Lithospheric structure from forward and inverse modeling of satellite gravity and magnetic data

Dissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Wolfgang Szwillus

Kiel, Juni 2018

Institut für Geowissenschaften
Abteilung Geophysik

Referent:	Prof. Dr. Jörg Ebbing
Korreferent:	Professor Richard England
Tag der mündlichen Prüfung:	17. Juli 2018
Zum Druck genehmigt:	
	Die Dekanin

Abstract

Satellite missions have provided the Earth's gravity and magnetic field at resolutions sufficient for large-scale geophysical applications. While satellite data do not possess the same resolution as ground data, their homogeneous coverage and low error make them ideal for studying large-scale lithospheric structure and processes. In this thesis, the sensitivity of satellite potential field data to structures in the lithosphere is investigated. Furthermore, a crustal model is derived based on a data base of active seismic experiments. The calculation of the crustal model is data-driven, unlike previous crustal models, which typically rely on expert knowledge. Finally, a simple probabilistic joint inversion of gravity gradients and topography is developed, that is capable of estimating a two-layer density model of the lithosphere and its uncertainty.

Since satellite measurements are highly accurate, even at long wavelengths, the gravity data processing techniques, in particular the topographic correction, also have to be sufficiently accurate. One possible source of long-wavelength error are far-field topographic effects. However, most topography is in isostatic balance. Thus, far-field topographic and isostatic effects tend to balance out. Exploiting this balance, several methods to minimize far-field effects are discussed.

Satellite magnetic data are an ideal addition to satellite gravity data, because they share the same global coverage and homogeneous data quality. Since minerals lose their magnetic properties above their Curie temperature, the thickness of the magnetized layer is influenced by the temperature distribution in the lithosphere. Thus, satellite magnetic data potentially reflect the temperature distribution inside the lithosphere. However, sensitivity tests using susceptibility inversion show that the magnetic field can be explained equally well using thickness variations or lateral susceptibility variations. Thus, satellite magnetic data alone cannot be used to constrain the temperature distribution.

To optimally use satellite data, available "ground truth" should be included. For the lithosphere, this ground truth mainly comes in the form of crustal structure acquired using active source seismology. However, due to the heterogeneous distribution of seismic profiles, expert knowledge is typically used to extrapolate the seismic results. In contrast, the crustal model developed here is generated by applying geostatistical interpolation to a data base of seismic studies. This approach permits assessing the reliability of the obtained results.

In the last part of the thesis, satellite gravity gradient data, topography data and the new crustal model are combined in a joint probabilistic inversion. The goal of the inversion is to estimate a two-layer density model (crust and lithospheric mantle) that is consistent with gravity and topography data, as well as the prior constraints from the crustal model. As a first test, the new algorithm is applied to the north mid-Atlantic region. In this region, a clear signature in the mantle lithospheric density is expected due to ocean floor spreading and the influence of the Azores hotspot. Indeed these signatures are resolved by the inversion, but severe trade-offs between the different layers occur and there is limited resolution of crustal structure. The algorithm could be improved in the future by including additional data (e.g. higher resolution gravity data or passive seismology) and a more detailed sub-surface parametrization (e.g. in terms of temperature and composition).

Zusammenfassung

Die global verfügbaren Messungen des Schwere- und Magnetfeldes mithilfe von Satelliten liegen heute in einer Auflösung vor, die großräumige geophysikalische Anwendungen ermöglichen. Zwar ist es mit Satelliten nicht möglich, die Auflösung von Bodenmessungen zu erreichen. Um aber großkalige lithosphärische Strukturen zu untersuchen, sind Satellitendaten ideal, da die Datenqualität sehr homogen ist und sie einen geringen Messfehler aufweisen. In dieser Dissertation wird untersucht, wie groß die Sensitivität von Satellitendaten tatsächlich ist. Darüber hinaus wird ein globales Krustenmodell basierend auf einer Datenbank aktiver seismischer Profile erstellt, das – anders als vorherige Krustenmodelle – weitgehend ohne manuelle Eingabe von Expertenwissen auskommt. Zuletzt werden die Ergebnisse der verschiedenen Kapitel in einer gemeinsamen, probabilistischen Inversion von Schweregradienten und Topografie kombiniert, mit der ein einfaches Zwei-Schicht Dichtemodell der Lithosphäre und dessen Unsicherheit bestimmt werden kann.

Da Satellitenmessungen des Schwerefeldes auch die langen Wellenlängen sehr genau auflösen, müssen die Konventionen der Schweredatenverarbeitung an diese Anforderungen angepasst werden. Insbesondere die topografische Korrektur kann eine Quelle von langwelligen Fehlern sein, da der Schwereeffekt von Topografie außerhalb des Untersuchungsgebiets die langen Wellenlängen systematisch verzerren kann. Allerdings sind die meisten großskaligen topografischen Massen aufgrund des isostatischen Ausgleichs kompensiert. Infolgedessen besteht eine Tendenz der gegenseitigen Aufhebung von topografischen und isostatischen Effekten, vor allem in großer Entfernung. Dies lässt sich ausnutzen um den Fernfeldeffekt infolge weit entfernter Topografie innerhalb eines Untersuchungsgebiet effektiv zu minimieren. Verschiedene Ansätze dies zu erreichen, werden vorgestellt.

Im Prinzip sind Satellitenmessungen des Magnetfeldes eine ideale Ergänzung für Schweremessungen, da sie ebenfalls in hoher Qualität global verfügbar sind. Da jedes magnetische Mineral über eine spezifische Curie-Temperatur verfügt, oberhalb der es seine magnetischen Eigenschaften verliert, besteht ein Zusammenhang zwischen der Mächtigkeit der magnetisierten Schicht und der Temperaturverteilung in der Lithosphäre. Folglich spiegelt auch das magnetische Feld der Lithosphäre die Temperaturverteilung wider. Allerdings zeigen Sensitivitätstest, die auf einer Inversion der Suszeptibilitätsverteilung innerhalb der Lithosphäre beruhen, dass zwischen einer Variation der Mächtigkeit der magnetisierten Schicht und einer lateralen Variation der Suszeptibilität nicht zuverlässig unterschieden werden kann. Somit ist das Magnetfeld alleine nicht geeignet, um die Temperaturverteilung zu bestimmen.

Für eine optimale Verwendung von Schweredaten sollten diese mit zusätzlichen Informationen kombiniert werden. Für die Lithosphäre stellen Krustenmodelle basierend auf aktiven seismischen Messungen eine wichtige Quelle solcher Information dar. Da allerdings die Verteilung aktiver seismischer Profile global äußerst ungleichmäßig ist, wird bei der Extrapolation zur Erstellung flächenhafter Krustenmodelle typischerweise auf Expertenwissen zurückgegriffen. Im Gegensatz dazu wird in dieser Arbeit ein Krustenmodell entwickelt, das auf geostatistischer Interpolation einer Datenbank aktiver seismischer Untersuchungen beruht. Der Vorteil dieses Vorgehens ist, dass die Genauigkeit des erhaltenen Krustenmodells abgeschätzt werden kann.

Im letzten Teil dieser Arbeit werden Schweregradienten, Topografie und das neue Krustenmodell gemeinsam in einer probabilistischen Inversion verwendet. Das Ziel dieser Inversion ist es, ein Zwei-Schicht Modell (Kruste und lithosphärischer Mantel) der Dichteverteilung zu erstellen, das mit den Schweredaten, der Topografie und dem a priori Krustenmodell kompatibel ist. Als Testregion für das Inversionsverfahren wird der zentrale Nord-Atlantik gewählt. In dieser Region ist eine systematische Signatur in der Dichte des lithosphärischen Mantels zu erwarten. Zum einen aufgrund der Ozeanbodenspreizung, zum anderen aufgrund des Einfluss des Azoren-Hotspots. In

der Tat gelingt es dem Inversionsverfahren den erwarteten Dichteverlauf im lithosphärischen Mantel zu bestimmen. Allerdings gibt es erhebliche Äquivalenzen zwischen den verschiedenen Schichten und die Auflösung innerhalb der Kruste ist sehr begrenzt. Verbessern lassen sich die Ergebnisse dieses Algorithmus indem zusätzliche Daten berücksichtigt werden (z.B. höher aufgelöste Schweredaten oder passive Seismologie) oder indem der Untergrund detaillierter parametrisiert wird (z.B. als Temperatur und Komposition).

Contents

1 Introduction

It was known at least since the end of the 19th century that major topographic features do not disturb gravity measurements as much as theoretically expected (Pratt, 1855; for a historical overview see Watts, 2001, ch. 1). Today this is confirmed by satellite-derived gravity models: The free-air anomaly at satellite height (225 km) is less than ± 100 mGal (Pail et al., 2010), whereas the expected gravity effect of topography is ± 500 mGal (see chapter 2). Thus, the topographic mass surplus needs to be compensated by a mass deficit at depth.

The principle of **isostasy** provides an explanation for this mass deficit. Analogous to an iceberg floating in water, the lighter crust 'floats' on the dense mantle. In this model, the crust-mantle system is in isostatic balance if there is a compensation depth at which pressure is constant. Airy (1855) and Pratt (1864) proposed two competing models of isostasy. In Airy's model the crust beneath high topography is thickened, whereas in Pratt's model the crust beneath high topography is less dense than regular crust. However, there is no need to strictly separate the two models in practice, because both thickness and density variations can contribute to isostatic compensation simultaneously.

Apart from explaining the small free-air anomaly, the principle of isostasy is also justified from a dynamical point of view. In simple rheological terms, the upper earth consists of an elastic lithosphere that overlays a viscous mantle. In an idealized case, the elastic lithosphere reacts to topographic loading by deforming, which causes internal stresses in the lithosphere that balance the external loading (Turcotte and Schubert, 2002, ch. 3). Importantly, a perfect elastic medium can maintain such stresses indefinitely without moving (Jordan, 1978). In contrast, the viscous mantle reacts to external loading by flowing. As a result, the lithosphere will sink or rise according to the flow in the sub-lithospheric mantle. The flow only stops once isostatic balance is attained in the sub-lithospheric mantle. The ongoing uplift observed in Fennoscandia, Siberia and Canada (Cathles, 1975) is caused by Glacial Isostatic Adjustement (GIA) and illustrates how quickly the lithosphere-asthenosphere system tends towards isostatic balance.

However, crustal isostasy is not the only manner in which topography can be maintained. Firstly, there can be an isostatic contribution from the mantle lithosphere, secondly topography can be maintained by the elastic strength of the lithosphere and thirdly topography can be maintained by convective forces acting on the base of the lithosphere leading to **dynamic topography**.

Equating lithosphere and crust is often a viable approximation – at least for the continents – because the density difference between crust and mantle is much larger than between lithospheric and sub-lithospheric mantle. However, density variations in the upper mantle caused by cooling are crucial for explaining ocean floor topography (Turcotte and Schubert, 2002, ch. 4.16). In the continental mantle lithosphere, density variations are not only caused by temperature variations, but also compositional differences. Jordan (1978) proposed that the continental lithosphere is neutrally buoyant, because it is depleted of iron-rich (dense) minerals. Other studies found that the continental mantle lithosphere does contribute up to 1 km of topography in some regions (Zoback and Mooney, 2003; Kaban et al., 2003). Thus, a more complete model of isostasy should include the mantle lithosphere.

Elastic support of topography is typically limited to wavelengths < 100-200 km. At larger wavelengths, the lithosphere hardly resists the topographic loading, so that large scale topography is mainly maintained isostatically (Turcotte and Schubert, 2002, ch. 3; Watts, 2001). However, some regions seem to possess anomalously strong lithosphere (e.g. Watts and Burov, 2003), where topographic loads with a wavelength of a several hundred kilometers can be maintained by elastic stresses (e.g. Royden, 1993). Note that the spectral methods used to estimate the lithospheric strength are somewhat controversial (e.g. McKenzie, 2016).

Convective motions in the viscous part of the mantle can deform the surface of the Earth and lead to **dynamic topography**. Mantle flow is driven by lateral density variations in the mantle, which are in turn caused by thermal or compositional variations. The speed of seismic shear waves is typically used as proxy for temperature and hence density (Karato, 1993). Global seismic tomography of the mantle systematically indicates that the mantle is heterogeneous (i.e. French and Romanowicz, 2015; Schaeffer and Lebedev, 2015). Thus, there plausibly is a contribution from mantle convection to surface topography.

Since the theoretical foundations of isostatic and dynamic topography are well developed, their respective contributions to topography can be quantified using different techniques. Note that the two types of topography are inherently complementary, i.e. any topography not explained by isostatic topography should be due to dynamic topography and vice versa. However, there is considerable variability caused by different modeling techniques and incomplete knowledge of sub-surface properties such as crustal thickness.

Isostatic topography can be calculated from seismological determinations of crustal structure and density. Typically, this is done using active source seismology, since it provides the best constraint on crustal structure and velocity (Prodehl and Mooney, 2012). Furthermore, P-wave velocity estimates can be converted to density estimates (Christensen and Mooney, 1995). However, passive seismic methods, particularly receiver functions (Kind et al., 2012) and ambient noise tomography (Shapiro and Campillo, 2004), can also be used. Based on the seismic density and structure estimates, one can calculate how much topography would be needed to achieve isostatic balance. This *crustal* isostatic topography can be subtracted from the observed topography to give *residual* topography. Thus, residual topography is the part of topography not explained by crustal isostatic equilibrium (Kaban et al., 1999).

Note that in many regions, seismological findings are indeed compatible with Airy-type isostasy: Many major mountain chains are underlain by thick crustal roots that compensate the high terrain above (e.g. Andes, Alps, Tibetan Plateau, Himalaya). However, there are also areas of considerable disagreement, like cratonic regions with thick crust but hardly any surface topography.

Flament et al. (2013) compared four different estimates of residual topography and conclude that the long-wavelength (>10,000 km) features are stable. At the same time, the predicted amplitudes (peak-to-peak) vary between 2.8 km and 4.2 km, depending on the crustal models and assumptions used. In contrast, a recent study focused on the oceanic domain (Hoggard et al., 2016) concluded that residual topography at wavelengths of 10,000 km or more is less than 1 km. Thus, there is considerable disagreement, particularly regarding the magnitude of residual topography.

Dynamic topography was first calculated by Richards and Hager (1984) using tomography-derived density variations to predict the instantaneous flow pattern in the mantle and the corresponding dynamic topography. Many authors have built on the original approach of Richards and Hager (Flament et al., 2013). Estimates of dynamic topography depend on the input seismic tomography model, how seismic velocity is converted to density, the assumed viscosity structure in the mantle, the parameterization of the lithosphere and the numerical approach used to calculate mantle flow. Flament et al. (2013) also compared five different dynamic topography calculations and found agreement in the long-wavelength structure, but significant disagreement in terms of peak-to-peak amplitude (from 2.2 km up to 5.8 km).

Even though isostatic and dynamic topography should be complementary, agreement between estimates of residual and dynamic topography is generally poor and limited to the longest wavelengths (Steinberger, 2007). To some extent this is caused by conflicting definitions of dynamic and residual topography, but even if complementary definitions are used, large differences remain (Molnar et al., 2015). In addition, the amplitude of dynamic topography results tend to be systematically higher than residual topography (Molnar et al., 2015). Recently, the agreement between residual and dynamic topography was improved by using a velocity-density conversion that reflects the chemically depleted continental lithosphere (Steinberger, 2016). In the model of Steinberger (2016), dynamic topography is only 30 % larger than residual topography, compared to factor 2 or 3 in previous calculations, but still large differences remain.

While the disagreements between dynamic and residual topography are large, there is also considerable uncertainty affecting these estimates. On a global scale, limited availability of data is a major source of uncertainty. For example, active seismic determinations of crustal structure are distributed very heterogeneously (Mooney, 2015), potentially leading to large uncertainties of isostatic topography in poorly surveyed areas. Uncertainties persist in well surveyed regions as well, such as the uncertainty of the velocity to density conversion (Herceg et al., 2015). Calculations of dynamic topography are likewise affected by several potential sources of uncertainties, including the viscosity distribution in the mantle, the conversion of seismic velocities to densities and limited resolution of seismic tomography (Flament et al., 2013). Thus, it is not surprising that even in regions with excellent data coverage the same topographic features are seen by different authors as either mainly isostatic or mainly dynamic (cf. Afonso et al., 2016b).

The gravity field might be able to shed some light on this discrepancy, since both forms of topography ultimately depend on the density structure. Recently, Molnar et al. (2015) challenged results of 1 km or more of dynamic topography derived from mantle flow calculations. They use analytical models of flow to derive a rule of thumb that 1 km of dynamic topography creates a free-air anomaly of about 50 mGal (30 mGal in the oceans). According to this rule of thumb, the maximum long wavelength free-air anomaly (c. 50 mGal) limits dynamic topography to less than 1 km.

However, the “50 mGal” rule of thumb only holds if viscosity is constant in the entire mantle and there is no deformation on the core-mantle boundary (Colli et al., 2016). There is evidence that the lower mantle is much more viscous than the upper mantle (e.g. King, 2016). In that case, numerical flow simulations predict that positive dynamic topography is associated with no or even negative free-air anomaly depending on the depth and wavelength of the causative density anomaly (Colli et al., 2016). Thus, the relation between small free-air anomalies and small values of dynamic topography does not hold in all situations (Yang and Gurnis, 2016). To fully exploit the possibilities of the gravity field, a more comprehensive approach than the rule of thumb is needed.

The goal of density modeling is to determine a density distribution in the sub-surface (typically the crust and upper mantle) that explains the measured gravity field and is consistent with constraining geophysical and geological knowledge. Such a density distribution can be found manually through forward modeling (e.g. Götze and Lahmeyer, 1988; Prezzi et al., 2009) or using automated inversion (e.g. Uieda and Barbosa, 2016). In a wider sense, density modeling can be included into *joint* inversion of several data sets (cf. Moorkamp et al., 2016). Such a joint approach can reduce ambiguities, but requires a relation between physical parameters in the sub-surface.

In the last years, new opportunities for large scale density modeling have arisen thanks to satellite gravity data. The satellite gravity missions GRACE (Tapley et al., 2004; Tapley et al., 2005) and GOCE (Floberghagen et al., 2011) have provided the gravity field at resolutions of c. 70 km nearly globally with high accuracy (Pail et al., 2010; Mayer-Guerr, 2015). Thus, satellite gravity data are now sufficient for geophysical modeling of major structures (Holzrichter and Ebbing, 2016; Köther et al., 2012). A crucial factor for the improved accuracy and resolution was the Global Ocean Circulation Explorer (GOCE) mission, carried out by the European Space Agency (ESA). GOCE was the first mission to measure not only the vector of gravitational acceleration, but also its spatial gradient, represented as the gradient tensor.

The specific sensitivities of the different representations of the gravity field to the density structure have long been exploited in different ways. These representations are all ultimately derivatives of the potential and are inherently sensitive to different depths (Bouman et al., 2016). The geodynamics community mostly relied on the geoid as representation of the gravity field because it is sensitive to density anomalies in the mantle down to the core-mantle boundary (Ricard et al., 1988; Richards and Hager, 1984). For the crust and upper mantle, the vertical gravity component has typically been used instead, since it can be measured using portable gravimeters or aircraft (Nabighian et al., 2005). While the gravity gradient tensor was already measured in the 1890s by Eötvös (Shaw and Jones, 1923), interest in gravity gradients has recently renewed. Gravity gradients were found sensitive to crustal density structure (e.g. Ebbing et al., 2014). In addition, using the complete gradient tensor is helpful for lithospheric forward modeling (Holzrichter and Ebbing, 2016). In exploration applications, gravity gradients have been shown to reduce interpretation ambiguity (Paoletti et al., 2016). For lithospheric modeling, jointly considering gravity, geoid and gradients is also considered beneficial (Fullea et al., 2009; Afonso et al., 2013b).

Like the gravity field, the magnetic field is available globally from recent satellite missions (e.g. Olsen et al., 2017). The magnetic field can be composed into the core field and the lithospheric field. At wavelengths of less than ~2500 km, magnetized rocks in the lithosphere are the main source of the magnetic field (Thébault et al., 2010). Since magnetized minerals become practically non-magnetic above their Curie temperature, the lithospheric field strongly depends on the temperature distribution in the lithosphere (Reynolds et al., 1990). Thus, the magnetic field could provide information about temperature in the lithosphere, which can be used to constrain lithospheric density.

In this thesis, I investigate if and how satellite gravity and magnetic data can be used to better constrain large-scale lithospheric structure. As discussed above, it is important to consider as much geophysical information as possible, since all methods are affected by inherent ambiguities. Thus, the uncertainties of the individual results are also important to allow a reasonable combination of different methods. The following chapters contain one published paper, two submitted manuscripts and one unpublished chapter that address different topics relating to lithospheric structure.

Paper I investigates how to incorporate far-field topographic effects in regional gravity modeling. The main tool of density modeling is the Bouguer anomaly (Nabighian et al., 2005), which is calculated by subtracting the gravity effect of topography from the free-air anomaly (Blakely, 2009). Due to the linearity of the gravity field, the Bouguer anomaly should only reflect density variations in the sub-surface, provided the gravity effect of topography has been estimated accurately.

However, the procedures for calculating the gravity effect of topography have been derived in the context of surface measurements, exploration purposes (i.e. Hammer, 1939) and limited computational resources (i.e. Hayford and Bowie, 1912). Often only topography within a radius of 1.5 degrees (~167 km) around the measurement point is considered. Mikuska et al. (2006) showed that Distant Terrain Effects of several hundred mGal can be caused by terrain outside of the 1.5 degree radius. The Distant Terrain Effects mainly influence the long wavelengths (>1000 km) of the gravity field, so they can be safely ignored for smaller study areas. However, these long wavelengths are related to deep (sub-lithospheric) sources. Thus, accurately accounting for far-field effects might be important large-scale investigations with satellite data.

The radius of 1.5 degrees has been used for decades in density modeling (i.e. Prezzi et al., 2009) – apparently without causing any problems. The most likely savior is isostasy. After all, not only the gravity effect of topography is neglected in the far-field, but also its compensating mass at depth.

In paper I the importance of far-field topographic-isostatic is systematically investigated. A global model of isostatic compensation is calculated. In the continents, an isostatic Moho depth is calculated from topography directly, under the assumption of crustal Airy-type isostasy. Since the subsidence of ocean floor with age results from cooling of the entire lithosphere, and crustal thickness is nearly constant in 'normal' oceans (McKenzie and Bickle, 1988), isostatic equilibrium in the oceans is attained by adjusting the density of the lithospheric mantle, down to a depth of 120 km.

Forward calculation of the gravity effect of this model show that far-field topographic and isostatic effects indeed tend to balance out. The combined Topographic-isostatic far-field effects are an order of magnitude smaller than topographic far-field effects alone.

Paper II is not directly connected to the topography or density structure, but uses the magnetic field.

Typically, either spectral techniques (Spector and Grant, 1970) or equivalent dipoles (Purucker et al., 1998) are used to directly estimate the *lower boundary of magnetization* from magnetic data. The lower boundary of magnetization can be interpreted as the Curie isotherm (Wasilewski and Mayhew, 1992). This is especially appealing in remote areas such as Antarctica where measurements of surface heat flow are extremely challenging (Maule et al., 2005), so that the magnetic field might provide important information on the thermal state.

However, directly determining the lower boundary of magnetization from the magnetic field has two kinds of limitations: On the one hand, the lower boundary of magnetization does not necessarily correspond to the Curie isotherm, because the lower boundary might reflect differences in composition or oxidation conditions (Pilchin and Eppelbaum, 1997). On the other hand, the magnetic field is not only affected by variations of the lower boundary of magnetization, but also reflects lateral variation in magnetic susceptibility. Rock susceptibility in the crust is related to its composition (Reynolds et al., 1990; Hunt et al., 1995). In addition, magnetic susceptibility affects only one geophysical observable, namely the static magnetic field. Thus, the only other way to constrain susceptibility is prior information, i.e. rock samples. Direct methods are hence affected by strong ambiguities.

To investigate the sensitivity of magnetic data to thermal structure, satellite magnetic data from the model MF7 (Maus, 2010) are inverted for a laterally varying susceptibility distribution. During the inversion three different geometries of the magnetized layer are used: The entire crust, only the crust colder than 580 °C (Curie temperature of magnetite), both crust and mantle colder than 580 °C are magnetic. The temperature distribution in the lithosphere is calculated independently of the magnetic field using a simple thermal model based on heat flow measurements and a seismological model of lithosphere-asthenosphere boundary (LAB) depth.

While these three geometries do lead to different susceptibility values, the differences are small compared to the prior uncertainties of susceptibility. Thus, magnetic data cannot distinguish between the different geometries, because the magnetic field could be equally well explained by lateral variation of susceptibility.

In paper III the accuracy of global crustal models is investigated. Errors in global crustal models directly propagate to the calculated isostatic topography, making uncertainty estimations of crustal models essential to assess the role of dynamic and isostatic topography. This uncertainty is mainly caused by interpolation of data acquired along profiles and reduced accuracy of older data. The popular crustal model Crust1.0 (Laske et al., 2013), which is a higher resolution update of Crust5.1 (Mooney et al., 1998), is based on an extensive catalogue of active seismic surveys (mostly refraction). The design philosophy of Crust1.0 is based on pre-defined tectonic domains (i.e. Orogen, extended crust, rift) to ensure geologically plausible results. However, by including expert knowledge in this way, a quantitative assessment of model accuracy becomes difficult.

To allow a quantitative estimation of uncertainty, a geostatistical interpolation technique is developed. The interpolation is a form of kriging, that allows for non-stationarity and works in spherical coordinates. As input I use a catalogue of active seismological surveys to derive a global map of crustal thickness and average crustal P-wave velocity. Furthermore, the uncertainties of the interpolated quantities are derived. Propagating the uncertainties to crustal residual topography, I find that the uncertainty is negligible in well-studied areas (North America, Europe, parts of East Asia), but crippling in many frontier areas.

In chapter 5, the crustal model derived in Paper III is employed in a probabilistic joint inversion. The inversion combines satellite gravity gradients (Bouman et al., 2016) and topography (Amante and Eakins, 2009) to invert for a density model of the lithosphere, consisting of the crust and mantle lithosphere. The mid-Atlantic region is chosen as test area for the inversion, since the lithospheric cooling should produce a clearly resolvable trend.

The important role of the lithospheric mantle for isostatic compensation was shown previously by probabilistic inversion for lithospheric models that include the effect of chemical composition of the lithospheric mantle on mineral phases and physical properties (Fullea et al., 2009; Afonso et al., 2013a; Afonso et al., 2016a). Such models have many degrees of freedom and thus require as much complementary data as possible, i.e. elevation, gravity, geoid, surface heat flow, surface wave dispersion curves, receiver functions (i.e. Tork Qashqai et al., 2016). While elevation and gravity data are available nearly globally in good quality, other data are much more limited. Thus, in many regions of the world, only simpler models (than the thermo-chemical models) can be estimated and these models might prove insightful, despite their simplicity (Fullea et al., 2007).

The inversion results confirm that satellite gravity gradients and topography are sufficient to estimate a first-order lithospheric model. However, the sensitivity to crustal structure is limited.

Contributions

Paper I My contribution to the paper are all calculations, all figures and the compilation of the initial manuscript.

Paper II I re-implemented the inversion routine, based on previous work by Baykiev (2017). All figures were created by myself and I compiled the initial version of the manuscript.

Paper III I implemented the non-stationary kriging routine and adapted it to spherical coordinates. I created all figures and the initial version of the manuscript.

2 Paper I - Importance of far-field topographic and isostatic corrections for regional density modelling

Wolfgang Szwillus[1], Jörg Ebbing[1,2], Nils Holzrichter[1]

[1] Department of Geosciences, Kiel University

[2] Department of Petroleum Technology and Applied Geophysics, Trondheim, Norway

Published in *Geophysical Journal International*, Vol. 207, Issue 1, Pages 274 – 287

2.1 Introduction

Nowadays global gravity models open up new possibilities for studying deep processes like dynamic topography and overall mantle dynamics (Panet et al., 2014). Recent gravity satellite missions like GOCE and GRACE provide a nearly global gravity field with a resolution of up to 80 km and uncertainties of less than 1 mGal (Floberghagen et al., 2011).

To isolate the sub-lithospheric part from the long-wavelength gravity field, careful processing is required. Conventionally, it is assumed that the Spherical Harmonic degrees < 12 reflect the sub-lithospheric field. However, Root et al. (2015) demonstrated that the long-wavelength gravity field not only reflects dynamic processes in the sub-lithospheric mantle but is also influenced by the internal lithospheric and more specifically crustal architecture, as well as glacial isostatic adjustment. Thus, simple wave-length filtering is not sufficient to isolate the sub-lithospheric field. Rather the Bouguer gravity anomaly must be stripped of the lithospheric field (Kaban et al., 1999).

Conventionally, the Bouguer anomaly is used for interpretation and modeling (Nabighian et al., 2005). This implies that topography is removed from measured gravity data to reveal sub-surface density inhomogeneities. The Bouguer correction is two-fold and the complete Bouguer correction accounts for the masses of a (spherical) Bouguer plate and the terrain correction. In the era of satellite gravity and airborne gradiometry such a definition is somewhat outdated and instead often a topographic mass reduction is applied taking into account the actual distribution of topographic masses (Hirt et al., 2012).

An open question is how masses outside the study area have to be considered or even whether a global correction is necessary. The terrain correction is commonly limited to a radius of 1.5° (~167 km) around a gravity station; a convention probably going back to Hayford and Bowie (1912). Topography beyond this radius was considered to be too distant to have any considerable gravity effect.

Mikuska et al. (2006) demonstrated that topographic masses beyond 1.5° can have an important contribution for ground surveys. They found that it is necessary to account for global topography to prevent introducing long wavelength errors into the Bouguer anomaly. But if a global correction for the topographic masses outside a study area is needed, how do we treat the other masses within the Earth outside our study area? One concept is to use a global topographic-isostatic model for correction (Grombein et al., 2014; Kaban et al., 1999).

In this contribution we will quantify the errors caused by neglecting distant effects both for topography and internal masses. This analysis will be performed for ground and satellite gravity as well as satellite gravity gradients.

We first outline how our isostatic model was constructed and how gravity was calculated. Next, the results of these calculations are given and their practical implications discussed.

2.2 Data and methods

All gravity fields in the following are calculated on a global 1°x1° grid of stations at the surface and a height of 225 km, which is similar to the flight height of the GOCE satellite during its latest mission phase. In this study we will use the spherical approximation, with a constant Earth radius of 6371 km. Thus, a height of h km will be converted to a radius of 6371 km + h. While GOCE-based gravity models provide information at higher resolutions, our chosen resolution of 1° is sufficient, because distant effects have much longer wavelengths than 1°.

The station height at ground level was extracted from the ETOPO1 grid nodes to achieve consistency between the station and topographic data and was at the ellipsoid for the offshore stations.

2.2.1 Topographic Effect

The topographic model is based on ETOPO1 (Amante and Eakins, 2009), which we sampled down to a resolution of 5 arc minutes (Figure 2.1). Topography of higher resolution has a significant effect on the topographic gravity effect (TE), but only at small wavelengths, which cannot be detected using a 1°x1° grid of stations.

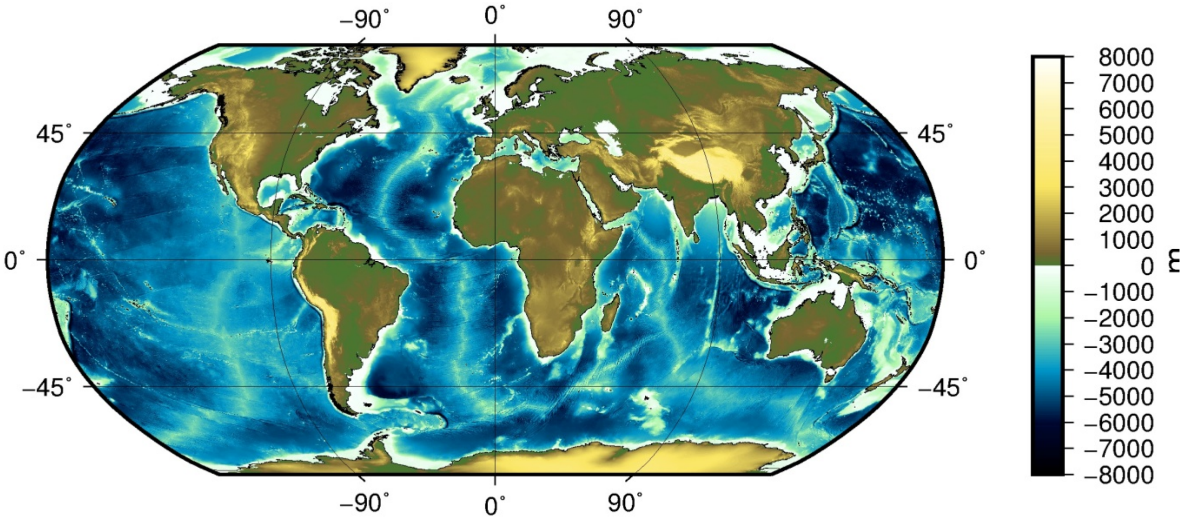

Figure 2.1: Map of global topography (ETOPO1) at a resolution of 5 arc minutes.

For both heights, we calculated the vertical gravity field component. At satellite height we additionally computed the full gravity gradient tensor as measured by the GOCE mission.

We assumed a constant topographic density of 2670 kg m^{-3} and a density of 1040 kg m^{-3} for seawater. To be consistent with our isostatic calculations (see next section), we used a reference crustal density of 2800 kg m$^{-3.}$ This gives a correction density for the oceans of 1760 kg m^{-3}. Different densities due to ice masses, sediments or crustal inhomogeneities are not considered.

For ground station, a radius of 1.5° is an established standard for topographic correction. The gravity effect of masses beyond 1.5° will be called the distant part of that effect. Thus, the topographic effect is split into an Inner Topographic Correction (ITE) and a Distant Topographic Correction (DTE), with TE = ITE+DTE. Thus DTE represents the error that is made when using a correction radius of 1.5°. For satellite gravity we will use the radius of 5° instead. The choice of 5° is justified in section 4.

We will also make use of a more general expression for arbitrary correction radius where ITE(r) refers to the topographic effect calculated using a correction radius of r. Remaining quantities are used analogously: DTE(r) = TE – ITE(r).

2.2.2 Isostatic Effect

In the second set of calculations, we tried to estimate the internal mass distribution by assuming an isostatic model similar to Grombein et al. (2014). We use an Airy-type crustal isostatic model onshore which is a good approximation over continents for long wavelengths. This was verified using regression analysis between Bouguer anomaly and topography (Braitenberg, 2015; Braitenberg et al., 2013). In oceanic areas, crustal thickness is nearly constant and thus cannot compensate topography (McKenzie and Bickle, 1988). Instead, ocean floor topography is caused by secular cooling of the lithosphere. Therefore, lateral variations of density exist in oceanic lithosphere, which we approximate by using Pratt-type isostasy. Furthermore, Airy-isostasy can result in Moho depth above the ocean sea floor if the water is very deep (Figure 2.2).

