S-wave tomography has more damped anomalies at depth than the P-wave tomography. The apparent loss of amplitude in the S-wave tomography below depths of 200 km (Fig. 5) might therefore be partly due to loss of resolution at this depth.

Chequerboard tests were also performed using squares of 150x150x150 km^3 and 100x100x100 km^3, and with rectangular boxes where the horizontal or vertical length scale was kept constant at 200 km in order to investigate the horizontal and vertical resolution separately. Boxes of size 100x100x200 km^3 could be well resolved in the P-wave tomography, also below 250 km depth. For S-waves, the horizontal limit of resolution was found to be about 150 km, best in the upper part of the model. Keeping the horizontal length scale constant at 200 km, we found that the limit of vertical resolution was about 150 km in the upper part of the model and 200 km in the lower part, again with the P-wave tomography having the best resolution. Hence, we can resolve features with a horizontal scale of about 100(150) km quite well in the P(S)-wave tomography, but vertical resolution is limited to about 150-200km.

The limits of resolution inferred from the chequerboard tests are well below the main features of the models in Figs 4 to 7. The channel-like low velocity region is on its thinnest around 150 km wide, though, and the deep cylindrical low velocity anomaly is quite narrow. In addition, the chequerboard tests created in all cases strong artificial anomalies in the V_P/V_S ratio. We analyse therefore in the following sections some synthetic models that simulate naturally occurring anomalies to a larger degree.

4.4.2 The low velocity channel extending from Norway and Denmark

Regional tomographies on a scale larger than ours (Weidle & Maupin 2008; Rickers et al. 2013) show a low-velocity channel from Iceland to southern Norway and Denmark at shallow depths. The ray-theoretical S-wave tomography of Wawerzinek et al. (2013) shows this channel-like feature at depths of 250-410 km, and the P-wave tomography of Medhus et al. (2012) shows a tendency of a channel at depths of 100-300 km, but much more blurred than in our images. As the channel imaged in this study might be part of a larger regional feature with geodynamic implications (Weidle & Maupin 2008; Rickers et al. 2013), it is important to ascertain whether the channel could result from horizontal smearing of separate anomalies in the uppermost mantle below Norway and Denmark. We place therefore two cylinders extending from 100 to 200 km depths with $\delta lnV_P = -3\,\%$ and $\delta lnV_S = -3\,\%$ or $-5\,\%$. The cylinder below Norway has a diameter of 180 km and the cylinder below Denmark a

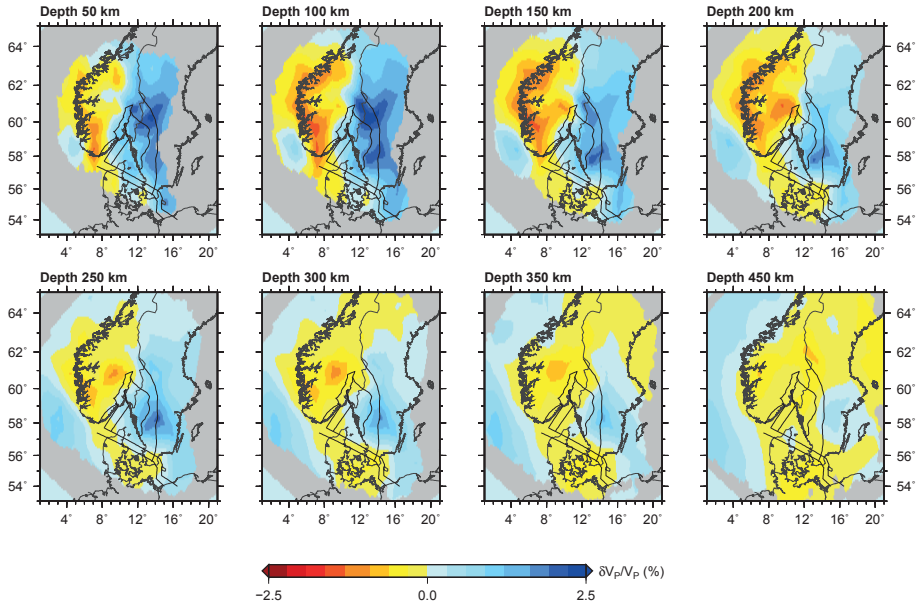


Figure 11. The same as Fig. 4 using the hybrid parameterisation.

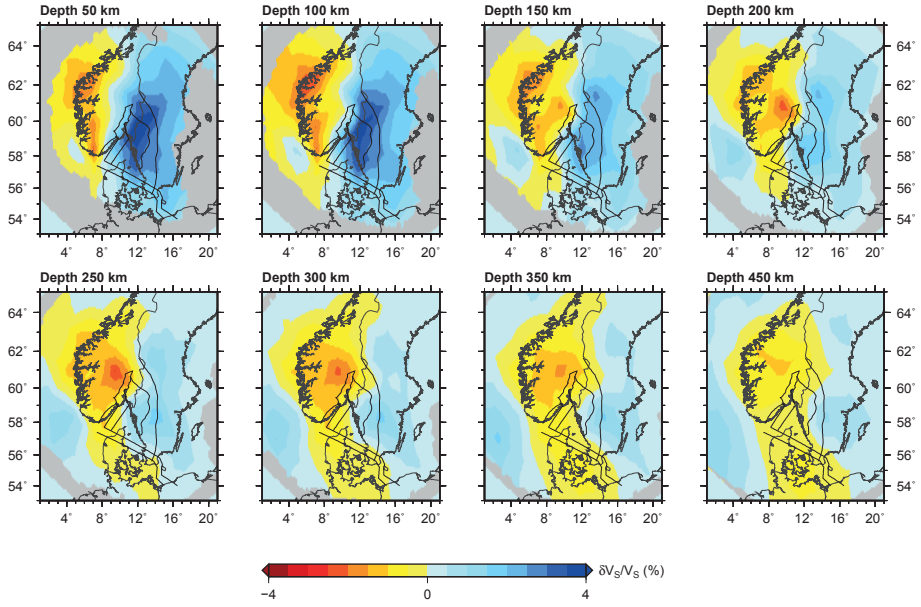


Figure 12. The same as Fig. 5 using the hybrid parameterisation.

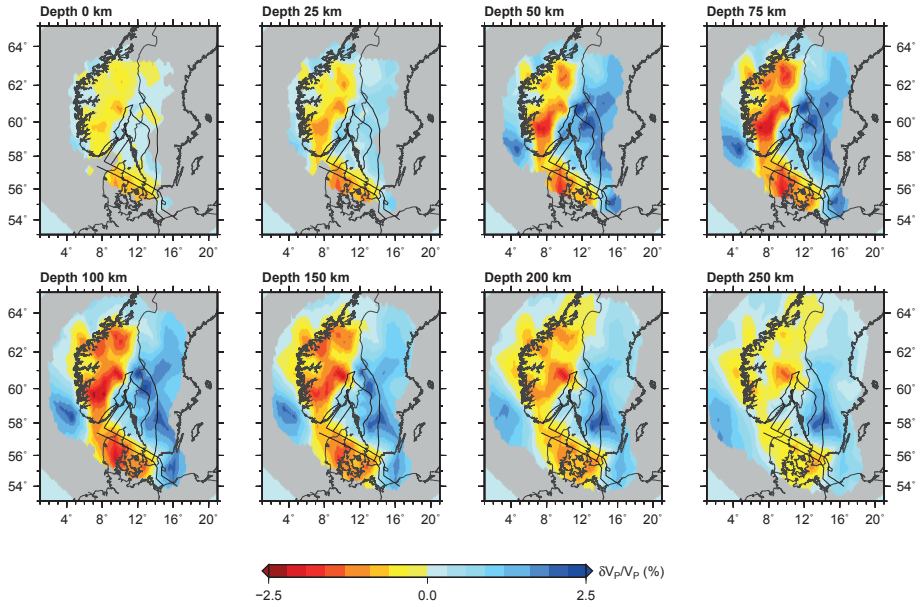


Figure 13. Images of δlnV_P in the uppermost layers. Input data are the raw travel time residuals without the crustal and topographic corrections from Kolstrup (2015).

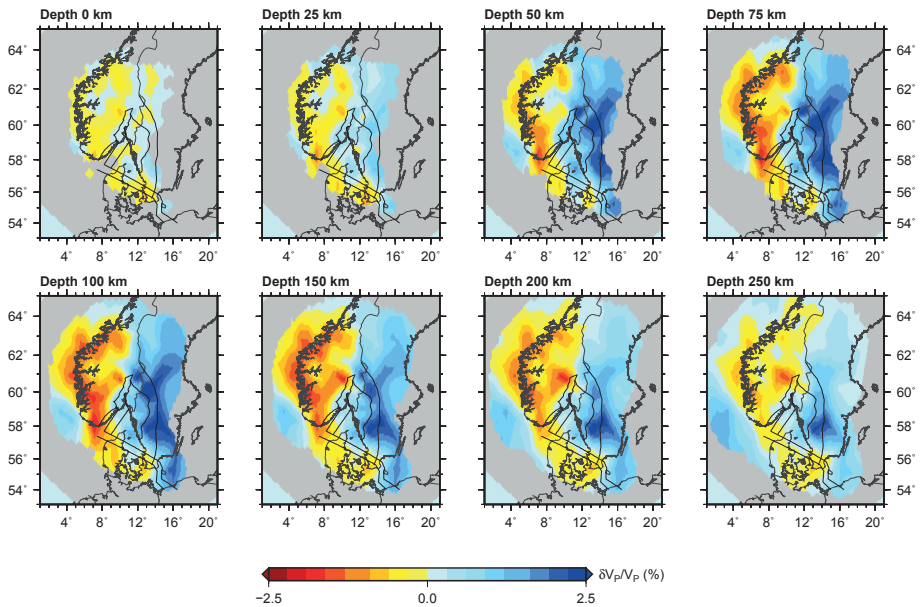


Figure 14. As Fig. 13 but with input data corrected for topography and crustal structure.

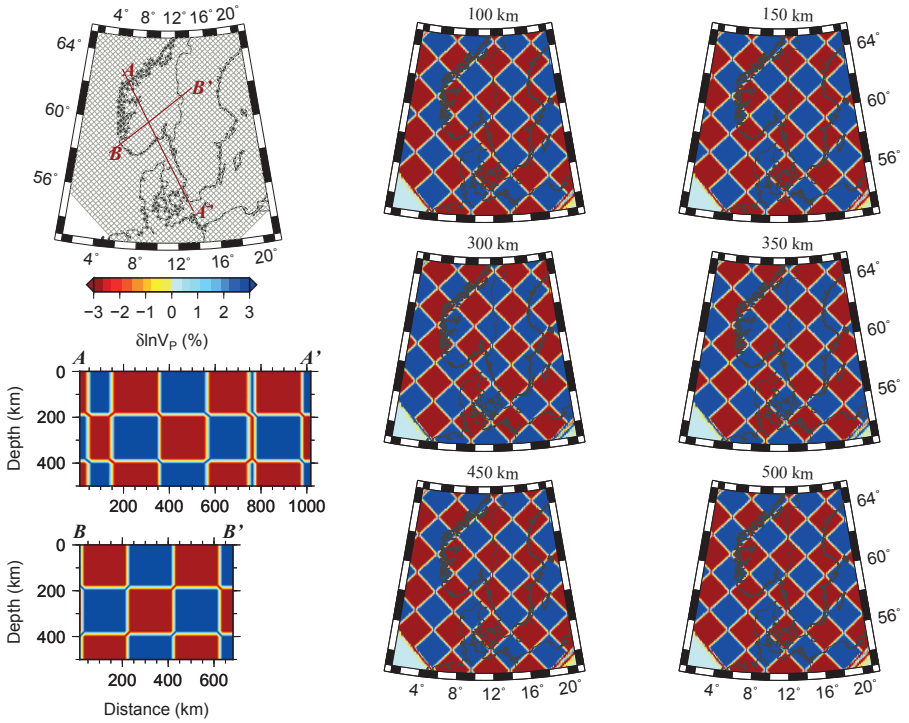


Figure 15. Synthetic input for a chequerboard test with squares of size 200x200x200km^3, $\delta lnV_P = \pm 3\%$, and $\delta lnV_S = \pm 3\%$ or $\pm 5\%$. Upper left map shows location of profiles and grid discretisation. The apparent lateral variation of the box size in the profiles is due to the obliquity of the profiles with respect to the chequerboard pattern.

diameter of 150 km. From the chequerboard tests we know that anomalies of this size are on the lower limit of resolution, and hence the risk of smearing the synthetic input is large.

The synthetic P-wave model is shown in Fig. 19 and the multiscale inversions for δlnV_P and δlnV_S are displayed in Figs 20 and 21. The inversion clearly separates the two cylindrical anomalies and does not smear them into a channel. This indicates that the channel-like feature is indeed a channel and does not represent smearing of separated anomalies in the shallow mantle. The hybrid inversions (not shown) have weaker amplitudes and give a more channel-like appearance of the recovered anomalies.

Using equal δlnV_P and δlnV_S of $-3\,\%$ shows that damping values around 10000 for both P and S give the least amount of artificial fluctuations in $\delta ln(V_P/V_S)$, with $\delta ln(V_P/V_S)$ being close to zero. These damping values are also reasonable with respect to the tradeoff between model variance and data misfit (Fig. 10). Figure 22 shows the resulting $\delta ln(V_P/V_S)$ for a synthetic input of $+2\,\%$ ($\delta lnV_P = -3\,\%$ and $\delta lnV_S = -5\,\%$). The recovered $\delta ln(V_P/V_S)$ is below $1\,\%$ and much smaller than the anomalies in the real tomography (Fig. 9), implying that the real anomalies have a larger contrast in V_P/V_S, or that they are better recovered than the small synthetic objects.

The vertical smearing of $\delta ln(V_P/V_S)$ is also much more severe than for V_P and V_S separately, due to the larger degree of vertical smearing of the S-wave tomography and the combined uncertainties of the P- and S-wave data sets. This test, along with the chequerboard tests, shows that we cannot interpret depth variations of the V_P/V_S ratio, but that we have robust estimates of horizontal variations averaged over depth. Hence, we can interpret the strong west to east variation in $\delta ln(V_P/V_S)$ in Fig. 9.

4.4.3 The deep cylindrical low velocity anomaly below central southern Norway

The cylindrical low velocity anomaly in the center of the study area extends from depths of approximately 150 to 350 km (Figs 4 and 5). As the resolution decreases at greater depths, we test if a cylinder located from 150 to 500 km depth gives a similar image as a cylinder confined to depths of 150 to 350 km. The synthetic cylinder has a diameter of 100 km and a velocity perturbation of $\delta lnV_P = -3\,\%$ and $\delta lnV_S = -3\,\%$ or $-5\,\%$.

The synthetic input model is shown in Fig. 23 and the inverted P-wave model is shown in Fig. 24. Obviously, the P-wave anomaly is recovered down to depths of around 450-500 km. The results for a similar cylinder extending only to

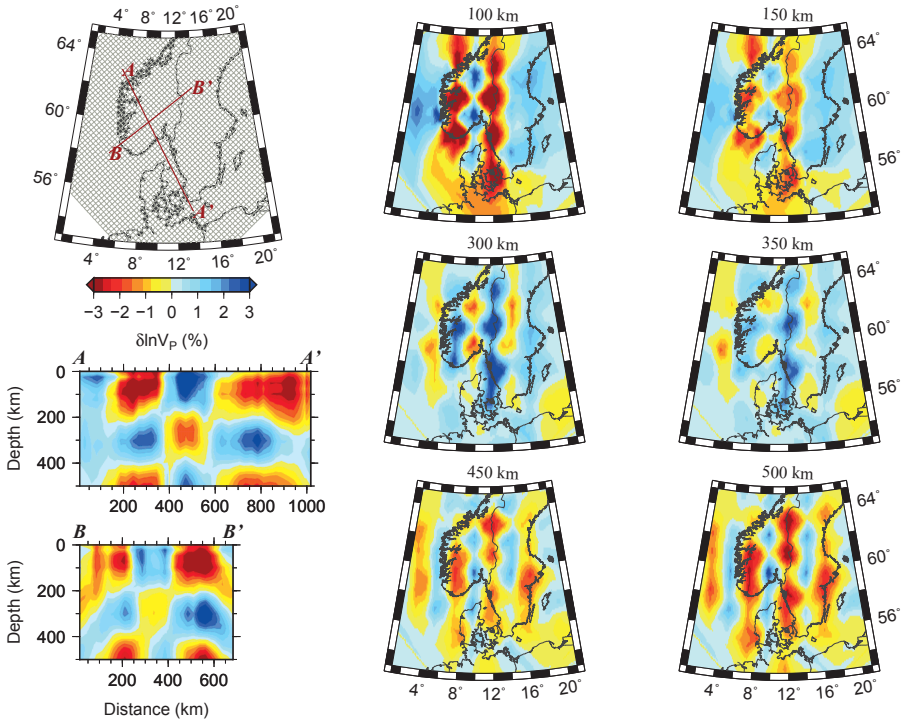


Figure 16. Recovered P-wave model using the 3D multiscale parameterisation for the chequerboard test shown in Fig. 15 and synthetic velocity perturbations $\delta lnV_P = \pm 3\%$. Upper left map shows location of profiles and grid discretisation.

a depth of 350 km (not shown) are much more similar to the anomaly found in the tomographic images in Figs 4 and 5. Similar tests were also performed using broader cylinders with diameters of 120 and 150 km, giving similar results but much stronger recovered anomalies. Hence, we can be quite confident that the cylindrical low velocity anomaly does not extend below 350-400 km depth and that it is indeed quite narrow with a diameter of about 100 km.