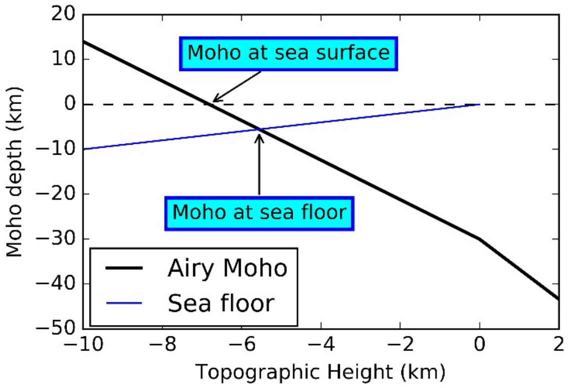

Figure 2.2: Moho depth calculated using Airy isostasy. If water depth is more than 5.5 km (which is 8 % of the oceans), the predicted Moho depth lies above the sea floor. Furthermore, if water depth is greater than 6.8 km, the predicted Moho lies above the sea surface. Thus, Airy isostasy is not useful for very deep water, at least with the chosen density contrast.

Our reference model (see Figure 2.3) has no topography and a crustal thickness of $c = 30$ km. The reference crustal density is 2800 kg m^{-3}, consistent with the average density of the crust (Christensen and Mooney, 1995). The reference mantle density is 3200 kg m^{-3}, (Lachenbruch and Morgan, 1990) thus the contrast between crust and mantle is 400 kg m^{-3}.

Onshore, the additional thickness of the crust r required for isostatic compensation, can be

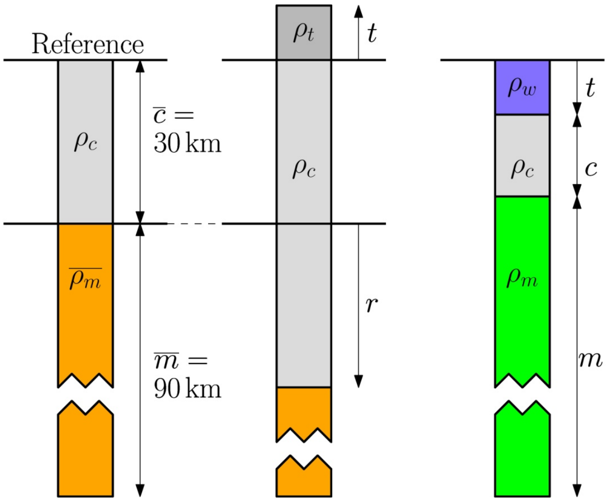

Figure 2.3: Sketch of our isostatic model. Left column: Reference column with zero topography. Middle column: Onshore, topography is compensated by a crustal root of thickness r. Right column: Offshore, bathymetry is compensated by variation of mantle lithosphere density.

computed directly from the topographic height, by balancing the mass in a column with the mass in the reference column (see Figure 2.3). This gives:

$$\mathbf{r} = \frac{\rho_\mathbf{t}}{\rho_\mathbf{c} - \rho_\mathbf{m}}\mathbf{t} \tag{2.1}$$

Where $\rho_t = 2670$ kg m^{-3} is topographic density, $\rho_c = 2800$ kg m^{-3} is the reference crustal density and $\rho_m = 3200$ kg m^{-3} is the reference mantle density. The resulting Moho depth distribution is shown in Figure 2.4 (a).

The thickness of oceanic crust is controlled mostly by mantle potential temperature underneath the spreading ridge, where the crust was formed (McKenzie et al., 2005). In oceanic areas, we emulate secular cooling without explicitly modeling the thermal structure. To this end, we split the oceanic plate into the crust and the mantle lithosphere. The density of the crust is constant and its thickness is derived from Crust1.0 (Laske et al., 2013), which is based on a compilation of active seismic experiments. The density of the mantle lithosphere is adjusted to provide isostatic equilibrium. By balancing the mass in each oceanic column with the mass in the reference column, the density of the oceanic mantle lithosphere is calculated (see Figure 2.3):

$$\rho_\mathbf{m} = \frac{\rho_\mathbf{c}\mathbf{c} + \rho_\mathbf{m}\mathbf{m} - \rho_\mathbf{w}\mathbf{w} - \rho_\mathbf{c}\mathbf{c}}{\mathbf{m}} \tag{2.2}$$

Here, c and m are crustal thickness and mantle lithosphere thickness; while ρ_c, ρ_m and ρ_w are the densities of crust, mantle lithosphere and water; w is the thickness of the water layer. The parameters of the reference model are given by the quantities marked with overbars.

We set the compensation depth to 120 km, which corresponds roughly to the deepest lithosphere-asthenosphere boundary in the oceans (Priestly and McKenzie, 2006). The resulting density distribution is shown in Figure 2.4 (b).

Some regions covered by water are in fact continental (i.e. the continental shelf) and have thick crust (compared to oceanic crust). Therefore, we have only applied equation (2.2) where the

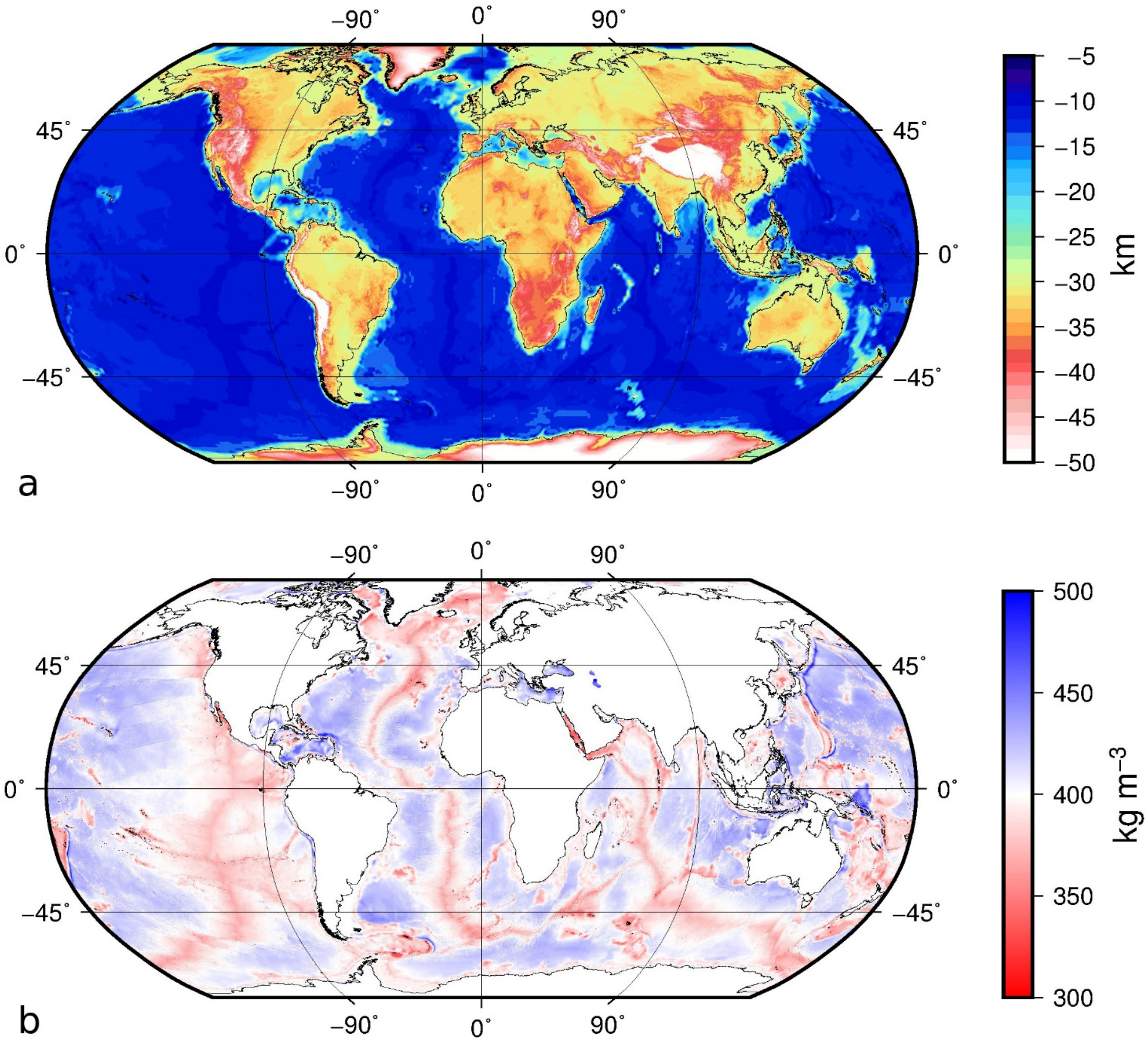

Figure 2.4: (a) Depth to Moho. This is computed using Airy isostasy onshore (see equation (2.1)), whereas the Moho depth from Crust1.0 is used offshore. To have the same resolution as the topography data, the resolution of Crust1.0 was increased to 5 arc minutes. (b) Density contrast between in-situ mantle lithosphere and reference crust. This density contrast is set to a constant value of 400 kg mˆ-3 in the onshore regions. Offshore it is computed using equation (2.2). Note that the youngest oceanic lithosphere is associated with lowest density and the density increases with the age of the plate. This is consistent with the thermal development of oceanic plates

water depth is greater than 350 m. In regions with shallower water we applied a modified version of equation (2.2), which accounts for the presence of water:

$$\mathbf{r} = \frac{\rho_\mathbf{m} - \rho_\mathbf{w}}{\rho_\mathbf{m} - \rho_\mathbf{c}} \tag{2.3}$$

The isostatic effect (IE) is calculated relative to the reference model. Therefore, onshore the IE is caused by the crustal roots with a negative density contrast of 400 kg m^{-3}. In oceanic areas with a normal, thin crust, there are two opposing density contrasts: The first contrast is due to the shallow Moho in offshore areas and always positive. The second contrast is between the reference mantle and the actual mantle densities and is always negative. The sum of TE and IE will be called the topographic-isostatic effect (TIE).

Unlike the topographic effect, the gravity effect of internal masses, in our case the isostatic effect, is normally not calculated with a limited correction radius. For instance, Hinze et al. (2005) compute the topographic effect using a correction radius of 1.5°, while the isostatic effect is computed globally. Another possibility is to compute the effect of the subsurface from the entire study area. Here, we will use the same radius-based approach for IE as we use for TE, to study the magnitude of far-field effects on IE.

Like TE, we will split IE into an Inner and a Distant Isostatic Effect, IE = IIE + DIE.

The combination of both effects is TIE and also split into an Inner Topographic-Isostatic Effect and a Distant Topographic-Isostatic Effect TIE = ITIE + DTIE.

2.2.3 Gravity Calculation

Gravity was calculated using tesseroids. ‘Tesseroids result from a subdivision of the [...] spherical reference surface into elements bounded by geographical grid lines and have a constant [...] height.’ (Heck and Seitz, 2007). Tesseroids thus arise naturally when using equi-angular grids in geographical coordinates. In low latitudes tesseroids resemble prisms with a curved surface. Close to the pole their shape becomes trapezoidal.

No closed-form expression for the gravity effect of a tesseroid can exist (Heck and Seitz, 2007). However, there are different approaches for approximating a tesseroid’s gravity effect (Grombein et al., 2013; Uieda et al., 2011; Wild-Pfeiffer, 2008). We choose to approximate the gravity effect of a tesseroid using numerical integration of Newton’s integral. We will use the vertical component for ground and satellite stations. Additionally, we compute the vertical and north-east gravity gradient component for satellite stations.

Our coordinate system is East-North-Up. We use this coordinate system, because it follows naturally from spherical coordinates, where the radius base vector is pointing outward. Since we are calculating gravity gradients in spherical coordinates, the East-North-Up coordinate system allows us to relate the components of the Marussi tensor directly with derivatives in spherical coordinates (Grombein et al., 2013). Note, that a positive mass has a negative gravity effect in the East-North-Up system, since it is creating a downward acceleration (in direction of negative radius).

For each station we stored TE and IE as a function of great circle distance between station and source in steps of 0.1° of distance. This way we can easily determine the contribution of far-field effects for individual stations without unnecessarily repeating calculations.

From this we computed statistics of the distant effect as function of correction radius. E.g., let DTE $(r)_i$ be the distant topographic effect computed using radius r at the i-th station. The shift affecting all stations is then:

$$\mathrm{Mean}\left(\mathbf{DTE}(\mathbf{r})\right) = \frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{i}} \mathrm{DTE}\left(\mathbf{r}\right)_{\mathbf{i}} \tag{2.4}$$

The standard deviation is

$$\mathrm{STD}\left(\mathrm{DTE}\left(r\right)\right) = \sqrt{\frac{1}{n} \sum_{i} \left(\mathrm{DTE}\left(r\right)_i - \mathrm{Mean}\left(\mathrm{DTE}\left(r\right)\right)\right)^2} \tag{2.5}$$

The root-mean square (rms) can then be computed as

$$\mathrm{RMS}\left(\mathrm{DTE}\left(\mathbf{r}\right)\right) = \sqrt{\mathrm{STD}\left(\mathrm{DTE}\left(\mathbf{r}\right)\right)^{\mathbf{2}} + \mathbf{Mean}\left(\mathrm{DTE}\left(\mathbf{r}\right)\right)^{\mathbf{2}}} \tag{2.6}$$

These statistics describe the average magnitude of distant effects as a function of correction radius. Isolating the shift is important because a constant global shift can normally be disregarded in regional density modeling. The rms is useful, because it is a measure of the expected magnitude of distant effects that incorporates both the spatially constant and varying contributions.

2.2.4 Regional test cases

In addition to the radius based calculations, we chose a test region in North America. It extends from -135°E to -60°E and from 25°N to 50°N. There has been considerable debate on what role dynamic topography plays for example in the Colorado plateau (Liu and Gurnis, 2010). Since the long-wavelength gravity field contains important constraints on dynamic topography (Molnar et al., 2015), it is interesting to investigate how different treatments of topographic and isostatic correction affect the long-wavelength gravity field.

For the regional test area, we calculated topographic and isostatic effects on grids that were found by extending the study area by 2.5°, 5°, 7.5° and 10° respectively (see Figure 2.5). We also used a different approach, where topographic and isostatic corrections are calculated using different areas. This occurs when the entire study area is used for isostatic correction, while topographic correction is limited to some radius around each station (e.g. Prezzi et al., 2009). Note that Prezzi et al. (2009) did not explicitly carry out an isostatic correction, but forward calculated the gravity effect of a regional density model. However, this is comparable to calculating the isostatic effect, because the Moho is the strongest density contrast in any lithospheric density model.

2.3 Results

2.3.1 Ground-based gravity

Both TE and IE (Figure 2.6 (a) and (b)) are closely correlated to topography, but there are three differences. First, the signs are opposite because the sign of the source masses is different. Second, the amplitude of the IE is lower (-600 to +450 mGal) than that of TE (-500 to +700 mGal). Third, IE is smoother than TE. The two latter differences are due to the greater source depth of IE.

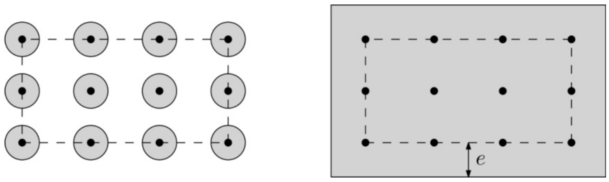

Figure 2.5: Difference between local and regional modeling. Each dot represents a gravity station. In local modeling (left figure), only a circle around each station is considered for gravity calculation. When using regional modeling (right figure) each station uses the same model for gravity calculation. The extent of this model is found here by extending the area covered by station (dashed line) by a certain distance *e*. Note that the distant effect in the case of regional modeling is always less than local modeling with a correction radius of *e*.

This difference in source depths prevents TE and IE from canceling perfectly (Figure 2.6 (c)). Maximum values of combined topographic-isostatic effect (TIE) are +/- 200 mGal, although these extreme values are very rare. At 90 % of the stations the values are between -50 and +40 mGal, and about 60 % of the stations show a TIE of less than +/- 10 mGal.

The strongest negative values of up to -200 mGal are located over the large orogens like the Himalaya or the Andes. However, the impact of these orogens extends beyond the areas of high terrain. Around the edges of these mountain belts a positive TIE of several tens of mGal is observed. This ‘halo’ around massive mountains indicates a far-field effect of the isostatic compensation masses.

TIE is closely related to bathymetry in the oceans. At the mid-oceanic ridges, TIE is slightly negative (< -20 mGal). As the water depth increases, TIE increases as well, reaching maximum values of +200 mGal at the deep ocean trenches. Note that the sign of the correction changes from negative to positive as the water depth increases.

When the topographic-isostatic effect is limited to a radius of 1.5° two changes occur (Figure 2.7 (a)). Firstly, there is a systematic increase of the amplitude compared to a global computation. This means that correcting with a radius 1.5° has a tendency to over-compensate. Secondly, the ‘halos’ observed at the edges of massive mountains like the Tibetan Plateau disappear.

The impact of a limited correction radius can be quantified by inspecting the distant topographic-isostatic effect (DTIE). In general, (Figure 2.7 (b)), DTIE is small where TIE is small as well. Thus, the majority of stations have DTIE between +/- 10 mGal. However, extreme values of -60 to +100 mGal occur. Distinct patches of positive anomalies cover the major orogenies but extend well beyond the area of high terrain. Similar patches exist around the deep ocean basins, especially in the Pacific Ocean and at the mid-oceanic ridges. The standard deviation of DTIE is 22 mGal.

When the correction radius is increased, far-field effects on TIE decrease significantly (Figure 2.8). The rms of DTIE(r) drops from around 45 mGal at a radius of 1° to 6.5 mGal at 5°. Between radii of 1° and 30° the relation between correction radius r and DTIE(r) is approximately a power-law with an exponent of ~ -1.1. Therefore, DTIE(r) reduces by a factor of ~2.2 if the correction radius is doubled. For any radius, the DTIE(r) is not merely a constant global shift, because the mean over all stations is for the most part five times smaller than the rms.

Note, that TIE is less affected by far-field effects than TE alone (Figure 2.7 (c)). DTE lies between 60 and 200 mGal, with a standard deviation of 28 mGal. Also, DTE introduces a long-wavelength trend in essentially all regions of the world. Furthermore, the rms of DTE decreases slower than DTIE when the correction radius is increased (Figure 2.8). Thus, topography and its isostatic

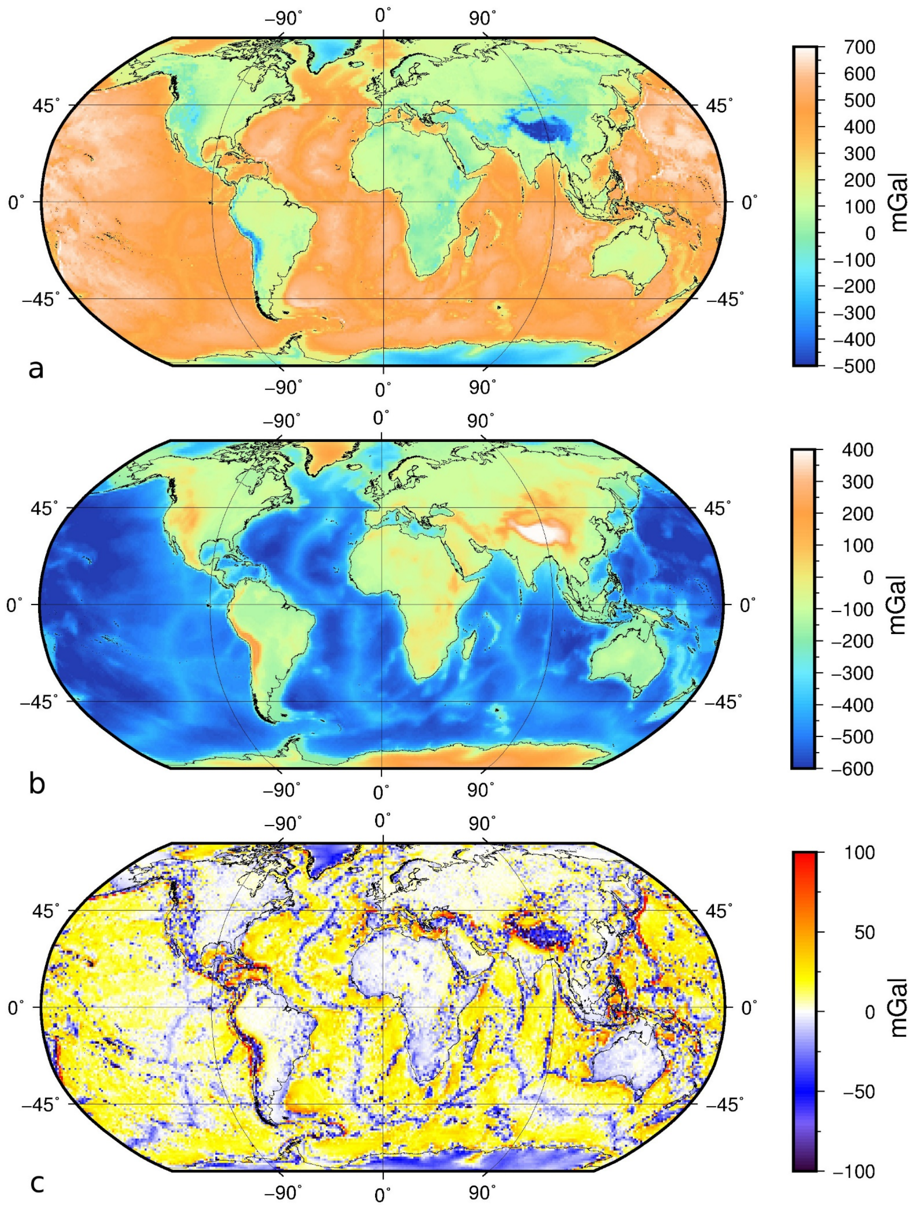

Figure 2.6: (a) Topographic effect (TE) for stations at ground level calculated using global topography. TE correlates strongly with topography. Maximum positive (negative) values are attained over the Pacific Ocean (Tibetan Plateau) respectively. Note that highest positive amplitude is 200 mGal larger than the highest negative amplitude, implying a dominance of the bathymetric effect over the topographic effect. (b) Isostatic effect (IE) for stations at ground level calculated using global isostatic model. Like the RE the IE correlates strongly with relief, but with opposite sign. Thus, IE and RE are anti-correlated; however, this correlation is not perfect. IE shows slightly smaller amplitudes than TE (-600 to +450 mGal, vs. -500 to +700 mGal). Furthermore, IE is smoother than TE due to the greater source depth of TE. (c) . Combined Topographic-Isostatic Effect (TIE) for stations at ground level calculated using global data. This is obtained by adding topographic and isostatic Effect. Thus TIE is also correlated with relief. However, IE has deeper sources than TE, so that often TE is greater than IE. Another noteworthy feature are the positive ‘halos’ around the massive orogens. These imply a net positive effect from the isostatic masses underneath the orogens. Therefore, they would necessitate an increased correction radius at the edges of these orogens.

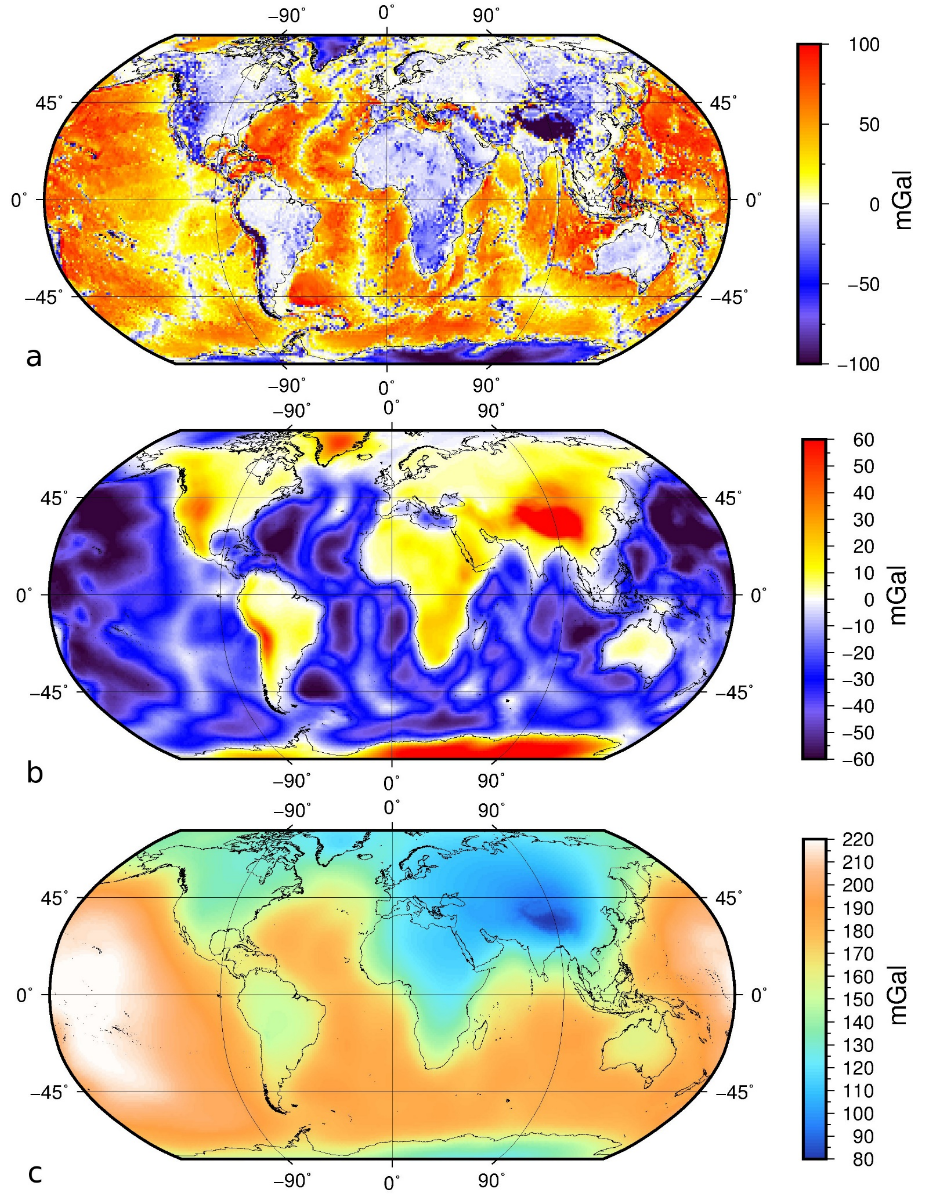

Figure 2.7: (a) Inner Topographic-Isostatic Effect (ITIE) for ground stations computed with a correction radius of 1.5°. Compared to Topographic-Isostatic Effect (TIE, Figure 2.6 (c)) the amplitudes increase, thus when using a limited correction radius there is a tendency to over-estimate TIE. Note the ‘halos’ around the massive orogens have essentially disappeared. (b) Distant topographic-isostatic Effect (DTIE) for ground stations. Compared to TIE, the features are associated with continental-scale relief features, such as orogens and extensive ocean basins. The amplitude of DTIE is smaller than TIE, but can still account for up to 50 % of the amplitude. This map can be interpreted as a map of the error incurred by limiting corrections to a radius of 1.5°. Note, the sign of DTIE and ITIE is opposite. This is due to the fact that ITIE always tends to overestimate the magnitude of TIE. Hence, DTIE needs to have the opposite sign of ITIE, so that the sum of DTIE and ITIE has a lower value.(c) Distant Topographic Effect (DTE) for stations at ground level.

compensation should always be considered together, to avoid introducing long-wavelength bias.

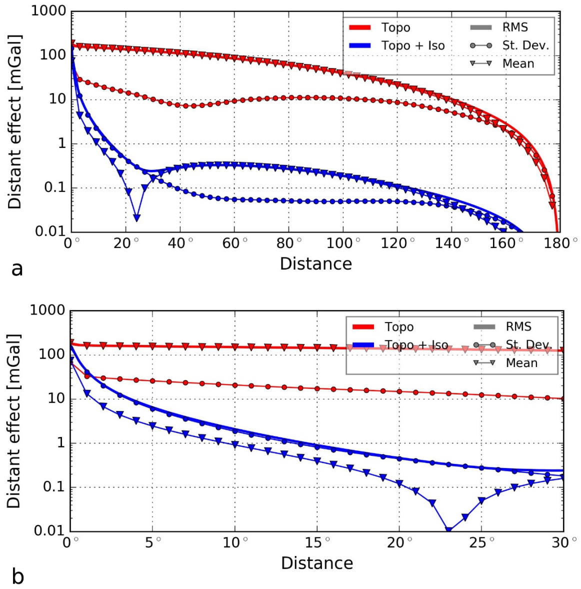

Figure 2.8: Remaining distant TE (red) and TIE (blue) as function of correction radius for ground stations. RMS (solid line), standard deviation (circles) and mean (triangles) are given. (a) gives the complete range of correction radii, while (b) is a zoom on radii between 0 and 30 °. For TE there is a significant amount of shift so that the standard deviation is mostly several times lower than the mean. In contrast, TIE has a lower mean than the standard deviation, as long as the correction radius is less than 30°. For radii of less than 30°, the standard deviation of TIE decays much faster than the standard deviation of TE. This highlights that it is important to consider isostatic and topographic effects together. The local minimum at a distance of around 23° distance is due to the fact that the mean changes its sign.

2.3.2 Satellite-based gravity

TE and IE for satellite stations are essentially low-pass filtered versions of the respective fields at ground level. We therefore do not discuss them further here.

Far-field effects have a much stronger effect on satellite stations than on stations at ground level (compare Figure 2.9 (a) and Figure 2.10 (a)). For radii less than 3°, the distant effect first decreases and then increases with radius (Figure 2.10 (a)), so clearly the correction radius should be larger than 3°. For radii above 3°, the rms remaining effect decreases, approximately obeying a power-law with an exponent of -2, similar to ground stations. To reduce the rms remaining distant effect to 10 % of the rms global effect, a correction radius of 15° would be necessary. For correction radii above 30°, DTIE(r) starts to be dominated by shift, which tends to be several times larger than the standard deviation for radii above 30 ° (Figure 2.10 (a)).

A correction radius of 5° gives DTIE(r) for satellite stations that are very similar to DTIE (1.5°) obtained for ground stations. The radius of 5° gives the maximum correlation coefficient between the DTIE(r) for satellite stations and DTIE (1.5°) for ground stations. This motivates our choice of anything beyond 5° as 'distant' for satellite stations.

DTIE (5°) for satellite stations is shown in Figure 2.9 (b). DTIE is between -9 and +10 mGal. Overall the distribution is skewed towards negative value since 90 % of the values lie between -8

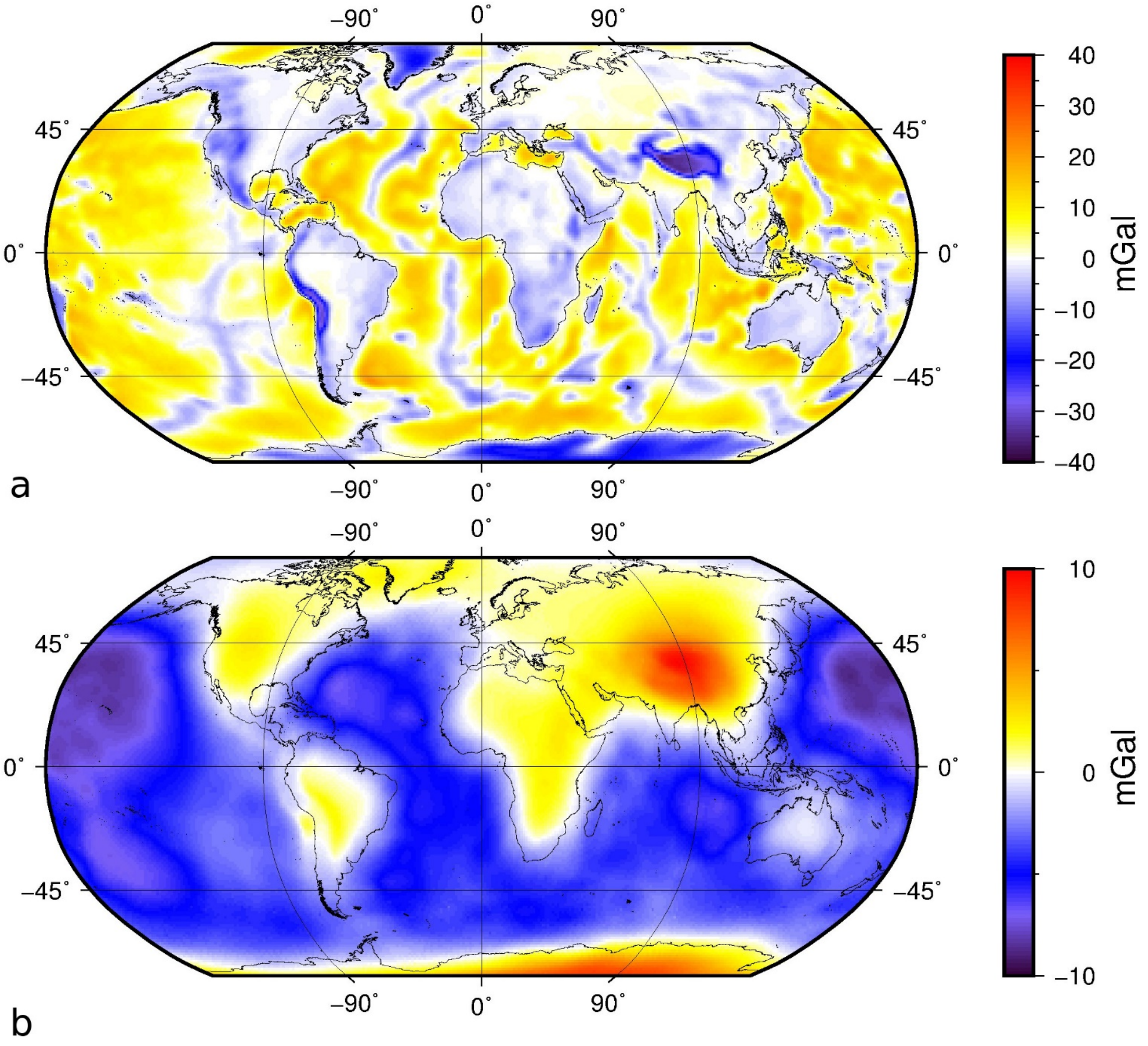

Figure 2.9: (a) Topographic-isostatic effect (TIE) for stations at a height of 225 km. This is the field continuation of TIE at ground level (Figure 6 (c)). The large-scale features observed are similar. (b) Distant (Distance > 5°) TIE for stations at a height of 225 km.