We also tested if the cylindrical low-velocity anomaly could result from vertical smearing of separate anomalies by placing two cylinders at depths from $100 - 200$ km and $300 - 400$ km. Both cylinders have a diameter of 100 km and velocity perturbations of $\delta lnV_P = -3\,\%$ and $\delta lnV_S = -5\,\%$ (supplementary Fig. 1). The recovered anomalies (supplementary Figs. 2 and 3) are smeared into a continuous cylinder, but with much smaller amplitudes than the input anomalies: about $-1\,\%$ in Vp and $-1.5\,\%$ in Vs. It is therefore more likely that the strong cylindrical anomaly in the tomography represents a continuous feature than the result of vertical smearing.

4.5 Ray-theoretical multiscale tomographies

As an additional check of our results, we also perform an inversion of the data with a ray-theoretical approach instead of the finite-frequency kernels, facilitating comparison of our results with the recent ray-theoretical tomographies of Medhus et al. (2012) for P-waves and Wawerzinek et al. (2013) for S-waves. These two studies use the same databases as we do, but different methods to measure travel times. We invert the high frequency band of our data sets using the same multiscale parameterisation and gridding as in the finite-frequency tomography. Resolution tests were also performed but are not shown.

The resulting δlnV_P model is displayed in Fig. 25 and shows more similarity with the relative tomography of Medhus et al. (2012) than the finite-frequency model in Fig. 4. The important difference between the ray-theoretical images (Fig. 25) and the finite-frequency images (Fig. 4) is that the inclusion of the low frequency data increases the sensitivity to velocity anomalies off coast and across the Skagerrak Sea between Norway and Denmark. In our ray-theoretical inversions, we lose the clear image of a channel and only sense the western high velocity anomaly in a small area close to the southwestern coast of Norway. This is very similar to the relative tomography of Medhus et al. (2012), and confirms that the inclusion of low frequency data increases resolution due to the different spatial sensitivity of finite-frequency measurements in different pass bands.

A ray-theoretical S-wave tomography is shown in Fig. 26. Compared to the finite-frequency images in Fig. 5 the ray-theoretical tomography has smaller anomalies and

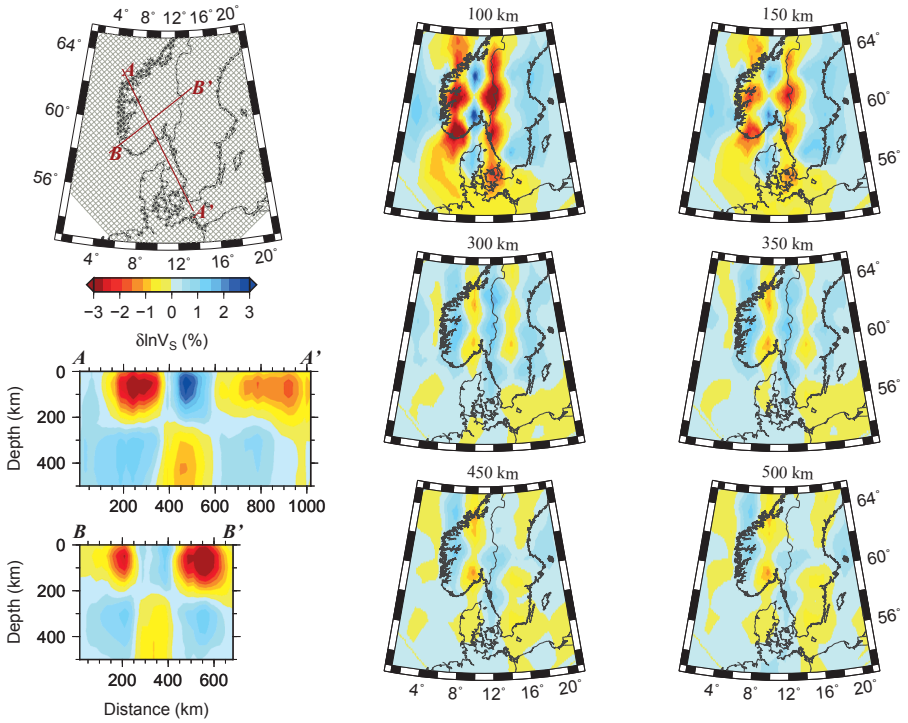


Figure 17. Recovered S-wave model using the 3D multiscale parameterisation for the chequerboard test shown in Fig. 15 and synthetic velocity perturbations $\delta ln V_S = \pm 3\%$. Upper left map shows location of profiles and grid discretisation.

less resolution off coast. The main features of the V_{SH} images in Figs 5 and 26 are similar to the images of Wawerzinek et al. (2013) but the amplitudes are much larger. Wawerzinek et al. (2013) infer a quite deep low velocity anomaly down to depths of 410 km where our images confine this deeply going anomaly to depths of around 350 km and to a smaller horizontal area. The increased recovery of amplitudes in our images can be ascribed to both the multiscale parameterisation and the finite-frequency kernels.

5 DISCUSSION

Resolution tests show that we have three main well-resolved features in our models: a low velocity channel extending from western Norway to Denmark, a deeper cylindrical low velocity anomaly below central southern Norway, and a belt of high velocities below Sweden. The lateral boundary between the high and low velocities is imaged with great detail in the horizontal directions, and its development with depth is also robustly constrained. For V_P/V_S ratios, resolution tests show a high degree of vertical smearing, implying that vertical variations are very poorly constrained, but (depth-averaged) lateral variations are robustly constrained down to 250-300 km depth.

5.1 Comparison to previous seismological studies

Our study reproduces the low velocities below southern Norway and Denmark and high velocities of the Fennoscandian Shield found by recent local geophysical investigations (Maupin 2011; Medhus et al. 2012; Köhler et al. 2012; Wawerzinek et al. 2013; Maupin et al. 2013) and regional studies of western Europe (Weidle & Maupin 2008; Zhu et al. 2012; Fichtner et al. 2013; Rickers et al. 2013; Zhu et al. 2013). Our study adds to this large-scale picture some important details that have not been imaged in previous studies. The low velocities below southern Norway and Denmark consist of two clearly distinct features: a quite narrow channel at shallow depths to the west, and a deeper, cylindrically shaped anomaly centered at the northern end of the Oslo Graben.

On a larger regional scale, Zhu et al. (2012, 2013) image a low velocity channel in V_{SV} and isotropic V_S at shallow depths below southern Norway, and quite interestingly, also an inversion of the low velocities at depths below 475 km where a high velocity anomaly is placed below southern Norway and Denmark. The low velocities below southern Norway are also seen in the regional tomographies of Weidle & Maupin (2008) and Fichtner et al. (2013).

Rickers et al. (2013) focuses on the North Atlantic region and image a channel with approximately −4 % anomaly

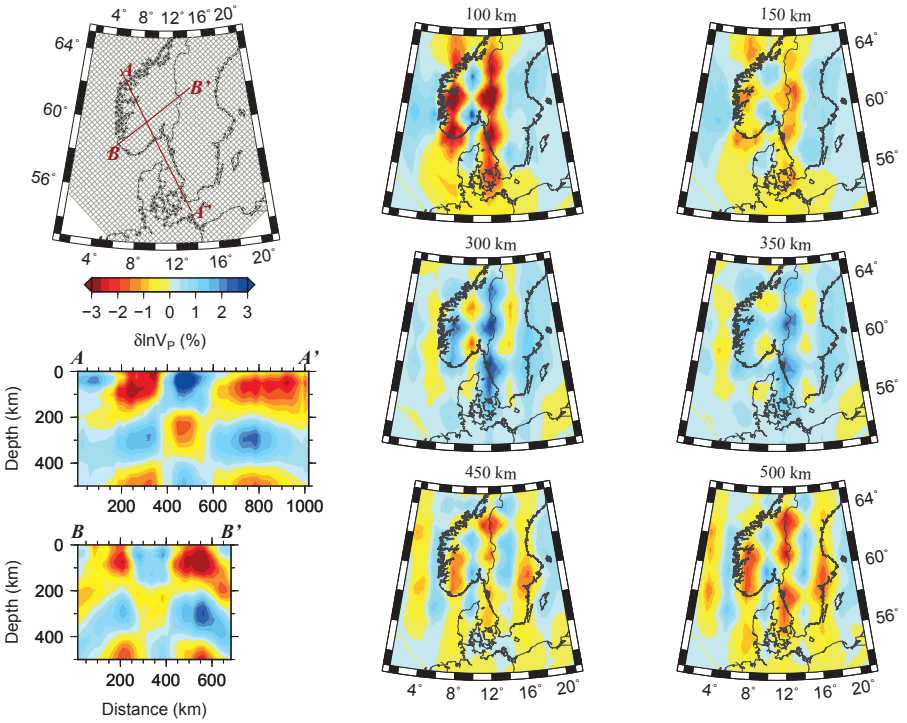


Figure 18. The same as Fig. 16 using the hybrid parameterisation.

in V_{SH} (with respect to a European 1D average model) connecting the Iceland hotspot to southern Norway and Denmark. The low velocity channel has its maximum between depths of 100-200 km but extends to a depth of 300 km. In the North Sea, northwest of southern Norway, the anomaly increases in strength and depth. This is also the region where we infer the lowest velocities, especially in the S-wave tomography (Fig. 5). In the Central Graben area, southwest of Norway and Denmark, Rickers et al. (2013) find high velocity anomalies between 2 to 3 %, confirming the high velocities off coast southern Norway and Denmark in our images. Although the relative magnitudes of anomalies in our images cannot be compared directly to the absolute velocity anomalies in the tomography of Rickers et al. (2013), we find a very similar low velocity channel in V_{SH}, the only difference being the higher degree of detail in our tomography.

Local studies based on surface waves (Maupin 2011; Köhler et al. 2012) also confirm this general picture of low velocities below southern Norway, but on this scale it is even more natural to compare our results with the recent P- and S-wave tomographies of Medhus et al. (2012) and Wawerzinek et al. (2013) already mentioned in the previous sections.

Medhus et al. (2012) were the first to clearly image the lateral boundary between low and high P-wave velocities that continues quite surprisingly from the Tornquist-Teisseyre Zone (TTZ) and Sorgenfrei-Tornquist-Zone (STZ) in northern Germany and Denmark into southern Norway. Our tomographic images show a boundary between low and high velocities in both V_P and V_{SH} that follows the STZ (Fig. 1) to an even larger degree than in the tomographies of Medhus et al. (2012) and with a correlation that extends to depths of around 300 km. The greatest difference between our study and the work of Medhus et al. (2012) is that we infer high velocities off coast western Denmark and Norway, which confines the low velocities to a shallow channel, instead of a boundary between two halfspaces. In addition, we image much more clearly the presence of a deeper low velocity structure below the northwestern end of the Oslo Graben, and infer in general higher magnitudes of velocity anomalies.

Where the P-wave study of Medhus et al. (2012) and S-wave study of Wawerzinek et al. (2013) estimate velocity anomalies of about equal magnitudes, our study images V_{SH} anomalies with magnitudes about twice the size of the V_P anomalies, as expected from both theoretical calculations (e.g. Goes et al. 2000) and seismological models (e.g. Kennett et al. 1998). In addition our P- and S-wave tomographies show much less difference in the spatial distribution of anomalies than the tomographies of Medhus et al. (2012) and Wawerzinek et al. (2013), showing the importance of regularization and parameterisation for seismic models (e.g. Trampert & Snieder 1996; Chiao & Kuo 2001;

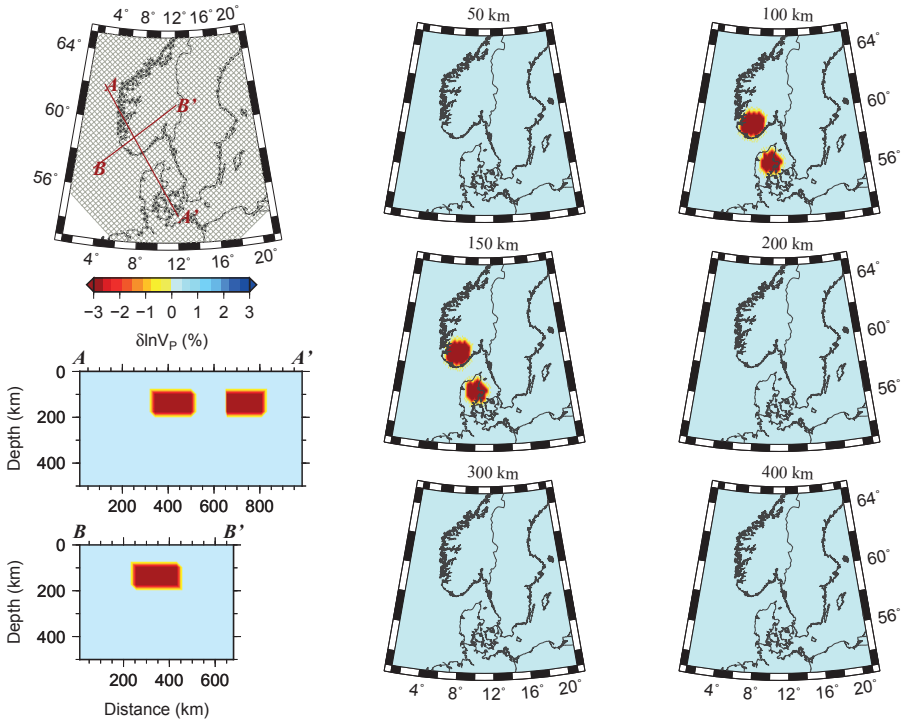


Figure 19. Synthetic input for a test with two cylinders in the depth range from 100-200 km and with diameters of 150 and 180 km. The velocity perturbations are $\delta V_P = -3\%$ and $\delta V_S = -3\%$ or -5%. Upper left map shows location of profiles and grid discretisation.

Nolet 2008; Chiao et al. 2010). Estimating robust perturbations of V_P/V_S is therefore only possible when using similar parameterisation and regularization of the P- and S-wave models, or when making joint inversions of the P- and S-wave data sets.

5.2 Comparison to non-tomographic studies

It seems clear that main structures we image are concordant with previous seismological studies but we recover more realistic magnitudes of the P- and S-wave anomalies.

It is very interesting to note that the gravity modelling of Ebbing et al. (2012) shows a quite strong contribution to surface topography from some non-crustal source, e.g. the upper mantle, in the location where we image the low-velocity channel, but they do not estimate any contribution to topography where we image the deep circular anomaly. This could mean that the causes of the two low velocity anomalies are different. Care should be taken that part of this estimated buoyancy from the mantle could be caused by uncertainties in Moho depth, as receiver function studies and active seismics infer different Moho depths, with up to 8 km difference, in the Western Gneiss Region (WGR, Fig. 1) at the nortwestern coast of southern Norway (Svenningsen et al. 2007; Stratford & Thybo 2011; Frasetto & Thybo 2013; Kolstrup & Maupin 2013).

Another gravity study that estimates a contribution to the surface topography from the mantle is the study of Jones et al. (2002). They use the long-wavelength (> 750 km) free-air gravity anomaly field to estimate the dynamic contribution to topography and infer a contribution on the order of 500-1000 m in southern Norway. Rickers et al. (2013) noted a strong correlation between the low velocity anomalies in their North-Atlantic tomography with both the areas of positive residual topography in Jones et al. (2002) and with continental regions estimated to have experienced major Neogene uplift (e.g. Japsen & Chalmers 2000). This suggests that the correlation between the low velocity anomalies in our images and the mantle contribution to topography in Ebbing et al. (2012) is not coincidental.

It is also interesting that the low-velocity channel is associated with a high degree of intraplate seismic activity (e.g. Lindholm et al. 2000; Bungum et al. 2010) that continues from the North Sea along the coast of western Norway and to northwestern Denmark (Fig. 1). The cause of the seismic activity is not believed to be post-glacial rebound and is more likely caused by a combination of for example ridge push, sedimentation and/or lateral lithospheric differences (Bungum et al. 2010). A similar concentration of intraplate stresses and seismicity in lithospheric "thin spots" has been described in the case of SE Brasil (Assumpção et al. 2004). Our study clearly images lateral lithospheric differences and

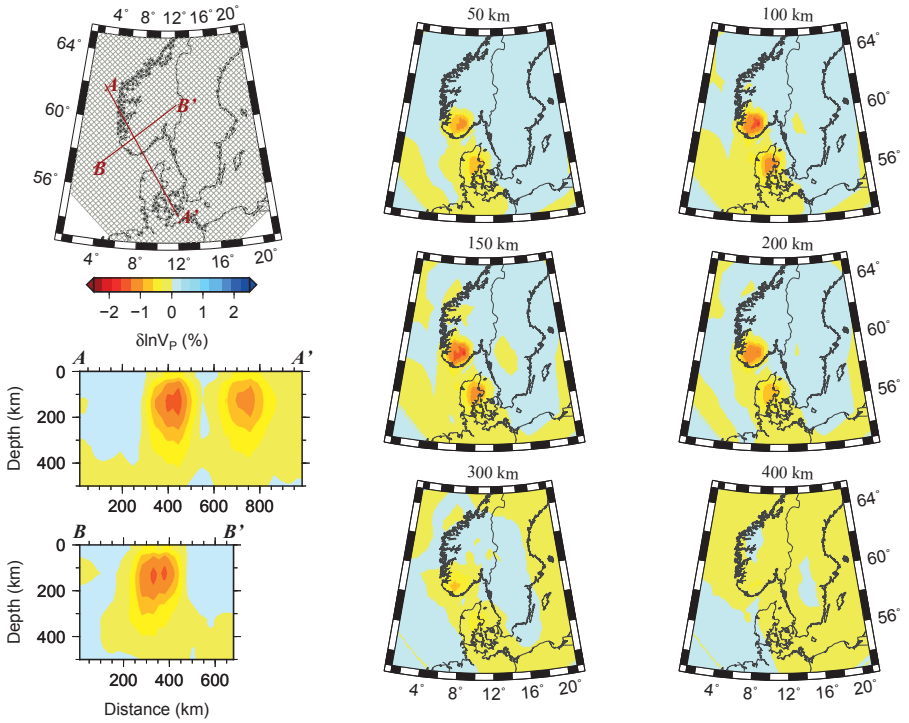


Figure 20. Recovered P-wave model using the 3D multiscale parameterisation in a test with the synthetic model shown in Fig. 19 and velocity perturbation $\delta lnV_P = -3\%$. Upper left map shows location of profiles and grid discretisation.

suggests that intraplate stresses concentrate in areas with a thinner lithosphere. The narrow and deep cylindrical low velocity anomaly in the northwestern end of the Oslo Graben is not associated with any centre of seismic activity, and this again suggests that it is different from the more shallow channel-like anomaly.

5.3 Tectonophysical implications - δlnV_P, δlnV_S and temperature variation

If low seismic velocities below the passive margins of the North Atlantic are indeed associated with dynamic support of topography from the mantle and Neogene uplift, as suggested by Rickers et al. (2013), it is quite important to quantify the anomalies in terms of temperature and/or composition and investigate their tectonic implications.

Within the upper subcontinental mantle and away from the hydrated regions close to subduction zones, seismic velocity anomalies are best explained by variations in temperature (from either convection, changes in lithospheric thickness and/or changes in heat production) and iron content (from melt extraction or metasomatic refertilisation) (e.g. Goes et al. 2000; Poudjom Djomani et al. 2001; Griffin et al. 2009). Separating the effects of the two parameters is notoriously difficult (Cammarano et al. 2003) and perhaps only possible in extreme cases where Archean and tectonically young lithospheres are compared (Goes et al. 2000).

As anhydrous variations in composition give seismic velocity variations smaller than 1-2 % in Proterozoic and younger tectonic regions (Goes et al. 2000; Cammarano et al. 2003; Griffin et al. 2009; Hieronymus & Goes 2010), we can make an end-member interpretation of the high and low velocity anomalies in terms of temperature. Using e.g. derivatives of V_S velocity with respect to temperature from the literature ($\partial lnV_S/\partial T \approx -1\%/100K$, Lee 2003), V_S velocity anomalies of $\pm 3\,\%$ translate into temperature anomalies about ∓ 300°C.

The main problem with this linear interpretation of seismic high and low velocity anomalies into hot and cold, is that shear anelasticity makes the temperature derivatives of V_P and V_S strongly temperature-dependent and the absolute value of the derivatives increases with increasing temperature (Cammarano et al. 2003). Velocity anomalies of different sign therefore give temperature anomalies of different magnitude, and absolute seismic velocities are needed to make a thermal interpretation (Goes et al. 2000; Cammarano et al. 2003). In addition, the anelasticity parameters used in calculating velocities and velocity derivatives (e.g. Berckhemer et al. 1982; Sobolev et al. 1996; Jackson et al. 2002; Shapiro & Ritzwoller 2004) also have a significant effect on the magnitude of estimated velocities (e.g.

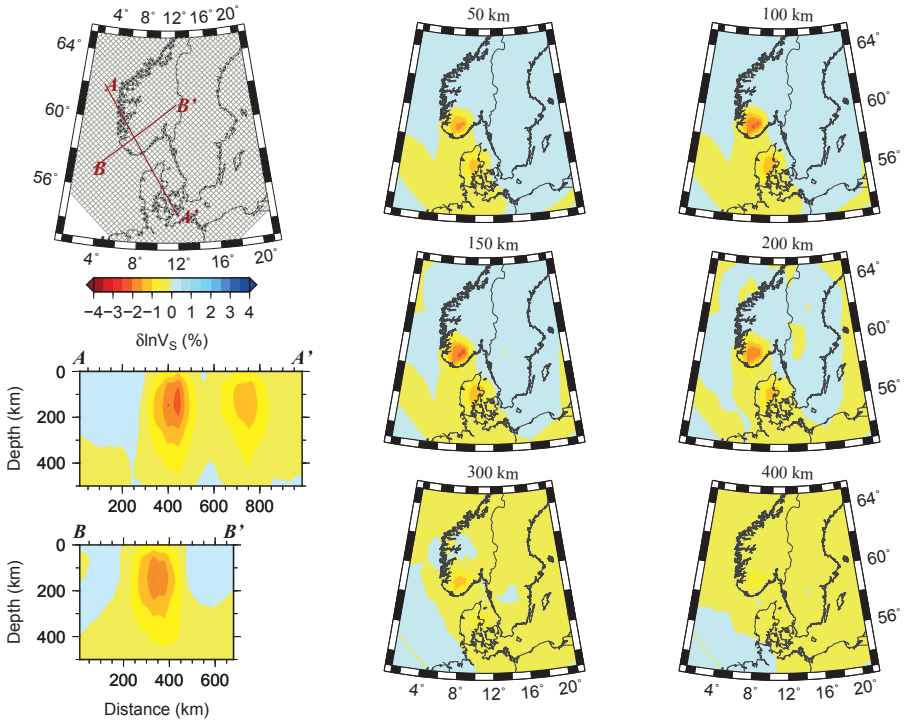


Figure 21. The same as Fig. 20 for $\delta ln V_S = -5\%$.

Goes et al. 2000; Cammarano et al. 2003; Kolstrup et al. 2012).

Due to these complications we take a more indirect approach in the interpretation of the velocity anomalies in Figs 4 to 7 and compare our results to those obtained by thermal lithosphere modelling in the study area (Kolstrup et al. 2012; Gradmann et al. 2013).

Kolstrup et al. (2012) modelled the temperature field and predicted seismic velocities in southern Norway using several geophysical data sets (Moho depth, geoid, surface heat flow), taking into account the dependence of seismic waves on temperature, composition and anelasticity at high temperature and pressure. To limit the range of possible models fitting the data sets, they used a constraint of local isostatic equilibrium for the lithospheric column overlying an adiabatic asthenosphere. Kolstrup et al. (2012) inferred a thin and warm lithosphere (∼100 km) in western Norway and a thick and cold lithosphere in eastern Norway and Sweden (∼200 km), and additionally a slightly thinner and warmer area associated with the Oslo Graben. The maximum relative difference in synthetic seismic shear wave velocity between the thin western lithosphere and the thick eastern lithosphere in Kolstrup et al. (2012) is up to 8.5 % at 100 km depth ($\Delta T \approx 400$°C) and 5.5 % at 150 km depth ($\Delta T \approx 300$°C). In our tomographic images of V_S we find a maximum relative anomaly between southeast and northwest of around 7.5% at 100 km depth and around 5% at 150 km depth (Fig. 5).

Using a more complete 3D modelling approach of the lithosphere in southwestern Scandinavia, Gradmann et al. (2013) estimated an abrupt change in lithospheric thickness from 100 to 200 km across the Oslo Graben, but also needed a change in composition between a fertile southern Norwegian mantle and a depleted Fennoscandian Shield to fit absolute V_{SV} velocities in southern Norway (Maupin 2011) and Sweden (Cotte & Pedersen 2002).

Both modelling approaches show that a lithosphere in local isostatic equilibrium can exhibit strong variations in seismic velocities just from variations in lithospheric thickness and associated temperature variations. It is therefore not necessary to invoke a plume or diapir in the uppermost mantle below southern Norway (Rohrman & van der Beek 1996; Rohrman et al. 2002) to explain the low seismic velocities. A direct translation of the relative velocity anomalies in Figs 4 and 5 into relative temperature anomalies could easily be taken as evidence for anomalously hot temperatures and diapiric upwelling of the asthenosphere into the lithosphere.

Neither the lithospheric modelling of Kolstrup et al. (2012) nor Gradmann et al. (2013) is able to explain the details in our tomographic images, but suggests that the low velocities in the channel-like structure between Denmark and Norway may be due to a thin but stable lithosphere of

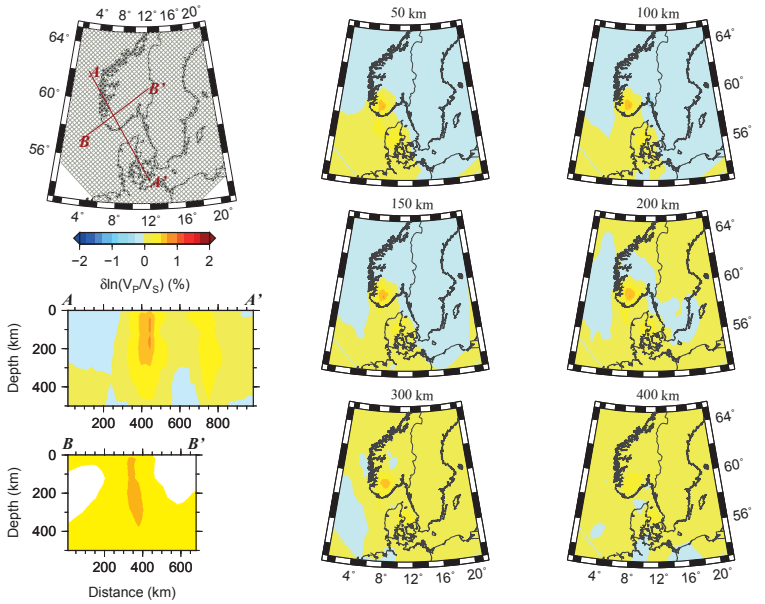


Figure 22. The same as Figs 20 and 21 for $\delta ln(V_P/V_S) = 2\%$.

about 100 km overlying convecting asthenosphere. The high velocities in Sweden and the sharp transition from low to high velocities can likewise be explained by a rapid increase in lithospheric thickness, possibly associated with a higher degree of depletion of the lithospheric mantle below Sweden.

What we cannot explain by variations in lithospheric thickness and mantle depletion is the deep cylindrical low velocity anomaly that does not fit at all with the uniform asthenosphere assumed below the lithosphere in Kolstrup et al. (2012) and Gradmann et al. (2013). Above 200 km depth, the cylinder merges with the low velocity channel and could be ascribed to lithospheric thickness variations, but between depths of 200 and 350 km it is a much stronger feature than the deep part of the channel and it has its own distinct geometry. The location of the structure is slightly west of the northern end of the Oslo Graben and it might therefore have an origin in the extensive Permian magmatism of the Oslo Graben (Neumann et al. 2004; Larsen et al. 2008; Torsvik et al. 2008).

The perhaps most natural explanation for such a structure is a centralized small-scale upwelling, but numerical simulations using a uniform composition for the asthenosphere indicate that upwelling would not give a velocity anomaly as strong as $-3\,\%$ in V_S and $1.5 - 2\,\%$ in V_P (Hieronymus et al. 2007). Hence, the narrow cylinder needs to be anomalously hot or to have an anomalous composition, with for example a higher water content or increased heat production.

The interpretations in this section are based mainly on the anomalies in V_S. The covariation of V_P and V_S in $\delta ln(V_P/V_S)$ provides additional information on the causes of the velocity anomalies, especially on compositional variations, and will be explored in the following section.

5.4 Tectonophysical implications - $\delta ln(V_P/V_S)$ and compositional variation

Many workers infer a variation of the subcontinental mantle composition with time; the oldest Archean compositions being more depleted in basaltic components (especially iron) than the Proterozoic, and the younger Phanerozoic compositions being the most fertile. Variations in iron content, or Mg# (atomic ratio 100*Mg/Mg+Fe), are caused by different degrees of melt extraction and refertilisation from metasomatism (e.g. Poudjom Djomani et al. 2001; Griffin et al. 2009). The depletion of old cratonic mantle contributes its higher velocities and to its tectonic stability (e.g. Jordan 1988; Poudjom Djomani et al. 2001).

Modelling also shows that the compositional difference between depleted and fertile lithospheric mantle, though it does not contribute much in itself to velocity variations, can be crucial for the long-term stability of variations in lithospheric thickness and hence contribute to lateral contrasts in temperature and seismic velocities (Hieronymus et al. 2007). In addition, variations in heat production between fertile and depleted mantle might be an overlooked contributor to seismic velocity anomalies in the subcontinental lithospheric mantle (Hieronymus & Goes 2010).

The variation in V_P/V_S is estimated to be a fairly strong indicator of compositional variations in Mg#, with high values of V_P/V_S being associated with low Mg# and a fertile mantle, and with low values of V_P/V_S being associated with high Mg# and depleted mantle (Lee 2003). We can there-

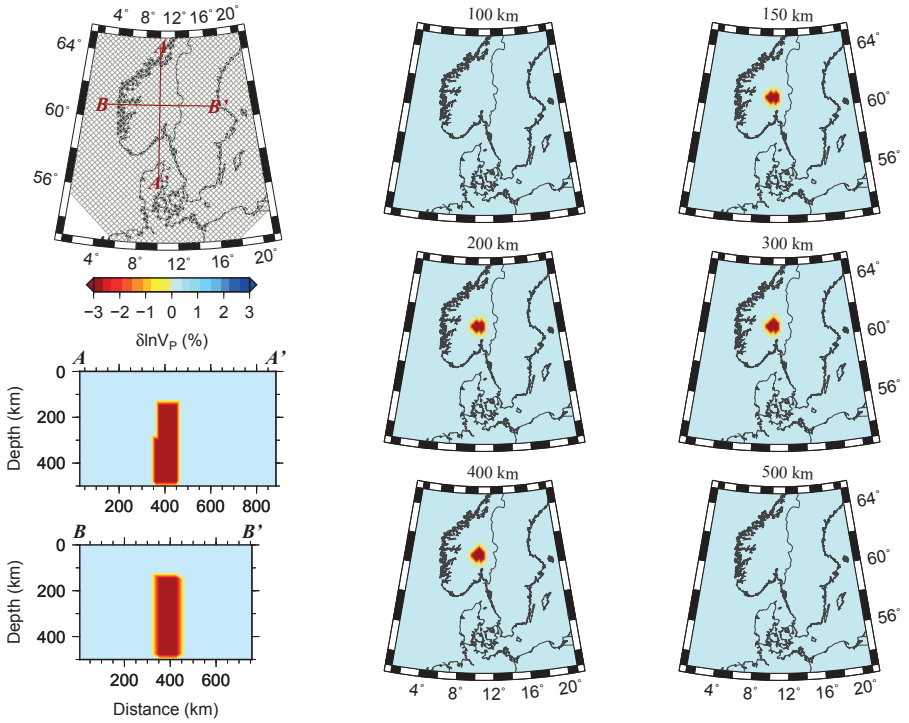


Figure 23. Synthetic input for a test with one cylinder extending from 150 to 500 km depth and with a diameter of 100 km. The velocity perturbations are $\delta lnV_P = -3\%$ and $\delta lnV_S = -3\%$ or -5%. Upper left map shows location of profiles and grid discretisation.

fore use the covariation of δlnV_P and δlnV_S to get an independent estimate of the degree of depletion inferred by Gradmann et al. (2013).

We estimate negative anomalies in $\delta ln(V_P/V_S)$ of up to $-3\,\%$ in Sweden and positive anomalies up to $+2\,\%$ in southern Norway and Denmark, but most of the values are in the range $\pm 1.5\,\%$ (Fig. 9). Below the Oslo Graben we find a positive anomaly of $+2\,\%$ at depths from 200-350 km.

The anomalies in V_P/V_S are relative variations with respect to the unknown 1D mean in our study area. We assume that we can use the regional mean V_P/V_S for Western Europe estimated by Artemieva (2007) from global absolute tomographies (Shapiro & Ritzwoller 2002; Bijwaard & Spakman 2000). At 150 km depth, where the resolution of both the P- and S-wave tomography is good, the average V_P/V_S ratio is estimated to be 1.73 ± 0.04 (Artemieva 2007). Taking this value as our regional mean at 150 km depth, the minimum anomalies of $\pm 1\%$ give a minimum change in V_P/V_S from 1.71 below Sweden to 1.75 below southern Norway. Using the relationship in Lee (2003), this change corresponds to a change in Mg# from a depleted composition (Mg# 93) to an unrealistically fertile composition (Mg# 84).

The relationship of V_P/V_S found in Lee (2003) has been questioned as it is an extrapolation based on data measured at standard pressure and temperature and does not take into account anelastic effects and phase transitions (Afonso et al. 2010). Afonso et al. (2010) find that the relationship between V_P/V_S and Mg# is valid only for garnet bearing assemblages below 900°C as anelastic effects makes the V_P/V_S ratio strongly dependent on temperature at high temperatures. The high V_P/V_S anomalies below southern Norway can therefore be explained by temperature alone and following Afonso et al. (2010) by a temperature anomaly of about +300°C. In Sweden though, the anharmonic V_P/V_S relationship should still hold down to depths of 100-150 km, as temperatures are estimated to be below 1000°C and garnet stable (Afonso et al. 2010; Gradmann et al. 2013). The negative anomalies in $\delta ln(V_P/V_S)$ between $-1\,\%$ and $-3\,\%$ at 50-150 km depth are therefore indicative of a depleted mantle with high Mg#.