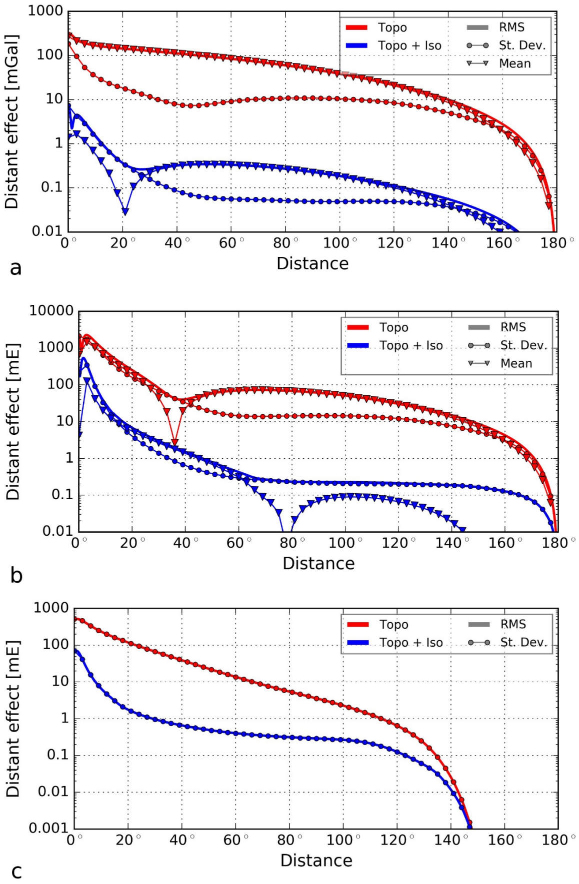

Figure 2.10: Remaining distant TE and TIE for all stations at satellite height. The subfigures refer to g_r (a), g_{rr} (b) and g_{ne} (c) respectively. In each case the red line gives the distant effects for TE alone, while the blue line gives distant effects for TIE. Different statistics are represented by solid lines (RMS), circles (standard deviation) and triangles (mean). Note TE alone is normally one or two orders of magnitude larger than TIE. (a) Remaining distant effect (gravity) as a function of correction radius for satellite stations. For small correction radii the remaining distant TIE increases with increasing correction radius. Decay of the remaining distant effect begins at a radius of 2.5°, which is similar to the height of the stations (225 km). At a radius of around 30°, standard deviation and mean of DTIE are roughly equal. For larger radius the mean is about five times larger than the standard deviation, so there is a significant amount of shift. (b) Vertical gravity gradient component. The local minima in the mean of are caused by a change of sign. (c) North-east gradient component. Note, the mean is completely negligible.

and +5 mGal and the mean value is -2 mGal.

2.3.3 Satellite-based gravity gradients

The vertical gravity gradient component (g_{rr}) is very closely related to vertical gravity. Thus, TIE for the g_{rr} is a high-pass filtered version of the g_r-correction (Figure 2.9 (a)). Compared to vertical gravity, distant sources are not as important for g_{rr} (see Figure 2.10 (b)). To reduce the remaining rms distant effect to 10% of the global rms, a radius of 10° is sufficient. The spatial distribution of distant effects for a correction radius of 5° is very similar to the distribution for vertical gravity (Figure 2.11 (b)).

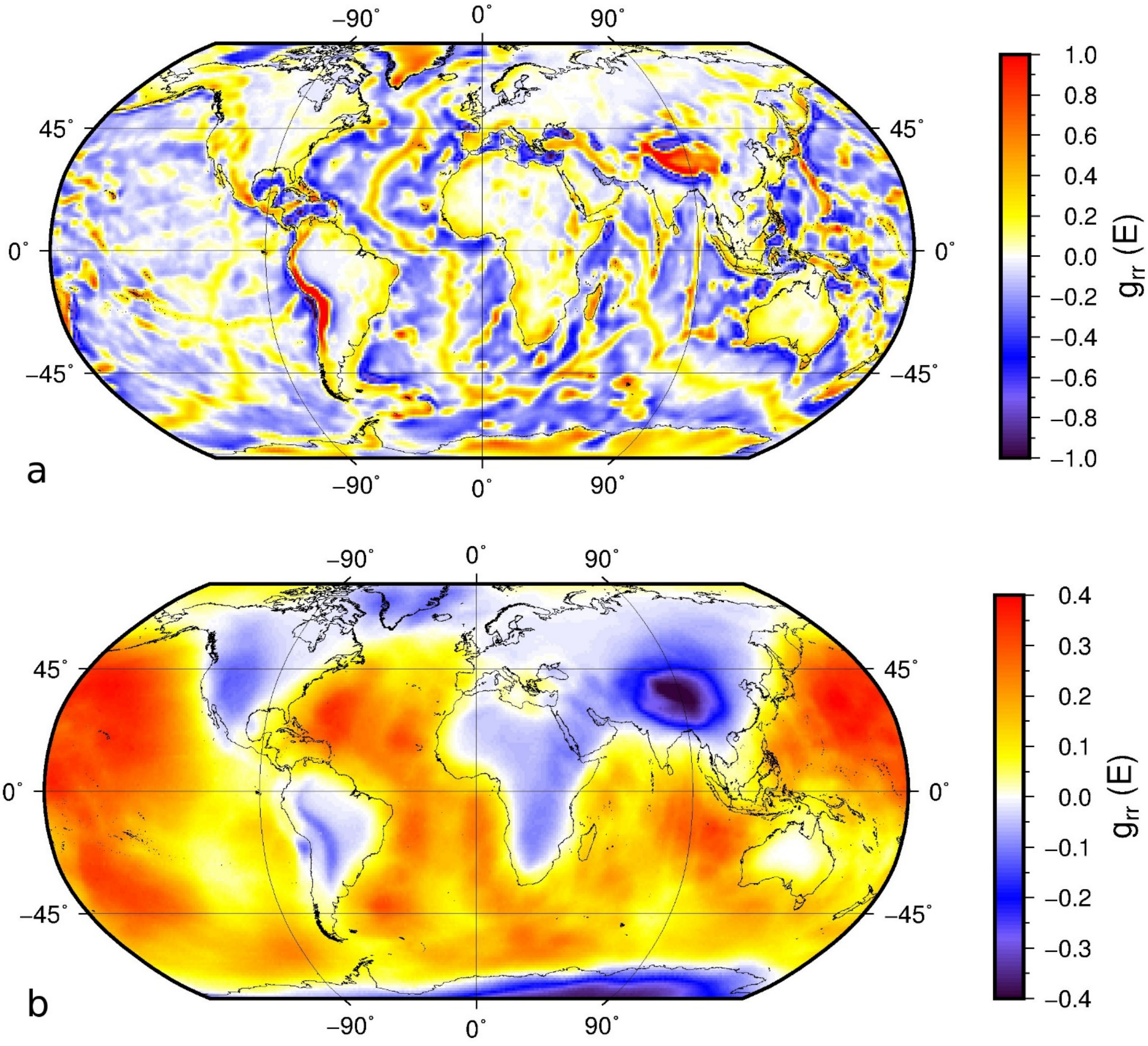

Figure 2.11: (a) Topographic-isostatic effect for stations at 225 km height, vertical gradient component. This component is essentially a high-pass filtered version of the vertical gravity component (Figure 9). (b) Distant topographic-isostatic effect (DTIE), distance > 5°, for stations at 225 km height, vertical gradient component.

TIE for the east-north gradient correlates with the edges of major mountain belts that are oriented at 45°, 135°, 225° or 315° N , i.e the Andes, Himalayan Belt. Additionally, the mid-oceanic ridges can cause a response of up to 0.2 E. Consequently, 90 % of the stations have values of less than 0.1 E and 50 % of the stations less than 0.03 E. Locally, however, the effect can be up to 0.8 E (Figure 2.12 (b)).

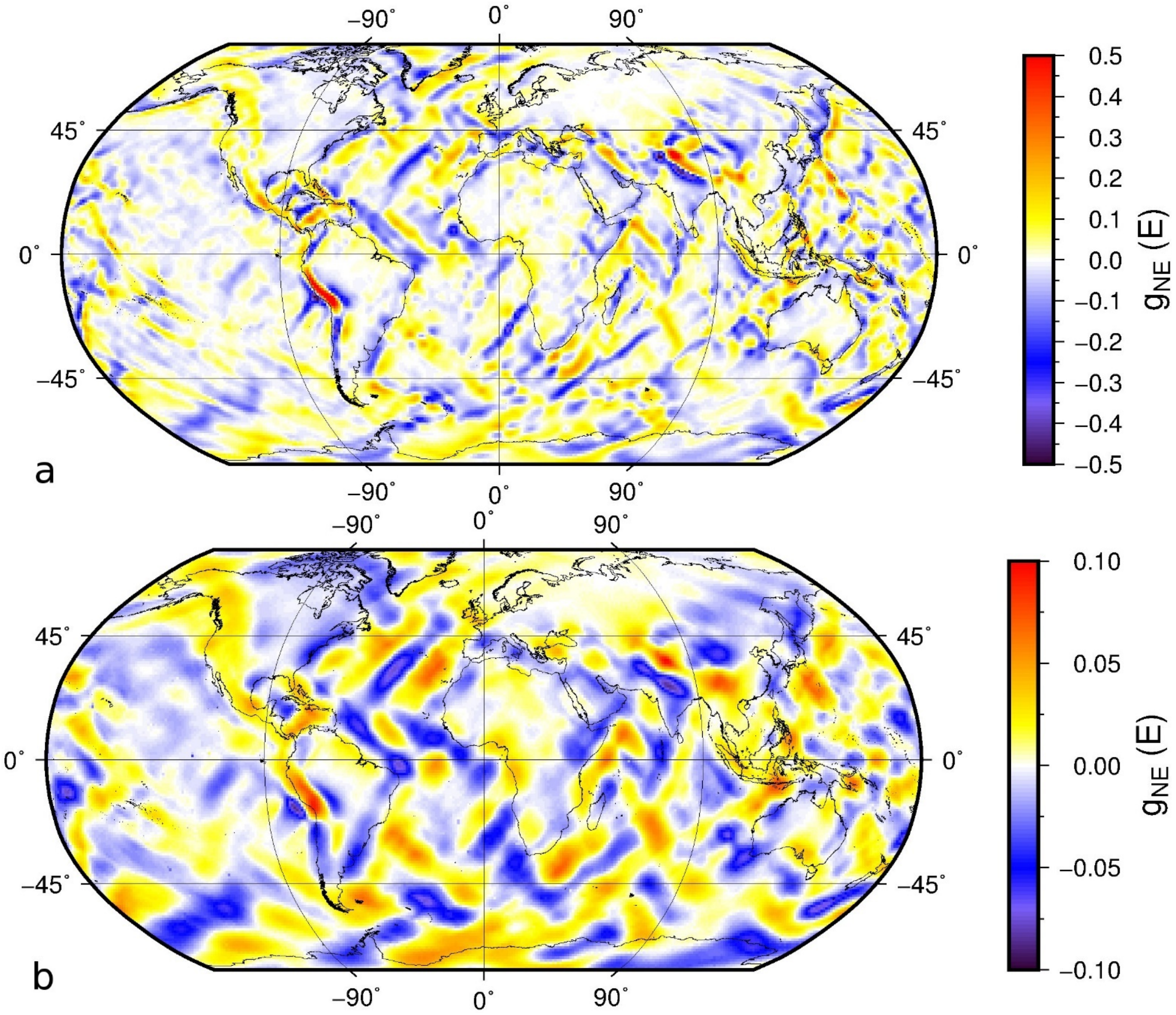

Figure 2.12: (a) Topographic-isostatic effect for stations at 225 km height, north-east gradient component. (b) Distant topographic isostatic effect for stations at 225 km height, north-east gradient component.

As the correction radius increases, the remaining distant effect reduces rapidly for the east-north gradient component (Figure 2.10 (c)). To reduce the remaining rms distant effect to 10% of the global rms, a radius of 8.5° is sufficient. The spatial distribution of the distant effect beyond 5° shows almost the same pattern as the global correction, except for a slight broadening of the features. This implies that a correction radius of 5° will slightly – but systematically – underestimate the global correction.

2.3.4 Regional test cases

Using the TIE obtained from a global calculation as reference, we compared different ways of calculating the TIE. For this, the mean of each grid was subtracted from each grid, to make the different results more easily comparable.

We compare the TIE obtained using different approaches for ground (Figure 2.13) and satellite (Figure 2.14) stations. A global topographic-isostatic correction (a) is used as reference to compare the other approaches.

The approach where different areas are used for topographic and isostatic correction (Figure 2.13 (c)), gives errors with respect to a global calculation between +/- 20 mGal. Using a correction radius of 1.5° for both topographic and isostatic (e) causes larger errors of +/- 40 mGal. However, by increasing the radius to 5°, the errors are reduced to less than 15 mGal. The purely grid-based approach (g) comes closest to the global calculation. The errors associated with this approach are less than 5 mGal, and this error is concentrated near the edges.

Satellite stations are affected more strongly by an inadequate calculation of topographic-isostatic effects. The 'mixed area' approach (Figure 2.14 (c)) causes the largest errors, of -25 to +60 mGal. This error is so substantial that in fact the actual TIE signal is overwhelmed, especially near the edges. The purely grid-based and purely radius-based approaches both approximate the results of a global calculation quite well. The highest magnitude of the error is 1 and 4 mGal respectively, with standard deviations of less than 1 mGal.

2.4 Discussion

The results of the previous section clearly show that a significant part of the long-wavelength gravity field is caused by distant sources associated with topography and its isostatic compensation. We thus confirm the finding of Root et al. (2015), that wavelength-filtering is insufficient to separate the lithospheric and sub-lithospheric field. However, far-field effects are weaker than calculated previously by Mikuska et al. (2006), because topographic and isostatic effects balance each other out to a certain degree.

This balancing only occurs however, if topographic and internal masses are considered over the same area. In our regional test case we found that using different areas for topographic and isostatic correction causes errors because of the uncompensated (or missing) distant topography. These errors are in the order of 25 % to 50 % of the actual topographic-isostatic effect for ground stations (Figure 2.13) and up to 100 % for satellite stations (Figure 2.14).

Radius-based approaches are quite effective if the same radius is used for topographic correction and modeling/isostatic correction. However, they fit poorly into existing modeling frameworks, which are based on using grids to represent the studied area.

We therefore propose to change the topography correction instead, i.e. calculate the topographic effect using a grid that coincides with the study area (plus an extension of a few degrees). This reduces far-field effects coming from outside the study area reasonably well. The topographic correction can thus not be considered independently from the area of study.

Calculating corrections globally in principle eliminates far-field effects completely. However, meaningful gravity modeling can only be done where good constraints from other methods, i.e. seismology, are available. Global reference models, like Crust1.0, are only well constrained in some areas of the world (Laske et al., 2013). By including such data in the modeling processes additional uncertainties are introduced. Also, these models were never intended for use with gravity. Most importantly, these models are not in isostatic balance.

For these reasons we propose 'nested' modeling instead of global reference models. First, a global topographic correction is applied to the free-air gravity data, giving a global Bouguer anomaly. Then, the study area is cut out from a global isostatic model. This is done by setting the crustal structure in the study area to the reference model. The gravity effect of this modified model represents all isostatic effects from outside the study area. This isostatic effect is subtracted from the Bouguer anomaly. This corrected anomaly can then be used for modeling inside the study

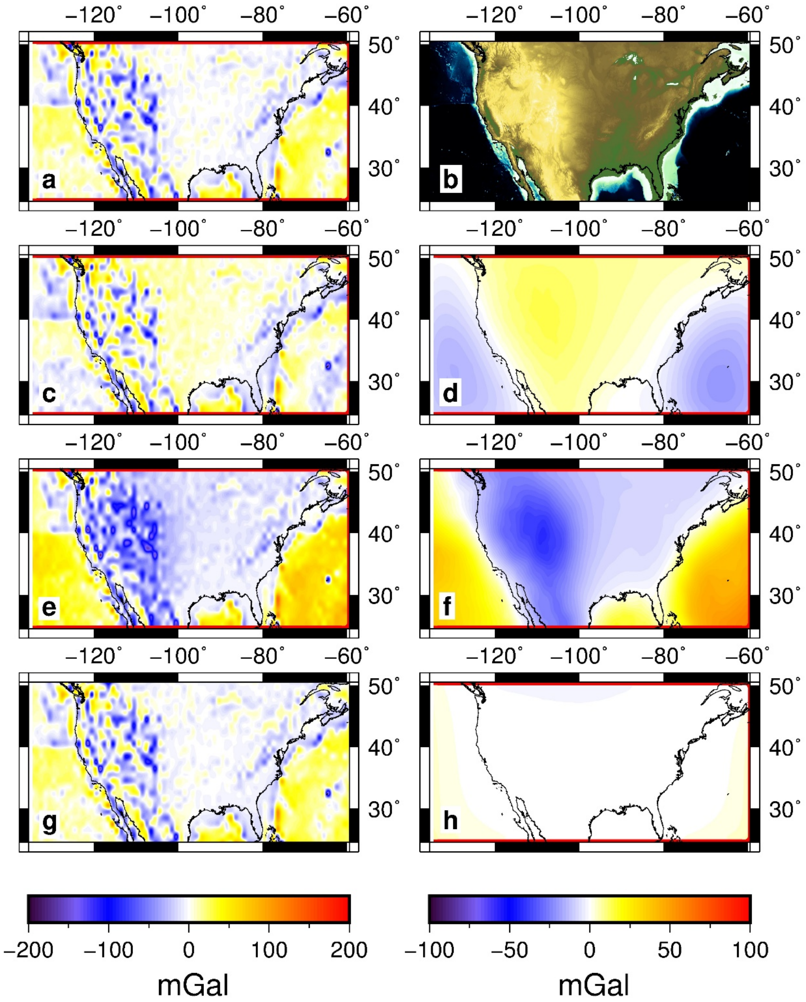

Figure 2.13: Results from the regional test case for ground stations. The left column gives the Topographic-Isostatic effect calculated using different approaches. (a) gives the result using global calculation and is considered as the reference, to which all other results are compared. (c) gives the result when a local topographic correction (correction radius 1.5°) is combined with an isostatic correction calculated from a grid corresponding to the study area extended by 5°. (e) is the result using a correction radius of 1.5° for both topographic and isostatic correction. Finally, (g) gives the result obtained using a grid corresponding to the study area extended by 5° for both topographic and isostatic correction. In (b) the topography in the test area is shown, while the remaining images in the right column give the difference of the respective approaches compared to the global topographic-isostatic correction, shown in (a). The range of values is reduced to +/- 100 mGal. Note the calculated fields show strong short wavelength features associated with topography. These are caused by the fact that the topographic grid has a resolution of only 5 arc minutes, which is normally insufficient for calculation of topographic effects at ground level. Also, the resolution of the station grid is 1°, so there is some aliasing as well. However, the errors displayed in the right column are much smoother, because they are caused by distant masses. Hence, the chosen resolution is adequate for judging the magnitude of distant effects.

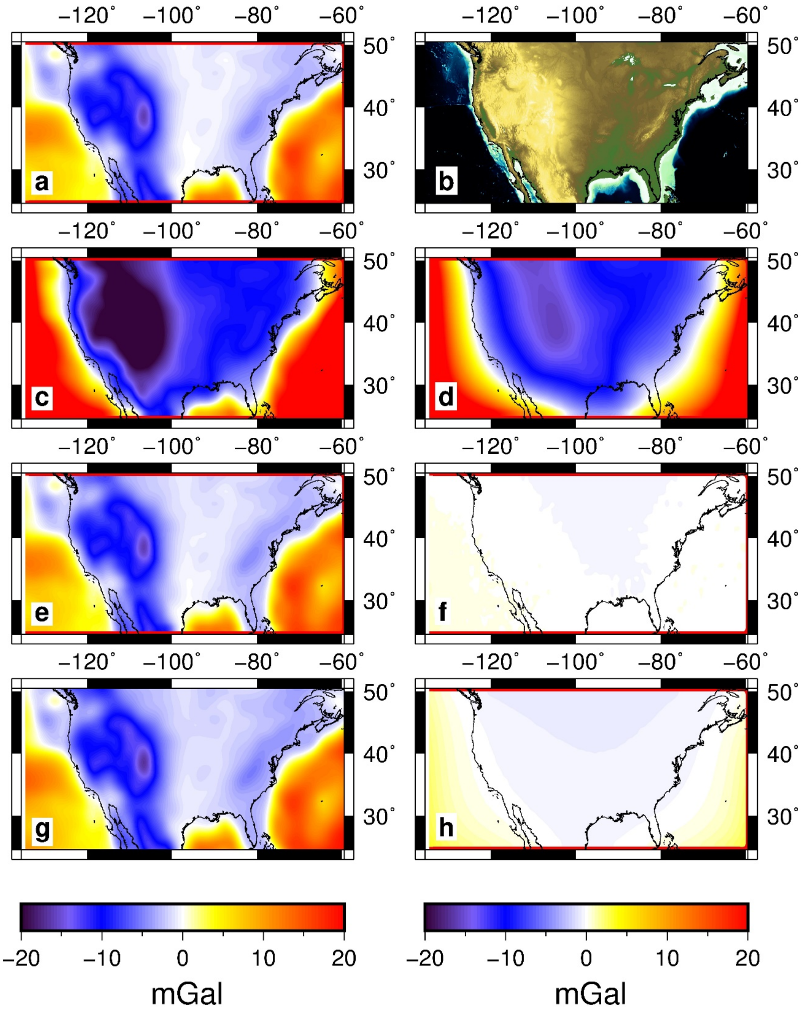

Figure 2.14: Results from the regional test case for satellite stations. The left column gives the Topographic-Isostatic effect calculated using different approaches. (a) gives the result using global calculation and is considered as the reference, to which all other results are compared. (c) gives the result when a local topographic correction (correction radius 15°) is combined with an isostatic correction calculated from a grid corresponding to the study area extended by 5°. (e) is the result using a correction radius of 15° for both topographic and isostatic correction. Finally, (g) gives the result obtained using a grid corresponding to the study area extended by 5° for both topographic and isostatic correction. In (b) the topography in the test area is shown, while the remaining images in the right column give the difference of the respective approaches compared to the global topographic-isostatic correction, shown in (a). These results clearly demonstrate that approaches where topographic and isostatic effects are not considered over the same regions are inferior to any approach where both effects are calculated from the same region. In (c) and (d) the negative topographic-isostatic effects associated with the high topography in the west are exaggerated by up to 20 mGal. Also, there are edge effects of more than 20 mGal. (g), (h),(i) and(j) are very similar to the result from global calculation (a). For the radius-based approach (g) and (h), errors are less than 1mGal, while the extension-based approach gives errors of less than 5 mGal, which are also limited to the edges of the study region.

area, like a normal Bouguer anomaly. In this way, global isostatic compensation is accounted for to first order. Of course this is an approximation, but it rests on the physical principle of isostasy.

To illustrate the impact of far-field effects, we will use uncertainties introduced by Moho error. This can be estimated using the expression for the gravity effect of a thin sheet:

$$\delta \mathbf{g} = \frac{4\pi \mathrm{G}}{\mathbf{2l+1}} (\mathbf{l+1}) \left(\frac{\mathbf{R+c}}{\mathbf{R+h}} \right)^{\mathbf{l+2}} \boldsymbol{\Delta} \rho \delta \mathbf{m} \tag{2.7}$$

Where G is the gravitational constant, l is the spherical harmonic degree, corresponding to a wavelength of $\lambda = \frac{40000 \text{ km}}{l}$, $\Delta\rho$ is the Moho density contrast, δm is the Moho uncertainty and h is the measurement height (Braitenberg et al., 2010).

At a wavelength of 1000 km (typical for distant effects) and a density contrast between lower crust and mantle of 200 kg m^{-3} we find that each kilometer of Moho uncertainty corresponds to an error of about 6 mGal at Earth's surface. In the height of GOCE, each kilometer of uncertainty results in 1.2 mGal of gravity error, and 0.01 E of vertical gravity gradient error. Even in well-studied areas like North-America or Europe, the Moho depth is not known everywhere with better accuracy than 5 km (Tesauro et al., 2014). This gives expected long-wavelength modeling uncertainties of 30 mGal for ground stations and 6 mGal (0.05 E) for satellite measurements. For simplicity's sake we have assumed that the Moho error is limited to a single wavelength. However, for the wavelengths in question the scaling factor between Moho error and gravity error does not depend strongly on the wavelength.

For ground stations a correction radius of 2.5 ° is sufficient to reduce the rms of distant effects of the topographic-isostatic effect to less than 16 mGal, which is 10 % of total rms effect (Figure 2.8). Using a correction radius of 12 °, this can be reduced to 1.5 mGal, corresponding to a relative accuracy of 1 %. Thus, a global correction is clearly not needed for ground stations, but correction radii above 1.5° are. Also, this is well below the errors associated with typical Moho uncertainties.

Satellite stations require larger correction radii, because of the different measurement geometry. To reach a rms relative accuracy of 10 % (0.7 mGal), a correction radius of 15° is required for gravity (Figure 2.10 (a)). An rms error threshold of 1 % (0.07 mGal) requires a correction radius of more than 130°. However, if shift is neglected, a radius of 40° is sufficient for 1 % relative accuracy,

For gradients, a radius of about 10 ° is sufficient for vertical and north-east gravity gradient to reach a rms relative accuracy of 10 %. A stricter error threshold of 1 % would lead to radii of about 35° for vertical and 30° for the north-east gravity gradient (Figures 2.8 (a) and (b)). This shows that gradients are less affected by distant effects, making them well-suited for regional modeling.

The required correction radii given above represent global averages. Depending on the location of the study area, larger or smaller correction radii might be required. Furthermore, some regions might show high values of the distant effects, but these might essentially be constant within that region. Thus, the values given above should be taken as guidelines that need to be evaluated on a case-by-case basis. For North America we find that for satellite stations a radius of 15° gives an RMS of 0.33 mGal, which is one third of the global average.

Our results rely on the assumption that all topography is compensated isostatically. This view has been challenged recently by a number of authors (see Flament et al., 2013). They claim that up to 2 km of Earth's topography is compensated dynamically. If this is true, we could no longer assume that distant topographic effects are compensated by distant isostatic effects.

However, it is a matter of discussion whether dynamic topography of this magnitude exists. One kilometer of dynamic topography should be associated with a free-air anomaly of 50 to 100 mGal, depending on the depth of the density anomaly driving the flow. The observed free-air anomaly thus limits dynamic topography to a few hundreds of meter (Molnar et al., 2015). This means isostatic compensation should be treated as the null hypothesis of topographic compensation. Consequently, global isostatic effects have to be addressed before conclusions about non-isostatic contributions to topography can be made.

Balmino et al. (2012) have computed a global topographic-isostatic correction assuming Airy-type isostasy both onshore and offshore. Onshore, we assumed Airy-type isostasy as well. Hence, it is no surprise that the isostatic effects differ only by up to +/- 30 mGal onshore. This difference can be explained by the different parameters (densities) used in our calculations. Over the ice sheets large differences of up to 100 mGal occur, because our isostatic model neglects the effect of ice, whereas Balmino et al. (2012) have included it.

In the oceans, there is a difference of several tens of mGal that is due to the different isostatic models we used. In fact, the sign of the combined topographic-isostatic correction is different between our results and those of Balmino et al. (2012). Their assumptions imply that large-scale bathymetry is caused by variations of crustal thickness in the oceans. However, bathymetry is caused by secular cooling of the lithosphere, whereas the crustal thickness is (apart from anomalous regions) nearly constant at 7-10 km (McKenzie and Bickle, 1988).

Our model tries to mimic this by changing the density of the mantle lithosphere to achieve isostatic equilibrium. This is a strong approximation, because the density variation is of thermal origin. Therefore, density anomalies should be concentrated near the top of the mantle lithosphere. The deeper parts of the lithosphere have not been affected by cooling from the top, because thermal conduction is slow. The half-space cooling model predicts that it takes around 80 million years until the thickness of the lithosphere has reached 120 km. This means that our model is better suited for old oceans.

Both our and the model by Balmino et al. (2012) are valid isostatic models in the sense that they are in isostatic equilibrium. However, they are opposing end members with respect to the depth of the compensation masses: Balmino et al. (2012) place the compensating masses as close to the surface as possible, whereas our model distributes the compensating masses over the entire depth of the lithosphere.

2.5 Conclusion

We computed global topographic corrections and inferred how much of topographic gravity effect is due to far-field effects. We demonstrate that the classic scheme of using a Bouguer plate and then accounting for deviations between actual topography and the plate is not needed anymore. Instead, both simple Bouguer correction and the terrain correction can be applied in one step.

We applied the same analysis to the gravity effect of isostatic compensation. This gravity effect was derived from a global isostatic model, based on Airy-isostasy on the continents, and Pratt-isostasy offshore.

Both the topographic and isostatic gravity effect are strongly influenced by distant sources, i.e. sources more far away than 1.5° (ground stations) and 5° (satellite stations). However, when topographic and isostatic effects are considered together, far-field effects become much weaker, because topographic and isostatic effects cancel each other out.

However, it is critical that topographic and isostatic corrections are calculated using the same area. We propose three different ways to achieve this:

1. Use the same correction radius for topographic correction and modeling (isostatic correction).
2. Use same grid area as the study area (extended by a few degrees) for topographic correction.
3. Use a global topography correction, and account for isostatic effect from masses outside of the study area.

Gravity gradients provide an additional way of reducing far field effects. Being a measured quantity, they are well-suited for regional modeling without requiring any supporting assumptions.

Acknowledgments

This study is funded by the German Research Council (DFG) as part of the Special Priority Programme SPP 1788 “Dynamic Earth”.

All figures in this work were created with the Generic Mapping Toolbox (GMT) (Wessel et al., 2013).

The authors would like to thank Carla Braitenberg for inspiring discussion about the interplay between topographic and isostatic effects. Also, two anonymous reviewers helped to improve an earlier version of the manuscript.

3 Paper II - Sensitivity of satellite magnetic field to crustal and upper mantle temperature structure

Wolfgang Szwillus[1], Eldar Baykiev[2], Jörg Ebbing[1]

[1] Department of Geosciences, Kiel University

[2] School of Cosmic Physics, Dublin Institute for Advanced Studies

Submitted to *Journal of Geophysical Research: Solid Earth*

3.1 Introduction

The magnetic field associated with the lithosphere depends on the temperature structure. Thus, a number of studies have used global or large-scale regional magnetic field data to estimate temperature distribution in the lithosphere. (Maule et al., 2005; Tanaka, 2017; Martos et al., 2017). Temperature affects the lithospheric magnetic field because ferromagnetic minerals become effectively non-magnetic above their Curie temperature. Pure Magnetite, which is the dominant magnetic mineral in crustal rocks, has a Curie temperature of c. 580 °C (Hunt et al., 1995).

The Curie isotherm is associated with the deepest magnetic sources, which in turn have a strong effect on the long wavelengths of the magnetic field. The deepest magnetic sources were previously estimated directly from the magnetic field using dipole-based (e.g. Maule et al., 2005) or spectral methods (e.g. Spector and Grant, 1970; Vervelidou and Thébault, 2015; Tanaka, 2017). Both satellite and aeromagnetic data can be used for this estimation.

The lower limit for magnetic sources is often associated with the Moho boundary (e.g. Wasilewski and Mayhew, 1992). This is based on the low susceptibility of mantle xenoliths samples. However, Ferré et al. (2014) reported on mantle xenolith samples with susceptibilities between 0.001 and 0.01 SI, which offers the possibility for a significant contribution of magnetization in the upper mantle. In addition, many of these samples show strong remanent magnetization. The remanent component might plausibly be oriented parallel to the modern magnetic field, due to viscous re-orientation of the remanent magnetization above the blocking temperature (Shive, 1989). This gives a much higher “effective” susceptibility because the induced and remanent magnetization are parallel. In fact, viscous remanent magnetization can be stronger than induced magnetization (Worm, 1989). Therefore, magnetized minerals in the upper mantle might contribute more to the lithospheric magnetic field than previously assumed.

The lithospheric magnetic field not only depends on the thickness of the magnetized lithosphere, but also on susceptibility. Separating susceptibility and thickness variations is challenging (Thébault et al., 2010). To circumvent this issue, many workers have used vertically integrated susceptibility (VIS), which is defined as the product of thickness and susceptibility (Purucker et al., 1998). Each crustal column is then replaced by an equivalent dipole with a dipole strength chosen according to the VIS of the crustal column. While this approach simplifies the calculations and eliminates trade-offs between thickness and susceptibility, it also does not possess any ability to resolve the source depth with magnetic data only. Adopting volumetric sources (Baykiev et al., 2016) instead of equivalent dipoles might provide some depth sensitivity, although the inherent resolution limitations of satellite magnetic data must be considered.

In this paper we investigate how sensitive the lithospheric magnetic field is to the temperature structure of the lithosphere. We estimate the temperature distribution using a simple thermal model based on heat flow and seismological data and derive at which depth the Curie temperature of magnetite is reached. We then invert satellite magnetic data for the laterally variable susceptibility distribution inside the crust/lithosphere for three different geometries of the magnetized layer: A complete magnetic crust, a magnetic crust that is thinned due to thermal effects and a two-layer model consisting of a magnetic crust and magnetic mantle, where the temperature is below the Curie temperature of magnetite. We then compare the inversion results for the different geometries and discuss if sufficient sensitivity to the thermal structure exists.

3.2 Methods and data

3.2.1 Magnetic field observations

The Earth's magnetic field has been investigated using satellites during the last decades. Since the MAGSAT mission in 1980 and the later missions Orsted and CHAMP (Langel and Hinze, 1998; Reigber et al., 2005) the accuracy and coverage have continuously improved. ESA's Swarm mission consists of a constellation of three satellites measuring the vector magnetic field (Friis-Christensen et al., 2008; Olsen et al., 2016). By exploiting the orbit geometry of the constellation, vector gradients of some magnetic field components can be estimated, which allows more accurate modeling of the magnetic field (Olsen et al., 2016).

Observations from the CHAMP satellite have been used to derive satellite-only models of the lithospheric magnetic field (Maus et al., 2008). We use the latest available version of the data set MF 7 (Maus, 2010) at an altitude of 400 km (Figure 3.1). MF 7 contains only the wavelengths of spherical harmonic degree 16 and higher because the core field dominates the long wavelength magnetic field. The maximum formal degree of MF7 is 133, corresponding to 300 km wavelength. Recently, Olsen et al. (2017) have derived a higher resolution models of the lithospheric magnetic field based on satellite gradient data increasing the available wavelength to 250 km. In the future this model will be preferred over MF7. However, for the purpose of this work we use MF 7 (Figure 3.1), also to be compatible with the results of Hemant and Maus (2005).

For our analysis, we prefer to use satellite data and not aeromagnetic surveys or a combined model because the latter are less capable of resolving the longest wavelengths and do not provide global coverage, despite massive improvements in the last decade (Lesur et al., 2016). Furthermore, the upper crust is characterized by small-scale variations of remanent magnetization, which have a stronger effect on aeromagnetic than satellite data (e.g. Hamoudi et al., 1998).