It should be noted that Afonso et al. (2010) estimate the compositional difference between a standard tectonic and Proterozoic lithosphere to be quite small, giving variations in $\delta(V_P/V_S)$ of approximately 0.01 that corresponds to anomalies of about $0.6\,\%$ in $\delta ln(V_P/V_S)$. Hence, the V_P/V_S anomalies in this study are too large to be explained by variations in Mg# alone, and temperature must account for some of the variation. As depletion in iron content (increase in Mg#) is associated with depletion in heat productive radiogenic elements leading to thermal anomalies of up to 100-300°C (Hieronymus & Goes 2010), variations in the V_P/V_S ratio due to the total effect of compositional variations in a

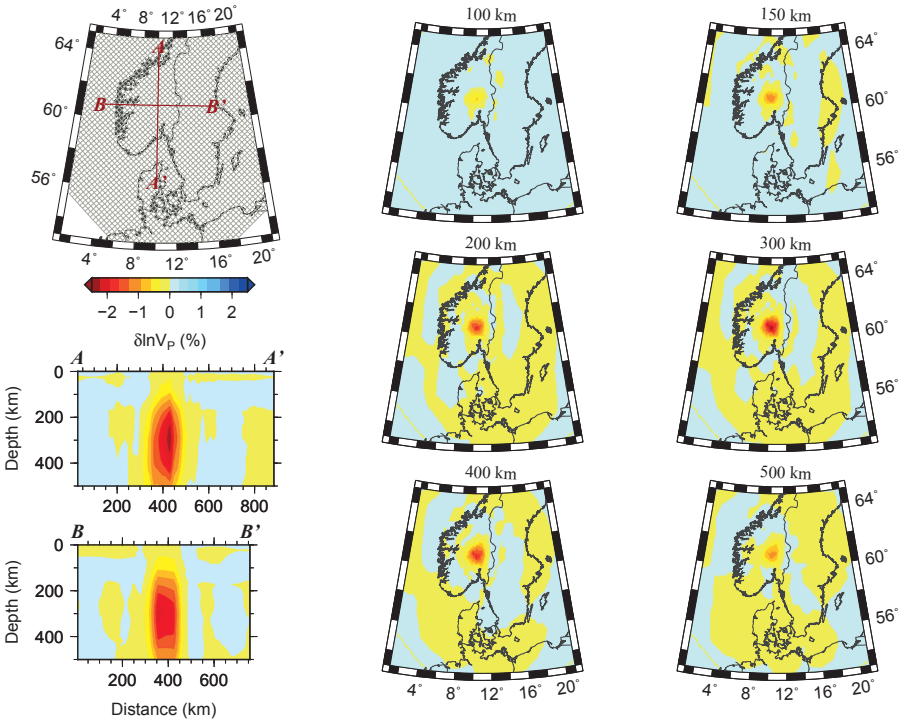


Figure 24. Recovered P-wave model using the 3D multiscale parameterisation with the synthetic model shown in Fig. 23 and a velocity perturbation $\delta lnV_P = -3\%$. Upper left map shows location of profiles and grid discretisation.

heterogeneous mantle might be underestimated in the modelling of Afonso et al. (2010). Variations in heat production associated with depletion of the mantle might therefore explain some of the west to east variation in V_P/V_S in our images.

Artemieva (2007) points out that the ratio V_P/V_S estimated from global studies is not sensitive to compositional variations between lithospheres of age older than 0.9 Ga, but is effective in delineating the outline of the continental shield mantle. Hence, despite the limitations in use of the V_P/V_S ratio, the covariation of δlnV_P and δlnV_S in our images effectively places southern Norway and Denmark outside the Fennoscandian Shield and indicates a strongly depleted shield mantle in Sweden.

5.5 Relation to topography and dynamic models

We find, as previous studies, a lateral mantle boundary coinciding with the Sorgenfrei-Tornquist Zone (STZ) in Denmark and southern Sweden (Fig. 1). The STZ continues as the Tornquist-Teisseyre Zone (TTZ) southeast of our study area, where it separates Phanerozoic Europe from the Precambrian Easteuropean Platform. Both the STZ (Cotte & Pedersen 2002; Shomali & Roberts 2002; Shomali et al. 2006) and TTZ (Zielhuis & Nolet 1994; Janutyte et al. 2014) are associated with a strong lateral boundary in the upper mantle.

Across the STZ, crustal studies show an increase in Moho depth, both in the Skagerrak Sea (Lie & Andersson 1998), in Denmark (Thybo 2001), and in southern Sweden (Thybo 2001). The STZ is, despite the change in crustal thickness, not associated with any marked change in topography.

The STZ and the Skagerrak and Oslo Grabens have a common origin in Carboniferous-Permian extensional tectonics and extensive magmatic activity (Thybo 2001; Neumann et al. 2004; Larsen et al. 2008). Despite this, the lateral mantle boundary was expected to continue westwards and not to take a detour into southern Norway following the graben structures, as the extensively reworked Danish Basin (e.g. Thybo 2001) has a very different crustal structure than the thicker crust of southern Norway (e.g. Stratford & Thybo 2011). Where the lateral mantle boundary branches north from the Oslo Graben, it is no longer associated with any major crustal structures at the surface. On the contrary, the boundary crosses the north-east trending Caledonides and the Scandinavian Mountains (Fig. 1).

Using the STZ as an example, Hieronymus et al. (2007) modelled how edge-driven convection would evolve between lithospheres of contrasting thickness (100 and 250 km) and composition (pyrolite and harzburgite) since the last Trias-

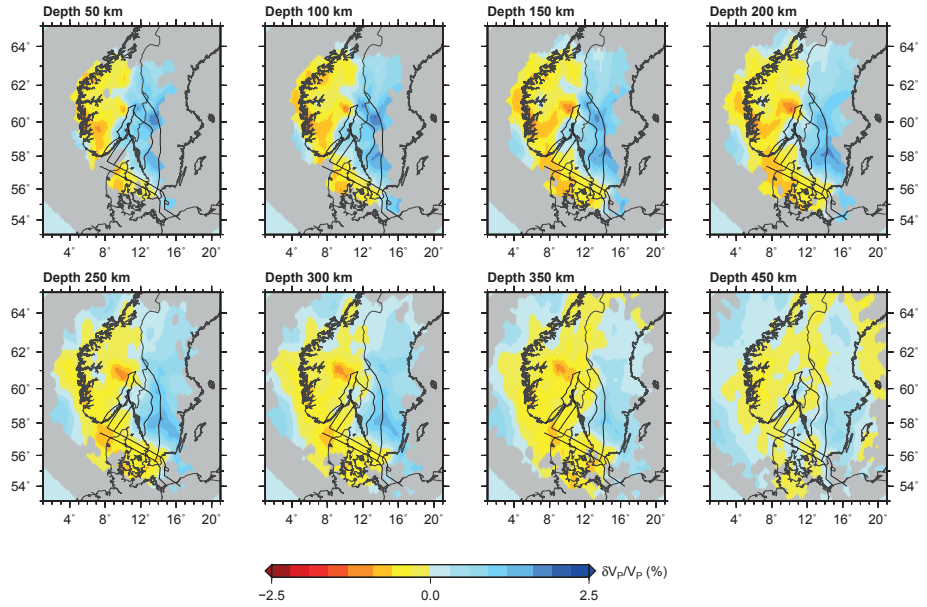


Figure 25. The same as Fig. 4 for the δlnV_P model obtained using only the high frequency P-wave data set and a ray-theoretical approach. The same 3D multiscale parameterisation and damping value as in Fig. 4 is used in the inversion.

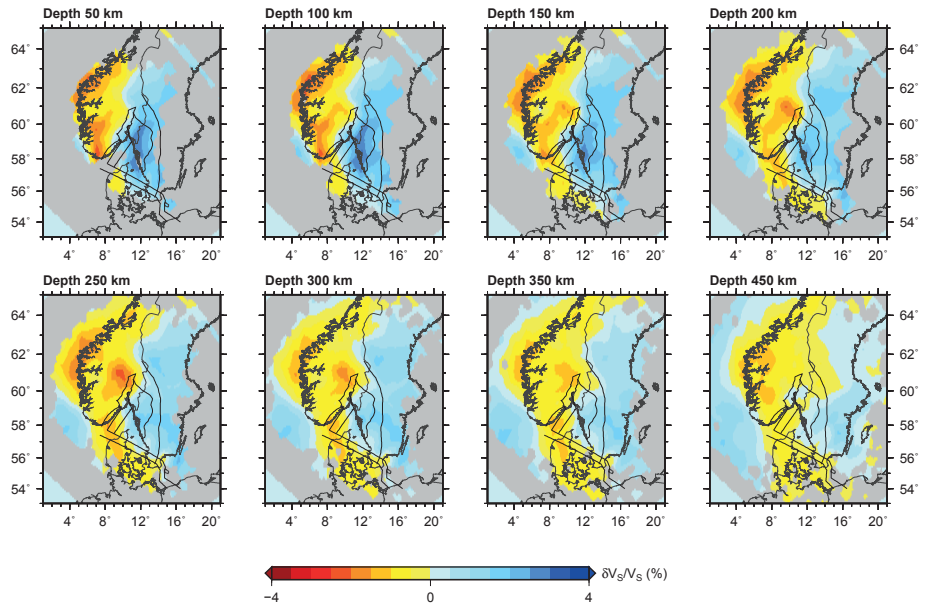


Figure 26. The same as Fig. 5 for the δlnV_S model obtained using only the high frequency S-wave data set and a ray-theoretical approach. The same 3D multiscale parameterisation and damping value as in Fig. 5 is used in the inversion.

sic rifting events 220 Ma. The modelling is able to explain the sharp lateral contrast in seismic velocities across the STZ due to a stable convection pattern. The modelling also shows that the difference in lithospheric thickness is not necessarily associated with any discernible change in surface heat flow, and that the high viscosity of the thicker and older lithosphere stabilizes it against erosion.

The model of Hieronymus et al. (2007) is compatible with the increase in lithospheric thickness inferred from lithospheric modelling in southern Norway (Kolstrup et al. 2012; Gradmann et al. 2013) and explains how such a configuration of the lithosphere can be stable over time. Pascal et al. (2004) suggest that the Oslo Graben developed at the edge of an already existing stepwise increase in lithospheric thickness. It is therefore possible that the sharp contrast in seismic velocities in the upper 50-250 km dates back to even older tectonic events than the Carboniferous-Permian rifting.

Another possible scenario is that the rifting left the originally thick (200 km?) shield mantle below southern Norway partly rehydrated and refertilised. Rehydration can effectively weaken lithospheric mantle, as seen in for example mantle xenoliths of the Proterozoic Colorado Plateau (Li et al. 2008). A lithospheric step would in this case have existed between the Danish Basin and southern Norway, facilitating edge-driven convection. Dynamic modelling shows that small-scale edge-driven convection can erode fertile and hydrated lithosphere (Hieronymus et al. 2007; Van Wijk et al. 2010; Kaislaniemi & van Hunen 2014). Lithospheric erosion by edge-driven convection has been used to explain Cenozoic episodic uplift and magmatism in for example the Atlas Mountains (Kaislaniemi & van Hunen 2014) and the Colorado Plateau (Van Wijk et al. 2010). Furthermore, the opening of the North Atlantic and the Icelandic plume could also have triggered convection and erosion of the southern Norwegian mantle, as regional tomographies point to a connection between the Iceland plume and southern Norway (Weidle & Maupin 2008; Rickers et al. 2013).

Hence, episodic erosion of an initially thick shield mantle below southern Norway to its current thickness of about 100 km could explain some of the episodic uplift inferred from geological observations throughout the Mesozoic and Cenozoic in this area (Gabrielsen et al. 2010). Nielsen et al. (2002) estimate that delamination of 35 km of lithospheric mantle will result in a long-lived uplift greater than 500 m of the average surface. A thin and weak lithosphere is also a necessary prerequisite for uplift from other mechanisms such as lithospheric folding due to far-field compressional stresses (Doré et al. 2008; Cloetingh & Burov 2011).

Japsen et al. (2012) reject edge-driven convection as a possible contributor to the topography of elevated passive continental margins but do not consider newer modelling (e.g. Hieronymus et al. 2007; Van Wijk et al. 2010; Kaislaniemi & van Hunen 2014) performed after the original work of King & Anderson (1998). They find that only far-field compressive stresses and lithospheric folding can explain their observations of uplift in areas such as southern Norway and eastern Greenland. Our images of strongly varying lithospheric properties suggest that the structure and evolution of the mantle must be taken into account as well, and modelling of edge-driven convection points to the importance of erosive weakening of continental margins. Episodic erosion of the lithosphere could act as a trigger of uplift events and weaken the lithospheric margin, facilitating uplift due to compressive stresses.

6 CONCLUSIONS

We present the first finite-frequency P- and S-wave tomographies in southwestern Scandinavia, using a wavelet-based multiscale parameterisation that greatly increases the recovery of amplitude and location of seismic velocity anomalies.

We image, as previous studies, low velocities in both V_P and V_S below southern Norway and Denmark compared to high velocities below Sweden and towards the central Fennoscandian Shield, but add to this picture important details. The low velocity region below southern Norway consists of two adjacent anomalies: a shallow channel-like feature extending from the coastal and western part of southern Norway into Denmark, and a cylindrically shaped anomaly that emerges from the channel further east at 150-200 km depth and extends to a depth of around 350 km. Furthermore, we find that the lateral boundary between the low and high velocities closely follows Carboniferous-Permian rift structures and centres of magmatic activity, i.e. the Sorgenfrei-Tornquist Zone and the Oslo Rift. Anomalies in the V_P/V_S ratio are robustly constrained in the upper part of the model but depth variations cannot be imaged with confidence. Below Sweden we estimate negative anomalies in the V_P/V_S ratio down to $-3\,\%$ and below southern Norway positive anomalies up to $+2\,\%$.

The variations in δlnV_P, δlnV_S and $\delta ln(V_P/V_S)$ point consistently to higher temperatures below southern Norway and Denmark compared to Sweden and also to a higher degree of depletion in the lithospheric mantle below Sweden. The low velocities and higher temperatures in Denmark and Norway can be explained to a large part simply by a thinner lithosphere, but the deeper cylindrical low velocity anomaly below the northwestern end of the Oslo Graben is more difficult to explain.

The channel-like low velocity region is associated with a high degree of intraplate seismicity and an estimated contribution to the support of the high topography by low density material in the uppermost mantle. It is possible that the thin lithosphere below southern Norway and the stepwise increase in thickness towards the central Fennoscandian Shield is older than the Carboniferous-Permian rifting. On the other hand, it is equally possible that an originally thick but weakened lithosphere has been eroded to its current thickness of about 100 km after the Permian, causing episodic uplift of the area.

ACKNOWLEDGEMENTS

This work has been done in the framework of the ESF EUROCORES TOPO-EUROPE Program 07-TOPO-EUROPE-FP-014 and was supported by MOST grant,102-2116-M-002-025, in Taiwan (S.-H. Hung). MAGNUS waveforms were recorded with the mobile KArlsruhe BroadBand Array (KABBA) of the Karlsruhe Institute of Technology, Germany as well as with permanent stations of the NORSAR array and the Norwegian National Seismological Net-

work. We thank Andy Frassetto for providing us with the DANSEIS data (University of Copenhagen) and Marie Keiding (Geological Survey of Norway) for help with the FENCAT database. We are grateful to Niels Balling (Univerity of Aarhus) for sharing data from the CALAS experiment with us. Seismological stations of the CALAS project, applied in this study, were from the University of Aarhus instrument pool and from the NERC Geophysical Equipment Facility (SEIS-UK). That project was financially supported by the Danish National Science Research Council. Financial support for the MAGNUS experiment was provided by the Universities of Aarhus, Copenhagen, Karlsruhe and Oslo as well as NORSAR. This work greatly improved from the constructive reviews of Karin Sigloch and an anonymous reviewer. Figures have been prepared using the Generic Mapping Tools (Wessel & Smith 1998) and M_Map (Pawlowicz 2005).

REFERENCES

Afonso, J. C., Ranalli, G., Fernàndez, M., Griffin, W. L., O'Reilly, S. Y., & Faul, U., 2010. On the Vp/Vs–Mg# correlation in mantle peridotites: Implications for the identification of thermal and compositional anomalies in the upper mantle, *Earth planet. Sci. Lett.*, **289**(3), 606–618.

Aki, K., Christoffersson, A., & Husebye, E. S., 1977. Determination of the three-dimensional seismic structure of the lithosphere, *J. geophys. Res.*, **82**(2), 277–296.

Artemieva, I. M., 2007. Dynamic topography of the East European craton: shedding light upon lithospheric structure, composition and mantle dynamics, *Glob. Planet Change*, **58**(1), 411–434.

Assumpção, M., Schimmel, M., Escalante, C., Barbosa, J. R., Rocha, M., & Barros, L. V., 2004. Intraplate seismicity in SE Brazil: stress concentration in lithospheric thin spots, *Geophysical Journal International*, **159**(1), 390–399.

Berckhemer, H., Kampfmann, W., Aulbach, E., & Schmeling, H., 1982. Shear modulus and Q of forsterite and dunite near partial melting from forced-oscillation experiments, *Phys. Earth Planet. Inter.*, **29**(1), 30–41.

Bijwaard, H. & Spakman, W., 2000. Non-linear global P-wave tomography by iterated linearized inversion, *Geophys. J. Int.*, **141**(1), 71–82.

Bingen, B., Skår, Ø., Marker, M., Sigmond, E., Nordgulen, Ø., Ragnhildstveit, J., Mansfeld, J., Tucker, R., & Liégeois, J., 2005. Timing of continental building in the Sveconorwegian orogen, SW Scandinavia, *Norw. J. Geol.*, **85**, 87–116.

Bungum, H., Olesen, O., Pascal, C., Gibbons, S., Lindholm, C., & Vestøl, O., 2010. To what extent is the present seismicity of Norway driven by post-glacial rebound?, *J. Geol. Soc.*, **167**(2), 373–384.