3.2.2 Magnetic calculations

We parametrize the crust as tesseroids (spherical blocks) with a size of 2° x 2°. The nominal maximum resolution of MF 7 is 300 km, so at this model resolution we are capable of resolving all features in MF 7. Each tesseroid represents a vertical crustal column. Within each column, the susceptibility is constant, because the vertical variation of susceptibility is likely beyond the resolving power of magnetic satellite data. To magnetize the tesseroids we use effective

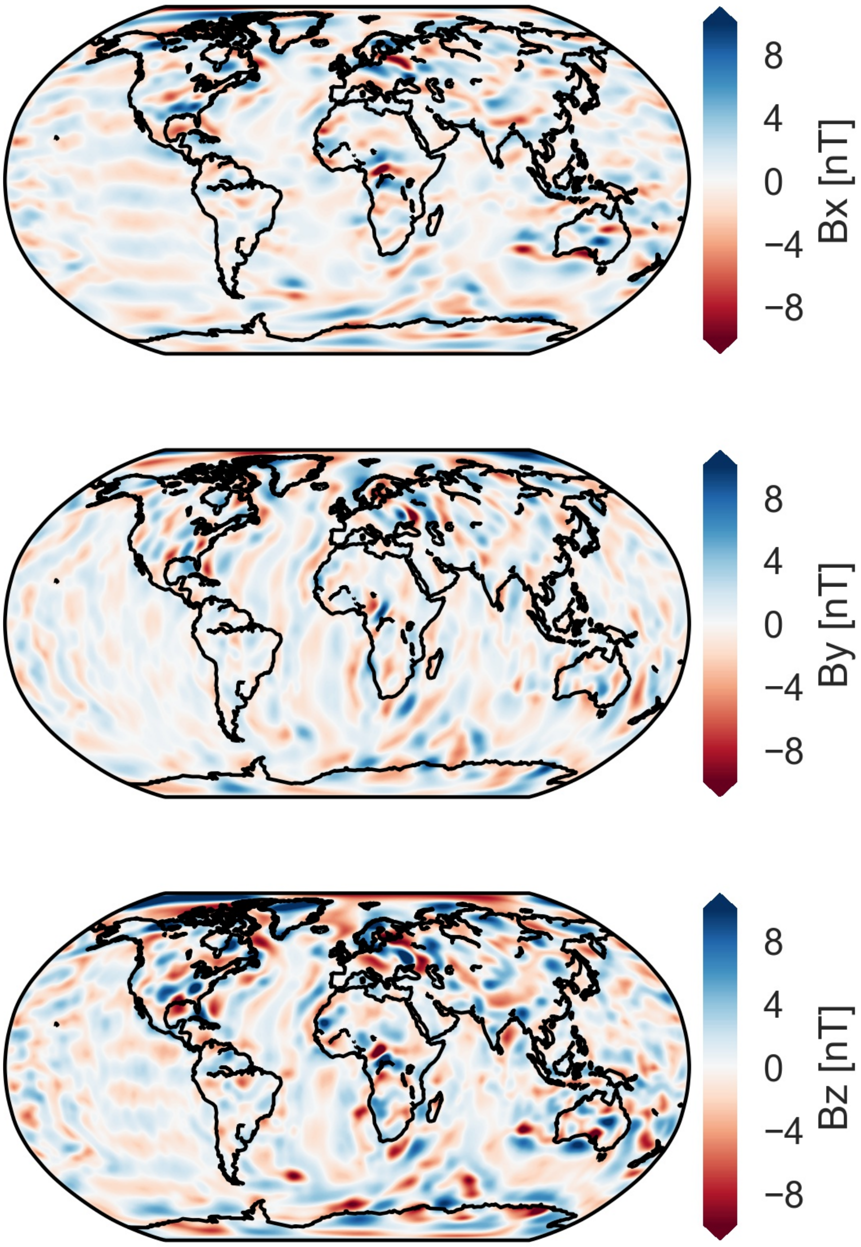

Figure 3.1: Vector components of the lithospheric magnetic field from MF 7 model at an altitude of 400 km.

susceptibility. Magnetization is then given by $\vec{M} = \vec{B_c}\chi$, where $\vec{B_c}$ is the core field and χ is the effective susceptibility. Effective susceptibility accounts for both induced and permanent magnetization parallel to the inducing field. The core field is calculated from DGRF (Thébault et al., 2015) at the date 2014/01/01 and contains up to spherical harmonic degree and order 13. A single direction and strength of the inducing field is used by evaluating the DGRF at the geometric center of each tesseroid. The magnetic field response of each tesseroid is then calculated using the software of Baykiev et al. (2016), which relies on adaptive numerical integration.

We use three different scenarios of the magnetized layer's geometries. The first scenario assumes the entire crystalline crust is magnetic. Crustal thickness is taken from the Litho1.0 model (Pasyanos et al., 2014). Scenario 2 relies on a simple thermal model based on surface heat flow, seismological lithosphere-asthenosphere (LAB) depth map and ocean floor age (see Appendix B). In the continents, we assume steady state heat conduction (e.g. Artemieva and Mooney, 2001) and adjust the total amount of heat production in each crustal column to make the heat flow data and the seismological LAB estimate compatible. In the oceans we use the half-space cooling model (Turcotte and Schubert, 2002). From the temperature distribution we obtain the Curie depth and use it as the lower boundary of magnetization in the crust (Figure 3.2a).

Scenario 3 is a two-layer model, consisting of a crustal and mantle layer. The lower boundary of magnetization is the Curie depth and we use two separate susceptibilities for the crust and mantle layer. In most continental areas the Moho boundary is deeper than the Curie isotherm (Figure 3.2b and c), whereas the Curie isotherm lies in the mantle in most oceanic regions.

Scenario #	Crustal magnetization	Mantle magnetization
1	Complete crystalline crust.	Non-magnetic
2	Crystalline crust colder than 580 °C	Non-magnetic
3	Crystalline crust colder than 580 °C	Mantle colder than 580 °C

In all scenarios the sedimentary part of the crust as given by the Litho1.0 model (Pasyanos et al., 2014) is assumed to be entirely nonmagnetic.

As a first step, we calculated the model response for constant susceptibility in the crust and mantle, respectively. This gives some limits on the possible values of magnetization and the quality of the data fit. Next, we inverted the magnetic field for the susceptibility distribution inside the crust/mantle.

3.2.3 Inversion strategy

To invert for the susceptibility distribution, we use a standard linear approach. The magnetic effect of tesseroid j on observable i (which iterates over all grid points and the three vector components) is assembled in a matrix $\mathbf{A_{ij}}$ and the unknown susceptibilities in a vector χ, so that the crustal magnetic field at all observation locations $\mathbf{B}$ caused by a susceptibility distribution χ is found as matrix product $\mathbf{B} = \mathbf{A}\chi$.

The software of Baykiev et al. (2016) operates in space domain and produces the complete, unfiltered magnetic field. However, the long wavelengths of the crustal field are masked by the core field. We thus apply a high-pass filter described by a matrix $\mathbf{F}$ to the design matrix $\mathbf{A}$, such that only degrees 16 and higher are contained in the forward calculated magnetic field. The final forward equation is then:

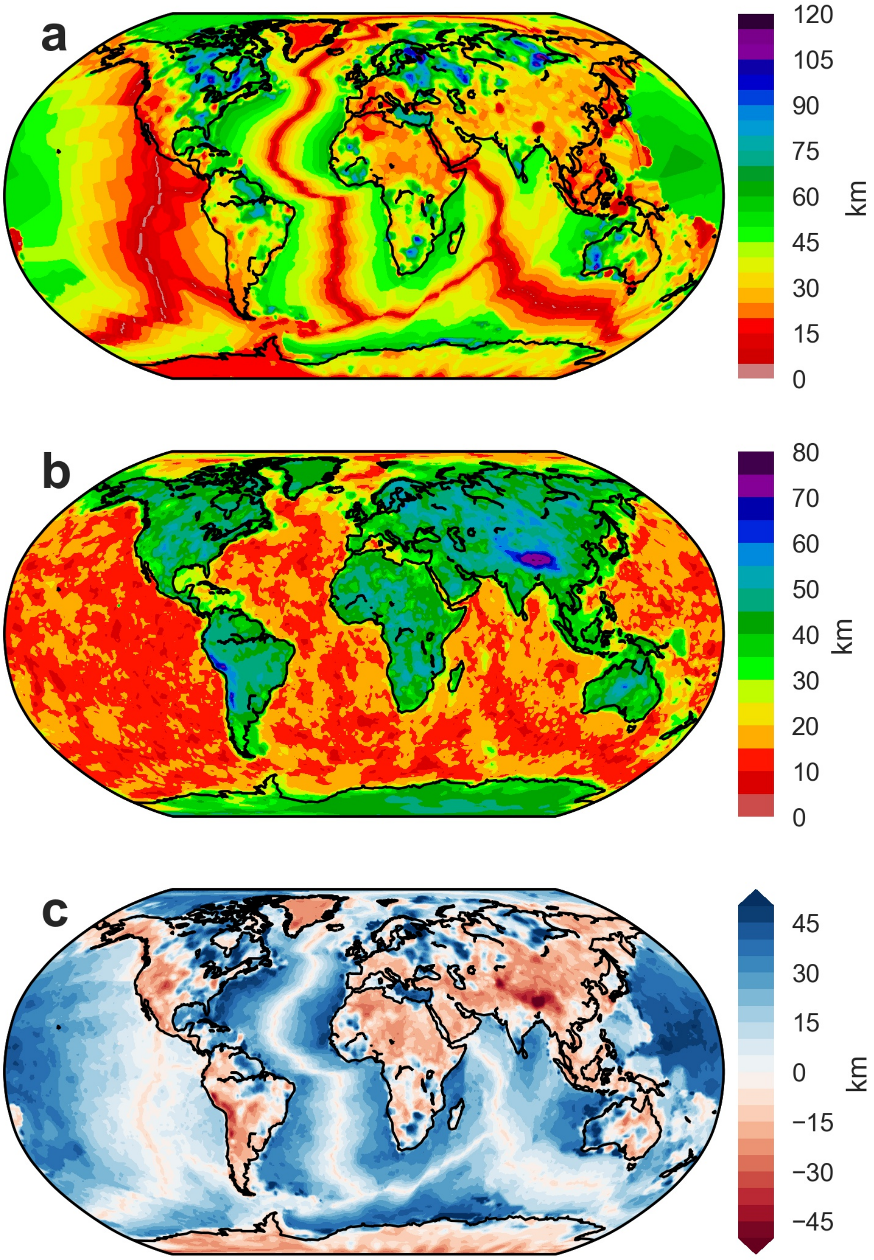

Figure 3.2: Geometries used for calculating the magnetic field. a: Curie depth estimate from thermal modeling. b: Moho depth from Litho1.0 model. c: Difference between Curie depth and Moho depth. Positive values mean the Curie isotherm lies in the mantle, possibly allowing for magnetization in the mantle. Where the difference is negative the Curie isotherm lies in the crust, so that part of the crust is non-magnetic.

$$\mathbf{B_{obs}} = \mathbf{FA}\chi = \hat{\mathbf{A}}\chi$$

The filtering matrix is derived by first defining spherical harmonic synthesis as a matrix product:

$$\mathbf{B} = \mathbf{SC}$$

Where **B** contains all vector magnetic components at all observation locations and **C** contains both the cosine and sine Gauss coefficients. Thus, spherical harmonic analsysis can be carried out by multiplying **B** with the pseudo-inverse of **S**. Now define **T** as a diagonal matrix, where the diagonal is one if the corresponding spherical harmonic coefficient is of degree 16 or higher, and zero if it is degree 15 or less. The filtering matrix can then be found thus:

$$\mathbf{F} = \mathbf{STS}^{\dagger}$$

Where $\mathbf{S}^{\dagger}$ is the pseudo-inverse of **S**.

To obtain an estimate of χ we use a standard least-squares approach (Tarantola, 2005). We assume covariance matrices for the susceptibilities $\mathbf{\Sigma}_{\chi}$ and the measured magnetic field $\mathbf{\Sigma_B}$ that reflect the scale of variation and uncertainty respectively. The a priori susceptibility distribution χ_0 is used to ensure that most of the susceptibility values are positive (Purucker et al., 1998). The least-squares estimate of χ can then be found by solving this system of linear equations:

$$(\hat{\mathbf{A}}^{\mathbf{T}}\mathbf{\Sigma}_{\mathbf{B}}^{\mathbf{-1}}\hat{\mathbf{A}} + \mathbf{\Sigma}_{\chi}^{\mathbf{-1}})\chi = \hat{\mathbf{A}}^{\mathbf{T}}\mathbf{\Sigma}_{\mathbf{B}}^{\mathbf{-1}}\mathbf{B} + \mathbf{\Sigma}_{\chi}^{\mathbf{-1}}\chi_{\mathbf{0}}$$

The covariance matrices regularize the inversion. Here, we use diagonal covariance matrices, which correspond to variance of the quantity in question. The standard deviation of the susceptibility is set to 0.01 SI and the standard deviation of the magnetic field components is set to 0.5 nT, which is less than the data residual of the MF models of about 1 nT (Maus et al., 2008). However, we chose a more optimistic standard deviation for the purpose of this sensitivity study.

The starting susceptibility distribution $\vec{\chi_0}$ is set to 0.025 SI in continental areas and to 0.04 SI in oceanic areas (Purucker et al., 1998). While this is not strictly a positivity constraint, this value is sufficient to ensure that most areas have positive susceptibility. For the mantle magnetization scenario (scenario 3), we set the mean susceptibility in the mantle to 0.0 and the standard deviation to 0.01 SI. In addition, we test the impact of the starting susceptibility using the susceptibility map of Hemant and Maus (2005).

3.3 Results

3.3.1 Constant susceptibility magnetic forward calculation

Before using inversion to get a laterally varying susceptibility, we first investigate the attainable fit with constant susceptibility. We include the effect of variation of Curie isotherm depth and allow for a magnetic mantle in cold areas and calculate the magnetic field produced for constant susceptibility in crust and mantle. We used a spectral and spatial measure of fit. The degree variance of the magnetic field at the surface is given by (Maus, 2008):

$$R_l = (l+1)\sum_{m=0}^{l}(g_l^m)^2 + (h_l^m)^2$$

Here g_l^m and h_l^m are the Gauss coefficients derived by spherical harmonical analysis.

We define the following spectral measure of misfit:

$$\text{spectral misfit} = \sqrt{\sum_{l=16}^{l_{max}} (R_l(\chi_c, \chi_m) - \overline{R_l})^2} \tag{3.1}$$

Where $R_l(\chi_c, \chi_m)$ is the degree variance for a given crustal and mantle susceptibility and $\overline{R_l}$ is the degree variance of the input data (MF 7).

In addition, we define the spatial misfit as the RMS of the residual field divided by the RMS of the observed field.

We find a broad range of small values of the spectral misfit that almost forms an elliptic arc around the origin (Figure 3.3a). The lowest misfit is found with an average crustal susceptibility of around 0.03 SI. Addition of mantle susceptibility leads to a slight decrease of spectral misfit if it is less than 0.03 SI. Thus, the mantle could plausibly contribute to the signal content.

However, susceptibility models with laterally constant susceptibility are utterly incapable of reproducing the measured field (Figure 3.3b). The relative spatial RMS is almost always larger than 1, which corresponds to a relative error of 100 %. Clearly, crustal thickness variations alone are incapable of fitting the magnetic field and laterally variable susceptibility is required.

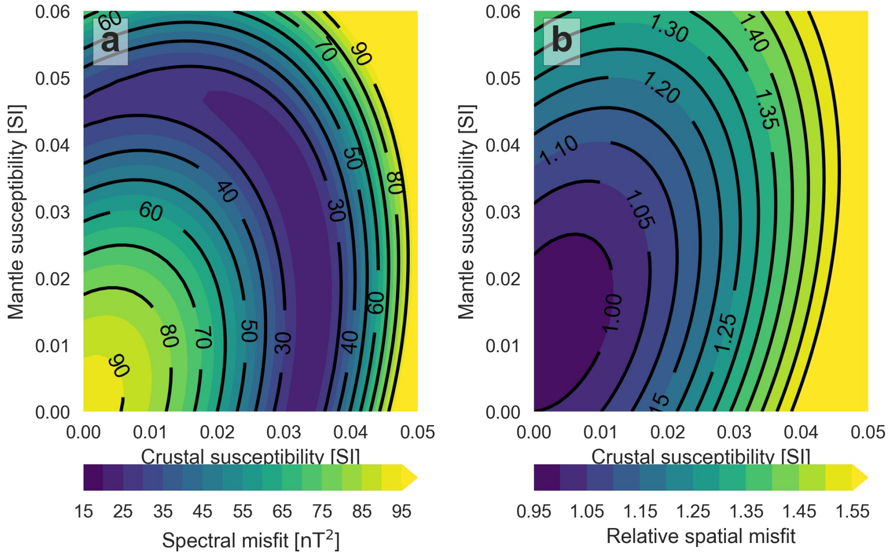

Figure 3.3: Impact of different constant susceptibility for crust and mantle using the second results of Curie depth calculation. a: Spectral misfit (eq. 3.1) b: Relative misfit in spatial domain for the vertical component.

3.3.2 Magnetic inversion

We applied the inversion strategy described above to the MF 7 data (Maus et al., 2008), as laterally constant susceptibility is not capable of explaining the measured field. All three inversion scenarios achieve a similar level of data fit. The root mean square of the residual is in all cases 0.20 nT, 0.20 nT and 0.27 nT for the North, East and Up component respectively.

The residuals display similar patterns in all scenarios. Thus, the predicted and measured fields are shown only for scenario 1 (wholly magnetized crust) (Figure 3.4) The residuals are much smaller inside the continents than over the Atlantic ocean. The patterns follow the ocean floor age trend, especially in the B_y (east) component. Thus, the residuals are probably related to remanence in the oceans, since we only use magnetization parallel to the inducing field.

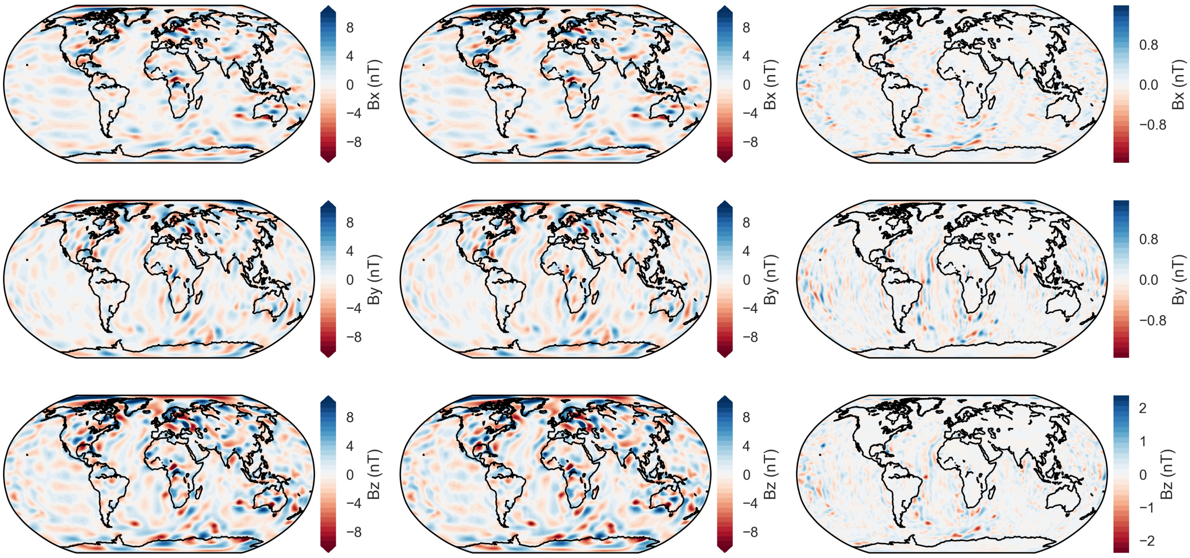

Figure 3.4: Calculated, observed and residual magnetic fields after inversion for the wholly magnetic crust scenario. The first column gives the calculated field from the susceptibility distribution in Figure 3.5. The second column gives the observed values from MF 7 and the last column gives the difference between observed and calculated fields. The rows respectively give the results for the north, east and up components.

The inverted susceptibility distribution shows a distinct separation with highly susceptible oceans and much less susceptible continents (Figure 3.5). This separation results from the apriori distribution of susceptibility imposed during the inversion of susceptibility, because the lithospheric magnetic field is not sensitive to continental-scale variation of susceptibility (Maus and Haak, 2003).

About 1.2 % of the grid points in the continents and 0.08 % of the oceans have negative susceptibility. The highest χ in the continent is found in the area of the Bangui anomaly in Central Africa, where values of about 0.1 SI are attained. The Kursk anomaly in Eastern Europe is spatially smaller but reaches similar values.

The susceptibility variations visible in the oceans is probably mostly related to remanent magnetization that the inversion attempts to fit using variable effective susceptibility. In the Atlantic Ocean, the cretaceous quiet zone (83-120 million years (Cande and Kent, 1995)) is characterized by increased susceptibility values.

A North-South striping pattern is visible in the inverted susceptibility in the South-Eastern Pacific Ocean. This striping pattern can also be seen very clearly in the B_x (North) component over the Pacific. Spectrally, it corresponds to degree 16, order 0, which is the first spherical harmonic component of MF 7 (Maus et al., 2008). Possibly this is a contribution from the core field, which becomes visible in the Pacific because the crustal field is too weak. Indeed, this region also possess a very shallow Curie depth due to the fast spreading velocity. Alternatively, it could also be long wavelength a artifact of the data.

The inversion results might depend on the starting susceptibilities χ_0. Hemant and Maus

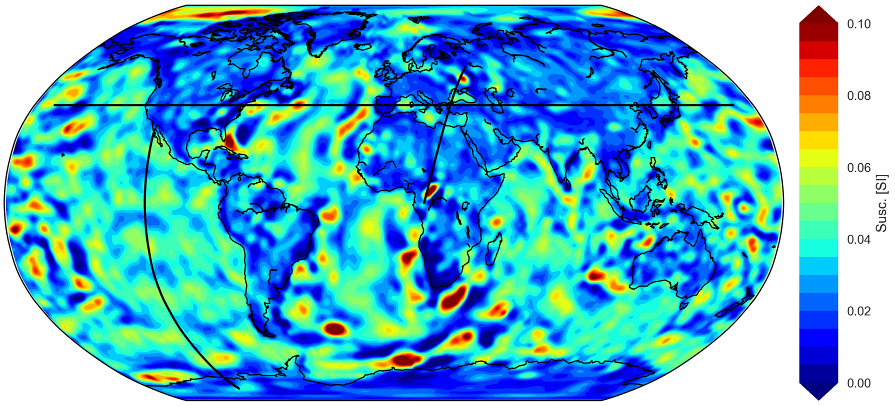

Figure 3.5: Inverted susceptibility distribution in scenario 1 (complete crystalline crust is magnetic). The black lines give the position of three profiles, which are shown in Figure 3.7 :

(2005) have derived global maps of susceptibilities based on geological provinces and typical susceptibility values for specific rock types. For the purpose of this comparison, we focus on the continental areas, because unlike our inversion approach, Hemant and Maus have explicitly considered remanence in the oceans.

Our inversion results for scenario 1 (Figure 3.6a) and the susceptibility distribution from Hemant and Maus (Figure 3.6b) agree qualitatively in the sense that many local extrema occur in both maps at the same location. For instance, the Tibetan plateau is visible as a clear local minimum in both cases. However, the shape of the features is often different, which can be partially explained by the inherent wavelength limitations of the satellite data used in the inversion. Furthermore, the amplitudes can differ substantially. In the inversion results, the continents have a systematically higher susceptibility, due to the starting susceptibility value of 0.025 SI. The mean difference between the results is 0.008 SI, and the RMS is 0.016 SI

Next, we used the Hemant and Maus map as starting susceptibility in the inversion. All other inversion parameters remain the same. The inversion residuals are slightly better than with constant starting susceptibilities. They are 0.17 nT, 0.17 nT and 0.23 nT respectively for the three vector components, which is a reduction by about 0.03 nT. The inversion outcome combines features of the previous inversion and Hemant and Maus's susceptibility map. For instance, the overall susceptibility resembles the value from Hemant and Maus, but the local patches of increased susceptibility from the previous inversion result are retained. Thus, the starting model clearly has an impact on the inversion results, but mainly in the longest wavelengths. However, this also depends on the weighting of data with compared to the starting susceptibility as controlled by the standard deviations chosen for the inversion.

The calculated Curie depth affects the inversion results relatively little, compared to a wholly magnetic crust. Both scenarios differ on average by 0.005 SI (RMS). Maximum differences of 0.03 SI occur, but they are limited to very small regions, because 90 % of the grid point have differences of less than 0.01 SI. There is no systematic shift between the two inversions (mean of differences is 0.0004 SI).

To highlight the differences, we focus on three profiles (Figure 3.7): An east-west profile at 40 °N, a north-south profile in the Pacific at 120 °W and a great circle profile connecting the Bangui and Kursk anomaly..

In the Northern Pacific and Atlantic, the inverted susceptibility is practically identical for both scenarios (Figure 3.7a). This is not surprising, since the models have the same lower boundary

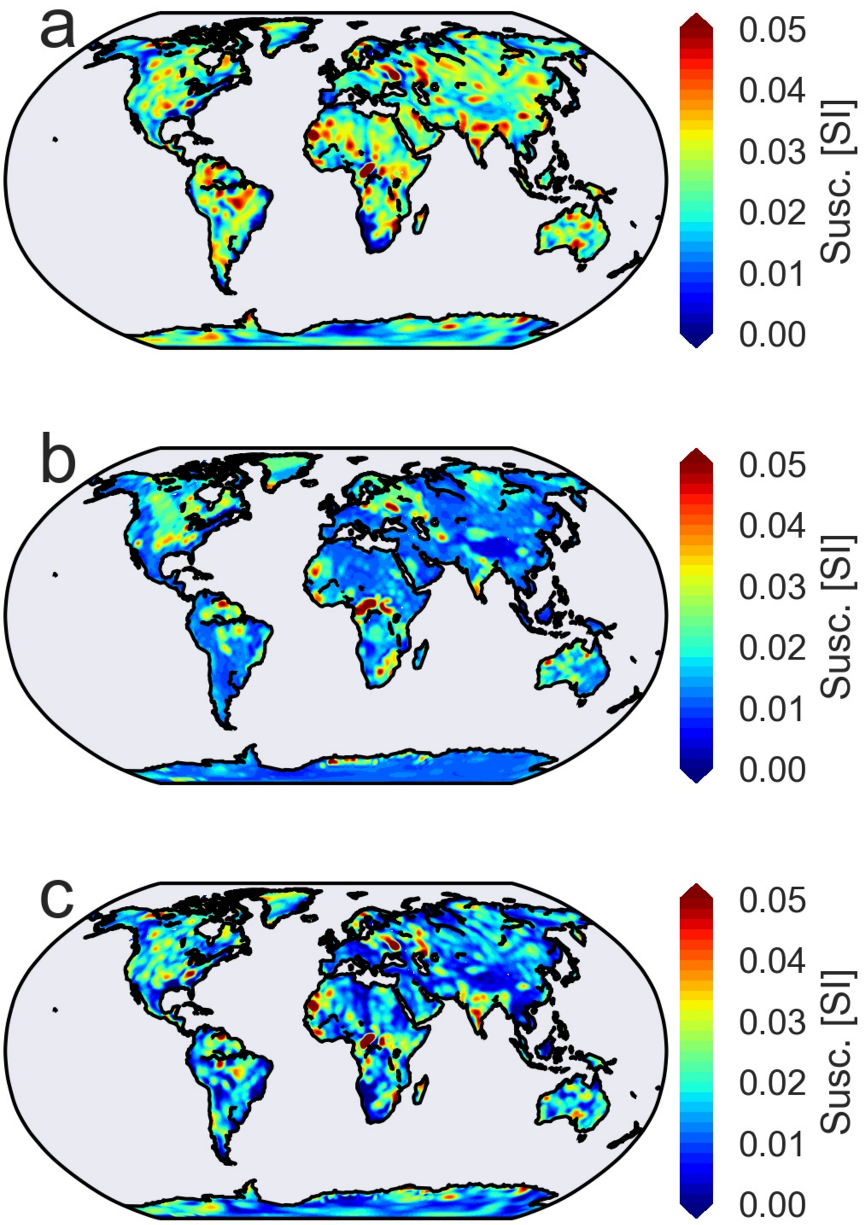

Figure 3.6: Effect of starting model on inversion results. a: Previous inversion results from scenario 1 (Figure 3.5), b: Susceptibility distribution from Hemant and Maus (2005), c: Inversion result with distribution in b as starting susceptibility.

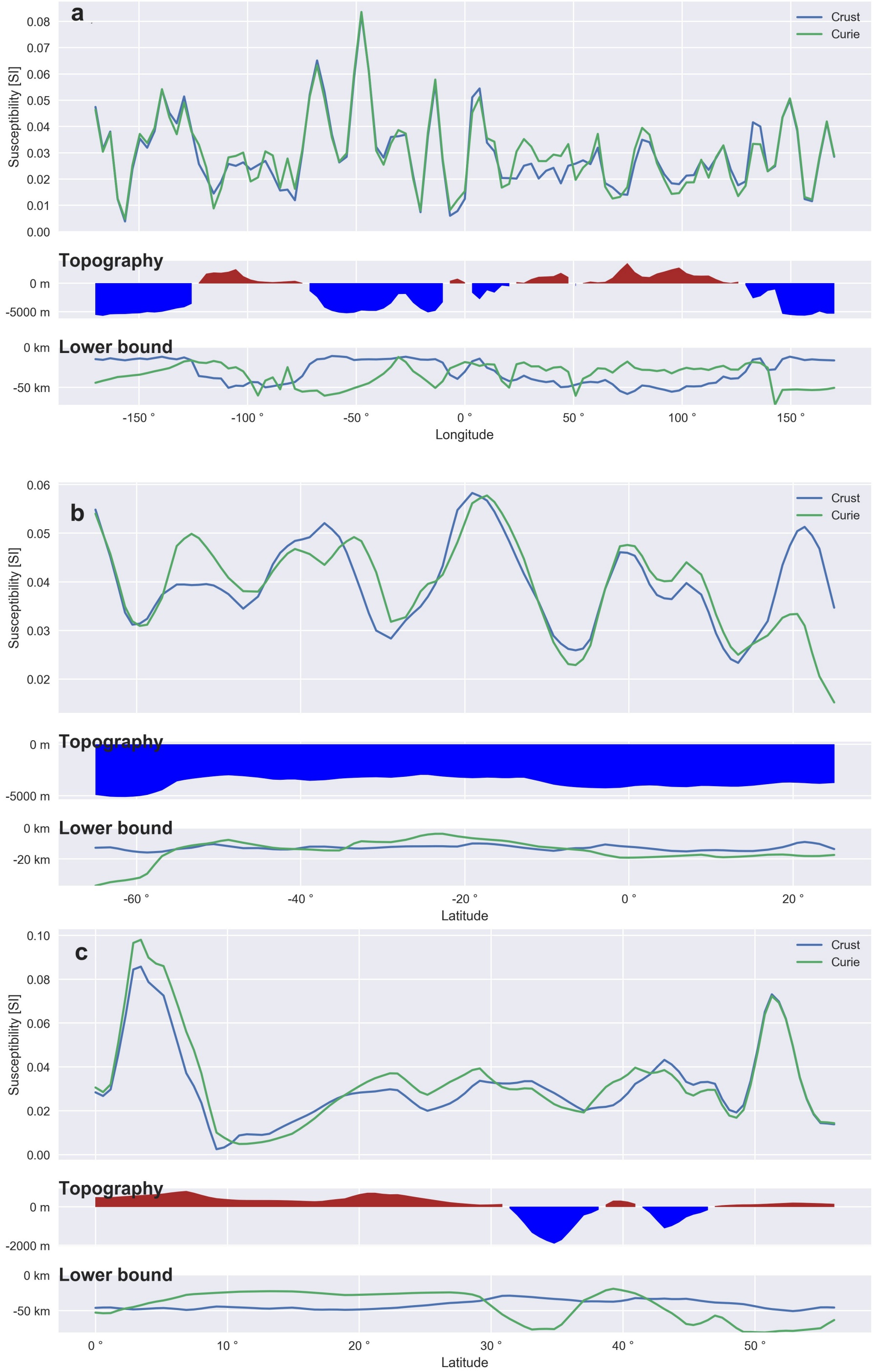

Figure 3.7: The location of the profiles is shown in Figure 3.5. a: East-west profile across the Atlantic Ocean at 40°N. b: North-south profile over the Pacific Ocean. Note the strong periodic variation of susceptibility starting at latitude 40°S which are not consistent with any feature seen in the topography, Moho depth or Curie isotherms. They are clearly visible in the magnetic field however. c: Great circle profile displayed from south to north connecting the Bangui anomaly (to the left), with the Kursk anomaly (last maximum to the right).

except near the spreading ridges, where the Curie depth is extremely shallow.

In continental areas, the Curie isotherm has a stronger effect on the magnetic field, because it mostly lies in the crust, which can lead to a much smaller magnetized layer. On this profile, differences are up to 0.02 SI between the two inversion scenarios. However, the overall position of relative extrema mostly agrees and the differences are often less than 0.01 SI. Furthermore, there is no one-to-one correlation between shallow Curie depth and differences in the inverted susceptibilities.

The north-south profile in the Southern Pacific (Figure 3.7b) shows that the striping pattern does not correlate with any feature in the bathymetry or crustal structure. In addition, the magnetized layer is extremely small south of 60 °S, because the crust is thin and the Curie isotherm lies at only a few kilometers depth.

The third profile (Figure 3.7c) shows the two most distinct anomalies in the continental crust, the Bangui and the Kursk anomaly. Both anomalies are clearly visible as localized susceptibility maxima. Thermal effects change the inverted susceptibility required to explain the Bangui anomaly in Central Africa considerably. Without thermal effects, it has a susceptibility of 0.08 SI, while the inclusion of thermal effects increases its susceptibilities to 0.1 SI. In contrast, the Kursk anomaly is not affected by thermal anomalies, because the Curie isotherm lies about 20 km below the Moho depth.

3.3.3 Mantle magnetization

Including mantle magnetization has a profound effect on the inverted susceptibility distribution, especially in the oceans (Figure 3.8). When mantle magnetization is included, the inverted susceptibility of oceanic crust is much more uniform and close to the starting value of 0.04 SI. Furthermore, the susceptibility is systematically lower than the starting value. In contrast, the mantle susceptibility is highly variable, reflecting the same patterns as seen in the crust-only inversions. The susceptibility of the oceanic mantle varies between -0.02 and 0.04 SI, but 90 % of the values are between +/- 0.01 SI.

In most continental regions, the Curie isotherm lies in the crust. Thus, only the coldest regions (Figure 3.2) permit magnetization in the mantle. One distinct feature is the increased susceptibility of 0.02 SI in the mantle underneath the Kursk anomaly.

However, the resulting mantle susceptibility depends strongly on the regularization parameters. Depending on the ratios of standard deviation of crustal and mantle susceptibility, more susceptibility variations can be placed in the mantle/crust. In the shown case, the standard deviation was set to 0.01 SI for both layers.

3.4 Discussion

3.4.1 Resolvability of thickness and susceptibility

We found that all three geometry scenarios differ in terms of the inverted susceptibility but are equally capable of explaining the observed magnetic field. However, the different inversion results might be exchangeable in the sense that they produce the same vertically integrated susceptibility (VIS). The vertically integrated susceptibility for scenarios 1 and 2 differ in terms of their long-wavelength structure only (Figure 3.10). As a result, their magnetic response is almost identical in the resolvable wavelength range above spherical harmonic degree 15.

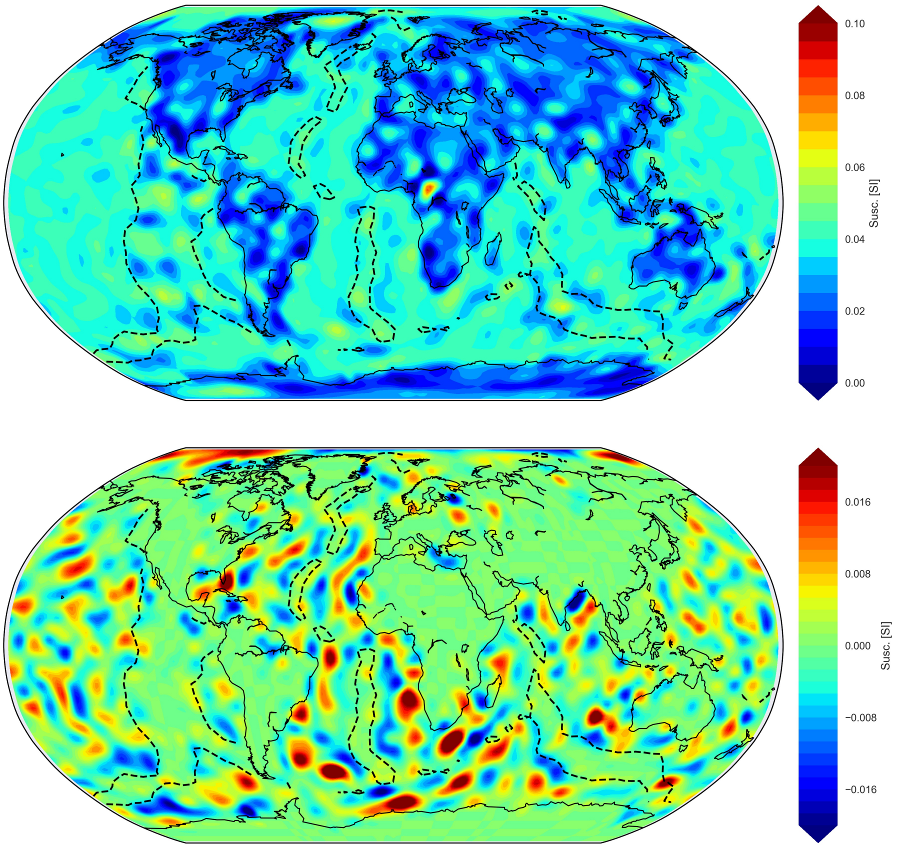

Figure 3.8: Inversion results with mantle magnetization. The upper figure gives crustal susceptibility, the lower figure gives susceptibility in the mantle. The dotted contour gives the ocean floor age of 20 million years, which is roughly the age, above which the oceanic plate has cooled significantly to allow mantle magnetization. Note that the susceptibility is set to zero in regions where the mantle is hotter than the Curie temperature. In some areas close to spreading ridges, the crust is extremely thin. As a result, there is magnetization in the mantle, despite a very shallow Curie depth.

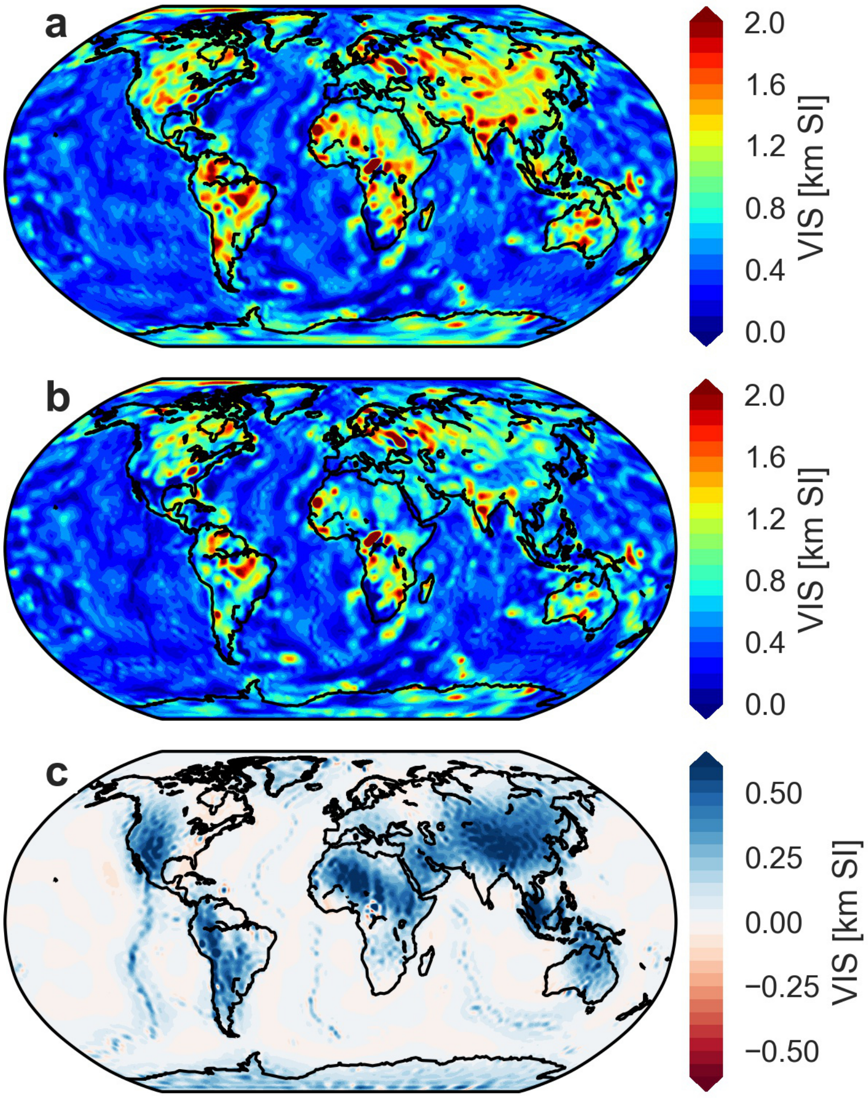

Figure 3.9: Comparison of the vertically integrated susceptibility obtained with a fully magnetic crust and temperature effects (scenarios 1 and 2). a: VIS from inversion scenario 1. b: VIS from inversion scenario 2. c: Difference between the two distributions of vertically integrated susceptibility. Note that the differences have a longer wavelength than the dominant features seen in the individual distributions. Furthermore the VIS for the model with magnetic crust (a) is systematically higher than with the inclusion of the Curie isotherm (b).

To determine to what extent variations of susceptibility can be separated from variations of thickness, we created a model consisting of a 1 km thick shell with variable susceptibility. We ran the same inversion as before, choosing the prior values for the shell's susceptibility to be the same as in scenario 1. The inversion result (Figure 3.10) agrees well with the results of Purucker et al. (1998).

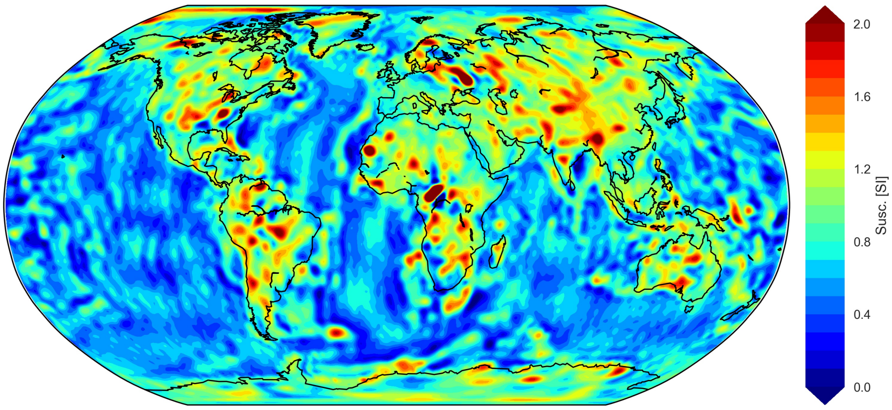

Figure 3.10: Inverted susceptibility for a homogeneous shell of 1 km thickness. The susceptibility values in SI can thus be read as vertically integrated susceptibility in SI $*$ km.

To test how good the shell approximation is, we divided the VIS values by the thickness of the crystalline crust to get a susceptibility distribution corresponding to the real crust. If vertically integrated susceptibility is a valid approximation, the forward calculation results using the thin shell and the crystalline crust should agree.

The differences between a constant and variable thickness shell are on the order of +/- 1 nT, which corresponds to about 10 % of the signal (Figure 3.11c). The RMS of the differences is 0.23 nT, 0.21 nT and 0.31 nT for the three vector components, respectively. Note that the differences are of the same magnitude as the residuals of our inversion results. A somewhat simplified interpretation of this result is that VIS explains 90 % of the modeled magnetic field and that the remaining 10 % of the signal require sources at adequate depths.

When comparing the different inversion results, the geometries are much more similar compared to the 1 km thick shell. When transferring the VIS values from scenario 2 to the geometry of scenario 1, we find that the RMS of the differences (for SH degree > 15) are 0.04 nT, 0.04 nT and 0.08 nT for the three vector components. Thus, from a VIS perspective the two models are interchangeable compared to the accuracy of the lithospheric field (1 nT) and the residuals of the inversion (0.2-0.3 nT)

Including mantle magnetization does not change the quality of the data fit. Thus, there is no necessity to invoke mantle magnetization to explain the measured field. The amount of susceptibility placed in the mantle by the inversion only depends on the specified standard deviation for the susceptibility in the crust and mantle. We set the standard deviation of mantle susceptibility to 0.01 SI which is the same as for the crust. Thus, we have equally allowed for crustal and mantle magnetization. However, if the standard deviation was reduced sufficiently, mantle magnetization would become insignificant, without affecting the data fit.

However, one might still prefer one model over the other if the susceptibility values of one model are more plausible. The inverted susceptibilities of all scenarios differ by less than 0.01 SI for 90 % of the grid points and all lie within the range of susceptibility between 0 and 0.1 SI expected for large scale average crustal magnetization (Hemant and Maus, 2005). As a result, there is no way to decide on a preferred model based on the susceptibility values, unless the susceptibility is

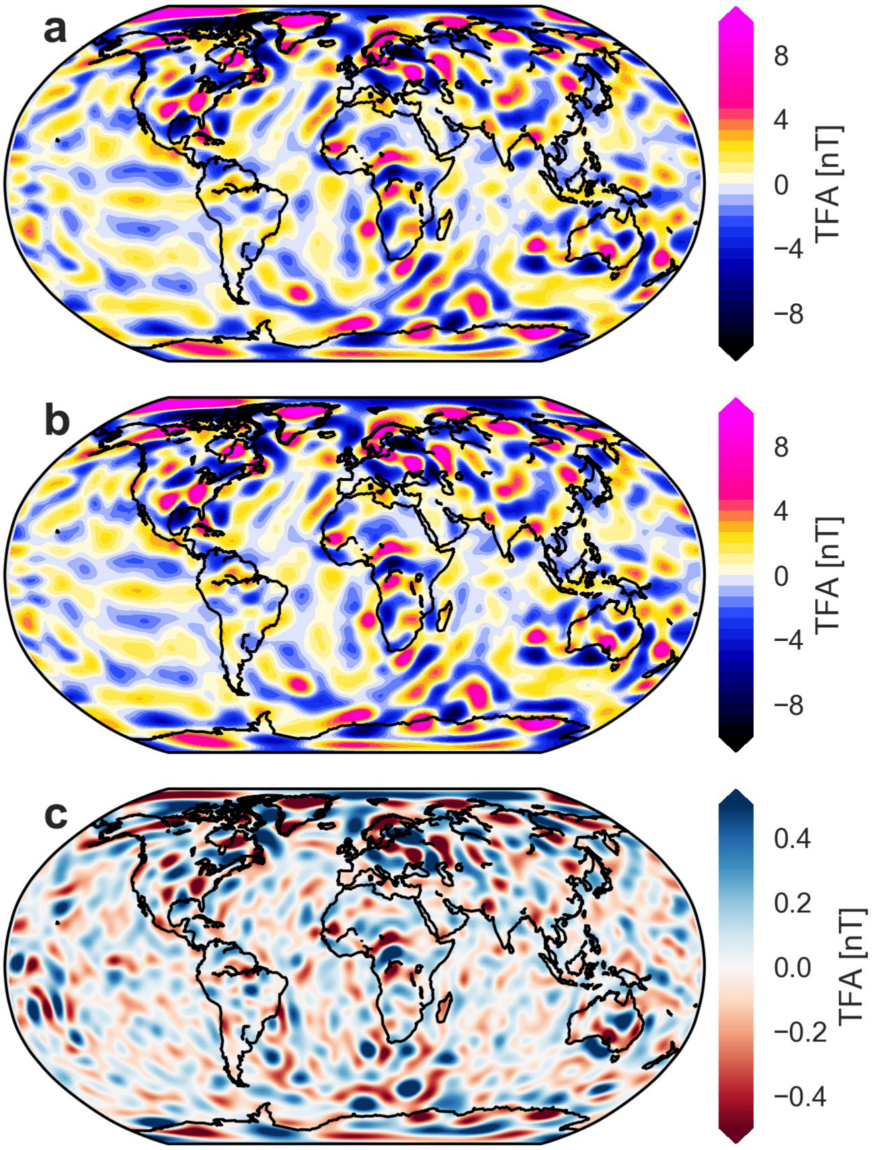

Figure 3.11: Total magnetic field anomaly calculated using the crustal geometry and an equivalent shell of 1 km thickness. a: Complete crust is magnetic. b: Equivalent shell, c: Difference between the two geometries.

known a priori within +/- 0.01 SI. Susceptibilities can easily vary by this amount even within a certain rock type (Hunt et al., 1995).

3.4.2 Impact of core field masking

Even if we concede that vertical resolution of the different sources is extremely difficult with satellite data, there remains the question of horizontal resolution. Satellite data are necessarily more capable of resolving the long wavelengths, because upward continuation acts as a low-pass filter. However, in the case of the magnetic field the core field masks the longest wavelengths of the lithospheric magnetic field so that only spherical harmonic degree 16 and higher are known. This wavelength filtering adversely affects the inversion.

In appendix A, we derive a numerical estimate of null space dimensionality by analyzing the singular value decomposition applied to the forward calculation matrix A with and without wavelength filtering at spherical harmonic degree 16. Filtering significantly increases the dimensionality of the null space. Each element from the null space is a susceptibility distribution that acts as a magnetic annihilator and produces no or only a very weak magnetic field. Magnetic annihilators can always be added to an inversion result without affecting the data fit (Gubbins et al., 2011), and can hence not be resolved by the inversion.

Next, we applied spherical harmonic analysis to elements from the null space. Spherical harmonic degrees of 10 or less (wavelength of more than 4000 km) in the susceptibility distribution are essentially not recoverable due to the masking effect of the core. Conversely, degrees 14 and higher are not strongly affected by the masking effect of the core. Degrees 11-13 are produce some magnetic response, but with suppressed amplitude. Thus, there exists some sensitivity to longer wavelengths of the susceptibility than those available for the magnetic field.

3.4.3 Composition effects on Curie temperature

Apart from temperature, rock composition is the main factor controlling rock magnetization. For example, pure Magnetite, which is regarded to be the dominant magnetic mineral in crustal rocks, has a Curie temperature of c. 580 °C (Hunt et al., 1995). However, titanium content has a systematic effect on the Magnetite-Ulvöspinel solid solution series. With increasing Ti content the Curie temperature decreases from 580° C to -153 °C for pure Ulvöspinel (Reynolds et al., 1990). Other compositional factors are the iron content and oxygen fugacity (Pilchin and Eppelbaum, 1997). Oceanic basalts displays a systematic change of natural remanent magnetization as a function of ocean floor age (Bleil and Petersen, 1983), probably related to low-temperature oxidation of titanomagnetites to titanomaghemites (Matzka et al., 2003).

Thus, even if it is possible to reliably estimate the lower boundary of magnetization, it does not necessarily represent an isotherm. In regional studies it is possible to verify the magnetic results by correlating them with additional information from other sources (e.g. Blakely, 1988). However, in frontier areas, such correlation is difficult.

3.4.4 Limitations of thermal modeling

Our thermal modeling approach outlined in Appendix A is extremely simplistic. Laterally variable heat production acts as a buffer that absorbs short wavelength variations in the heat flow data that is incompatible with the long wavelength structure as given by the seismological LAB. A more sophisticated combination of the two data sets would require information on the uncertainties of both datasets.

Artemieva et al. (2001) have argued that due to the sparse coverage of heat flow measurements meaningful derivation of thermal structure is only possible in certain well studied areas. Although additional data have been acquired since then, most of the continents are still poorly known (Davies, 2013) and a global map can only be derived by interpolating over large areas.

However, the thermal model is suitable for the purpose of this study, since it gives a reasonable range of variation of Curie depth. Even a much more refined thermal modeling approach is not expected to alter the Curie depth dramatically. As demonstrated above, even a 1 km thick shell agrees to 90 % with the more geological relevant scenarios in terms of the produced magnetic field.

3.5 Conclusions and outlook

We invert the global magnetic field model MF 7 (Maus et al., 2008) based on data from the CHAMP satellite for the susceptibility distribution inside the crust and upper mantle. We included the effect of temperature on magnetization by estimating the Curie isotherm depth using a simple thermal model.

We identify two main issues limiting the usefulness of satellite magnetic anomalies for the estimation of temperature structure. First, susceptibility and thickness variations are extremely difficult to disentangle, because 90 % of the signal at satellite height is explainable using vertically integrated susceptibility only. Second, the longest wavelengths of the susceptibility distribution are not resolvable, even if the thickness is known, due to the masking effect of the core field.

Both issues might be improved by including data from a lower altitude. However, accurate recovery of the long wavelengths from aeromagnetic surveys alone is challenging and a combination of aeromagnetic and satellite data is theoretically ideal. Alternatively, ever-improving global models of the magnetic field such as LCS-1 (Olsen et al., 2017) might permit synthesizing the data at lower altitudes.

One option to combine satellite and aeromagnetic data is to use our inversion results as starting model for a regional inversion of aeromagnetic data. This ensures that the global magnetic field is fit to an acceptable degree and also prevents far-field effects from negatively affecting the long wavelengths inside the regional study area.

Some recent mathematical developments promise to better separate the core field than wavelength filtering. Baratchart et al. (2017) have demonstrated that if the susceptibility is known in a finite area, the core and lithospheric field can be separated uniquely. Another option is to use use correlation-based modeling to separate different parts of the magnetic field (Holschneider et al., 2016). Another option is to exploit time-varying magnetic field, but this requires time series of the magnetic field of 50 years or more (Baykiev et al., 2016).

A completely different approach is to link the magnetic field with different geophysical data sets by exploiting relations between the rock properties that create the observations. Rock temperature exerts first-order control on many material properties, including density, seismic velocity and electrical conductivity, especially in the upper mantle (Afonso et al., 2008). This can be used to construct a thermodynamically consistent relation between different parameters by relying on mineral physics and petrology (Connolly, 2009).

A: Null space estimation

We can give a numerical estimate of the dimensionality of the null space by analyzing the forward calculation design matrix. Applying singular value decomposition to the design matrix A gives

$$U\Sigma V^T = A$$

Where the columns of U contain the left singular vectors of A, Σ is a diagonal matrix containing the singular values σ_k of A (ordered from large to small) and the columns of V contain the right singular vectors.

Taking only the k largest singular values gives a truncated low-rank approximation of A, called A_k. The approximate null space of the design matrix A can be found by finding a suitable k such that $(A - A_k)\chi \approx 0$. If the ℓ_2-norm is used and ϵ is an error tolerance for individual field observations and there are $3N$ observations in total, one finds a condition for the singular values:

$\sigma_{k+1} \leq \frac{(3N)^{1/2}\epsilon}{||\chi||_{\ell_2}}$

The approximate null space is spanned by all right singular vectors where the corresponding singular values is below the cut-off value σ_{k+1}. The complement of the null space is spanned by the right singular vectors larger than the cut-off value σ_{k+1}.

We use the scenario of the complete magnetic crust, but results for other scenarios did not deviate significantly. Due to memory requirement the singular value decomposition was calculated at a reduced tesseroid resolution of c. 4°, which leads to $N = 3200$. Assuming a tolerance of $\epsilon = 0.01$ nT, the threshold for σ_{k+1} is about $0.25\frac{\mathrm{nT}}{\mathrm{SI}}$. Note that the chosen tolerance is an extremely optimistic scenario, because the accuracy of the lithospheric field is on the order of 1 nT. Thus, any field with an RMS of less than 0.01 nT can be considered as effectively zero.

We applied the SVD analysis both to the unfiltered design matrix and to the filtered design matrix, where all magnetic field components below degree 16 are removed. In both cases the singular values drop off continuously in the beginning. However, roughly at $k = 1500$, the singular values of the filtered matrix start to decay more rapidly than the singular values of the unfiltered matrix. The cutoff is reached at $k = 1900$ in the filtered case and $k = 2900$ in the unfiltered case. (Figure 3.12a). Thus, the dimensionality of the null-space is 3200 - 1900 = 1300 in the filtered, and 3200 - 2900 = 300 in the unfiltered case. Therefore, filtering significantly enlarges the null space.

Each column of V corresponds to a susceptibility distribution (grid). Applying spherical harmonic analysis to all susceptibility distributions V_i contained in V gives the spherical harmonic coefficients $(C_i)_{lm}$, belonging to this distribution. From this the degree variances $(R_i)_l$ of each susceptibility distribution is derived. The spectral content of the null space and its complement can be estimated by separately averaging the degree variances $(R_i)_l$ above and below the cutoff-singular value. Next, we divided the degree variances by $2l + 1$, because the spectrum would otherwise artificially slope up, because there are $2l + 1$ spherical harmonics for a given degree l.

Directly interpreting the spectra is difficult, because the singular vectors are orthonormalized in the linear algebra sense, and not in the sense of the Hilbert space like spherical harmonics (Figure 3.12b). However, by comparing the filtered and unfiltered results, it is possible to see how filtering affects the spectral characteristics of the null space. In the unfiltered case, the null space complement has an almost flat spectrum, except for the first few spherical harmonic degrees. Thus all except the very long wavelength susceptibility structures produce a magnetic field above the threshold of 0.01 nT.

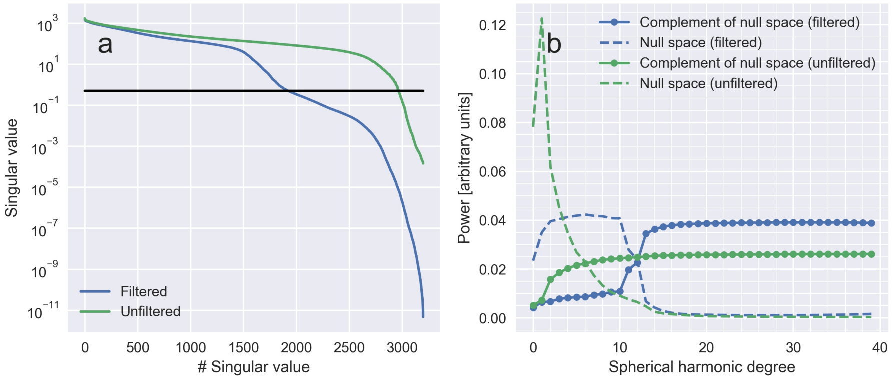

Figure 3.12: a: Singular values of the design matrix in the filtered and unfiltered case. The underlying geometry is a completely magnetic crust (scenario 1). The black line denotes the cut-off value for a tolerance of 0.01 nT per observation. b: Spherical harmonic spectrum of the null space for the filtered and unfiltered forward design matrix. Note that the units of the spectrum are arbitrary because the singular vectors are always orthonormalized.

In the filtered case, the spectrum of the null space complement is only flat above degree 13 and decreases very rapidly between degrees 10 and 13. The spectrum of the null space itself likewise increases strongly going from degree 13 to 10. This confirms that filtering of long wavelengths enlarges the null space because the wavelengths of the susceptibility below spherical harmonic degree 13 do not produce any measurable field after filtering.

Remarkably, the filtered null space does not start at degree 16, which is the filtering degree of MF 7, but at a lower degree of 13. This would indicate some sensitivity to degrees ~ 13 to 16 of the susceptibility distribution. A likely explanation for this is that the induced magnetization is a product of susceptibility and the inducing core field. Thus, the magnetization can contain spherical harmonic degrees not contained in the susceptibility.

B: Thermal modeling

We present a simple approach to model the temperature distribution inside the crust and upper mantle based on surface heat flow, a seismological Lithosphere-Asthenosphere boundary (LAB) estimate and the ocean floor age. We assume that the thermal field in the continental crust and upper mantle is stationary, so the 1-d heat conduction equation can be used

$$k\frac{\partial^2 T}{\partial z^2} = -h(z)$$

Where k is the heat conductivity and $h(z)$ is the heat productivity as a function of depth. z is positive downwards. Assuming a constant heat conductivity $k = 2.5\frac{\mathrm{W}}{\mathrm{m\,K}}$, the heat equation can be solved by simple integration. To this end, we define the first and second definite integral of h as follows

$$H(z) = \int_0^z h(z')dz'$$

and

$$\mathcal{H}(z) = \int_0^z H(z')dz'$$

Often, a parametric representation of the heat production is used, for example exponentially decaying heat production with depth (Turcotte and Schubert, 2002). However, we will not assume any specific distribution with depth, because we show below that there is no way of resolving the vertical distribution except in terms of total heat production in the entire crust when only using heat flow and LAB depth.

The heat flow q (measured positive upwards) and the temperature at an arbitrary depth are given by

$$q(z) = q_0 - H(z)$$

and

$$T(z) = T_0 + \frac{q_0\, z - \mathcal{H}(z)}{k}$$

Where q_0 is surface heat flow and T_0 is surface temperature. We assume furthermore that there is no heat production in the mantle. At the Moho boundary, heat flow and temperature must be continuous if crust and mantle are in thermal equilibrium. Thus, the temperature structure in the mantle is given by

$$T(z) = T(M) + \frac{q(M)}{k}(z - M)$$

where M is the Moho boundary depth.

In principle, $h(z)$ and q_0 are enough to estimate the temperature structure inside the lithosphere (Artemieva and Mooney, 2001).. However, these parameters are poorly known. Additional information can be included if temperature is known at another depth. Here, we will assume that the LAB depth L, corresponding to the temperature of $T_L = 1315$ °C is known (from seismology). This gives the following constraint:

$$T_L = T(M) + \frac{q(M)}{k}(L - M) = T_0 + \frac{q_0\, L + MH(M) - \mathcal{H}(M)}{k}$$

This equation is linear in terms of the heat production parameters and can be re-arranged to constrain the total heat production in each column $H(M)$ and second integral of heat production $\mathcal{H}(M)$. However, $H(M)$ and $\mathcal{H}(M)$ cannot be resolved independently, but only in relation to each other.

$H(M)$ should always be less than the total surface heat flow q_0. Otherwise, heat would be flowing from the crust into the mantle. The theoretically possible values of $\mathcal{H}(M)$ are bracketed by $H(M)$. The minimum value of $\mathcal{H}$ is zero (all heat produced at the bottom of the crust) and the maximum value is $\mathcal{H}(M) = H(M)M$ (all heat is produced at the surface), If heat production is vertically uniform, $\mathcal{H}(M) = \frac{1}{2}H(M)$. Thus, we chose to parametrize $\mathcal{H}$ in terms of a ratio η (between 0 and 1), such that $\mathcal{H}(M) = \eta H(M)M$. We constrain ourselves to the case $\eta \geq \frac{1}{2}$, which means that heat production tends to be concentrated more near the surface.

Note that other common models of heat production can also be formulated in terms of η. For the case of exponential decaying heat production $h(z) = h_0 e^{-z/D}$, one finds that $\eta = \frac{1}{1-e^{-M/D}} - \frac{D}{M} \approx 1 - \frac{D}{M}$, where the approximation only holds if $M \gg D$.

For the oceans, we applied the half-space cooling model based on the ocean floor age (Turcotte and Schubert, 2002). Temperature inside the lithosphere is given by:

$$T(z,t) = (T_a - T_0)\,\mathrm{erf}(\frac{z}{2\sqrt{\kappa t}}) + T_0$$

Where T_a is the mantle temperature, κ is the heat diffusivity, T_0 is the temperature at the surface and t is the ocean floor age. This equation can be easily re-arranged to give the depth belonging to the a specified temperature T_c.

$$z = 2\sqrt{\kappa t}\,\mathrm{erf}^{-1}\left(\frac{T_c - T_0}{T_a - T_0}\right)$$

3.5.1 Data

We use the heat flow data by Davies (2013). We re-gridded to a resolution of 1° to make them compatible with the seismological data used.

The Litho1.0 model (Pasyanos et al., 2014) is based on the Crust1.0 model (Laske et al., 2013) and can be seen as an update of Crust1.0. The Lithosphere-Asthenosphere boundary (LAB) is estimated using inversion of surface wave data. We interpret the seismological LAB given by the Litho1.0 model as the isotherm of 1315° C. For the oceanic parts of the world we use the Ocean floor age grid by Müller et al. (2008) to estimate the temperature structure. All input data are shown in Figure 3.13.

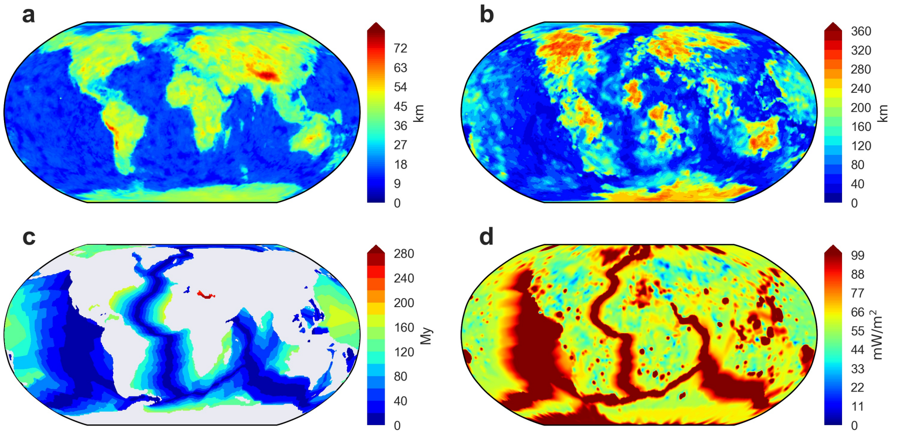

Figure 3.13: Input data used to construct the thermal model. a: Depth to Moho boundary of Litho1.0 model. b: Depth to LAB from Litho1.0 model. c: Ocean floor age map of Müller et al. (2013) d: Heat flow from compilation of Davies (2013).

We apply the approach above to estimate the heat production required to reconcile heat flow and Litho1.0 (Figure 3.14). At each point a maximum and minimum value of heat production is derived, depending on whether the heat is concentrated near the top ($\eta = 1$) or spread evenly throughout the crust ($\eta = 0.5$). There are some points where this approach fails and the heat flow data and lithospheric thickness are irreconcilable. These are regions with a thin lithosphere, but low surface heat flow. In these regions either the data are incorrect or the assumptions of our thermal model are not applicable.

3.5.2 Effect of heat production on Curie depth

Next, we created two different thermal models to show how heat productivity affects the Curie isotherm of magnetite (c. 580 °C). In the first, a constant heat productivity of $h = 1.0 \cdot 10^{-6}\mathrm{W/m^3}$

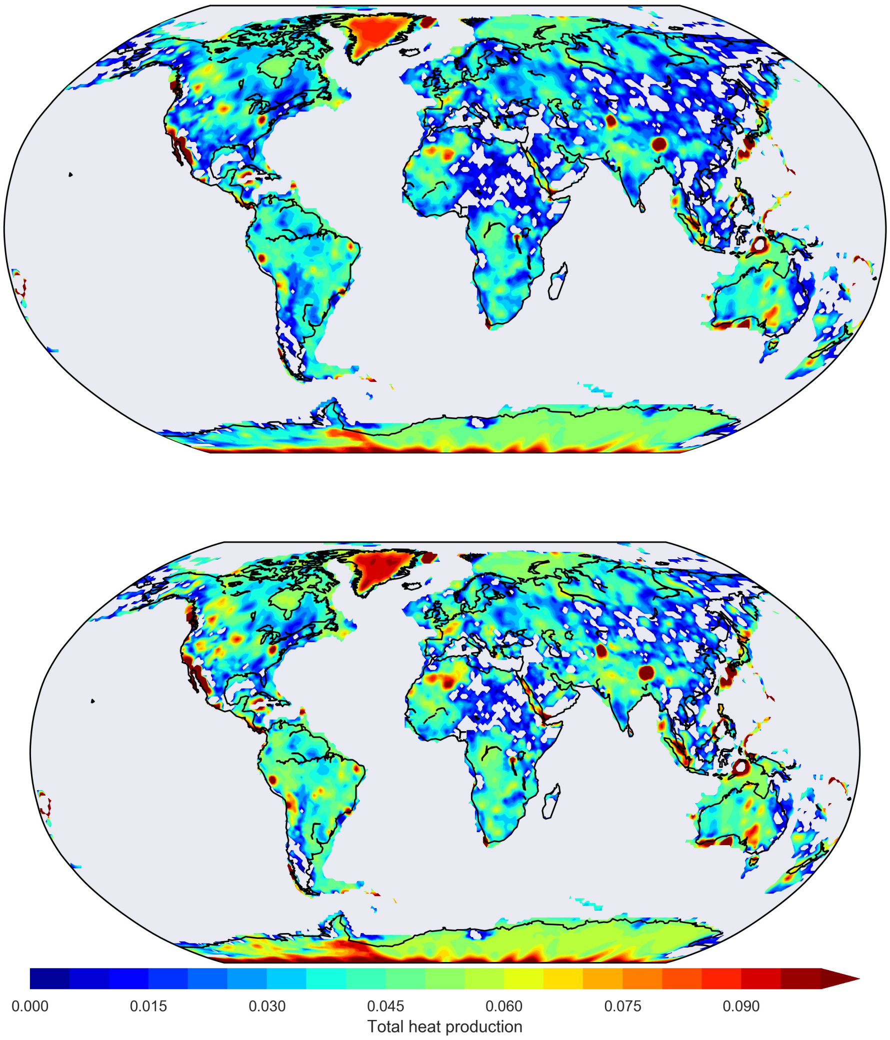

Figure 3.14: Range of possible values of total heat production required to reconcile heat flow data and the seismological LAB. The upper figure shows the minimum allowed of heat production and the bottom shows the maximum. Note that there are some points where no solution could be found, because a negative heat production would be required.

was assumed everywhere in the crust. In this case, the total heat production is proportional to overall crustal thickness. The second model assumes vertically uniform heat production with lateral variation derived using the method above. In the main text we have always used the second model of heat production. Both models use the same parameters for the thermal modeling in the oceans.

The two maps of Curie isotherm show similar patterns, but the absolute depths differ on average by about 11 km (Figure 3.15a,b). Furthermore, if the heat production varies laterally, the Curie depth becomes systematically more shallow in the continents and Curie depths of more than 75 or less than 20 km appear less often (Figure 3.15c). Additional variability of the Curie depth could be introduced, by changing the value of η.