Cammarano, F., Goes, S., Vacher, P., & Giardini, D., 2003. Inferring upper-mantle temperatures from seismic velocities, *Phys. Earth Planet. Inter.*, **138**(3), 197–222.

Charléty, J., Voronin, S., Nolet, G., Loris, I., Simons, F. J., Sigloch, K., & Daubechies, I. C., 2013. Global seismic tomography with sparsity constraints: Comparison with smoothing and damping regularization, *J. geophys. Res.*, **118**(9), 4887–4899.

Chevrot, S. & Zhao, L., 2007. Multiscale finite-frequency rayleigh wave tomography of the kaapvaal craton, *Geophys. J. Int*, **169**(1), 201–215.

Chevrot, S., Martin, R., & Komatitsch, D., 2012. Optimized discrete wavelet transforms in the cubed sphere with the lifting scheme – implications for global finite-frequency tomography, *Geophys. J. Int*, **191**(3), 1391–1402.

Chiao, L.-Y. & Kuo, B.-Y., 2001. Multiscale seismic tomography, *Geophys. J. Int.*, **145**(2), 517–527.

Chiao, L.-Y. & Liang, W.-T., 2003. Multiresolution parameterization for geophysical inverse problems, *Geophysics*, **68**(1), 199–209.

Chiao, L.-Y., Lin, J.-R., & Gung, Y.-C., 2006. Crustal magnetization equivalent source model of Mars constructed from a hierarchical multiresolution inversion of the Mars Global Surveyor data, *J. geophys. Res.*, **111**, E12010.

Chiao, L.-Y., Fang, H.-Y., Gung, Y.-C., Chang, Y.-H., & Hung, S.-H., 2010. Comparative appraisal of linear inverse models constructed via distinctive parameterizations (comparing distinctly inverted models), *J. geophys. Res.*, **115**, B07305.

Chou, H.-C., Kuo, B.-Y., Chiao, L.-Y., Zhao, D., & Hung, S.-H., 2009. Tomography of the westernmost Ryukyu subduction zone and the serpentinization of the fore-arc mantle, *J. geophys. Res.*, **114**, B12301.

Cloetingh, S. & Burov, E., 2011. Lithospheric folding and sedimentary basin evolution: a review and analysis of formation mechanisms, *Basin Research*, **23**(3), 257–290.

Cocks, L. & Torsvik, T., 2006. European geography in a global context from the Vendian to the end of the Palaeozoic, in *European Lithosphere Dynamics*, vol. 32 of **Geol. Soc. Lond. Mem.**, pp. 83–95, eds Gee, D. & Stephenson, R., Geological Society, London.

Cohen, A., Daubechies, I., & Feauveau, J.-C., 1992. Biorthogonal bases of compactly supported wavelets, *Communications on pure and applied mathematics*, **45**(5), 485–560.

Cotte, N. & Pedersen, H., 2002. Sharp contrast in lithospheric structure across the Sorgenfrei–Tornquist Zone as inferred by Rayleigh wave analysis of TOR1 project data, *Tectonophysics*, **360**(1), 75–88.

Dahlen, F., Hung, S.-H., & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes – Theory, I., *Geophys. J. Int.*, **141**(1), 157–174.

Doré, A., Lundin, E., Kusznir, N., & Pascal, C., 2008. Potential mechanisms for the genesis of Cenozoic domal structures on the NE Atlantic margin: pros, cons and some new ideas, *Geol. Soc. Lond. Spec. Publ.*, **306**(1), 1–26.

Ebbing, J., England, R., Korja, T., Lauritsen, T., Olesen, O., Stratford, W., & Weidle, C., 2012. Structure of the Scandes lithosphere from surface to depth, *Tectonophysics*, **536**, 1–24.

Euler, G., 2014. Project SEIZMO, http://epsc.wustl.edu/∼ggeuler/codes/m/seizmo/.

FENCAT, 2011. Institute of Seismology, University of Helsinki, http://www.helsinki.fi/geo/seismo/english/bulletins/index.html.

Fichtner, A., Saygin, E., Taymaz, T., Cupillard, P., Capdeville, Y., & Trampert, J., 2013. The deep structure of the North Anatolian Fault Zone, *Earth planet. Sci. Lett.*, **373**, 109–117.

Frasetto, A. & Thybo, H., 2013. Receiver function analysis of the crust and upper mantle in Fennoscandia – isostatic implications, *Earth planet. Sci. Lett.*, **381**, 234–246.

Frederiksen, S., Nielsen, S. B., & Balling, N., 2001. A numerical dynamic model for the Norwegian–Danish Basin, *Tectonophysics*, **343**(3), 165–183.

Gaál, G. & Gorbatschev, R., 1987. An outline of the Precambrian evolution of the Baltic Shield, *Precambrian Res.*, **35**, 15–52.

Gabrielsen, R. H., Faleide, J. I., Pascal, C., Braathen, A., Nystuen, J. P., Etzelmuller, B., & O'Donnell, S., 2010. Latest Caledonian to Present tectonomorphological development of southern Norway, *Mar. Petrol. Geol.*, **27**(3), 709 – 723, The link between shallow and deep processes in sedimentary basins.

Goes, S., Govers, R., & Vacher, P., 2000. Shallow mantle temperatures under Europe from P and S wave tomography, *J. geophys. Res.*, **105**(B5), 11153–11169.

Gradmann, S., Ebbing, J., Fullea, J., et al., 2013. Integrated geophysical modelling of a lateral transition zone in the litho-

spheric mantle under Norway and Sweden, *Geophys. J. Int.*, **194**(3), 1359–1374.

Griffin, W., O'Reilly, S. Y., Afonso, J. C., & Begg, G., 2009. The composition and evolution of lithospheric mantle: a re-evaluation and its tectonic implications, *J. Petrol.*, **50**(7), 1185–1204.

Hieronymus, C. & Goes, S., 2010. Complex cratonic seismic structure from thermal models of the lithosphere: effects of variations in deep radiogenic heating, *Geophys. J. Int.*, **180**(3), 999–1012.

Hieronymus, C. F., Shomali, Z. H., & Pedersen, L. B., 2007. A dynamical model for generating sharp seismic velocity contrasts underneath continents: application to the Sorgenfrei–Tornquist Zone, *Earth planet. Sci. Lett.*, **262**(1), 77–91.

Högdahl, K., Andersson, U., & Eklund, O., 2004. *The Transscandinavian Igneous Belt (TIB) in Sweden: a review of its character and evolution*, vol. 37 of **Special Paper**, Geological Survey of Finland.

Hung, S.-H., Shen, Y., & Chiao, L.-Y., 2004. Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach, *J. geophys. Res.*, **109**, B08305.

Hung, S.-H., Chen, W.-P., Chiao, L.-Y., & Tseng, T.-L., 2010. First multi-scale, finite-frequency tomography illuminates 3-D anatomy of the Tibetan Plateau, *Geophys. Res. Lett.*, **37**, L06304.

Hung, S.-H., Chen, W.-P., & Chiao, L.-Y., 2011. A data-adaptive, multiscale approach of finite-frequency, traveltime tomography with special reference to P and S wave data from central Tibet, *J. geophys. Res.*, **116**, B06307.

Jackson, I., Gerald, F., John, D., Faul, U. H., & Tan, B. H., 2002. Grain-size-sensitive seismic wave attenuation in polycrystalline olivine, *J. geophys. Res.*, **107**(B12), 2360.

Janutyte, I., Majdanski, M., Voss, P., Kozlovskaya, E., et al., 2014. Upper mantle structure around the Trans-European Suture Zone obtained by teleseismic tomography, *Solid Earth Discussions*, **6**(2), 1723–1763.

Japsen, P. & Chalmers, J. A., 2000. Neogene uplift and tectonics around the North Atlantic: overview, *Glob. Planet. Change*, **24**(3-4), 165–173.

Japsen, P., Chalmers, J. A., Green, P. F., & Bonow, J. M., 2012. Elevated, passive continental margins: Not rift shoulders, but expressions of episodic, post-rift burial and exhumation, *Glob. Planet Change*, **90**, 73–86.

Jones, S. M., White, N., Clarke, B. J., Rowley, E., & Gallagher, K., 2002. Present and past influence of the Iceland Plume on sedimentation, *Geol. Soc. Lond. Spec. Publ.*, **196**(1), 13–25.

Jordan, T. H., 1988. Structure and formation of the continental tectosphere, *J. Petrol.*, (1), 11–37.

Kaislaniemi, L. & van Hunen, J., 2014. Dynamics of lithospheric thinning and mantle melting by edge-driven convection: Application to Moroccan Atlas mountains, *Geochem. Geophys. Geosyst.*, **15**, 3175–3189.

Kennett, B. & Gudmundsson, O., 1996. Ellipticity corrections for seismic phases, *Geophys. J. Int.*, **127**(1), 40–48.

Kennett, B., Widiyantoro, S., & Hilst, R. v. d., 1998. Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. geophys. Res.*, **103**(B6), 12469–12493.

Kennett, B. L. N., Engdahl, E. R., & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, **122**(1), 108–124.

King, S. D. & Anderson, D. L., 1998. Edge-driven convection, *Earth planet. Sci. Lett.*, **160**(3), 289–296.

Köhler, A., Weidle, C., & Maupin, V., 2012. Crustal and uppermost mantle structure of southern Norway: results from surface wave analysis of ambient seismic noise and earthquake data, *Geophys. J. Int.*, **191**(3), 1441–1456.

Kolstrup, M., 2015. *Old sutures and young plumes? – New geophysical investigations of the crust and upper mantle in southwestern Scandinavia*, Ph.D. thesis, University of Oslo.

Kolstrup, M. L. & Maupin, V., 2013. A Proterozoic boundary in southern Norway revealed by joint-inversion of P-receiver functions and surface waves, *Precambrian Res.*, **238**, 186–198.

Kolstrup, M. L., Pascal, C., & Maupin, V., 2012. What compensates the topography of southern Norway? Insights from thermo-isostatic modeling, *J. Geodyn.*, **61**, 105–119.

Larsen, B. T., Olaussen, S., Sundvoll, B., & Heeremans, M., 2008. The Permo-Carboniferous Oslo Rift through six stages and 65 million years, *Episodes*, **31**(1), 52–58.

Lee, C.-T. A., 2003. Compositional variation of density and seismic velocities in natural peridotites at STP conditions: Implications for seismic imaging of compositional heterogeneities in the upper mantle, *J. geophys. Res.*, **108**(B9), 2441.

Li, Z.-X. A., Lee, C.-T. A., Peslier, A. H., Lenardic, A., & Mackwell, S. J., 2008. Water contents in mantle xenoliths from the Colorado Plateau and vicinity: Implications for the mantle rheology and hydration-induced thinning of continental lithosphere, *J. geophys. Res.*, **113**, B09210.

Lidmar-Bergstrom, K. & Bonow, J. M., 2009. Hypotheses and observations on the origin of the landscape of southern Norway - A comment regarding the isostasy-climate-erosion hypothesis by Nielsen et al. 2008, *J. Geodyn.*, **48**(2), 95–100.

Lie, J. & Andersson, M., 1998. The deep-seismic image of the crustal structure of the Tornquist Zone beneath the Skagerrak Sea, northwestern Europe, *Tectonophysics*, **287**(1), 139–155.

Lindholm, C. D., Bungum, H., Hicks, E., & Villagran, M., 2000. Crustal stress and tectonics in Norwegian regions determined from earthquake focal mechanisms, *Geol. Soc. Lond. Spec. Publ.*, **167**(1), 429–439.

Loris, I., Nolet, G., Daubechies, I., & Dahlen, F., 2007. Tomographic inversion using l1-norm regularization of wavelet coefficients, *Geophysical Journal International*, **170**(1), 359–370.

Loris, I., Douma, H., Nolet, G., Daubechies, I., & Regone, C., 2010. Nonlinear regularization techniques for seismic tomography, *Journal of computational physics*, **229**(3), 890–905.

Lou, X., van der Lee, S., & Lloyd, S., 2013. AIMBAT: A Python/Matplotlib tool for measuring teleseismic arrival times, *Seismol. Res. Lett.*, **84**(1), 85–93.

Martin, M., Ritter, J., et al., 2005. High-resolution teleseismic body-wave tomography beneath SE RomaniaŮI. Implications for three-dimensional versus one-dimensional crustal correction strategies with a new crustal velocity model, *Geophys. J. Int.*, **162**(2), 448–460.

Maupin, V., 2011. Upper-mantle structure in southern Norway from beamforming of Rayleigh wave data presenting multipathing, *Geophys. J. Int.*, **185**(2), 985–1002.

Maupin, V. & Kolstrup, M., 2014. Insights in P and S relative travel time tomography from analyzing finite-frequency Fréchet kernels, *Submitted to Geophysical Journal International*.

Maupin, V., Agostini, A., Artemieva, I., Balling, N., Beekman, F., Ebbing, J., England, R., Frassetto, A., Gradmann, S., Jacobsen, B., et al., 2013. The deep structure of the Scandes and its relation to tectonic history and present-day topography, *Tectonophysics*, **602**, 15–37.

Medhus, A., Balling, N., Jacobsen, B., Weidle, C., England, R., Kind, R., Thybo, H., & Voss, P., 2012. Upper-mantle structure beneath the Southern Scandes Mountains and the Northern Tornquist Zone revealed by P-wave traveltime tomography, *Geophys. J. Int.*, **189**(3), 1315–1334.

Medhus, A. B., Balling, N., Jacobsen, B. H., Kind, R., & England, R. W., 2009. Deep-structural differences in southwestern Scandinavia revealed by P-wave travel time residuals, *Norw. J. Geol.*, **89**, 203–214.

Menke, W., 1984. *Geophysical data analysis: discrete inverse theory*, Academic press, San Diego, Calif.

Meyerholtz, K. A., Pavlis, G. L., & Szpakowski, S. A., 1989. Convolutional quelling in seismic tomography, *Geophysics*, **54**(5),

570–580.

Neumann, E., Wilson, M., , Heeremans, M., Spencer, E., Obst, K., Timmerman, M., & Kirstein, L., 2004. Carboniferous-Permian rifting and magmatism in southern Scandinavia, the North Sea and northern Germany: a review, *Geol. Soc. Lond. Spec. Publ.*, **223**(1), 11–40.

Nielsen, S. B., Paulsen, G. E., Hansen, D. L., Gemmer, L., Clausen, O. R., Jacobsen, B. H., Balling, N., Huuse, M., & Gallagher, K., 2002. Paleocene initiation of Cenozoic uplift in Norway, *Geol. Soc. Lond. Spec. Publ.*, **196**(1), 45–65.

Nielsen, S. B., Gallagher, K., Egholm, D. L., Clausen, O. R., & Summerfield, M., 2009a. Reply to comment regarding the ICE-hypothesis, *J. Geodyn.*, **48**(2), 101–106.

Nielsen, S. B., Gallagher, K., Leighton, C., Balling, N., Svenningsen, L., Jacobsen, B. H., Thomsen, E., Nielsen, O. B., Heilmann-Clausen, C., Egholm, D. L., Summerfield, M. A., Clausen, O. R., Piotrowski, J. A., Thorsen, M. R., Huuse, M., Abrahamsen, N., King, C., & Lykke-Andersen, H., 2009b. The evolution of western Scandinavian topography: A review of Neogene uplift versus the ICE (isostasy-climate-erosion) hypothesis, *J. Geodyn.*, **47**(2-3), 72–95.

Nironen, M., 1997. The Svecofennian Orogen: a tectonic model, *Precambrian Res.*, **86**(1), 21–44.

Nolet, G., 2008. *A breviary of seismic tomography: Imaging the Interior of the Earth and Sun*, Cambridge University Press, Cambridge.

Obayashi, M., Suetsugu, D., & Fukao, Y., 2004. PP-P differential traveltime measurement with crustal correction, *Geophys. J. Int*, **157**(3), 1152–1162.

Pascal, C. & Olesen, O., 2009. Are the Norwegian mountains compensated by a mantle thermal anomaly at depth?, *Tectonophysics*, **475**(1), 160–168.

Pascal, C., Cloetingh, S., & Davies, G., 2004. Asymmetric lithosphere as the cause of rifting and magmatism in the Permo-Carboniferous Oslo Graben, in *Permo-Carboniferous Magmatism and Rifting in Europe*, vol. 223, pp. 139–156, eds Wilson, M., Neumann, E.-R., Davies, G., Timmerman, M., Heeremans, M., & Larsen, B. T., Geol. Soc. Lond. Spec. Publ.

Pawlowicz, R., 2005. M_Map: A mapping package for Matlab, http://www.eos.ubc.ca/∼rich/map.html.

Poudjom Djomani, Y. H., OŠReilly, S. Y., Griffin, W., & Morgan, P., 2001. The density structure of subcontinental lithosphere through time, *Earth planet. Sci. Lett.*, **184**(3), 605–621.

Rickers, F., Fichtner, A., & Trampert, J., 2013. The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: Evidence from full-waveform inversion, *Earth planet. Sci. Lett.*, **367**, 39–51.

Ritsema, J., Van Heijst, H., Woodhouse, J., & Deuss, A., 2009. Long-period body wave traveltimes through the crust: implication for crustal corrections and seismic tomography, *Geophys. J. Int*, **179**(2), 1255–1261.

Roberts, D., 2003. The Scandinavian Caledonides: event chronology, palaeogeographic settings and likely modern analogues, *Tectonophysics*, **365**(1-4), 283–299.

Rohrman, M. & van der Beek, P., 1996. Cenozoic postrift domal uplift of North Atlantic margins: an asthenospheric diapirism model., *Geology*, **24**, 901–904.