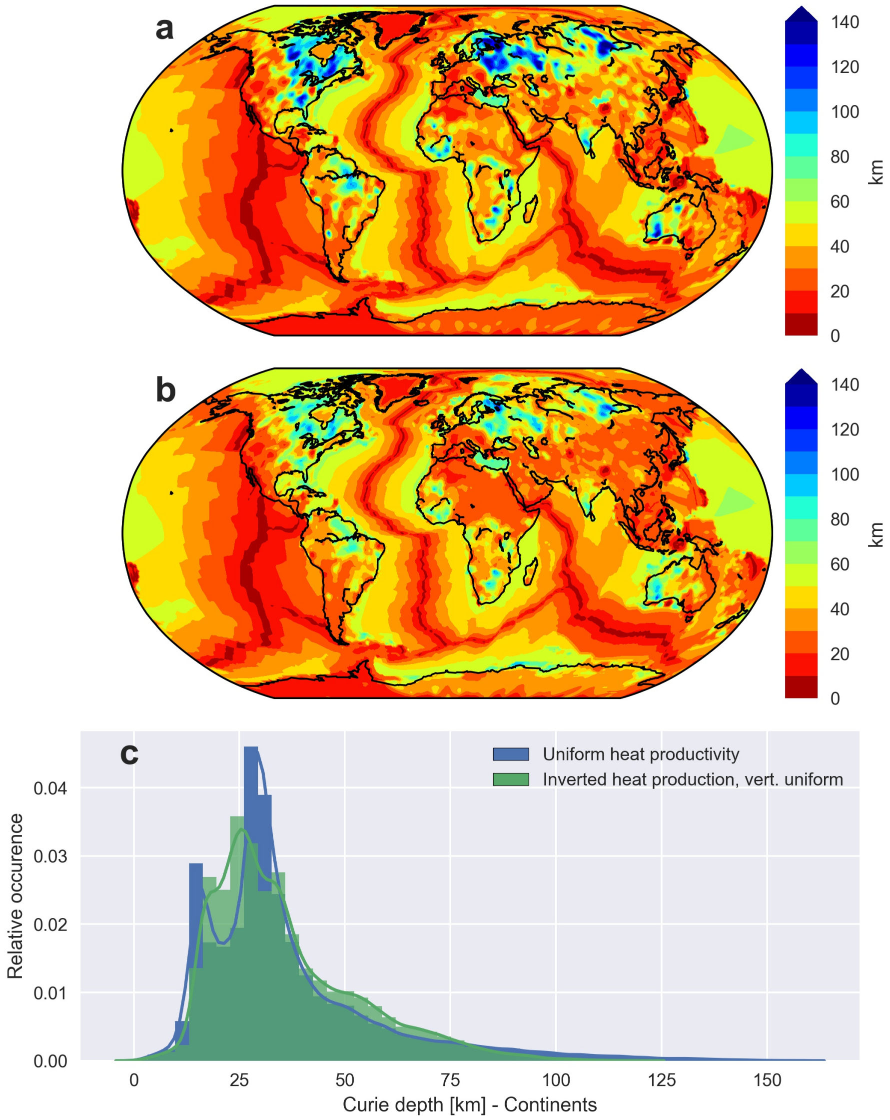

Figure 3.15: Distributions of Curie depth for different assumptions of heat production. Note that the Curie depth in the oceans is derived from the half-space cooling model and is the same in all three figures. a: Lateral and vertical uniform heat production. b: Vertical uniform heat production with total heat production inverted from heat flow and seismological LAB. c: Statistical distribution of Curie depths obtained using different models. Only continental points are included, because the two models are identical in the oceans.

Acknowledgments

We would like to thank Kumar Hemant Singh for providing the VIS map. This study was supported by the Deutsche Forschungsgemeinschaft (DFG) Priority Program DynamicEarth, SPP 1788 and the ESA STSE “Swarm magnetic gradients for lithospheric modeling”.

4 Paper III - Global crustal thickness and velocity structure from geostatistical analysis of seismic data

Wolfgang Szwillus[1], Juan Carlos Afonso[2], Jörg Ebbing[1], Walter D. Mooney[3]

[1] Department of Geosciences, Kiel University

[2] Macquarie University, Sydney

[3] USGS, Menlo Park

Submitted to *Journal of Geophysical Research: Solid Earth*

4.1 Introduction

Global models that describe the seismic velocity and density structure of the crust play a crucial role for many geophysical applications. For instance, crustal models are needed to determine crustal corrections for mantle seismic tomography (Schaeffer and Lebedev, 2015) or for calculating residual mantle gravity and the crustal contribution to isostatic topography (Kaban et al., 1999). Gravity inversions often use crustal models as additional constraints or as a starting model (e.g. Reguzzoni et al., 2013).

The most common global crustal models are those based primarily on seismic information. Specifically, active (controlled-source) seismic studies are carried out along regional-scale profiles and use refracted or reflected seismic waves to determine thickness and velocity of the crustal layers (Prodehl and Mooney, 2012). Although active source studies provide the most accurate information, the use of so-called passive methods (based on seismic energy generated by earthquakes) has become routine in the past decade. Receiver functions (Kind et al., 2012) and the ambient noise technique (Shapiro and Campillo, 2004) are two of the most popular techniques.

Creating a two-dimensional model of crustal properties from seismic information given along profiles or at individual points requires some form of interpolation. Most authors use geological a priori information during the interpolation process to ensure that a geologically plausible result is obtained. For instance, the Crust5.1 model (Mooney et al., 1998) uses tectonic domains based on broad age categories ranging from Archean to Proterozoic (Artemieva and Mooney, 2001) and crustal domains (Oceans, Rifts, Orogen, etc.). Crustal thickness from active seismology is then interpolated separately within each category. The P-wave velocity structure is described by three layers, although not all layers need to exist at every point. Each category is assigned an archetypical vertical velocity profile that is assigned to all points of this category. In this way no explicit interpolation of velocity information from individual points is necessary. The more recent models Crust2.0 (Bassin et al., 2000) and Crust1.0 (Laske et al., 2013) have essentially used the same idea. Continental-scale models that use variations of this technique have also been developed (e.g. Chulick et al., 2013; Stolk et al., 2013; Tesauro et al., 2014).

A major drawback of including predefined domains is that the uncertainty of the resulting crustal model is hard, if not impossible, to estimate. In fact, the plausible appearance of the model might be misleading in areas with sparse data coverage. Furthermore, the boundaries of the domains might be uncertain as well and contribute to total crustal model uncertainty.

In this work, we use a geostatistical interpolation approach to derive a crustal model and quantitative estimates of its uncertainty from a database of seismic points (from the USGS GSC database). The only additional data is the age of the ocean floor, which is needed to separate continental and oceanic domains. Since we do not include other additional data sets (e.g. gravity anomalies), we expect our results to be ambiguous in regions where data coverage is nil or sparse. However, we do not consider this a drawback of the method but an illustration of the limitations of the current database. The interpolation itself is a form of non-stationary kriging, which allows lateral variation of covariance parameters expected on a global scale (Risser and Calder, 2017).

We apply this method to the latest version of the Global Seismic Catalogue (GSC) collected by the USGS (Mooney, 2015). The GSC contains mainly active seismic measurements. However, the method is completely general and can be applied to any compilation of seismic points. We obtain global estimates of the depth to the Moho boundary and average crustal P-wave velocity on a global grid with 1° spacing. Estimates of crustal model uncertainty can be helpful for any geophysical application that relies on such models.

To demonstrate some of the advantages of having representative estimates of uncertainty, we apply our results to the estimation of residual topography. Here, we consider residual topography as that part of the surface topography which is not explained by crustal structure (Kaban et al., 1999; Molnar et al., 2015). By propagating the uncertainties of the crustal model to residual topography, we can derive a global map of residual topography uncertainty.

4.2 Data

4.2.1 Crustal database

Each data record in the USGS GSC (Global Seismic Catalogue) database provides information about crustal structure at a geographical location based on seismic surveys. From this database we select 6340 seismic observations that have detected the Moho boundary. The vertical structure at each point is described as layers of different types (sediments, crust or mantle) and their associated thickness and seismic velocities. Note that information about S-wave velocities is very sparse and hence we limit ourselves to P-wave velocities. In addition, the type of survey (mostly refraction) and the year when the results were published are given (Figure 4.1).

About 75 % of oceanic and 60 % of continental points are reversed refraction surveys (Figure 4.2), which are generally considered to be the most reliable method (Mooney et al., 1998). About 7 % of the oceanic and 28 % of the continental points are from Unknown/Other methods. This encompasses some passive seismological methods like earthquake models, earthquake tomography and/or waveform models.

The data were acquired over a period from 1950 to 2005. The distribution for oceanic points shows two distinct maxima in the 1970s and after 1995. There are relatively few data points acquired between 1980 and 1995. The distribution is more uniform for continental points, but also shows an increase in the 1970s and late 1990s (Figure 4.2).

Many crustal models give velocities and thicknesses separately for the upper, middle and lower crust. Interpolating this layering from the seismic points is challenging, because the number of layers is not consistent over all points. Thus, it is not clear how the layers at different points relate to each other. To circumvent this issue, we interpolate the average velocity of the crystalline crust directly.

The average velocity of the crystalline crust is calculated as the weighted average:

$$\overline{v} = \frac{\sum_i v_i t_i}{\sum_i t_i}$$

where t_i is the thickness of layer i, and v_i is its velocity.

After calculating the average velocities, we find that the database contained some points with very low average crustal velocities. At these points, low-velocity sediments are incorrectly classified as part of the crystalline crust in the database. Following Chulick et al. (2013), all points with average velocities of 5.8 km/s or less are excluded from the interpolation. In addition, we remove all oceanic points with average velocities ≤ 6.4 km/s. The oceanic crust consists mainly of gabbroic rocks with velocities ≥ 6.6 km/s, underlying a slower (~5 km/s) layer of basalts and sheeted dikes (White et al., 1992). Therefore, points with anomalously low velocities likely represent local features such as fracture zones that are not representative of large-scale oceanic crust. After removing these anomalous sites, 5197 points are retained.

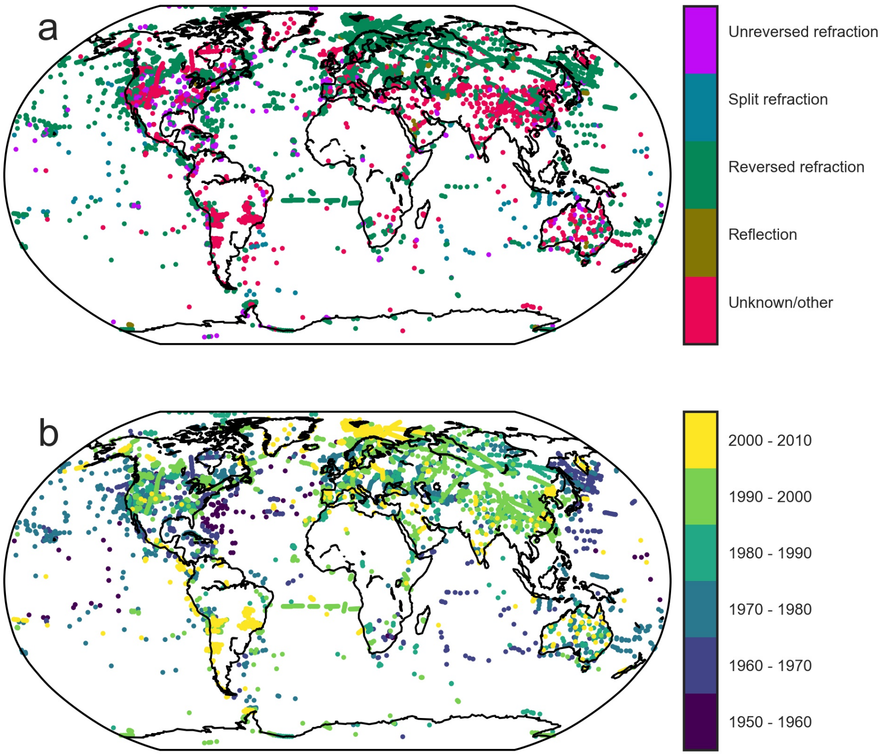

Figure 4.1: Moho depth points from the USGS GSC database: a: Seismic method used to determine Moho depth and P-wave velocity, b: Year of publication of results. Unknown/other contains points derived from earthquake models, waveform models, tomography, sonobuoy measurements and the time term method.

4.2.2 Ocean floor age map

During the interpolation procedure, oceans and continents are treated separately. This is necessary due to their different nature and evolution. The thickness and velocity structure

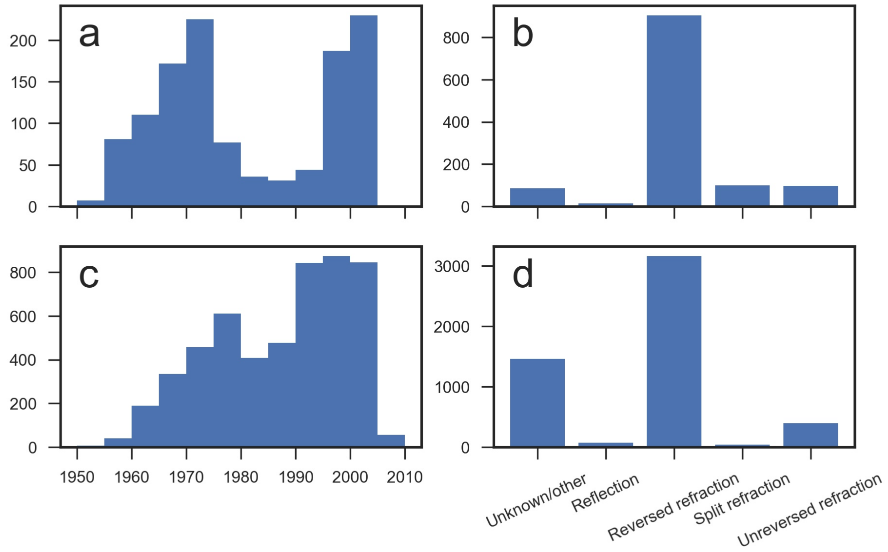

Figure 4.2: Frequency of used seismic methods and year of acquisition. a and b show the frequency of acquisition year and seismic method used for oceanic points, c and d give the frequencies for continental points.

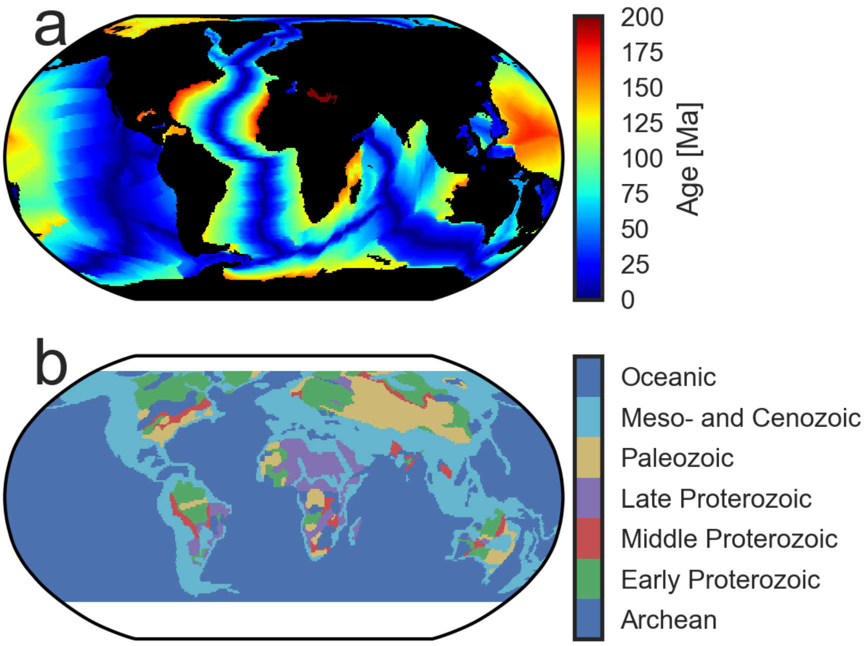

Figure 4.3: a: Ocean floor age (Müller et al., 2008) b: Tectonothermal/ Tectonic age categories from Goutorbe et al. (2011). Note that the tectonic age map stops at 60° latitude.

of unaltered oceanic crust mainly depends on the conditions at the mid-ocean ridge where it was formed (cf. McKenzie and Bickle, 1988). By contrast, the properties of continental crust are the result of a much longer evolution. Accordingly, continental and oceanic domains are defined based on the ocean floor age (Müller et al., 2008) (Figure 4.3a). A point, where an ocean floor age is defined is considered oceanic, whereas a point with no ocean floor age is considered continental. In total there are 554 oceanic and 4643 continental points.

This binary ocean-continent classification results in a sharp transition between these domains at the continental margin. For simplicity, we use the binary ocean-continent transition, but we acknowledge that the real crustal structure is certainly more complicated in the transitional region.

The ocean-continent classification employed here is a simplified version of the tectonic classification used in other crustal models. For example, the thermo-tectonic age map from Mooney et al. (1998) (as digitized by Goutorbe et al. (2011)) introduces additional spatial structures in the continental domains (Figure 4.3b). At the same time, the average crustal thickness for the different continental ages is quite similar, ranging from ~36 to 41 km (Figure 4.4). By comparison, the standard deviation within age categories is more than 6 km. Thus, giving up tectonic regionalization leads only to an acceptable loss of information.

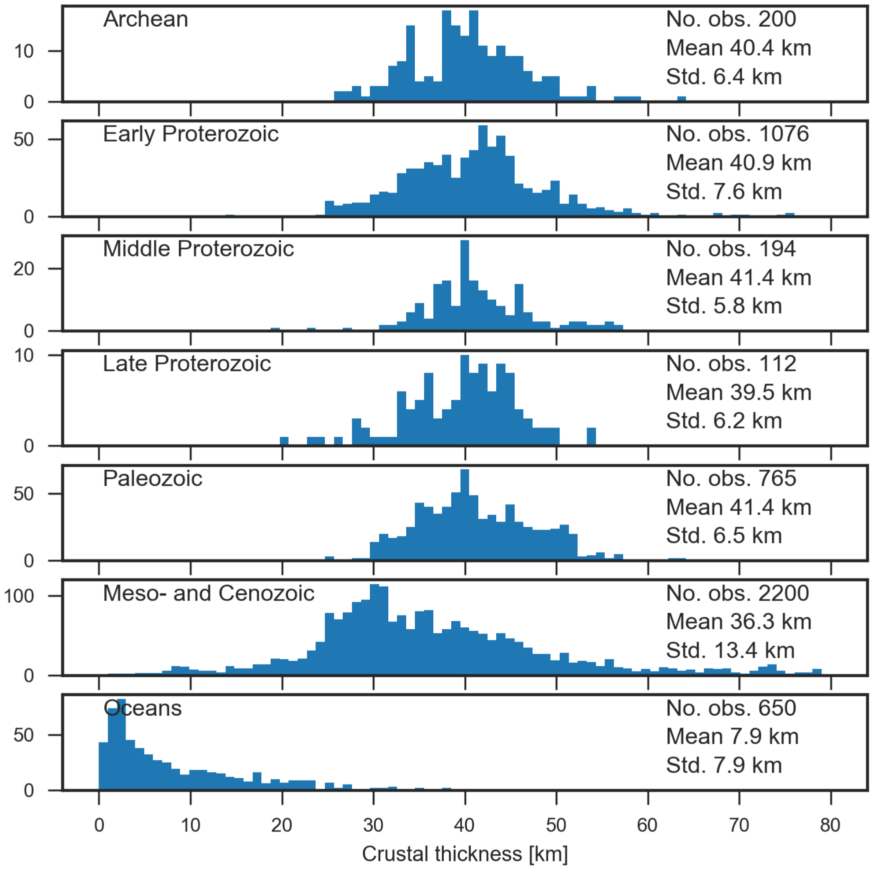

Figure 4.4: Crustal thickness of points from the GSC database falling in each of the thermo-tectonic age categories. Note that we only use points that lie inside the thermo-tectonic age map of Goutorbe et al. (2011).

4.3 Geostatistical analysis and interpolation

We use a non-stationary kriging method based on Gaussian fields to interpolate the Moho depth and average P-wave velocity. Let $Z(x)$ be the spatial field of interest (either the Moho depth or average crustal P-wave velocity), where x is the two-dimensional location in terms of geographical coordinates. Z is separated into three parts (Cressie, 2015):

$$Z(x) = \mu(x) + Y(x) + \epsilon(x)$$

Here, μ is a deterministic component (trend), which describes the first order spatial structure, Y is a Gaussian field and ϵ describes the spatially independent errors, which are also assumed to be Gaussian.

The trend component μ depends only on the classification as oceanic or continental:

$$\mu(x) = \begin{cases} \mu_{ocean} & \text{if } x \text{ in ocean domain} \\ \mu_{continent} & \text{if } x \text{ in continent domain} \end{cases}$$

The Gaussian field Y has zero mean and is described by a covariance function $C(x_1, x_2; \vec{\theta})$, which in turn depends on a set of parameters $\vec{\theta}$. Often, it is assumed that C is stationary, which implies that C only depends on the distance between x_1 and x_2.

$$\text{Cov}(x_1, x_2) = C(x_1, x_2) = C_s(d),$$

where $d = ||x_1 - x_2||$ denotes the **great-circle-distance** between the points x_1 and x_2.

Equivalently, the covariance structure can be described by the semivariogram function $\gamma(d)$, which is given by:

$$\gamma(d) = C(0) - C(d),$$

where $C(0)$ denotes the variance of the quantity in question.

The covariance structure is typically described by at least 3 parameters: the nugget τ^2 , the sill σ^2 and the range ρ. The nugget is often due to unsampled small-scale variability or noise in the data. In our case, the nugget effect is contained in ϵ. The sill describes the overall variance of the quantity of interest and the range is a measure of correlation distance. Points that are more distant than the range are assumed to be uncorrelated (or very nearly).

We chose to use a spherical covariance function, because it has a number of beneficial properties. It is an admissible covariance function on the sphere using great-circle distances (Guinness and Fuentes, 2016), it does not impose differentiability on the interpolated field and it is computationally efficient, because it gives exactly zero covariance for points which are more separated than the range, which allows us to use sparse covariance matrices. The covariance function is given by (Cressie, 2015)

$$C_s(d; \tau^2, \sigma^2, \rho) = \begin{cases} \sigma^2 \left(1 - \frac{3d}{2\rho} + \frac{1}{2}\left(\frac{d}{\rho}\right)^3\right) + \delta_{d,0}\tau^2 & \text{if } d < \rho \\ 0 & \text{if } d \geq \rho \end{cases} \tag{4.1}$$

Here $\delta_{d,0}$ is zero unless $d = 0$.

4.3.1 Non-stationarity

Stationarity might be a poor approximation, because the parameters of the covariance function represent a measure of roughness and scale of variability (Cressie, 2015). Thus, it is likely that different geologic regions are characterized by different parameters of the covariance function and their variation needs to be considered.

To test how much the semi-variogram properties vary spatially, we calculate semi-variograms for different portions of North America. We use North America, because the database contains enough points to permit calculating informative semi-variograms even for smaller subsets of data. The North American continent is split into four quadrants, where each quadrant has a size of 60° x 30°. Inside each quadrant we estimate an empirical semi-variogram using only continental points (Figure 4.5). We find pronounced differences between the quadrants, which motivates us to use a non-stationary approach.

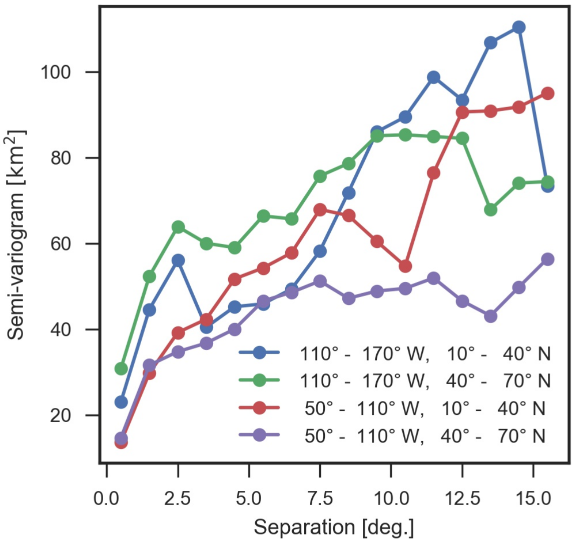

Figure 4.5: Empirical semi-variogram of Moho depth calculated of four quadrants of North America. The bin size for estimating the semi-variogram is 1° great-circle distance. The four quadrants show differences in terms of nugget, sill and range.

In the non-stationary case, the covariance function explicitly depends on the spatial locations x_1 and x_2. One way to derive a non-stationary covariance function from a stationary covariance function is to allow τ^2, σ^2 and ρ to vary spatially (cf. Paciorek and Schervish M. J., 2004). They are then given by 3 functions $\tau^2(x)$, $\sigma^2(x)$ and $\rho(x)$, which leads to a non-stationary covariance function:

$$C^{NS}(x_1, x_2; \theta) = \frac{\rho_{\text{eff}}}{\sqrt{\rho(x_1)\rho(x_2)}} C_s(||x_1 - x_2||; \tau(x_1)^2, \sqrt{\sigma(x_1)\sigma(x_2)}, \rho_{\text{eff}}) \tag{4.2}$$

Where

$\rho_{\text{eff}} = \sqrt{2}\frac{\rho(x_1)\rho(x_2)}{\sqrt{\rho(x_1)^2+\rho(x_2)^2}}$

is the effective range between the points x_1 and x_2.

To estimate the functions τ^2, σ^2 and ρ from data points, the functions need to be given a discrete representation. We use the mixture approach of Risser and Calder (2017), in which K master locations b_k with corresponding parameters τ_k^2, σ_k^2 and ρ_k are defined. The parameter at an arbitrary location is then found as a weighted sum of the mixture components

$\sigma(x)^2 = \sum_{k=1}^{K} w_k(x)\sigma_k^2,$

where each weight function w_k decreases exponentially with great-circle distance from the master location:

$$w_k(x) \propto \exp \frac{-||x - b_k||^2}{2\lambda_w}$$

under the condition that $\sum_{k=1}^{K} w_k(x) = 1$.

The parameter λ_w controls the behavior of the interpolation. If λ_w is small, the interpolation effectively becomes nearest neighbor interpolation, whereas large values of λ_w lead to a smooth distribution of the parameter. Thus, the parameter needs to be adjusted to the scale of the domain in question and the desired interpolation behavior.

We chose the discrete mixture approach, because it is directly applicable to a spherical geometry and can effectively adapt to a very heterogeneous distribution of points. Furthermore, the locations of the master locations can be determined using clustering techniques semi-automatically, which means no additional information is needed.

4.3.1.1 Clustering analysis

Let the value of Z be known at N locations x_i. We first construct the master locations using agglomerative clustering of the points based on their great-circle distance (see Appendix A for algorithm details). The clustering will then be used to estimate the nugget, sill and range from the data. How the points are assigned to clusters is controlled by the choice of the number of clusters to create. Here, we separately cluster oceanic and continental points.

We tested different numbers of clusters to optimize the spatial radius and the number of points in each cluster. The radius of a cluster is defined here as half the maximum distance of any two points that are a member of that cluster. We consider the following factors to choose the number of clusters. First, the radii of the clusters should be large enough to detect the long-range correlation between points. Second, there should be enough clusters to detect spatial variations of the covariance parameters. Third, the radius of all clusters should be similar, because otherwise the clusters might appear to possess very different ranges ρ, even though this is just a result of the different cluster radius. Last, each cluster should contain enough points to reliably estimate the covariance parameters. Fulfilling these conditions is not possible everywhere and some compromise needs to be made.

After numerous tests, we decided to use 35 oceanic and 50 continental clusters. For the oceans this is mainly motivated by preventing clusters containing only a single point (Figure 4.6). For both oceanic and continental domains, we try to choose the number of clusters so that the mean cluster radius is c. 10 ° (c. 1100 km). The interpolation parameter λ_w is set to $(1100\text{km})^2$, which is the squared mean cluster radius.

The master locations b_k are defined as the spherical average location of all members of each cluster. This is calculated by converting the location of the points to a Earth Centered Earth Fixed (ECEF) Cartesian coordinate system, taking the average location of all points in 3D space and then projecting back to the surface of the Earth. This avoids issues that arise when taking the average of points with positive and negative longitude or near the poles. The spatial

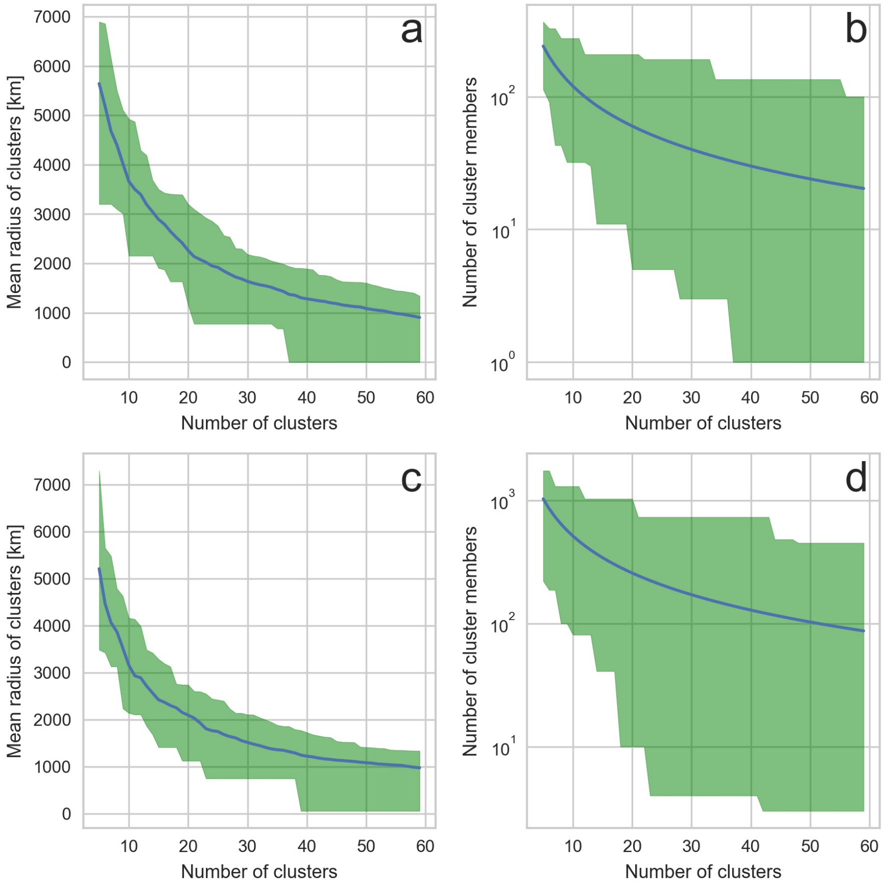

Figure 4.6: Results of clustering. Radius of clusters (a,c) and number of points in each cluster (b,d). The shaded area gives the range over all clusters, while the blue line gives the mean. Oceanic areas (a,b) were treated separately from continents (c,d). For oceans we selected 35 clusters, because this prevents single point clusters from appearing (b) and gives a mean radius of about 1100 km (corresponding to 10°). For the continents we used 50 clusters, because this gives the desired mean radius of 1100 km.

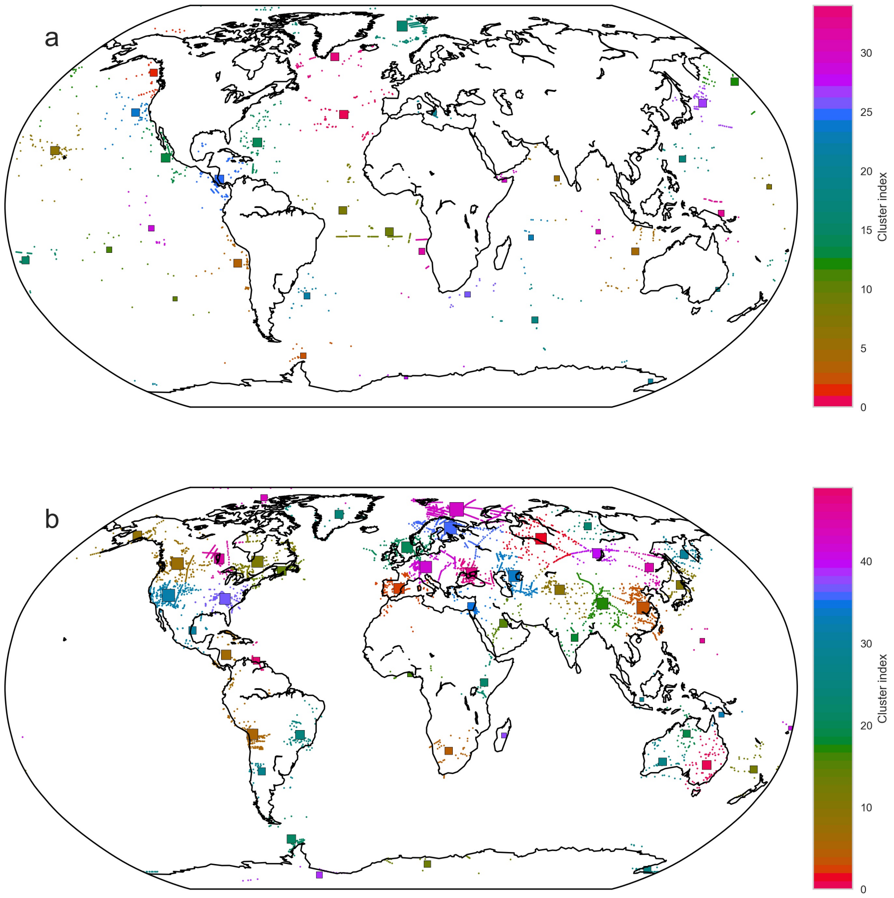

Figure 4.7: Results of clustering for oceanic and continental points. Small dots mark the location of points from the database. Colors indicate to which cluster each point belongs. The centroid of each cluster is indicated as a larger square, where the size is proportional to the number of points contained in each cluster. a: Oceanic clusters. b: Continental clusters

structure of the clusters (Figure 4.7) is reasonably adapted to the heterogeneous distribution of the data points.

4.3.1.2 Inference of covariance parameters

The non-stationary covariance function depends on the covariance parameters at each master location. These parameters have to be estimated from the points of the database. By assuming that the Gaussian field Y is stationary in the neighborhood of each master location b_k, the parameters can be efficiently estimated by using the local maximum likelihood estimation technique (Tibshirani and Hastie, 1987).

Let $\hat{x_i}$ be the points belonging to an arbitrary cluster k and let $\hat{Z_i} = Z(\hat{x_i})$ be the values at each location. Since Z is a Gaussian field, $\hat{Z_i}$ are elements of a multivariate normal distribution with covariance matrix Σ and mean μ. The entries of the covariance matrix are given by evaluating the stationary covariance function $C_s(d; \tau_k^2, \sigma_k^2, \rho_k)$, for all pairs of points inside the cluster (eq. 4.1).

The mean value inside each cluster is estimated using the sample mean of all cluster members. Our method is therefore slightly biased, because the Z_i are correlated.

The log-likelihood function is the probability density function of a multivariate Gaussian distribution with constant mean μ and covariance matrix Σ:

$$\log L(\vec{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\det\Sigma(\vec{\theta_k}) - \frac{1}{2}(\vec{Z}-\mu)^T\Sigma^{-1}(\vec{\theta_k})(\vec{Z}-\mu),$$

where n is the number of points in $\vec{Z}$.

We numerically maximize the value of the log-likelihood to obtain $\vec{\theta}$ using the L-BFGS-B algorithm (Zhu et al., 1997; Byrd et al., 1995) as implemented in the scipy package (Jones et al., 2001).