Rohrman, M., Van Der Beek, P. A., Van Der Hilst, R. D., & Reemst, P., 2002. Timing and mechanisms of North Atlantic Cenozoic uplift: evidence for mantle upwelling, *Geol. Soc. Lond. Spec. Publ.*, **196**, 27–44.

Shapiro, N. & Ritzwoller, M., 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J. Int.*, **151**(1), 88–105.

Shapiro, N. & Ritzwoller, M., 2004. Thermodynamic constraints on seismic inversions, *Geophys. J. Int.*, **157**(3), 1175–1188.

Shomali, Z., Roberts, R. G., & Pedersen, L. B., 2006. Lithospheric structure of the Tornquist Zone resolved by nonlinear P and S teleseismic tomography along the TOR array, *Tectonophysics*, **416**(1), 133–149.

Shomali, Z. H. & Roberts, R. G., 2002. Non-linear body wave teleseismic tomography along the TOR array, *Geophys. J. Int.*, **148**(3), 562–574.

Sigloch, K., 2008. *Multiple-frequency body-wave tomography*, Ph.D. thesis, Princeton University.

Simons, F. J., Loris, I., Nolet, G., Daubechies, I. C., Voronin, S., Judd, J., Vetter, P. A., Charléty, J., & Vonesch, C., 2011. Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity, *Geophys. J. Int*, **187**(2), 969–988.

Sobolev, S. V., Zeyen, H., Stoll, G., Werling, F., Altherr, R., & Fuchs, K., 1996. Upper mantle temperatures from teleseismic tomography of French Massif Central including effects of composition, mineral reactions, anharmonicity, anelasticity and partial melt, *Earth planet. Sci. Lett.*, **139**(1), 147–163.

Sørensen, K., 1986. Danish Basin subsidence by Triassic rifting on a lithosphere cooling background, *Nature*, **319**, 660Ű663.

Stratford, W. & Thybo, H., 2011. Seismic structure and composition of the crust beneath the southern Scandes, Norway, *Tectonophysics*, **502**(3), 364–382.

Svenningsen, L., Balling, N., Jacobsen, B., Kind, R., Wylegalla, K., & Schweitzer, J., 2007. Crustal root beneath the highlands of southern Norway resolved by teleseismic receiver functions, *Geophys. J. Int*, **170**(3), 1129–1138.

Sweldens, W., 1996. The lifting scheme: A custom-design construction of biorthogonal wavelets, *Applied and computational harmonic analysis*, **3**(2), 186–200.

Thybo, H., 2001. Crustal structure along the EGT profile across the Tornquist Fan interpreted from seismic, gravity and magnetic data, *Tectonophysics*, **334**(3), 155–190.

Torsvik, T. H., Smethurst, M. A., Burke, K., & Steinberger, B., 2008. Long term stability in deep mantle structure: Evidence from the¯ 300 Ma Skagerrak-Centered Large Igneous Province (the SCLIP), *Earth and Planetary Science Letters*, **267**(3), 444–452.

Trampert, S. & Snieder, R., 1996. Model estimations biased by truncated expansions: possible artifacts in seismic tomography, *Science*, **271**, 1257–1260.

Van Wijk, J., Baldridge, W., Van Hunen, J., Goes, S., Aster, R., Coblentz, D., Grand, S., & Ni, J., 2010. Small-scale convection at the edge of the Colorado Plateau: Implications for topography, magmatism, and evolution of Proterozoic lithosphere, *Geology*, **38**(7), 611–614.

VanDecar, J. & Crosson, R., 1990. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. seism. Soc. Am.*, **80**(1), 150–169.

Wawerzinek, B., Ritter, J., & Roy, C., 2013. New constraints on the 3D shear wave velocity structure of the upper mantle underneath Southern Scandinavia revealed from non-linear tomography, *Tectonophysics*, **602**, 38–54.

Weidle, C. & Maupin, V., 2008. An upper-mantle S-wave velocity model for Northern Europe from Love and Rayleigh group velocities, *Geophys. J. Int.*, **175**(3), 1154–1168.

Weidle, C., Maupin, V., Ritter, J., Kvarna, T., Schweitzer, J., Balling, N., Thybo, H., Faleide, J. I., & Wenzel, F., 2010. MAGNUS - A Seismological Broadband Experiment to Resolve Crustal and Upper Mantle Structure beneath the Southern Scandes Mountains in Norway, *Seismol. Res. Lett.*, **81**(1), 76–84.

Wessel, P. & Smith, W., 1998. New improved version of the generic mapping tools released, *EOS Trans. AGU*, **79**(33), 579.

Yang, T. & Shen, Y., 2006. Frequency-dependent crustal correction for finite-frequency seismic tomography, *Bull. seism. Soc. Am.*, **96**(6), 2441–2448.

Yuan, Y. O. & Simons, F. J., 2014. Multiscale adjoint

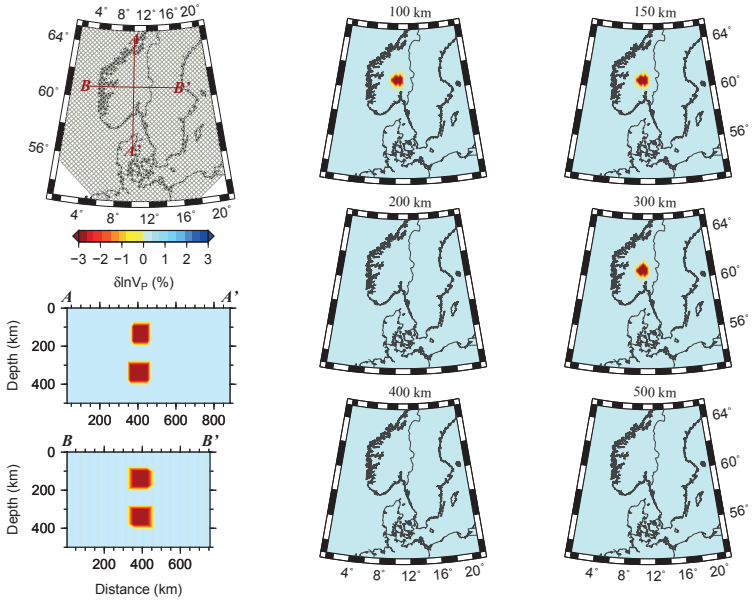
waveform-difference tomography using wavelets, *Geophysics*, **79**(3), WA79–WA95.

Zhu, H., Bozdağ, E., Peter, D., & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography, *Nature Geoscience*, **5**(7), 493–498.

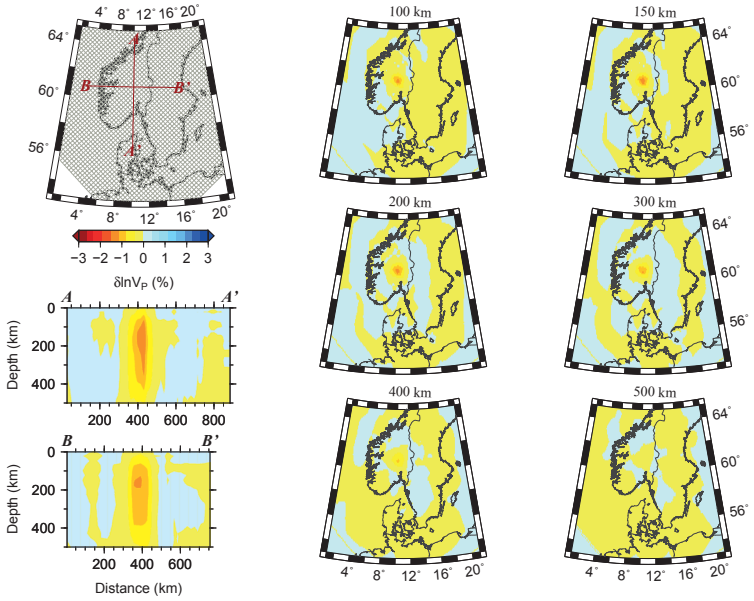
Zhu, H., Bozdağ, E., Duffy, T. S., & Tromp, J., 2013. Seismic attenuation beneath Europe and the North Atlantic: implications for water in the mantle, *Earth planet. Sci. Lett.*, **381**, 1–11.

Zielhuis, A. & Nolet, G., 1994. Deep seismic expression of an ancient plate boundary in Europe, *Science*, pp. 79–79.

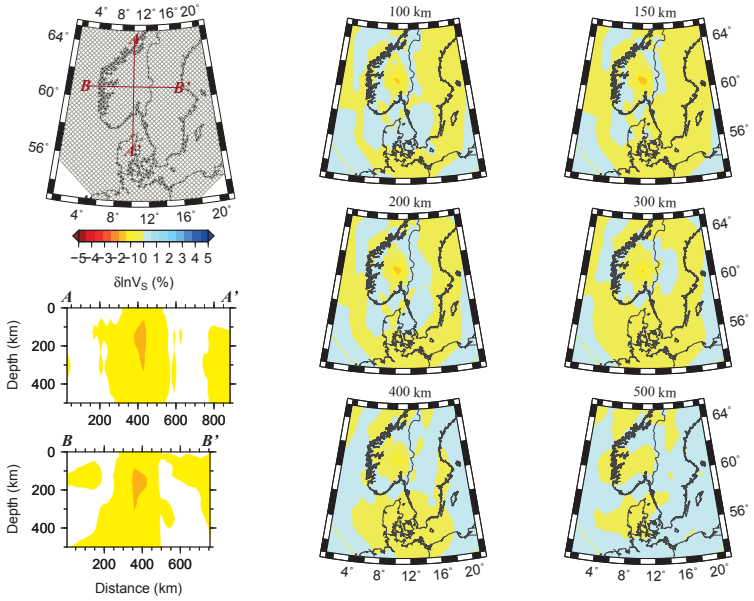
3.5 Appendix - Supplementary figures



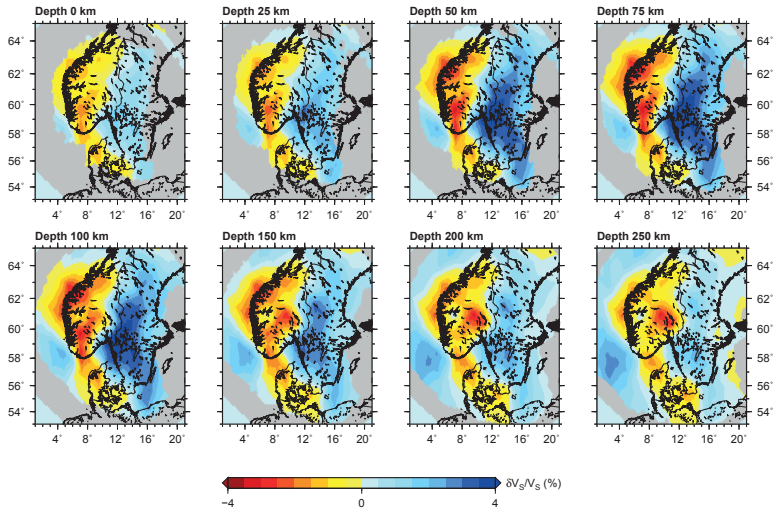
Supplementary Figure 1: Synthetic input for a test with two cylinders placed at $100 - 200$ km and $300 - 400$ km depth and with diameters of 100 km. The velocity perturbations are -3% for Vp and -5% for Vs. Upper left map shows location of profiles and grid discretisation.



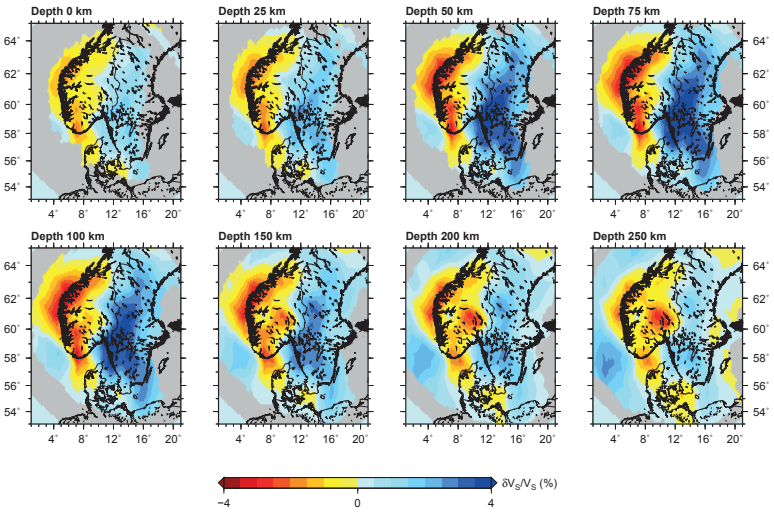
Supplementary Figure 2: Recovered P-wave model using the 3D multiscale parameterisation in a test with the synthetic model shown in Supplementary Fig. 1.



Supplementary Figure 3: Recovered S-wave model using the 3D multiscale parameterisation in a test with the synthetic model shown in Supplementary Fig. 1.



Supplementary Figure 4: Images of δlnVs in the uppermost layers, corresponding to Fig. 13 for P. Input data are the raw travel time residuals without the crustal and topographic corrections from Kolstrup (2015).



Supplementary Figure 5: As supplementary Fig. 4 but with input data corrected for topography and crustal structure, corresponding to Fig. 14 for P.

3.6 A narrow plume below southern Norway?

We image in chapter 3.4 a narrow plume-like feature below southern Norway at depths between 150 and 350 km, with anomalies in V_P of -2% and in V_S of -3%. It has a diameter of about 100 km and is located slightly west of the northern end of the Oslo Graben. The apparent plume is quite well constrained with depth, and does not extend below 350 km. At 150 km depth the plume seems to merge with the shallow channel-like anomaly that connects western Norway and Denmark, and possibly also connects Norway and Iceland (Weidle & Maupin, 2008; Rickers et al., 2013).

As mentioned in chapter 3.4, the plume-like structure is difficult to explain by upwelling from small-scale, edge-driven convection, as this process gives very small velocity contrasts in the sublithospheric mantle (Hieronymus et al., 2007). Hence, a more localized upwelling with stronger anomalies in temperature and/or composition is probably needed to explain the velocity structure.

It is important to understand that the lack of a seismic velocity anomaly deeper than 350 km does not exclude the presence of a thermal anomaly at depth, as the sensitivity of seismic velocities to temperature decrease with increasing temperatures (Cammarano et al., 2003; Goes et al., 2004). In other words, a larger temperature anomaly is needed at greater depths to produce the same seismic velocity anomaly as at shallower depths. This is one of the reasons why mantle plumes are difficult to image, and it may not be possible to image the continuity of a whole mantle plume (Goes et al., 2004; Hwang et al., 2011).

The plume-like structure does not have a shape similar to the classic, purely thermal, "head-and-tail" plumes (Ribe et al., 2007, and references therein), but a structure more similar that of a thermochemical "failed dome" (Ribe et al., 2007), or to a "tail-without-head" plume (Farnetani & Samuel, 2005). The latter structure has its origin from a thermochemical plume that partly ponds beneath the 660 km phase transition and partly penetrates into the upper mantle (Farnetani & Samuel, 2005). The modelling of Farnetani & Samuel (2005) shows that thermochemical plumes can have a variety of shapes, and that mantle winds driven by tectonic plate motion can alter the plume geometry significantly.

To thoroughly evaluate the causes of the plume-like structure in our tomographies, we need a quantification of the seismic structure of different types of small-scale plumes in the upper mantle. To our knowledge, quantification of seismic structure from dynamic modelling has only been done for thermal whole mantle plumes (Goes et al., 2004), or for the lower mantle (Farnetani & Samuel, 2005). Due to this lack of modelling results, we compare instead our plume-like feature to tomographic models of small-scale upper mantle plumes in Europe: the Eifel plume (Ritter, 2007) and the French Massif Central Plume (Barth et al., 2007).

The Eifel plume has been imaged with both P- and S-wave tomography and is estimated to extend from depths of 50-60 km to at least 410 km (Ritter et al., 2001; Keyser et al., 2002; Ritter, 2007). The plume is associated with relative velocity anomalies down to -2% for P and -3% for S at depths of 50-100 km and smaller velocity anomalies of about -1% for both wave types below 100 km (Ritter et al., 2001; Keyser et al., 2002). It is a quite narrow feature with a diameter of 100-120 km, and is interpreted as a small upper mantle plume that buoyantly upwells from the mantle transition zone (660 km) (Ritter, 2007).

The French Massif Central plume been imaged using several tomographic methods (Granet et al., 1995; Barth et al., 2007, and references within) and is inferred to be a narrow structure with a diameter of 100-120 km. It extends from depths of about 60 km to a least 330 km depth, and have relative V_P anomalies between -0.8 and -0.5% (Barth et al., 2007) and absolute V_P anomalies of about -2% with respect to the radial global Earth model *sp*6 (Morelli & Dziewonski, 1993; Piromallo & Morelli, 2003).

The main similarity between the plume-like structure below southern Norway and the plumes of central Europe is the narrow diameter of about 100-120 km and the depth extension to somewhere between 330 and 410 km. The most important difference is that the central European plumes extend to very shallow depths (50-60 km) and are associated with Cenozoic volcanism, whereas the plume-like structure below southern Norway merges with a more regional low-velocity anomaly at 150 km depth and is clearly not associated with recent volcanic activity. Relative velocity anomalies from different tomographies and different study regions are difficult to compare, but the anomalies of -2% for P and -3% for S below southern Norway are at least of the same order of magnitude as the anomalies found below the Eifel region.