When using the likelihood function directly, we find that the sill and nugget are systematically underestimated in data-sparse regions. This would lead to overly optimistic uncertainty estimates in regions with little data points. To somewhat alleviate this issue, we use a Bayesian approach and put independent inverse-gamma priors on the σ_k^2, τ_k^2 and ρ_k. The inverse gamma distribution has a log probability density function given by

$$\log f(x; \alpha, \beta) \propto -(\alpha+1)\log x - \frac{x}{\beta}.$$

The mean and standard deviation are respectively:

Mean $= \frac{\beta}{\alpha-1}$ Std.Dev. $= \frac{\beta}{(\alpha-1)\sqrt{\alpha-2}}$

The hyperparameters α and β where chosen for each covariance parameter such that the prior distribution agrees with a specified mean and standard deviation as given in Table 4.1. Note that we set the standard deviation of all properties to be at least as large as the mean value, thus the prior is relatively weak and only has an effect on clusters with very little data points. The nugget parameter is chosen according to previous accuracy estimates of active seismic studies (Christensen and Mooney, 1995; Mooney et al., 1998). The sill parameter choice is guided by the standard deviation of Moho depth and P-wave velocity in the database. For the range, we use an average value of 10° in agreement with the mean size of the clusters.

The modified posteriori likelihood becomes, after application of Bayes theorem (Tarantola, 2005):

$$\log \tilde{L}(\theta) = \log L(\theta) + \sum_{\phi \in \{\tau_k^2, \sigma_k^2, \rho_k\}} \log f(\phi; \alpha_\phi, \beta_\phi)$$

$\tilde{L}(\theta)$ is optimized using the L-BFGS-B algorithm as above.

Table 4.1: Parameters of the prior we put on the covariance parameters. We assign different parameters to continental and oceanic points for the Moho depth to account for the overall much less variable nature of Moho depth in the oceans. The parameters of the velocity prior are identical for ocean and continent.

	Moho		P-wave-velocity	
Parameter (cont.)	Mean	Std. Dev.	Mean	Std.Dev.
Nugget τ^2	$(4.0\,\mathrm{km})^2$	$(4.0\,\mathrm{km})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$
Sill σ^2	$(80.0\,\mathrm{km})^2$	$(80.0\,\mathrm{km})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$
Range ρ	10°	10°	10°	10°
Parameter (ocean)				
Nugget τ^2	$(1.0\,\mathrm{km})^2$	$(3.0\,\mathrm{km})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$
Sill σ^2	$(40.0\,\mathrm{km})^2$	$(40.0\,\mathrm{km})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$	$(0.2\,\mathrm{km\,s^{-1}})^2$
Range ρ	10°	10°	10°	10°

4.3.1.3 Interpolation and prediction

After the sill, nugget and range belonging to all clusters have been determined, the value of Z can be predicted at an arbitrary set of locations x_j^*. Let $\vec{Z}$ be the vector of known values, and let $\vec{Z}^*$ be the vector of unknown values at locations x_j^*. The complete covariance between two arbitrary points x_i and x_j is found by combining equations 4.1 and 4.2:

$$C(x_i, x_j; \vec{\theta}) = \begin{cases} \sigma(x_i)\,\sigma(x_j)\,\frac{\rho_{\mathrm{eff}}(x_i,x_j)}{\sqrt{\rho(x_i)\rho(x_j)}} \left(1 - \frac{3||x_i - x_j||}{2\rho_{eff}} + \frac{1}{2}\left(\frac{||x_i - x_j||}{\rho_{eff}}\right)^3\right) + \delta_{ij}\tau(x_i)^2 & \text{if } d < \rho_{\mathrm{eff}} \\ 0 & \text{if } d \geq \rho_{\mathrm{eff}} \end{cases}$$

Evaluating the equation above for all pairs of points (both observation and prediction points) gives the complete covariance matrix. The complete covariance matrix naturally separates into three block matrices: Σ_{ZZ} is the covariance between the observed points, Σ_{ZZ^*} is the cross-covariance between the observed and prediction locations and $\Sigma_{Z^*Z^*}$ is the covariance between prediction locations.

The prediction of Z^* is then given by (Risser and Calder, 2017)

$$\vec{Z}^* = \mu(x^*) + \Sigma_{ZZ^*}\Sigma_{ZZ}^{-1}(\vec{Z} - \vec{\mu}),$$

where $\vec{\mu}$ contains the mean value at each of the prediction locations, depending on its classification as oceanic or continental.

The posteriori covariance matrix of the prediction error at all prediction locations is:

$$\Sigma_{Z^*|Z} = \Sigma_{Z^*Z^*} - \Sigma_{ZZ^*}\Sigma_{ZZ}^{-1}\Sigma_{ZZ^*}^T \tag{4.3}$$

The diagonal of $\Sigma_{Z^*|Z}$ contains the error variances, i.e. the square of the estimated uncertainty. The off-diagonal entries of $\Sigma_{Z^*|Z}$ correspond to the error covariances at different locations.

4.3.1.4 Outlier removal and duplicate selection

We test all points for consistency with neighboring points using a hold-one-out cross validation approach similar to Stolk et al. (2013). The quantity of interest is predicted at each observation location, but without using that observation itself. If the difference between the measured and predicted value is larger than twice the estimated interpolation error and the difference is larger than 5 km (Moho depth) or 0.2 km/s (P-wave velocity), it is flagged as an outlier and not used in prediction. In addition, there are some locations with duplicate values. Here, we predict the value at the location of the duplicates, without using the duplicate values and choose the value that is closest to the interpolated value at that location.

The estimation of parameters σ_k^2, τ_k^2 and ρ_k is then repeated using only the chosen points. Next, a second round of outlier detection follows, which only uses the points not flagged as outliers and the chosen duplicate points. This gives the final set of points for the interpolation.

Apart from detecting outliers, the hold-out procedure also provides insights into the accuracy of the data set and our method. The difference between an observation z_j and the prediction $\hat{z}_j$ if that observation is held out is called the *prediction error*. The root mean squared prediction error gives an estimate of the overall accuracy of the prediction procedure and the internal consistency of the database:

$$RMSPE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(z_j - \hat{z}_j)^2}$$

Furthermore, we can also test the error estimation. Specifically, the prediction error should be similar to the estimated error (from equation 4.3). To quantify this, we use the deviation ratio

$$\mathrm{DR}_j = \left|\frac{z_j - \hat{z}j}{\hat{\sigma}_j}\right|,$$

where $\hat{\sigma}_j$ is the estimated error corresponding to $\hat{z}_j$. Thus, the optimal value of the deviation ratio should be one. If the deviation ratio is larger than one, the error has been underestimated, whereas values below one indicate an overestimation of error.

Note that under the assumption of normality, the deviation ratio should follow a folded normal distribution (Leone et al., 1961) with standard deviation one. If this assumption holds, the median deviation ratio should be $\frac{\mathrm{erf}^{-1}(\frac{1}{2})}{\sqrt{2}} \approx 0.67$. Around 68 % of the points should have a deviation ratio of one or less (i.e. the error is over-estimated) and 95 % of the points should have a deviation ratio of less than two.

4.4 Results

4.4.1 Outlier detection

Applying the hold-one-out cross-validation technique to the database, 310 points were flagged as outliers for interpolating the Moho depth. Another 321 points were removed due to the selection of outliers. This leaves 4566 points for the interpolation of Moho depth, corresponding to about 14 % removed points. Applying the same approach to P-wave velocity leads to a similar number of removed points: 255 points as outliers and 316 due to the selection of duplicate points.

4.4.2 Covariance parameters

The spatial distribution of the three covariance parameters nugget, sill and range has a strong impact on the behavior of the interpolation, particularly the estimated uncertainties. The nugget τ^2 at a point determines the minimum uncertainty at that location and the maximum uncertainty tends towards the sum of the nugget and the sill $\tau^2 + \sigma^2$. The range ρ controls how quickly the uncertainty increases with distance from the nearest prediction point. Moreover, larger values of ρ lead to a smoother interpolated field (Chiles and Delfiner, 1999).

The estimated covariance parameters of the Moho depth points display a strong dichotomy between oceans and continents (Figure 4.8). Nuggets and sills are higher in the continents than in the oceans, whereas the range tends to be higher in the oceans. This is due to the relatively uniform crustal thickness in the oceans, which does not change significantly, unless affected by magmatism (e.g. White et al., 1992).

The nugget and sill in the oceans are relatively uniform with values of around 1 km^2 and $9\,\text{km}^2$, respectively. The range in the oceans is more variable. It is more than 20° in the Western Pacific and Northern Atlantic but can be less than 5° in the Southern Atlantic and Indian Ocean. Overall the range tends to be higher in well-studied regions compared to more sparsely surveyed areas. This suggests that the observed differences do not reflect differences in oceanic crustal structure but differences in sampling density.

The continents are also characterized by relatively uniform nugget, sill and range. The most notable anomaly is the high sill over the Andes, with values up to $(12\,km)^2$. Note that other major orogens such as the Himalaya do not show this increased sill. In addition, the Andes have a reduced range. The different covariance parameters in the Andes might be an artifact caused by the large north-south extent of the Andes. Thus, in north-south directions the sill should be less than in East-West directions and vice versa for the range. It might hence be possible that the anomalous values in the Andes are the result of the inability of an isotropic covariance function to describe an anisotropic structure. As a result of the high sill, the uncertainty estimates in South America might be too high.

The covariance function parameters of the average P-wave velocity differ substantially from those of the Moho depth (Figure 4.9). Over both continents and oceans, sill and nugget are often of the same magnitude. This implies that small-scale variations or measurement errors are of similar importance as large-scale structural features. As a result, direct interpolation of the velocity might not be possible in all regions.

Over the continents nugget and sill are relatively uniform. Exceptions are Central Eurasia, Northern Europe, parts of North America and Brazil, since these regions are characterized by increased sill of up to $(0.3\,km/s)^2$, relatively low nugget and increased range. In these regions large-scale velocity anomalies exist that should be well suited for interpolation.

By contrast, the oceans possess almost everywhere similar nugget and sill values of about $(0.1km/s)^2$ and typical range between 5° and 7.5°. Thus, interpolation of velocity in the oceans will probably not be able to construct an extensive model, especially since the distribution of survey points is in many regions extremely sparse.

4.4.3 Interpolated Moho depth and average P-wave velocity

The interpolated depth of the Moho discontinuity below the geoid (Figure 4.10a) shows well-known large scale crustal features. The data are available in supporting information S1 and S2. We will discuss the continental features in section 4.4.5, when comparing with Crust1.0.

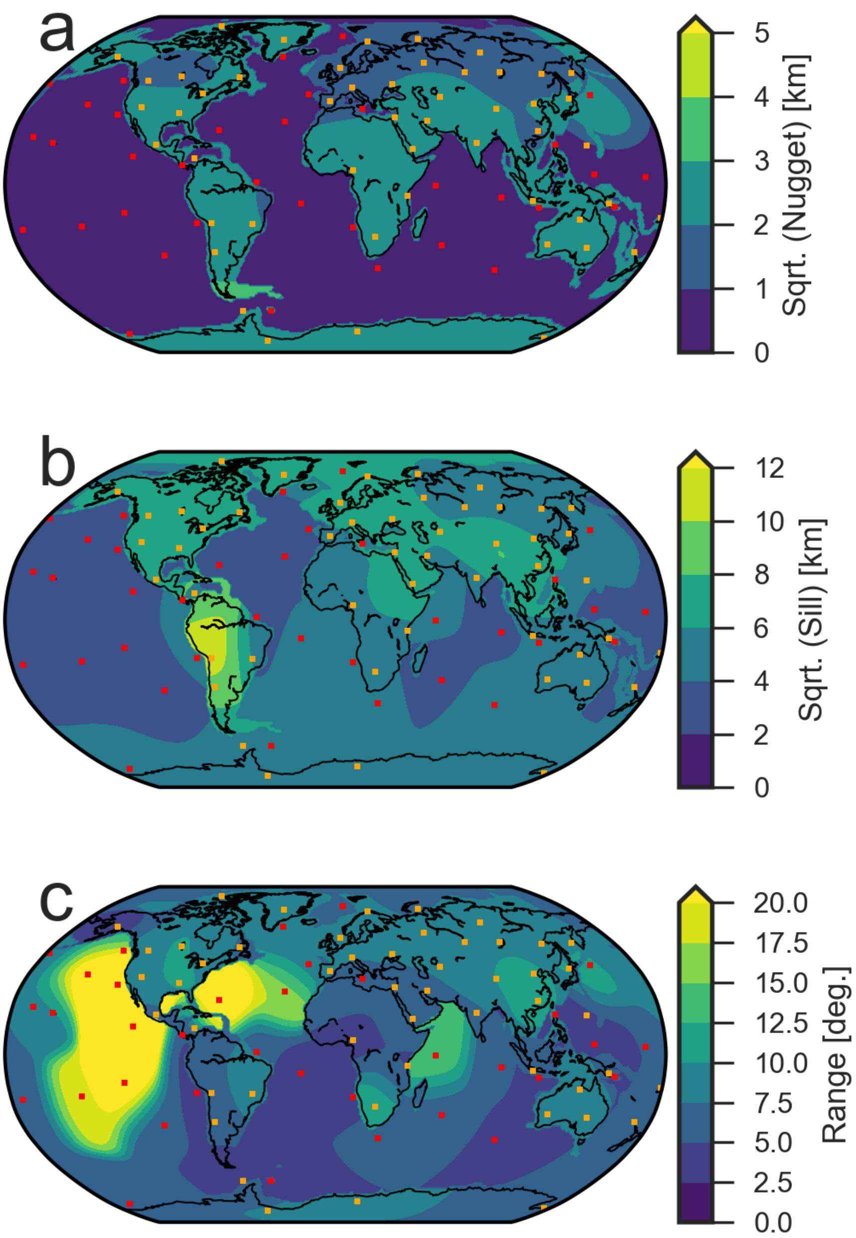

Figure 4.8: Estimated parameter distribution for Moho depth. The red and orange squares give the centers of the clusters used for interpolation in the oceans and continents respectively. Note we give the square root of nugget and sill.

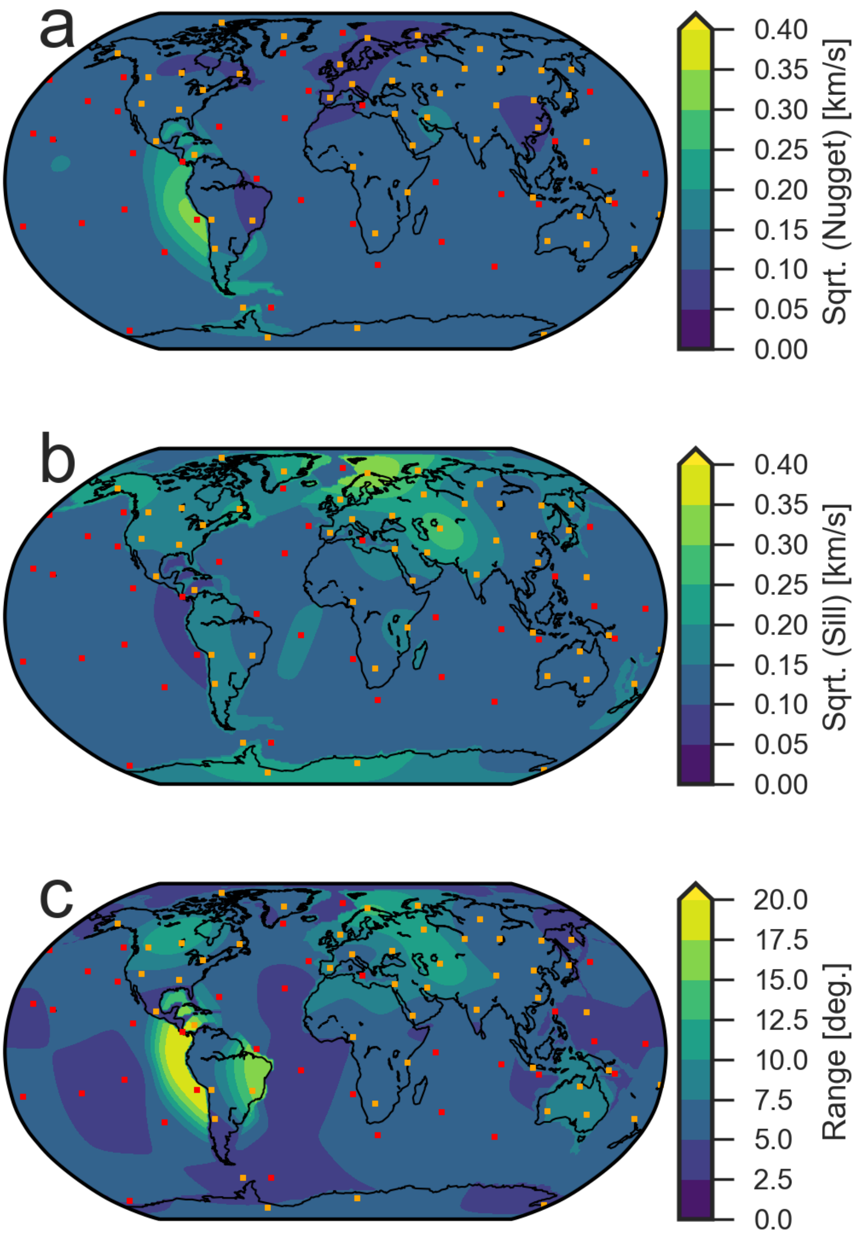

Figure 4.9: Distribution of covariance parameters for the average P-wave velocity of the crystalline crust.

The oceans are mostly flat with some single spots corresponding to individual data points. This occurs in areas with few points and estimated low range ρ. Structure is best resolved in the Western Pacific Ocean, where the trend of the pacific ridge is clearly visible. In the other oceans hardly any structure could be interpolated.

Our interpolation method produces 1-sigma estimates of uncertainty due to the underlying assumption of a Gaussian field. Furthermore, the uncertainties measure the internal consistency of the database and cannot describe the external accuracy of the database. The estimated uncertainties range from about 1 km to more than 12 km (Figure 4.10b). In agreement with the estimated distribution of nugget and sill (section 4.4.2), uncertainties in the oceans are systematically lower than in the continents. Unsurprisingly, uncertainty is negatively correlated with point density. However, the highest uncertainties are estimated for South America, even though Africa contains even less points. This is a result of the extremely high sill value for South America, which might not be realistic. The lowest uncertainties are estimated for North America, most of Eurasia and Australia.

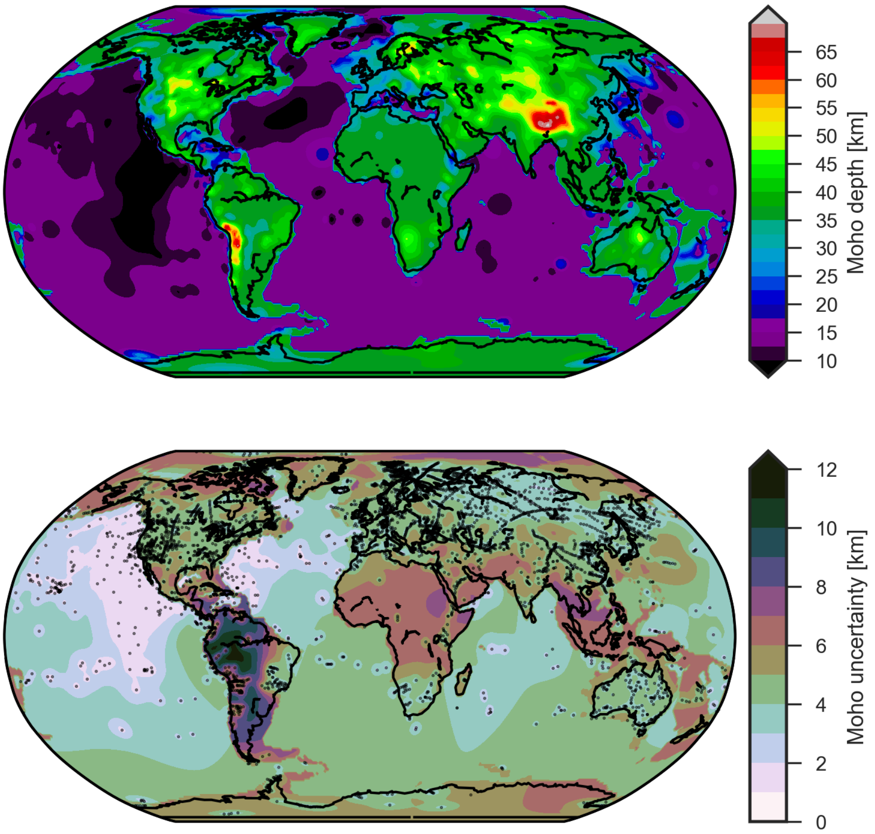

Figure 4.10: Global map of Moho boundary depth below the geoid and its estimated uncertainty.

The interpolated P-wave velocity of the crystalline crust and its uncertainty is shown in Figure 4.11. In the oceans hardly any structure is visible in the interpolation results except for some very localized velocity anomalies. The magnitude of these anomalies is typically less than 0.2 km/s, which is comparable to the typical estimated uncertainty. Hence, these localized features are probably of much smaller scale than the resolution of the data points or are the results of velocity inaccuracies.

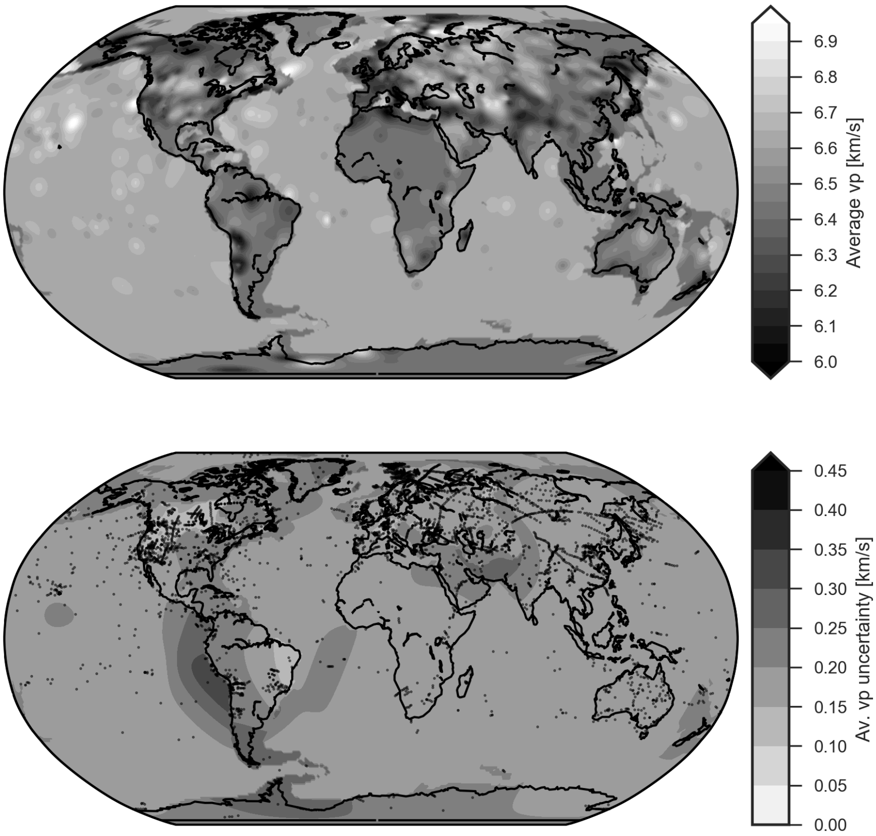

Figure 4.11: Interpolated distribution of P-wave velocity in the crystalline crust.

4.4.4 Quality of error estimation

4.4.4.1 Hold-one-out cross validation

The hold-one-out cross validation allows testing the prediction accuracy and error estimate. In the case of Moho depth, the RMS prediction errors for oceans and continents were 3.5 and 4.7 km respectively. The median absolute deviation ratio is 0.53 for oceans and 0.48 for continents, which is less than the theoretically expected value of 0.67 . Deviation ratios above one occur for 30 % of the oceanic points and 25 % of the continental points, which is comparable to the theoretical expectation of 32 %. Thus, overall the error estimate is adequate, albeit slightly too pessimistic.

The prediction error as a function of survey year reflects the increasing quality of instrumentation and processing over time. For continents (Figure 4.12b), newer surveys are clearly characterized by smaller prediction errors and deviation ratios than older surveys. The prediction error decreases from a median value of 3 km for surveys from the 1950s, to less than 2 km for surveys after the year 2000. For surveys acquired in oceanic domains, the prediction error stays more or less constant at about 1 km (Figure 4.12a). The deviation ratio shows the same trend (Figure 4.12b and d).

Thus, the interpolation results are more consistent with newer - more accurate - data than with older data. Note that the survey year was not explicitly included in the interpolation process in any way and this is a natural outcome of the geostatistical analysis and interpolation.

Repeating the same analysis for the survey type, shows a similar pattern (Figure 4.13): Reversed refraction profiles have the lowest prediction errors, followed by split refraction and unreversed refraction. Reflection profiles are on average least accurate, followed by the unknown/other category, which is also highly variable. Thus, overall the most accurate type of survey is most consistent with the interpolation results.

Analyzing the P-wave velocity in terms of survey year and type yields the same conclusions: Newer surveys systematically have lower prediction errors and deviation ratios. Likewise, reversed refraction profile is the type of survey with the lowest prediction errors and deviation ratios. The root mean squared prediction error is about 0.18 km/s for oceans and continents. The median absolute deviation ratios are 0.58 for oceans and 0.46 for continents. About 75 % of the points stay within the estimated error bounds.

Overall, the prediction errors are within the range of reported error estimates for crustal scale seismic profiles, which are about 10 % (Christensen and Mooney, 1995). Thus, our interpolation method and the chosen parameters are adequate for reproducing the seismic measurements.

However, holding out only a single observation might lead to optimistic error estimates, due to the very uneven distribution of points. For instance, many points are taken from continuous profiles, so that it should be relatively easy to reconstruct that single point using the neighboring points from the same profile. For this reason we also used a more challenging cross-validation test of the data set.

4.4.4.2 Median split cross-validation

We split the data set into a validation and prediction set. The points from the prediction set will be used to predict the points from the validation set. In this test, the older half of the data is used to predict the Moho depth at the points of the newer half. The split between older and newer half occurs at the median survey year 1989. Most parameters of the interpolation were left unchanged, with the exception of the number of clusters, which had to be reduced to 35

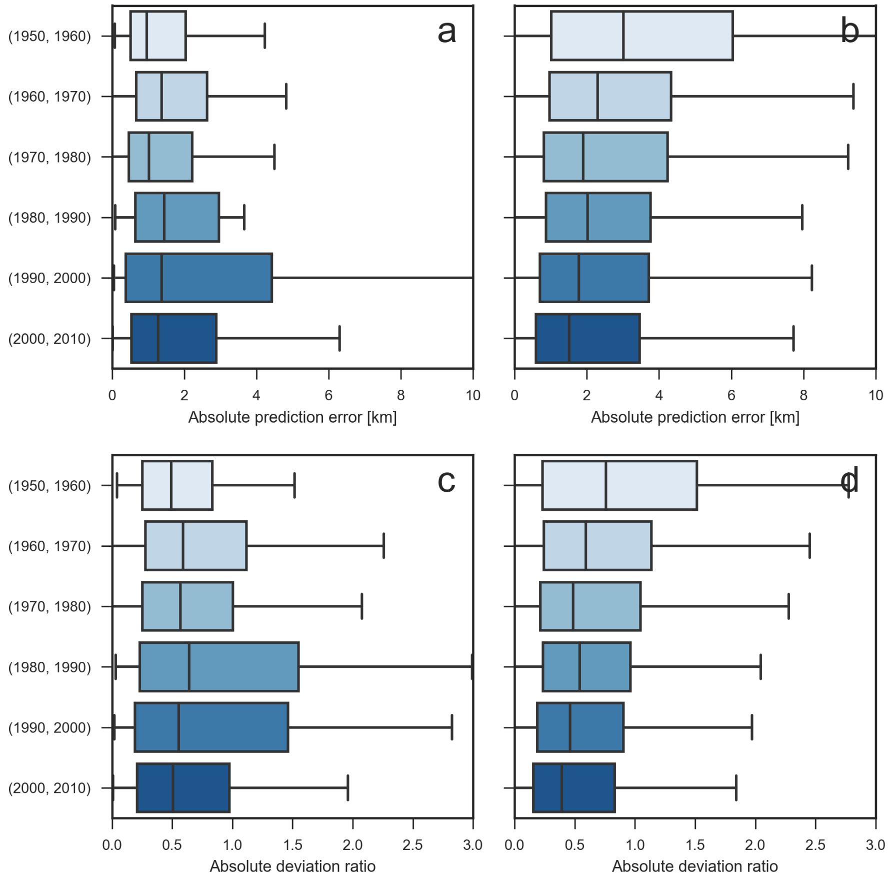

Figure 4.12: Prediction accuracy for points grouped according to decade in which the survey was carried out. a) and b) give the absolute prediction error, while c) and d) give the absolute deviation ratio. a) and c) refer to oceanic and points, while b) and d) refer to continental points. In all figures, the colored boxes cover the 25 % to 75 % percentile, while the Median is marked by a vertical line inside each box. The extent of the whiskers gives 1.75 of the interquartile range relative to the median.

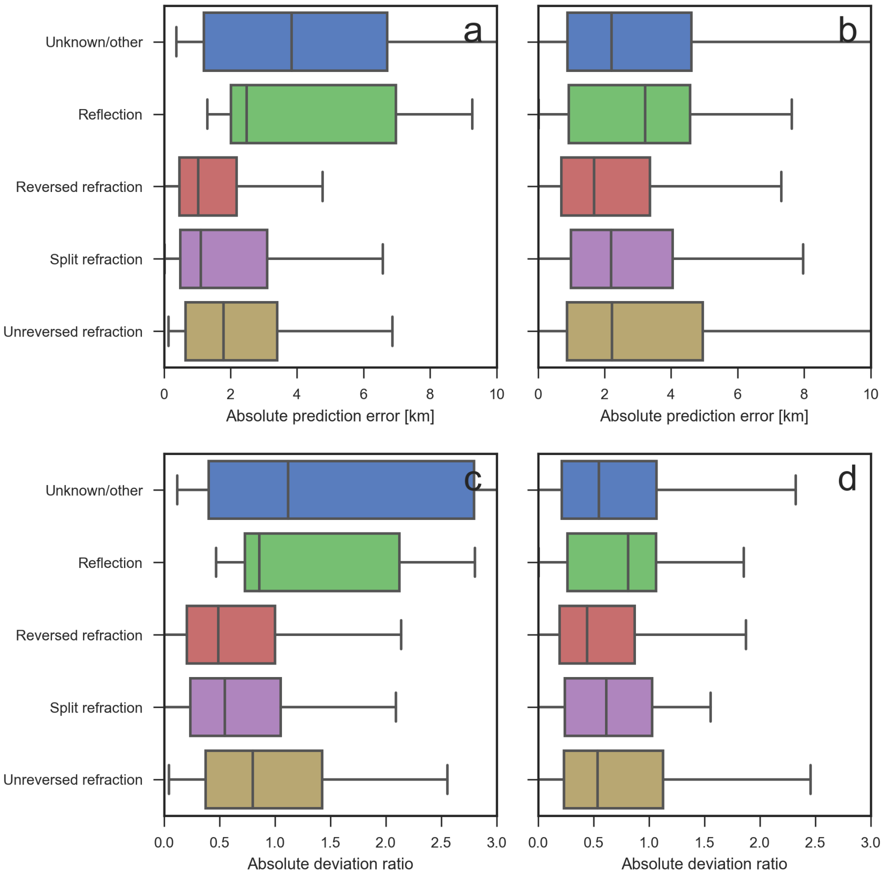

Figure 4.13: Prediction accuracy grouped according to survey type. a) and b) give the absolute prediction error, while c) and d) give absolute deviation ratio. a) and c) refer to oceanic points, while b) and d) correspond to continental points. For an explanation of the figure see caption of Figure 4.12.

continental and 30 oceanic clusters, in order to prevent single-point clusters and to maintain an average cluster radius of 10°.

The root mean prediction errors is 5.6 km for oceans and 8.1 km for continents, which is roughly twice as large as the prediction errors of the hold-one-out cross validation. While the oceanic error estimates are on average adequate (median deviation ratio of 0.68), the estimated error for continents is systematically too low (median deviation ratio of 0.83). In addition, the distribution of continental deviation ratios becomes skewed towards high values, because about 15 % of the continental points have deviation ratios of 2 or more, greatly surpassing the theoretical expectation of 5 %.

In general, the error estimate can be seen as adequate for this test case, because most points from the validation set stay within the error bounds. However the interpolation error estimate does not and cannot account for the possible presence of unexpected or rapid changes that do not conform to the underlying assumptions of the interpolation. For example, in sparsely surveyed area, the Moho depth distribution might appear to be more smooth than it actually is. This leads to an underestimation of sill and range and ultimately to an underestimation of uncertainty. This should be considered when interpreting the uncertainties over large poorly-covered regions.

4.4.5 Comparison with Crust1.0

4.4.5.1 Global

On a global scale, our interpolated Moho depth distribution agrees well with the Moho map of Crust1.0 (Laske et al., 2013), with a median absolute difference of 2.9 km. The spatial pattern of the disagreement between our model and Crust1.0 (Figure 4.14) shows that the largest differences occur in South America and Asia. While the overall patterns are similar, our results appear less sharp than Crust1.0 due to the tectonic regularization applied in Crust1.0.

In the oceans, Crust1.0 shows significantly more structure, since a constant crustal thickness in the unaltered oceans was assumed and the Moho depth was calculated from ocean floor depth. We have directly interpolated Moho depth and are thus not using the topographic information, leading to a flat Moho in most oceans. The continent-ocean transition is modeled explicitly in Crust1.0 as a smooth transition over the continental margin, whereas we have used a sharp jump from continental to oceanic crust. This leads to a systematic different between our results and Crust1.0 on the edges of the continents.

Statistical analysis of the differences between our results and Crust1.0 (Figure 4.15a), shows that the differences between the two models are typically 3 km on all continents except South America, where typical differences are 4.5 km. Importantly, better surveyed continents are not associated with less disagreement between our model and Crust1.0 (i.e. Africa and North America are similarly different). However, our estimated uncertainties (Figure 4.15b) differ substantially for the different continents and tend to be higher than the average differences between our model and Crust1.0. This suggest that taking the difference between different crustal models is an overly optimistic indicator of model uncertainty, especially if the models are based on similar methods and/or database.

Our interpolation results for average P-wave velocity differ from Crust1.0 substantially. Overall, the scale of velocity variations in Crust1.0 is much less, since velocities are average values for complete domains. In most of the oceans the two models differ mostly by a shift, because our average velocity in the oceans is 6.65 km/s, whereas Crust1.0 uses 6.75 km/s.