If the plume-like feature below southern Norway is indeed a plume from either the transition zone (660 km) or from the lower mantle, it is important to estimate where it was back in time, especially in the Neogene where southern Norway probably experienced a tectonic uplift (e.g. Anell et al., 2009; Gabrielsen et al., 2010; Japsen et al., 2012). According to the ITRF2008 model (Altamimi et al., 2011), the Eurasian plate is moving about 25 km/Myr in a northeastern direction. The plate velocities in the ITRF models (Altamimi et al., 2002, 2007, 2011) are relative to the mean lithosphere reference frame (the no-net-rotation reference frame) (e.g. Argus & Gordon, 1991), which does not take into account the westward drift of the mean lithosphere with respect to the mantle (Bostrom, 1971; Ricard et al., 1991; Gripp & Gordon, 2002; Argus & Gross, 2004; Altamimi et al., 2007). Hence, in a hotspot reference frame that represents the lower mantle, southern Norway moves in a southwestern direction, with a velocity of about 10-20 km/Myr (Gripp & Gordon, 2002). The proper choice of reference frame for the

plume-like structure in southern Norway is probably the hot spot reference frame, as it is unlikely that a deep structure (150-350 km) moves along with the lithospheric plate. Since the upper mantle may have a net eastwards flow (Doglioni, 1993; Doglioni et al., 2005), it is possible that the velocity of 10-20 km/Myr towards southwest is a lower estimate of the relative motion between the lithosphere of southern Norway and the underlying upper mantle.

Using the velocity of 10-20 km/Myr towards the southwest implies that the might-be-plume was about 300-600 km southwest from its current position 30 Ma, relative to the Eurasian plate plate. During the Neogene, the plume should therefore have moved from somewhere in the North Sea to central southern Norway (Fig. 1.1), following approximately the profile B from west to east (Figs 6 and 7 in section 3.4). This comparison of the plume-like structure with plate motion therefore suggests an alternative scenario to explain the present-day seismic structure: that the passage of the lithospheric plate above a plume-like upwelling heated the plate and triggered erosion of the mantle lithosphere below southern Norway. The two apparently separate low velocity anomalies (the shallow channel and deep cylinder) might therefore be linked through dynamic processes. This scenario is off course very speculative and in high need of testing from for example dynamic modelling.

Chapter 4

Thermo-isostatic lithosphere modelling

4.1 Introduction

This chapter presents the results of thermal and isostatic lithosphere modelling in southern Norway. The basic assumptions behind the applied method are that the lithosphere is in both thermal and isostatic equilibrium, and that the asthenosphere follows an adiabatic temperature profile. The numerical method was adapted from work of Pascal (2006) by Kolstrup (2010) and uses three fully coupled equations to describe temperature, pressure and density of the lithospheric mantle. Synthetic seismic velocities are calculated from the distributions of mantle temperature, pressure and density, taking into account anelastic attenuation.

The method can be described as a non-linear iterative inversion for mantle temperature and temperature-related density, but we used the term modelling in the work, as the method is very different from seismic inversion methods. First of all, because a unique solution of mantle temperature and density is found, and secondly because the input data are non-seismological data sets: topography, Moho depth, crustal density, surface heat flow and geoid undulations.

The seismic velocities predicted from this lithosphere model of southern Norway are very useful in the interpretation of the tomographic results in chapter 3.4 as they show the the range of velocity variations possible in a steady-state lithosphere. Seismic low-velocity anomalies are often taken as evidence for temperature anomalies, but there is nothing anomalous in the temperature distribution in this lithosphere model. There are clearly lateral temperature *variations* but they are the results of variations in lithosphere thickness, not indicators of diapirism and thermal disequilibrium.

This type of isostatic modelling therefore gives important insights in the structure of

the lithosphere in the study area. On the other hand, we find that the imposed a priory assumptions of thermal and isostatic equilibrium put very strong bounds on the inverted models and represents a rather static view of the world. Hence, the inverted lithosphere model is rather an end-member model of several possible subsurface scenarios that can be used to compare with seismological models and to inspire further investigations.

The modelling in this chapter was done with a newer compilation of crustal structure than in the work of Kolstrup (2010), but before the joint inversion for crustal S-wave velocities described in chapter 2. The S-wave velocity models and Moho depths from chapter 2 are therefore not included in the compilation of crustal structure used in this chapter.

4.2 Manuscript - Journal of Geodynamics

Chapter 5

Conclusions and outlook

The seismic models of the crust and upper mantle presented in this dissertation show several intriguing features: an old Proterozoic boundary preserved in the lower crust; an upper-mantle boundary between high and low velocities following Paleozoic rift structures; and a young plume-like feature, most likely of Cenozoic age. We summarize in this chapter our findings on crust and mantle structure and discuss the results in relation to geological processes at passive continental margins. To keep the text short and easily readable, we only include the most important references, or references that are not given in the previous chapters.

5.1 Crust and Moho structure

The main findings on the crust and Moho structure from P-wave receiver functions and surface waves are summarized below, with an outlook for future research.

Main findings

- We quantify the character of the crust-mantle transition by calculating its maximum velocity gradient with depth at each station. The gradients are plotted in map views, facilitating geological interpretations. The analysis shows that the crust-mantle transition varies from a sharp discontinuity to a gradual transition in different geological blocks in the study area.

- We find a sharp and shallow (28-42 km) Moho discontinuity in all the areas that where heavily deformed in the Sveconorwegian orogeny (1.2-0.9 Ma). These parts of southern Norway are found southwest of the Mylonite Zone and the Sveconorwegian Front (Fig. 5.1). The sharp and clear Moho is probably related to extension in the last stages of the Sveconorwegian Orogeny.

- Northeast of the Mylonite Zone and Sveconorwegian Front, the crust-mantle transition is deep, gradual, and located between depths of 40 and 50 km. The central part of the area is covered by Caledonian Nappes, and the age and geochemistry of the crust below is therefore not well known. The southeastern part comprises granitoids of the Transscandinavian Igneous Belt (TIB, Fig. 5.1) of late Paleoproterozoic age (1.71-1.65 Ga) that most likely formed in an Andean type subduction setting with subduction towards the northeast. The presence of large-scale magmatic underplates that are only partially eclogitized (Leech, 2001), i.e. a small-scale mixture of metastable gabbro and eclogite, can be an explanation of the deep and gradual crust-mantle transition.

- The lateral boundary between a sharp Moho and a gradual crust-mantle transition has a northwest strike similar to that of the major tectonic boundaries in the Archean and Paleoproterozoic provinces of the Fennoscandian Shield. This northwest strike is approximately perpendicular to the trace of the Caledonides (Fig. 5.1). This suggests that the main features of the lower crustal velocity structure are Proterozoic in age and survived the Caledonian orogeny.

- The area where a discrepancy exists between the Moho depth inferred from active refraction seismics (Stratford & Thybo, 2011a) and receiver function imaging (Frasetto & Thybo, 2013) coincides with the area where we infer a deep and gradual crust-mantle transition. This suggests that the active seismics, based on refracted waves from above, are more sensitive to the upper part of the gradual crust-mantle transition, whereas the receiver functions, based on teleseismic waves from below, have a more event sampling of the transition.

An outlook for future work

- Many studies focus on properly locating the depth of the Moho interface, but we suggest that a quantitative analysis of the characteristics of the crust-mantle transition in a geological context is equally important. It would be very interesting to extend the analysis to the rest of the Fennoscandian Shield, especially the Svecofennian areas, in order to map the extent of the area having a gradual crust-mantle transition. Such a mapping could give important insights into the geological history of the shield and the nature of the continental crust-mantle transition.

- Another important question is, what causes diffuse and gradual crust-mantle transitions? The hypothesis of partial eclogitization in Leech (2001) is very attractive and reasonable, but we find it could improve from a quantification of seismic velocities

resulting from different mixtures of mafic gabbro and eclogite, especially taking into account the different scale levels of crustal heterogeneities and seismic waves.

5.2 Upper mantle structure

The results regarding both mantle structure and methods for mantle investigations are summarized in the following, with suggestions for further work.

Main findings

- We present an automated data processing procedure tailored for measuring relative, finite-frequency travel-time residuals. The procedure combines the ICCS and MCCC algorithms, and uses the low-frequency residuals as input to the processing of higher frequency bands. This reduces the risk of cycle skipping inherent in cross-correlation methods like ICCS and MCCC.

- The comparison of absolute and relative finite-frequency Fréchet kernels shows that their differences are negligible for high-frequency waves in a typical regional network configuration. The differences are not negligible for low-frequency waves, as the relative low-frequency kernels show a sensitivity reduction with depth ($\sim$25 % at 400 km depth with respect to the absolute kernels). Using absolute kernels with relative travel times therefore implies less recovery of deep anomalies and a slightly higher data misfit due to the inflicted discrepancy between the high- and low-frequency kernels.

- Calculation of crustal delay times with finite-frequency kernels including both near- and far-field terms shows that ray-theoretical crustal corrections are good approximations of finite-frequency corrections, and better than corrections calculated with kernels based on a far-field approximation. Long-period waves are sensitive to a smoothed version of the crustal structure, and a filtered Moho map (wavelengths $\gtrsim$ 100 km) can be used to calculate low-frequency ray-theoretical corrections that approximate well the finite-frequency corrections.

- The main structures of the upper mantle found with P- and S-wave tomography are:

  - A low-velocity channel extending from off coast western Norway to Denmark, strongest at depths between $50 - 200$ km for S and $100 - 200$ km for P, with anomalies of -2% in V_P and -3.5% in V_S.

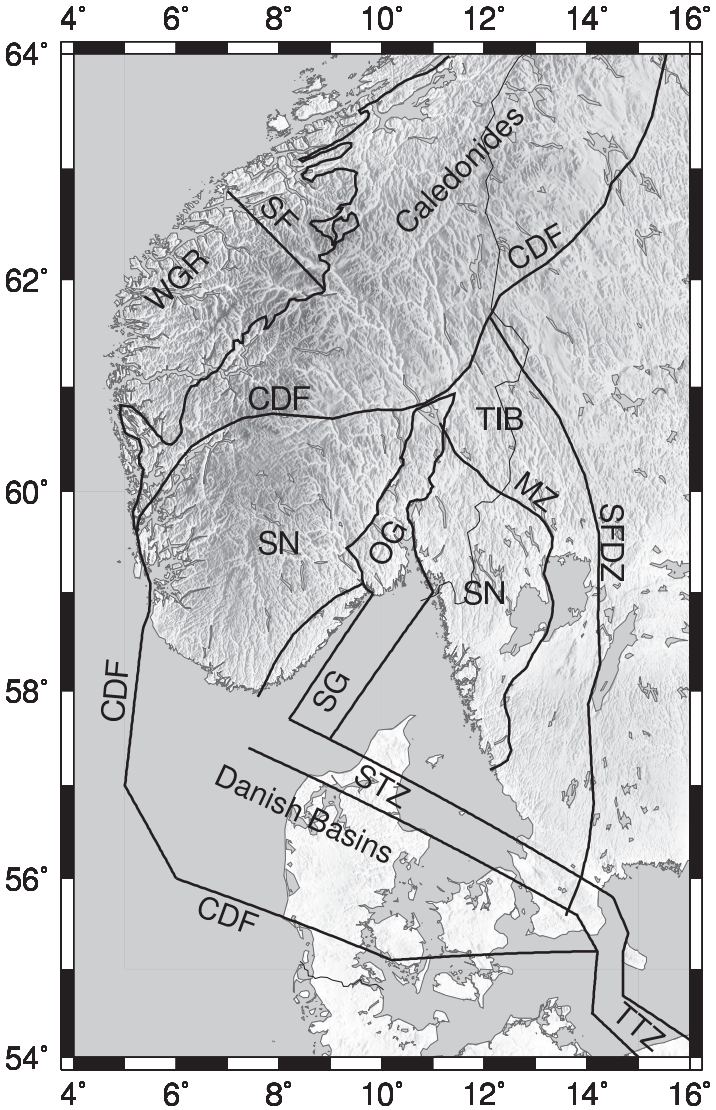


Figure 5.1: Tectonic and topographic map of the study area. CDF, Caledonian front; MZ, Mylonite Zone; OG, Oslo Graben; SG, Skagerrak Graben; SN, Sveconorwegian; SF, Sveconorwegian front; SFDZ Sveconorwegian Frontal Deformation Zone; STZ, Sorgenfrei-Tornquist Zone; TIB, Transscandinavian Igneous Belt; TTZ, Teisseyre-Tornquist Zone; WGR, Western Gneiss Region.

  - A narrow cylindrically shaped low-velocity anomaly located northwest of the Oslo Graben. The plume-like anomaly emerges from the low-velocity channel at a depth of 150 km and extends to depth of 350 km, with a diameter of about 100 km, and anomalies of -2% in V_P and -3% in V_S.
  - A belt of high-velocity anomalies in Sweden, located primarily east of the Mylonite Zone for P and east of the Oslo Graben for S, and with values of $+2.5\%$ in V_P and $+4\%$ in V_S.
  - The covariation of V_P and V_S is quantified in the anomalies of their ratio, $\delta ln(V_P/V_S)$. These anomalies are robustly constrained in the upper part of the models (< 250 km). Positive anomalies are found below southern Norway and negative anomalies are found below Sweden, with magnitudes in the range of -3% and $+2\%$, respectively.

- The shallow low-velocity channel can be explained by a relatively thin lithosphere in Southern Norway (~ 100 km) juxtaposed to the thick shield lithosphere in Sweden. The low velocity channel coincides with an estimated contribution to the topography of western Norway from low-density material in the uppermost mantle (Ebbing et al., 2012), and also with a high degree of intraplate seismicity off coast western Norway. Thermo-isostatic modelling suggests that the high topography of western Norway is partially supported by low densities in a relatively thin lithospheric mantle. A thin lithosphere can explain the presence of low density material, and can also explain a concentration of intraplate stress and seismicity in lithosphere thin spots (Assumpção et al., 2004).

- The most natural explanation for the deeper low-velocity anomaly is a localized plume-like upwelling. The imaged structure has similar seismic characteristics as the Eifel plume and French Massif Central plume (Barth et al., 2007; Ritter, 2007), but does not reach as shallow depths. If the plume-like structure has remained stationary with respect to the lower mantle the last 30 Myr, it implies that it has moved relative to the Eurasian plate from somewhere below the North Sea to central southern Norway during the Oligocene and Neogene.

- The positive anomalies in $\delta ln(V_P/V_S)$ below southern Norway and Denmark can be explained by elevated temperatures alone, whereas the negative anomalies in southern Sweden also indicate a high degree of depletion in iron (high Mg#) of the lithospheric mantle. The positive anomalies effectively places southern Norway and Denmark outside the Fennoscandian Shield.

- The wiggly boundary between the low-velocity channel below southern Norway and

the high velocities of Sweden follows zones of Carboniferous-Permian rifting and magmatism very closely. This leaves to main possibilities for the development of the upper mantle in the area:

- It is possible that the thinner lithosphere below southern Norway and stepwise increase in lithospheric thickness towards the Fennoscandian Shield predates the Permian rifting and caused the focusing of rifting and magmatism at the location of the Oslo Graben, as suggested by Pascal et al. (2004).
- Another possible scenario is that the Permian rifting left the originally thick Shield lithosphere below southern Norway partly refertilized and rehydrated. This modified shield lithosphere would be prone to erosion triggered by small-scale convection and/or plume interaction. Episodic erosion of an originally thick lithosphere can provide a mechanism to explain the episodic uplift inferred from geological observations in the region.

An outlook for future work

- The multiscale, finite-frequency tomographic method of Hung et al. (2011) is able to image details in subsurface structure that we do not find in our more traditional tomographic models derived using a ray-theoretical approach or regularizations with a higher degree of smoothing (vertical convolutional quelling). The application of the method to more of the Fennoscandian Shield or other cratonic areas could therefore give important information about the subcontinental mantle.
- The thermo-isostatic modelling in this work is quite sensitive to uncertainty in Moho depth and crustal density. The method should therefore preferably be applied to a network of stations, where the crustal structure is well known, and not to interpolated grid points in a regular grid. If a very detailed and well-constrained crustal model was used, it might be possible to resolve variations in the mantle reference density ρ_0. With the current input data, the small-scale variations of ρ_0 only reflect the uncertainty of the input data.

5.3 Mantle processes and topography of passive continental margins - suggestions for modelling

Geological processes happen over a very long time span while seismic tomography gives a snapshot of the current subsurface structure. Linking tectonic history, long-term mantle processes and seismic tomography is therefore seldom straightforward. In our tomographic

models in chapter 3.4 we infer a lithospheric mantle below southern Norway and Denmark that is significantly different from the proper shield mantle below Sweden. The transition between the two types of lithospheres very clearly follows the Carboniferous-Permian rift structures, and this adds a time dimension to the seismic models.

As discussed previously (chapter 3.4), modelling of small-scale edge-driven convection suggests that the continental lithosphere can be eroded if it is not stabilized by high viscosity from dehydration (e.g. Hieronymus et al., 2007; Van Wijk et al., 2010; Kaislaniemi & van Hunen, 2014). This type of lithosphere erosion has been suggested as the cause of uplift and formation of elevated topography in several places (e.g. Van Wijk et al., 2010; Kaislaniemi & van Hunen, 2014). The geological evidence of substantial uplift in southern Norway in the Neogene seems to be solid, (Anell et al., 2009; Gabrielsen et al., 2010; Japsen et al., 2012), although the accurate magnitude of uplift is difficult to measure (Anell et al., 2009). We suggest to perform numerical modelling like Van Wijk et al. (2010) and Kaislaniemi & van Hunen (2014) to quantify if episodic erosion of an initially thick mantle lithosphere below southern Norway could have caused the uplift inferred from geological observations.

In southern Norway, lithosphere erosion could also have been triggered by the passage of a narrow, plume-like upwelling (section 3.4). It would therefore be important to quantify the effect of both small-scale convection and localized upwelling on lithosphere erosion. In the context of southern Norway, it would also be essential to model how preexisting mantle heterogeneity and zones of weakness, like the Permian rift structures, would affect lithosphere erosion, and also to take into account the opening of the North-Atlantic Ocean.

A wider perspective on the proposed modelling is that a thin and warm continental mantle lithosphere coupled with a thick continental crust seems to be the key prerequisite for subduction initiation at passive continental margins (Nikolaeva et al., 2010), in addition to a relatively straight margin (Marques et al., 2014). A lithosphere thickness of around 100 km, like in southern Norway, seems to limit the regimes of stability from overthrusting and subduction initiation (Nikolaeva et al., 2010). Nikolaeva et al. (2011) analyzed the subduction risk at the American Atlantic passive margins and found that subduction might initiate in the southern part of the Brazilian margin in a time span of 10-20 My. We suggest a similar analysis as Nikolaeva et al. (2011) for the eastern Atlantic Margin, but where Nikolaeva et al. (2011) use the thermal lithosphere compilations of Artemieva & Mooney (2001), we suggest to take into account seismic models of the upper mantle at the passive margins.

Previous modelling of small-scale convection has to a large part focused on modelling specific areas, e.g. the Colorado Plateau, Atlas Mountains, Sorgenfrei-Tornquist Zone,

to explain structural features of these areas. We therefore also suggest a more general modelling approach to investigate if the processes of small-scale, edge-driven convection and subduction initiation at the continental margins could be linked through the erosive weakening of the subcontinental lithospheric mantle from edge-driven convection.

Bibliography

Altamimi, Z., Sillard, P., & Boucher, C., 2002. ITRF2000: A new release of the International Terrestrial Reference Frame for earth science applications, *Journal of Geophysical Research*, **107**(B10), 2214.

Altamimi, Z., Collilieux, X., Legrand, J., Garayt, B., & Boucher, C., 2007. ITRF2005: A new release of the International Terrestrial Reference Frame based on time series of station positions and Earth Orientation Parameters, *Journal of Geophysical Research*, **112**, B09401.

Altamimi, Z., Collilieux, X., & Métivier, L., 2011. ITRF2008: an improved solution of the international terrestrial reference frame, *Journal of Geodesy*, **85**(8), 457–473.

Anell, I., Thybo, H., & Artemieva, I. M., 2009. Cenozoic uplift and subsidence in the North Atlantic region: geological evidence revisited, *Tectonophysics*, **474**(1), 78–105.

Argus, D. F. & Gordon, R. G., 1991. No-net-rotation model of current plate velocities incorporating plate motion model NUVEL-1, *Geophysical Research Letters*, **18**(11), 2039–2042.

Argus, D. F. & Gross, R. S., 2004. An estimate of motion between the spin axis and the hotspots over the past century, *Geophysical research letters*, **31**, L06614.

Artemieva, I. M. & Mooney, W. D., 2001. Thermal thickness and evolution of Precambrian lithosphere: a global study, *Journal of Geophysical Research*, **106**(B8), 16387–16414.

Assumpção, M., Schimmel, M., Escalante, C., Barbosa, J. R., Rocha, M., & Barros, L. V., 2004. Intraplate seismicity in SE Brazil: stress concentration in lithospheric thin spots, *Geophysical Journal International*, **159**(1), 390–399.

Barth, A., Jordan, M., & Ritter, J. R., 2007. Crustal and upper mantle structure of the French Massif Central plume, in *Mantle Plumes*, pp. 159–184, eds Ritter, J. R. R. & Christensen, U., Springer.

Bostrom, R., 1971. Westward displacement of the lithosphere, *Nature*, **234**, 536–538.

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.

Chiao, L.-Y. & Kuo, B.-Y., 2001. Multiscale seismic tomography, *Geophysical Journal International*, **145**(2), 517–527.

Chiao, L.-Y. & Liang, W.-T., 2003. Multiresolution parameterization for geophysical inverse problems, *Geophysics*, **68**(1), 199–209.

Cloetingh, S., Ziegler, P., Bogaard, P., Andriessen, P., Artemieva, I. M., Bada, G., Van Balen, R., Beekman, F., Ben-Avraham, Z., Brun, J.-P., et al., 2007. TOPO-EUROPE: the geoscience of coupled deep Earth-surface processes, *Global and Planetary Change*, **58**(1), 1–118.

Dahlen, F., Hung, S.-H., & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes – Theory, I., *Geophysical Journal International*, **141**(1), 157–174.

Doglioni, C., 1993. Geological evidence for a global tectonic polarity, *Journal of the Geological Society*, **150**(5), 991–1002.

Doglioni, C., Green, D. H., & Mongelli, F., 2005. On the shallow origin of hotspots and the westward drift of the lithosphere, *Geological Society of America Special Papers*, **388**, 735–749.

Doré, A., Lundin, E., Kusznir, N., & Pascal, C., 2008. Potential mechanisms for the genesis of Cenozoic domal structures on the NE Atlantic margin: pros, cons and some new ideas, *Geological Society, London, Special Publications*, **306**(1), 1–26.

Du, Z. & Foulger, G., 1999. The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves, *Geophysical Journal International*, **139**(2), 419–432.

Ebbing, J., England, R., Korja, T., Lauritsen, T., Olesen, O., Stratford, W., & Weidle, C., 2012. Structure of the Scandes lithosphere from surface to depth, *Tectonophysics*, **536**, 1–24.

England, R. W. & Ebbing, J., 2012. Crustal structure of central Norway and Sweden from integrated modelling of teleseismic receiver functions and the gravity anomaly, *Geophysical Journal International*, **191**(1), 1–11.

Farnetani, C. & Samuel, H., 2005. Beyond the thermal plume paradigm, *Geophysical Research Letters*, **32**, L07311.

Frasetto, A. & Thybo, H., 2013. Receiver function analysis of the crust and upper mantle in Fennoscandia – isostatic implications, *Earth and Planetary Science Letters*, **381**, 234–246.

Gabrielsen, R. H., Faleide, J. I., Pascal, C., Braathen, A., Nystuen, J. P., Etzelmuller, B., & O'Donnell, S., 2010. Latest Caledonian to Present tectonomorphological development of southern Norway, *Marine and Petroleum Geology*, **27**(3), 709 – 723.

Goes, S., Govers, R., & Vacher, P., 2000. Shallow mantle temperatures under Europe from P and S wave tomography, *Journal of Geophysical Research*, **105**(B5), 11153–11169.

Goes, S., Cammarano, F., & Hansen, U., 2004. Synthetic seismic signature of thermal mantle plumes, *Earth and Planetary Science Letters*, **218**(3), 403–419.

Gradmann, S., Ebbing, J., Fullea, J., et al., 2013. Integrated geophysical modelling of a lateral transition zone in the lithospheric mantle under Norway and Sweden, *Geophysical Journal International*, **194**(3), 1359–1374.

Granet, M., Wilson, M., & Achauer, U., 1995. Imaging a mantle plume beneath the French Massif Central, *Earth and Planetary Science Letters*, **136**(3), 281–296.

Gripp, A. E. & Gordon, R. G., 2002. Young tracks of hotspots and current plate velocities, *Geophysical Journal International*, **150**(2), 321–361.

Hieronymus, C. F., Shomali, Z. H., & Pedersen, L. B., 2007. A dynamical model for generating sharp seismic velocity contrasts underneath continents: application to the Sorgenfrei–Tornquist Zone, *Earth and Planetary Science Letters*, **262**(1), 77–91.

Hung, S.-H., Dahlen, F., & Nolet, G., 2000. Frechet kernels for finite-frequency traveltimese – Examples, II., *Geophysical Journal International*, **141**(1), 175–203.

Hung, S.-H., Chen, W.-P., & Chiao, L.-Y., 2011. A data-adaptive, multiscale approach of finite-frequency, traveltime tomography with special reference to P and S wave data from central Tibet, *Journal of Geophysical Research*, **116**, B06307.

Hwang, Y. K., Ritsema, J., van Keken, P. E., Goes, S., & Styles, E., 2011. Wavefront healing renders deep plumes seismically invisible, *Geophysical Journal International*, **187**(1), 273–277.

Japsen, P., Chalmers, J. A., Green, P. F., & Bonow, J. M., 2012. Elevated, passive continental margins: Not rift shoulders, but expressions of episodic, post-rift burial and exhumation, *Global and Planetary Change*, **90**, 73–86.

Julià, J., Ammon, C., Herrmann, R., & Correig, A., 2000. Joint inversion of receiver function and surface wave dispersion observations, *Geophysical Journal International*, **143**(1), 99–112.

Julià, J., Ammon, C., & Herrmann, R., 2003. Lithospheric structure of the Arabian Shield from the joint inversion of receiver functions and surface-wave group velocities, *Tectonophysics*, **371**(1-4), 1–21.

Julià, J., Ammon, C., & Nyblade, A., 2005. Evidence for mafic lower crust in Tanzania, East Africa, from joint inversion of receiver functions and Rayleigh wave dispersion velocities, *Geophysical Journal International*, **162**(2), 555–569.

Kaislaniemi, L. & van Hunen, J., 2014. Dynamics of lithospheric thinning and mantle melting by edge-driven convection: Application to Moroccan Atlas mountains, *Geochemistry, Geophysics, Geosystems*, **15**, 3175–3189.

Keyser, M., Ritter, J. R., & Jordan, M., 2002. 3D shear-wave velocity structure of the Eifel plume, Germany, *Earth and Planetary Science Letters*, **203**(1), 59–82.

Köhler, A., Weidle, C., & Maupin, V., 2011. Directionality analysis and Rayleigh wave tomography of ambient seismic noise in southern Norway, *Geophysical Journal International*, **184**(1), 287–300.

Köhler, A., Weidle, C., & Maupin, V., 2012. Crustal and uppermost mantle structure of southern Norway: results from surface wave analysis of ambient seismic noise and earthquake data, *Geophysical Journal International*, **191**(3), 1441–1456.

Kolstrup, M., 2010. *The lithospheric mantle below southern Norway – Temperature and density investigated using two complementary methods*, Master's thesis, University of Oslo.

Langston, C., 1979. Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *Journal of Geophysical Research*, **84**(B9), 4749–4762.

Leech, M., 2001. Arrested orogenic development: eclogitization, delamination, and tectonic collapse, *Earth and Planetary Science Letters*, **185**(1), 149–159.

Marques, F., Cabral, F., Gerya, T., Zhu, G., & May, D., 2014. Subduction initiates at straight passive margins, *Geology*, **42**(4), 331–334.

Maupin, V., 2011. Upper-mantle structure in southern Norway from beamforming of Rayleigh wave data presenting multipathing, *Geophysical Journal International*, **185**(2), 985–1002.

Maupin, V., Agostini, A., Artemieva, I., Balling, N., Beekman, F., Ebbing, J., England, R., Frassetto, A., Gradmann, S., Jacobsen, B., et al., 2013. The deep structure of the Scandes and its relation to tectonic history and present-day topography, *Tectonophysics*, **602**, 15–37.

Medhus, A., Balling, N., Jacobsen, B., Weidle, C., England, R., Kind, R., Thybo, H., & Voss, P., 2012. Upper-mantle structure beneath the Southern Scandes Mountains and the Northern Tornquist Zone revealed by P-wave traveltime tomography, *Geophysical Journal International*, **189**(3), 1315–1334.

Medhus, A. B., Balling, N., Jacobsen, B. H., Kind, R., & England, R. W., 2009. Deep-structural differences in southwestern Scandinavia revealed by P-wave travel time residuals, *Norwegian Journal of Geology*, **89**, 203–214.

Morelli, A. & Dziewonski, A. M., 1993. Body wave traveltimes and a spherically symmetric P-and S-wave velocity model, *Geophysical Journal International*, **112**(2), 178–194.

Nielsen, S. B., Paulsen, G. E., Hansen, D. L., Gemmer, L., Clausen, O. R., Jacobsen, B. H., Balling, N., Huuse, M., & Gallagher, K., 2002. Paleocene initiation of Cenozoic uplift in Norway, *Geological Society, London, Special Publications*, **196**(1), 45–65.

Nielsen, S. B., Gallagher, K., Leighton, C., Balling, N., Svenningsen, L., Jacobsen, B. H., Thomsen, E., Nielsen, O. B., Heilmann-Clausen, C., Egholm, D. L., Summerfield, M. A., Clausen, O. R., Piotrowski, J. A., Thorsen, M. R., Huuse, M., Abrahamsen, N., King, C., & Lykke-Andersen, H., 2009. The evolution of western Scandinavian topography: A review of Neogene uplift versus the ICE (isostasy-climate-erosion) hypothesis, *Journal of Geodynamics*, **47**(2-3), 72–95.

Nikolaeva, K., Gerya, T., & Marques, F., 2010. Subduction initiation at passive margins: numerical modeling, *Journal of Geophysical Research*, **115**, B03406.

Nikolaeva, K., Gerya, T. V., & Marques, F. O., 2011. Numerical analysis of subduction initiation risk along the Atlantic American passive margins, *Geology*, **39**(5), 463–466.

Özalaybey, S., Savage, M., Sheehan, A., Louie, J., & Brune, J., 1997. Shear-wave velocity structure in the northern Basin and Range province from the combined analysis of receiver functions and surface waves, *Bulletin of the Seismological Society of America*, **87**(1), 183–199.

Pascal, C., 2006. On the role of heat flow, lithosphere thickness and lithosphere density on gravitational potential stresses, *Tectonophysics*, **425**(1), 83–99.

Pascal, C., Cloetingh, S., & Davies, G., 2004. Asymmetric lithosphere as the cause of rifting and magmatism in the Permo-Carboniferous Oslo Graben, in *Permo-Carboniferous Magmatism and Rifting in Europe*, vol. 223, pp. 139–156, eds Wilson, M., Neumann, E.-R., Davies, G., Timmerman, M., Heeremans, M., & Larsen, B. T., Geological Society, London, Special Publications.

Piromallo, C. & Morelli, A., 2003. P wave tomography of the mantle under the Alpine-Mediterranean area, *Journal of Geophysical Research*, **108**(B2), 2065.

Ribe, N., Davaille, A., & Christensen, U., 2007. Fluid dynamics of mantle plumes, in *Mantle plumes*, pp. 1–48, eds Ritter, J. R. R. & Christensen, U., Springer.

Ricard, Y., Doglioni, C., & Sabadini, R., 1991. Differential rotation between lithosphere and mantle: A consequence of lateral mantle viscosity variations, *Journal of Geophysical Research*, **96**(B5), 8407–8415.

Rickers, F., Fichtner, A., & Trampert, J., 2013. The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: Evidence from full-waveform inversion, *Earth and Planetary Science Letters*, **367**, 39–51.

Ritter, J., 2007. The Seismic Signature of the Eifel Plume, in *Mantle Plumes*, pp. 379–404, eds Ritter, J. R. & Christensen, U. R., Springer.

Ritter, J. R., Jordan, M., Christensen, U. R., & Achauer, U., 2001. A mantle plume below the Eifel volcanic fields, Germany, *Earth and Planetary Science Letters*, **186**(1), 7–14.

Rohrman, M. & van der Beek, P., 1996. Cenozoic postrift domal uplift of North Atlantic margins: an asthenospheric diapirism model., *Geology*, **24**, 901–904.

Sigloch, K., 2011. Mantle provinces under North America from multifrequency P wave tomography, *Geochemistry, Geophysics, Geosystems*, **12**(2), Q02W08.

Stratford, W. & Thybo, H., 2011a. Seismic structure and composition of the crust beneath the southern Scandes, Norway, *Tectonophysics*, **502**(3), 364–382.

Stratford, W. & Thybo, H., 2011b. Crustal structure and composition of the Oslo Graben, Norway, *Earth and Planetary Science Letters*, **304**(3), 431–442.

Tian, Y., Sigloch, K., & Nolet, G., 2009. Multiple-frequency SH-wave tomography of the western US upper mantle, *Geophysical Journal International*, **178**(3), 1384–1402.

Turcotte, D. & Schubert, G., 2002. *Geodynamics*, Cambridge University Press.

Van Wijk, J., Baldridge, W., Van Hunen, J., Goes, S., Aster, R., Coblentz, D., Grand, S., & Ni, J., 2010. Small-scale convection at the edge of the Colorado Plateau: Implications for topography, magmatism, and evolution of Proterozoic lithosphere, *Geology*, **38**(7), 611–614.

Wawerzinek, B., 2012. *Untersuchung der elastischen Scherwellenstruktur unter dem Südskandinavischen Gebirge*, Ph.D. thesis, Fakultät für Physik des Karlsruher Instituts für Technologie (KIT).

Wawerzinek, B., Ritter, J., & Roy, C., 2013. New constraints on the 3D shear wave velocity structure of the upper mantle underneath Southern Scandinavia revealed from non-linear tomography, *Tectonophysics*, **602**, 38–54.

Weidle, C. & Maupin, V., 2008. An upper-mantle S-wave velocity model for Northern Europe from Love and Rayleigh group velocities, *Geophysical Journal International*, **175**(3), 1154–1168.

Wilson, J. T., 1966. Did the Atlantic close and then re-open?, *Nature*, **211**, 676–681.

Yoo, H., Herrmann, R., Cho, K., & Lee, K., 2007. Imaging the three-dimensional crust of the Korean Peninsula by joint inversion of surface-wave dispersion and teleseismic receiver functions, *Bulletin of the Seismological Society of America*, **97**(3), 1002–1011.