In the continents there is qualitative agreement between the two models in many areas. Examples include the separation between slow Western Europe and faster Central Europe visible in both

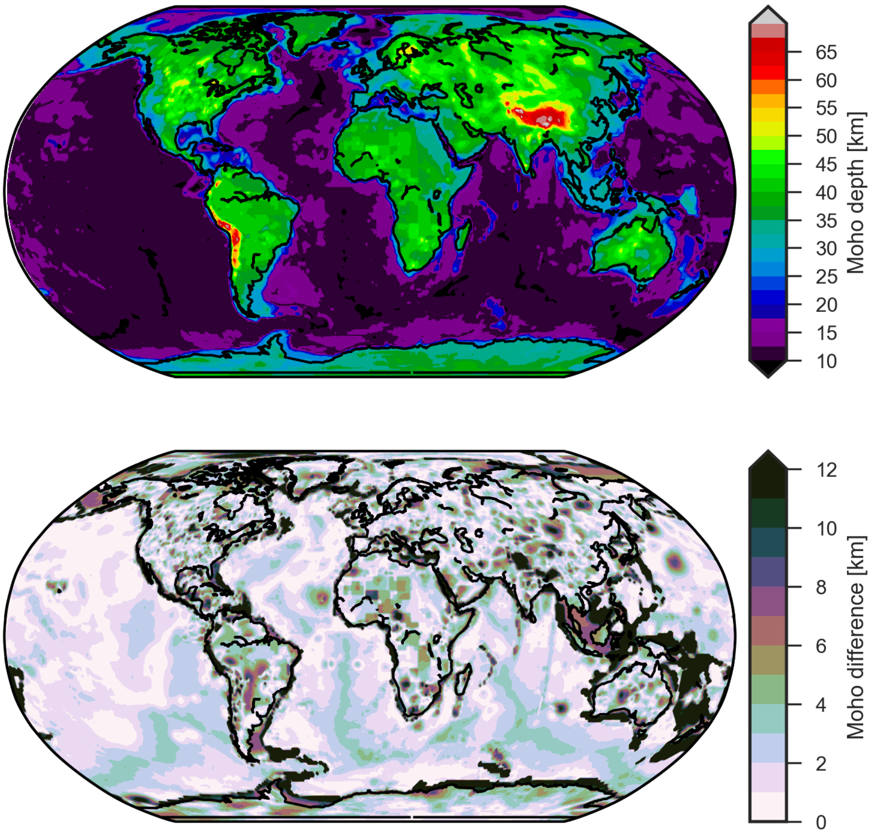

Figure 4.14: Depth to Moho according to Crust1.0 model and the differences with respect to our model.

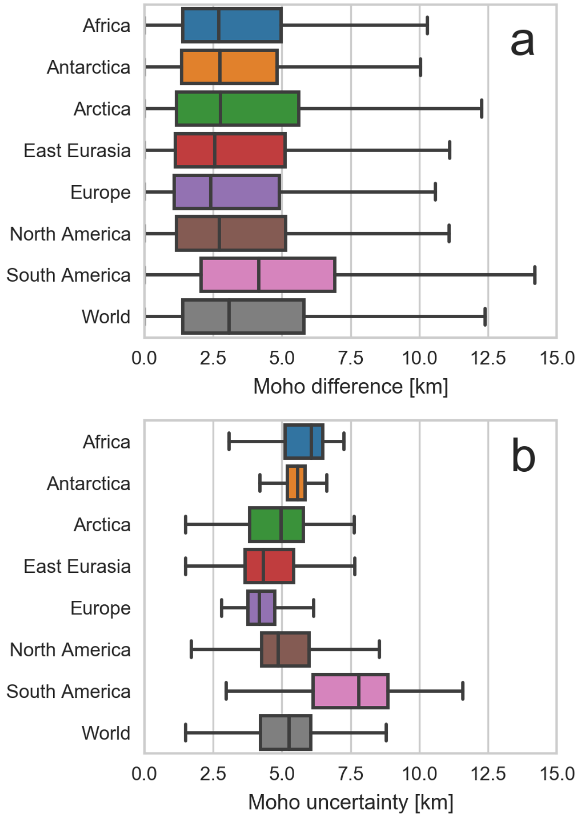

Figure 4.15: (a): Statistical distribution of differences between our model and Crust1.0. (b): Distribution of uncertainties in our model. The two models are more similar then the uncertainties would suggest, because they rely on the same data source. Note that the frequencies are calculated from the grid points and do not take into account the different sizes of grid cells at higher latitudes. For an explanation of the figure see caption of Figure 4.12.

models, or the reduced velocity in the Western United States. However, even in areas with qualitative agreement, the results of the direct interpolation are less defined spatially, but are characterized by larger velocity variations.

Disagreement between the models can be up to 0.5 km/s and the RMS difference between the models is 0.03 km/s. Thus, on average the models again agree better than expected based on the estimated interpolation uncertainty.

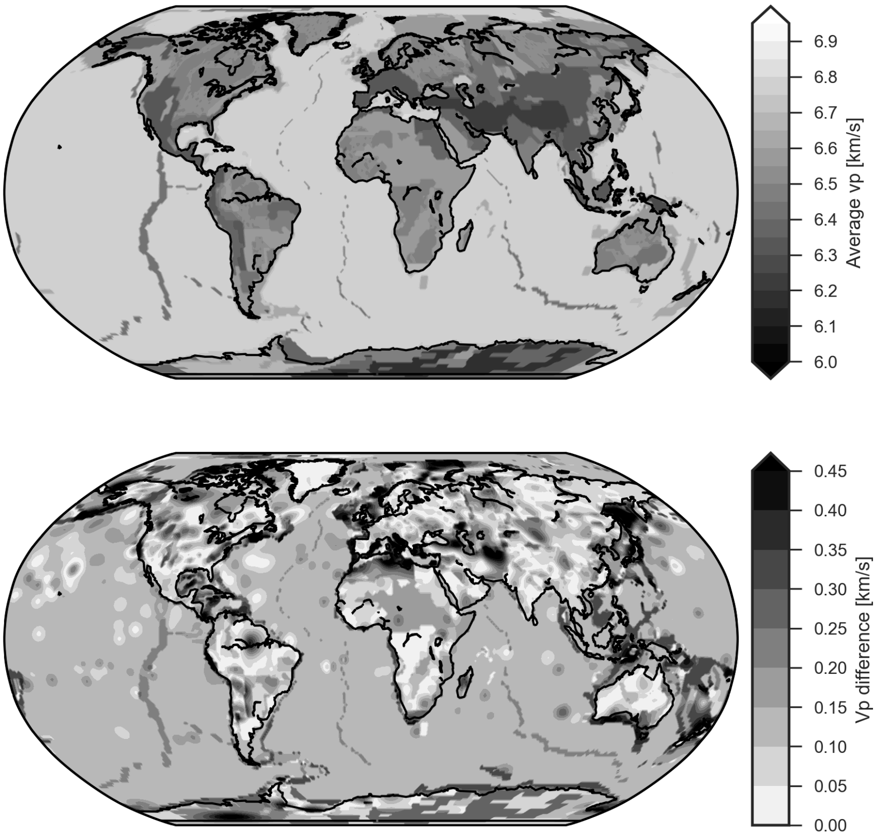

Figure 4.16: Average crustal P-wave velocity for Crust1.0 and the difference with respect to our results.

4.4.5.2 South America

On the global map, South America stands out as the area with the highest uncertainty, reaching values of 12 km or more (Figure 4.17). This is does not seem entirely reasonable, because South America contains much more data points than for example Africa.

The uncertainty values of more than 10 km occur only in areas that are not covered by any points. In regions with points, the uncertainty is decreased to realistic values of around 5 km. However, even in the densely studied areas in the Andes, the uncertainty is ~8 km and the thick crust underneath the Andes is not interpolated along the entire mountain chain. There are two explanations for this:

Firstly, the Moho depth is extremely variable, even for relatively short distances. The Moho boundary depth goes from up to 70 km in the Andean Cordillera, to less than 30 km in the hinterland basins. This leads to high estimated sill values and low estimated range values. Secondly, Moho depth variations are highly anisotropic, as variations along the orogen (roughly

N-S) are much less than perpendicular to the mountain chain (E-W). Thus, an anisotropic covariance function would be required to adequately model the spatial structure in this region.

The main difference with respect to previous models (Crust1.0 or the model of Chulick et al. (2013)) is the very thin crust in the hinterlands of the Andean orogeny. There are only a few points that show this thinning and this trend is interpolated along the entire north-south extent of the mountain chain. However, this feature is also associated with the highest uncertainties of 12 km. We note that Van der Meijde et al. (2013) detected a similar area of thinned crust using gravity inversion.

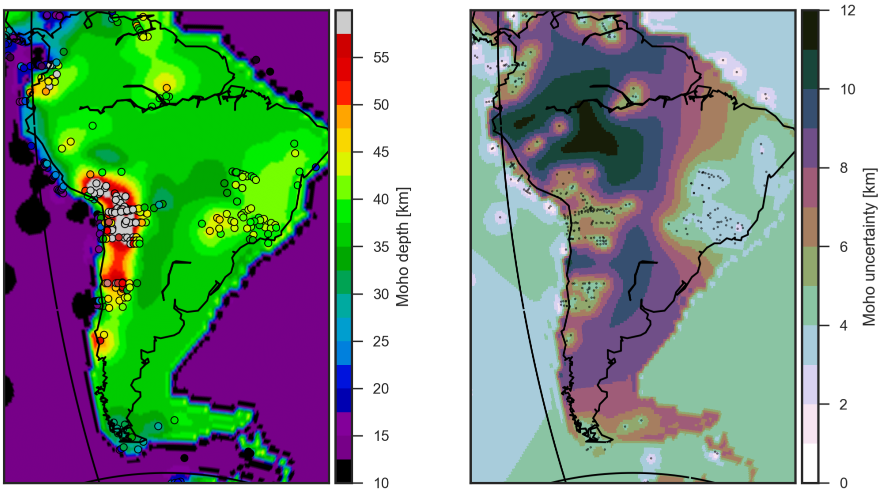

Figure 4.17: Interpolated Moho depth for South America and the estimated uncertainty from interpolation. The dots give the location of interpolation points.

4.4.5.3 Africa

We chose Africa to illustrate the behavior of the interpolation algorithm in a data sparse environment (Figure 4.18). Data points are concentrated in Southern Africa, the East-African rift system and the Mediterranean coast. The values of the range ρ in most of Africa are very low, so that the results are not interpolated very far away from the points. These few scattered points provide little constraint on the overall crustal structure, which can therefore be regarded as virtually unknown, as reflected in the high uncertainties of around 7 km.

Note that this pessimistic view only applies to our database, which does not contain important new results derived from passive seismic methods (Fishwick and Bastow, 2011). However, the crustal structure of Africa remains unclear, and non-seismic methods are still important to better determine it (Globig et al., 2016).

Crust1.0 shows a tile-like pattern of Moho depth variation (Figure 4.14), that reflects the pre-defined crustal types. Over South Africa, Crust1.0 disagrees with our results as well as the points in the GSC database.

4.4.5.4 Europe

Europe is among the most well surveyed areas in the world. Thus, comparing our results with Crust1.0 in this area highlights differences caused by different methodologies, because neither method is adversely affected by sparse data.

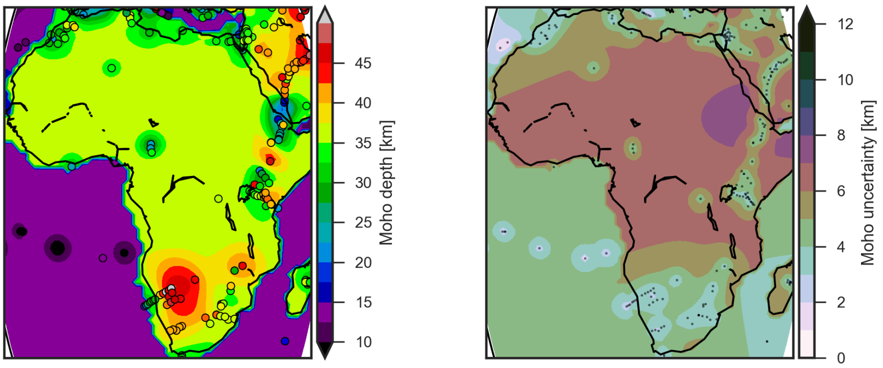

Figure 4.18: Interpolated Moho depth for Africa and the estimated uncertainty from interpolation.

In this region, there is substantial agreement of the Moho depth between our results and Crust1.0 (Figure 4.19), apart from the general differences arising from the different treatment of the continent-ocean boundary (see for example the Mediterranean coast of Africa). On the continent, substantial large-scale differences mostly occur where no data are available in our database (e.g. Baltic States, Figure 4.19d). There are some punctual large differences, for example in the Alps. However, given the coarse parameterization of both models, neither model is capable of resolving the crustal variations on the scale of the Alps satisfactorily. Overall, higher differences between the models are also associated with higher estimated uncertainties (Figure 4.19c and d).

While the main patterns are resolved by both for the Moho depth, this does not hold for average P-wave velocity. In continental Europe there tends to be agreement in areas covered by seismic points. However, in some areas, for example Finland, there are differences of up to 0.2 km/s, despite a sufficient number of data points. In less well covered areas, there are larger discrepancies.

The impact of the different methodologies is strikingly visible in the way the separation between West and Central Europe is represented in the Crust1.0 and our results. In Crust1.0 the velocity changes abruptly, due to the tectonic localization, whereas our results show a more diffuse boundary (Figure 4.20).

The diffuse appearance does not reflect the nature of the boundary itself, but is a result of the interpolation. The underlying assumption of a Gaussian field with slowly varying covariance parameters is not well suited to represent abrupt parameter jumps. Furthermore, the precise location of the boundary is not completely constrained from our data set, which contributes to a diffuse boundary.

4.5 Application of global maps to residual topography

We will use the calculation of residual topography as an example of how the uncertainties can be used in geophysical applications. Isostatic topography t_{iso} is given by

$$t_{iso} = \frac{M - M_{ref}}{\rho_t}(\rho_m - \rho_c)$$

Where M is Moho depth (positive downwards) and M_{ref} is Moho depth in the reference model, $\rho_t = 2670\frac{\mathrm{kg}}{\mathrm{m}^3}$ is the density of topography, ρ_c is the mean crustal density and $\rho_m = 3200\frac{\mathrm{kg}}{\mathrm{m}^3}$ is

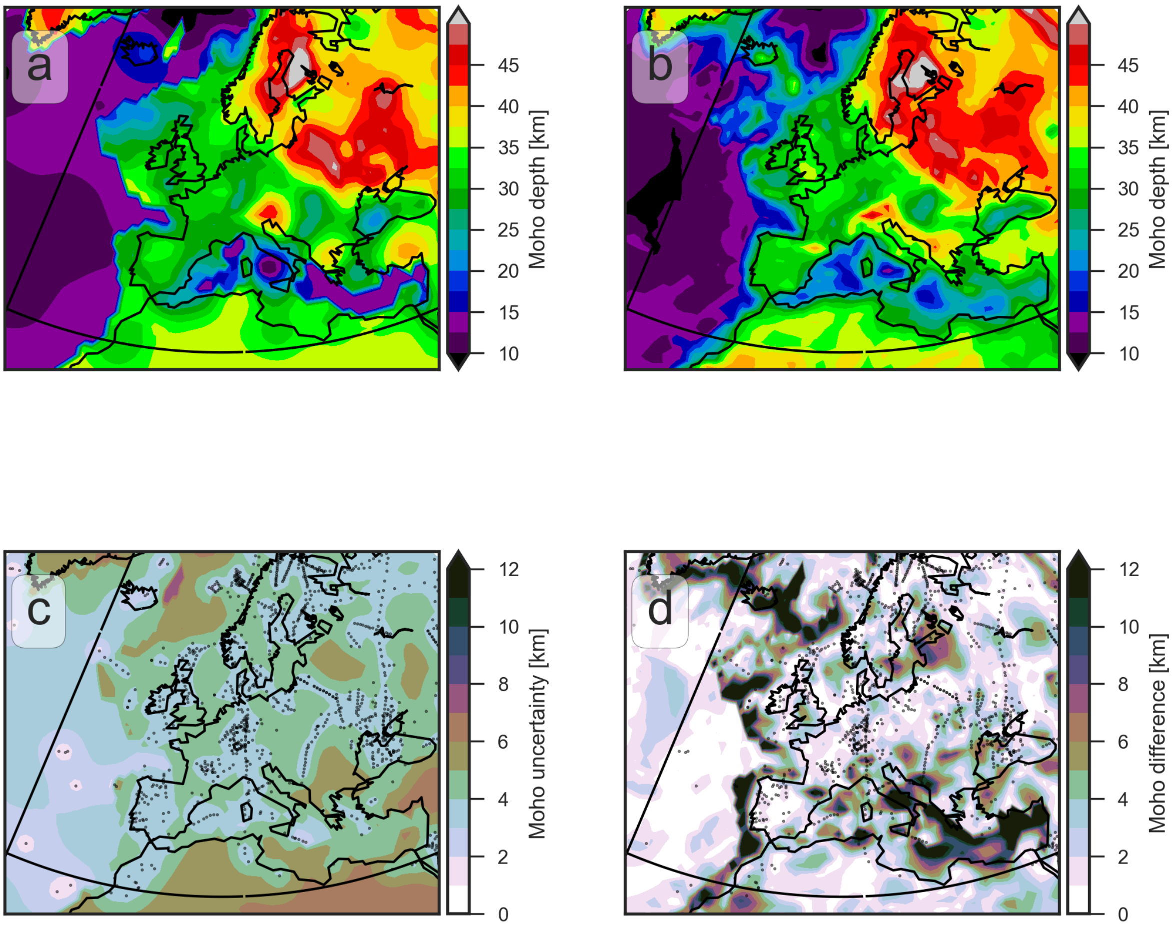

Figure 4.19: Moho depth interpolation results for Europe. (a), (b): Moho depth from our results and Crust1.0. (c): Estimated uncertainty (d): Difference between our results and Crust1.0

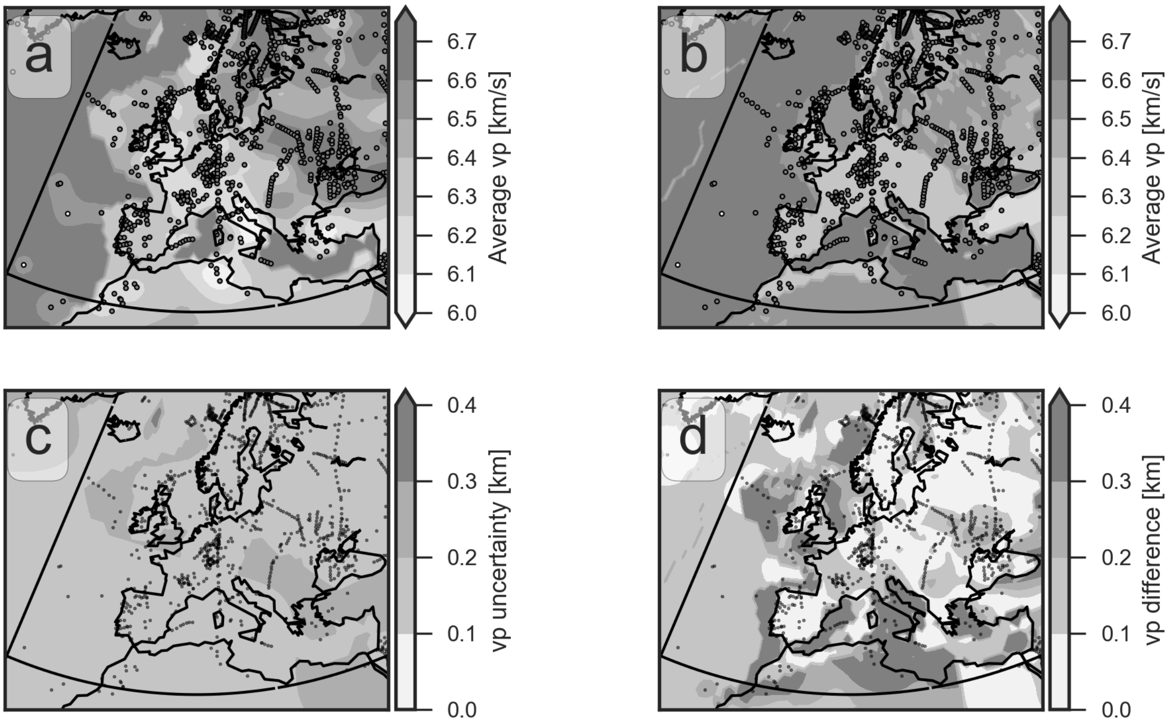

Figure 4.20: P-wave velocity of the crystalline crust for Europe. (a), (b): Comparison of our results and Crust1.0. The circles give the original points from the database (c): Estimated uncertainty (d): Difference between our results and Crust1.0

the reference mantle density. For simplicity, we will neglect the contribution of sediments and ice here.

Residual topography is defined as $t_{equiv} - t_{iso}$ (Kaban et al., 1999), where equivalent topography accounts for the reduced density of water:

$$t_{equiv} = \begin{cases} t & \text{if } t > 0 \\ t\frac{1640}{2670} & \text{if } t < 0 \end{cases}$$

Residual topography is possibly due to four causes: Errors in the crustal model, density variations in the lithospheric mantle, elastic flexure of the lithosphere and viscous stresses resulting from mantle convection that act on the base of the lithosphere (e.g. Molnar et al., 2015).

To find the mean crustal density ρ_c, a velocity-density relationship is required. Since the typical $v_p - \rho$ relations are pressure and thus depth-dependent (Christensen and Mooney, 1995; Brocher, 2005), we cannot apply them immediately to the mean crustal velocity. Instead, we use linear regression to the mean velocities and densities from Crust1.0 to derive an empirical relation between average crustal v_p and density. This leads to the following relations for oceans and continents

$$\rho[\frac{\text{kg}}{\text{m}^3}] = 350 + 385\, v_p[\frac{\text{km}}{\text{s}}]$$

$$\rho[\frac{\text{kg}}{\text{m}^3}] = 590 + 346\, v_p[\frac{\text{km}}{\text{s}}] \tag{4.4}$$

We propagate the uncertainties of Moho and P-wave velocity to the isostatic topography using Gaussian error propagation. Isostatic topography is a simple 1-D calculation, so spatial covariances between errors at different locations can be ignored. The contribution of a Moho uncertainty of δM is then

$$\delta t_M = \left| \frac{\rho_m - \rho_c}{\rho_t} \right| \delta M.$$

A velocity uncertainty of δv contributes

$$\delta t_v = \left| \frac{M - M_{ref}}{\rho_t} \right| \frac{\partial \rho_c}{\partial v_p} \delta v,$$

where the partial derivative of density with respect to density corresponds to the linear term in the empirical velocity-density relation (eq. 4.4).

The total uncertainty is the sum of the uncertainty variances resulting from Moho depth and P-wave velocity estimates, $\delta t = \sqrt{\delta t_M^2 + \delta t_v^2}$, assuming velocity and Moho depth errors are uncorrelated. Note that velocity and depth errors are in reality correlated because velocity is used to convert measurements of travel time to depth. This would tend to reduce the overall uncertainty, because depth and density have opposing effects on isostatic topography.

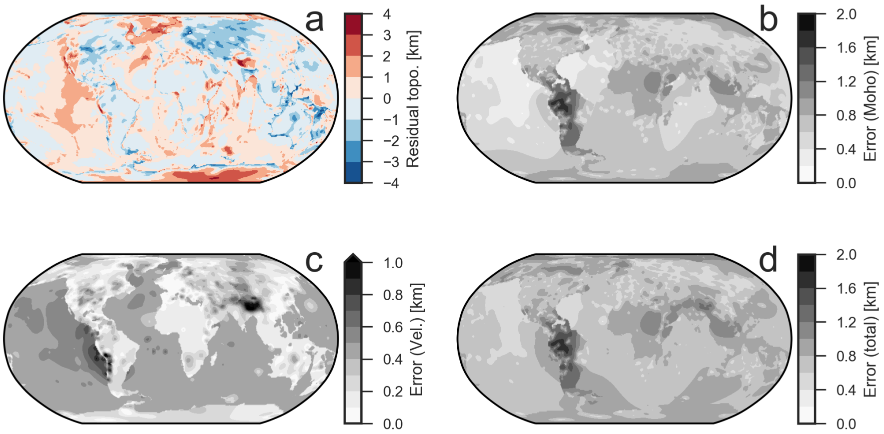

Figure 4.21: (a): Residual topography calculated from our crustal models. (b): Uncertainty of residual topography resulting from uncertainty of Moho depth. (c): Uncertainty caused by P-wave velocity uncertainty. (d): Total uncertainty assuming Moho depth and P-wave velocity uncertainties are uncorrelated.

The calculated residual topography (Figure 4.21a) has values between -4 and 6 km. On a large scale the patterns are as would be expected: Positive anomalies at mid-oceanic ridges, that systematically decrease as the oceanic plate gets older. Continents are generally associated with negative anomalies, that are probably related to thick, cold sub-lithospheric mantle, whose cold thick lithosphere pulls down topography (Zoback and Mooney, 2003)

In general, the uncertainties coming from Moho depth interpolation are larger than the errors coming from velocity in the continents (Figure 4.21b and c). Only in the regions with the thickest crust, velocity contributes more than a few 100 m to the uncertainty. The oceans have a very high velocity uncertainty, leading to systematic uncertainty of almost 1 km. Maximum total uncertainty is 2 km, which only occurs in South America and Himalayas (Figure 4.21d).

Note that, as discussed above, the uncertainties for South America are probably to pessimistic in the Andean region.

On average the total uncertainty is 1.1 km. For around 65 % of the points, the magnitude of estimated residual topography is larger than the estimated error and for around 40 % of the points, the value of residual topography is twice as high as the estimated error. This demonstrates mantle contributions to topography above the uncertainty threshold exist. However, the depth at which the mantle contribution originates cannot be constrained with this method. The sources might lie in the lithospheric or in the deeper mantle. Differentiating between lithospheric and deeper sources requires additional information, for example from gravity modeling (e.g. Afonso et al., 2016b).

4.6 Conclusions

We have constructed a global crustal model based on a USGS database of crustal seismic studies. Unlike previous studies we have tried to include as little additional information as possible, following a design philosophy that tries to respect the data as much as possible. We interpolated Moho depth and average P-wave velocity of the crystalline crust using a non-stationary kriging algorithm (Risser and Calder, 2017). A major advantage of our approach is that it is relatively straightforward to derive uncertainties.

The uncertainties we derived should be seen as a worst-case estimate, since we have only included a subset of seismic methods that sample the crust. For instance, short-period surface waves, ambient noise tomography (Shapiro and Campillo, 2004) and receiver functions (Kind et al., 2012) also provide information about crustal structure. However, we chose not to include them in our analysis. Mainly because the combination of information from different sources requires good knowledge of the uncertainties affecting either data source. Thus, the kind of analysis we carried out here will certainly benefit a combined analysis of several types of data. Furthermore, active seismic surveys, which make up the bulk of the GSC database, are generally seen as the most reliable way of imaging the crust.

However, uncertainties remain even in areas where a large number of profiles have been acquired with estimated Moho depth errors of around or less than 4 km over large areas. These uncertainties stem both from errors of individual data points, which are reflected in the nugget effect and the uncertainty introduced by interpolating. We find that the nugget effect of the Moho depth is generally less than the estimated accuracy of 10 % (4 km) given by Mooney et al. (1998), which is probably due to better accuracy of newer studies, as confirmed using hold-one-out cross validation.

Interpolation of average crustal v_p proved to be more challenging than interpolating the Moho depth. To some extent this is a bit enigmatic, since depths are in all seismic methods calculated from the velocities in some way. However, the relative accuracy of velocities is still better than the relative accuracy of Moho depth, which partly explains this apparent contradiction.

Are domain-averaged velocity profiles and other methods of avoiding direct interpolation of v_p preferable over direct interpolation? On the one hand this depends on the importance given to small-scale variations of velocity. Are they noise or incompletely imaged structure? In the latter case, direct interpolation would at least partially reflect them instead of averaging them out. On the other hand it depends on what a crustal model should represent: Is it one plausible model or is it a mean of all plausible models? Our method aims for the latter. Thus, a sharp velocity transition between velocity domains would be represented as a smooth transition, as long as the precise location of that transition is unknown. Both variants of crustal models are valid, but it needs to be clear what type a specific crustal model is. Ultimately, it might be worthwhile to find ways to incorporate the tectonic regularization into the interpolation to combine advantages of both approaches.

Combining the results of Moho depth and velocity interpolation, we could derive a map of residual topography, that shows the expected features. In addition, we could also estimate the uncertainty of our residual topography map. These uncertainties can be up to 2 km and are on average 1 km. Given the expected range of dynamic topography of a few hundred meters (Molnar et al., 2015), this raises the question if global statements about dynamic topography are even possible at this stage. Of course, crustal models incorporating additional information reduce this uncertainty. But even in already well-studied areas, uncertainties can easily reach 1 km. A more comprehensive approach that includes gravity in addition to topography is probably the most effective way to constrain residual topography, since gravity and topography are both sensitive to the density structure.

Future research about interpolation of crustal structure could improve our results in different ways. Constructing a layered model of crustal structure instead of a one-layer model might be a logical next step. In addition, the ocean-continent classification could be derived not only from additional data, but also using the crustal thickness data itself. The mathematical model underlying our interpolation can be improved by including anisotropy.

Since new seismic data are continually acquired, any crustal model is already outdated at the time of publication. In addition, global and regional catalogues of seismic determinations already exist (Baranov and Morelli, 2013; Assumpção et al., 2013), that rely on different selections of data. For this reason, the interpolation software is freely available here (supporting information S3) to allow users to apply them to any database of seismic studies.

Acknowledgments

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) Priority Program DynamicEarth, SPP 1788 and an Australia-Germany Joint Research Cooperation Scheme funded by the German Academic Exchange Service (DAAD) and Universities Australia.

A: Clustering algorithm

We use an agglomerative clustering algorithm (Hastie et al., 2009) as implemented in sci-kit learn (Pedregosa et al., 2011).

Let D be the distance matrix, such that D_{ij} is the great-circle distance between points i and j. The algorithm begins with every point in its own singleton cluster. Larger clusters are then created iteratively by repeatedly merging the two most similar clusters. The distance between two clusters of points G and H is in our case defined as the maximum distance between any points of pairs in G and H.

$$d(G,H) = \min_{i \in G, j \in H} d_{ij} \tag{4.5}$$

This procedure is repeated until the number of remaining groups is equal to the number specified by the user. When using eq. 4.5, the algorithm is called complete linkage agglomerative clustering. We prefer to use complete linkage over other methods, because it tends to create compact clusters (Hastie et al., 2009).

5 Probabilistic joint inversion of satellite gravity gradients and topography - setup and first tests

5.1 Introduction

This chapter explains the possibilities for using the data and models from the previous chapters in a probabilistic joint inversion context. Probabilistic approaches are pivotal for integrated petrological-geophysical modeling of the lithosphere and upper mantle (Afonso et al., 2013b; Khan et al., 2013). Integrated modeling describes the sub-surface in terms of composition and temperature instead of the "geophysical" description in terms of e.g. density and wave speed (Afonso et al., 2016a). The geophysical parameters are instead derived from composition and temperature using rock physics and petrology (Connolly, 2009). Traditional inversion methods, which rely on linearization, are not well suited for this type of problem, because of the inherent non-linearity and non-uniqueness of such problems (Afonso et al., 2013a).

Probabilistic inversion frameworks typically build on the Bayesian formalism and use Monte-Carlo-Markov-Chain methods to draw samples from the Bayesian posterior distribution (Afonso et al., 2013b). There are two main advantages to this approach: First, no derivatives need to be calculated, which means abrupt property changes - for example due to mineral phase changes - can be included. Second, MCMC gives an ensemble of possible solutions. This allows in-depth exploration of the inversion result uncertainty, for example when there are strong trade-offs between some of the parameters. The main drawback of MCMC is that is less computationally efficient than linearized methods (e.g. Ambikasaran et al., 2013).

In this chapter, a probabilistic approach is used to invert satellite gravity gradients and topography for the lithospheric density distribution. Note that this is not an integrated geophysical-petrological model, but uses density directly. This problem *could* also be tackled in a linearized way, but since ultimately the goal is to integrate the result into a more comprehensive approach, a probabilistic approach was chosen instead. The Kriging result from chapter 4 are used as starting values for the inversion and the information about uncertainty covariance is included in the inversion. Satellite gravity gradient data mainly based on measurements from the GOCE satellite are used, which are processed in agreement with the results from chapter 2. In lights of the results from chapter 3, the magnetic field is not included. However if tighter constraints on thermal structure or susceptibility are available the code could be extended to include the magnetic field as well.

The middle of the North Atlantic - extending from 45° W to 15° W and from 25° N to 55° N - was chosen as test region for the inversion algorithm for two reasons. Firstly, the secular cooling of the oceanic lithosphere should produce a well-resolvable density structure in the lithospheric mantle (Turcotte and Schubert, 2002). Secondly, the area is affected by hot spot volcanism originating from the Azores and the New England hotspot (Steinberger, 2000; Gente et al., 2003). The two hot spots give rise to elevated sea floor in the form of the Azores islands and the Great Meteor seamounts. This gives the opportunity to test both for "normal" oceans and anomalous oceanic structure in the same area.

5.2 Data

5.2.1 Topography data

The ETOPO1 model (Amante and Eakins, 2009) gives global topography and bathymetry at a resolution of 1 arc-minute (Figure 5.1a). We reduced this resolution to 1 degree by calculating the area-weighted average of all grid cells contained in a 1 degree cell. This ensures that the total topographic mass is conserved (as long as the grid cell contains only positive or negative topography), which is important for calculation of isostatic topography and the gravity effect.

To facilitate the calculation of isostatic topography, the measured topography was converted to equivalent topography (Kaban and Mooney, 2001), by multiplying negative topography with a scaling factor θ given by:

$$\theta = \frac{\rho_c - \rho_w}{\rho_t}$$

Here, $\rho_c = 2800\frac{\mathrm{kg}}{\mathrm{m}^3}$ is the reference crustal density, $\rho_w = 1030\frac{\mathrm{kg}}{\mathrm{m}^3}$ is the water density and $\rho_t = 2670\frac{\mathrm{kg}}{\mathrm{m}^3}$ is the assumed topographic density.

The isostatic effect of sediments of thickness L on topography is calculated using the following expression and removed from the observed equivalent topography:

$$t_{sed} = \frac{L \cdot (\rho_c - \rho_{mean}(L))}{\rho_t}$$

Sediment thickness is taken from the compilation of NOAA (Whittaker et al., 2013), which was downsampled to a resolution of 1 degree (Figure 5.1c).

A polynomial density function is used inside the sediments (Mooney and Kaban, 2010):

$$\rho(z) = \sum_{i=0}^{4} a_i z^i$$

Here $a_0 = 2008.37$, $a_1 = 373.4$, $a_2 = -84.14$, $a_3 = 7.94$, $a_4 = -0.2602$ and z is depth below the sea floor (in km). The resulting density is in $\frac{\mathrm{kg}}{\mathrm{m}^3}$. The function is limited to a maximum density of $2800\frac{\mathrm{kg}}{\mathrm{m}^3}$. Using this function, the average sedimentary density is calculated as a function of total sedimentary thickness L, $\rho_{mean} = \frac{1}{L}\int_0^L \rho(z)\,\mathrm{d}z$.

Most surface features strongly correlate with ocean floor age (Figure 5.1d)

5.2.2 Gravity gradient data

Typically, satellite gravity data are transferred into spherical harmonics, from which the gravity gradient components can be recovered (Pail et al., 2010; Mayer-Guerr, 2015). However, this approach leads to smoothing due to the properties of global spherical harmonic functions. In contrast, Bouman et al. (2016) have applied a combination of spectral and space domain processing techniques to GOCE data in an effort to maximize the signal content at the two mission altitudes of 225 and 255 km. We use the grid containing all 6 gravity gradient components at 225 km altitude.

To avoid edge and far-field effects, the gravity data do not cover the complete study region, but extend only from 40° W to 20° W and from 30° N to 50° N. This gives an extension of 5° in every direction. The topographic effect of the study area is calculated using tesseroids and subtracted from the measured free-air gravity gradients. Since our model will be nearly

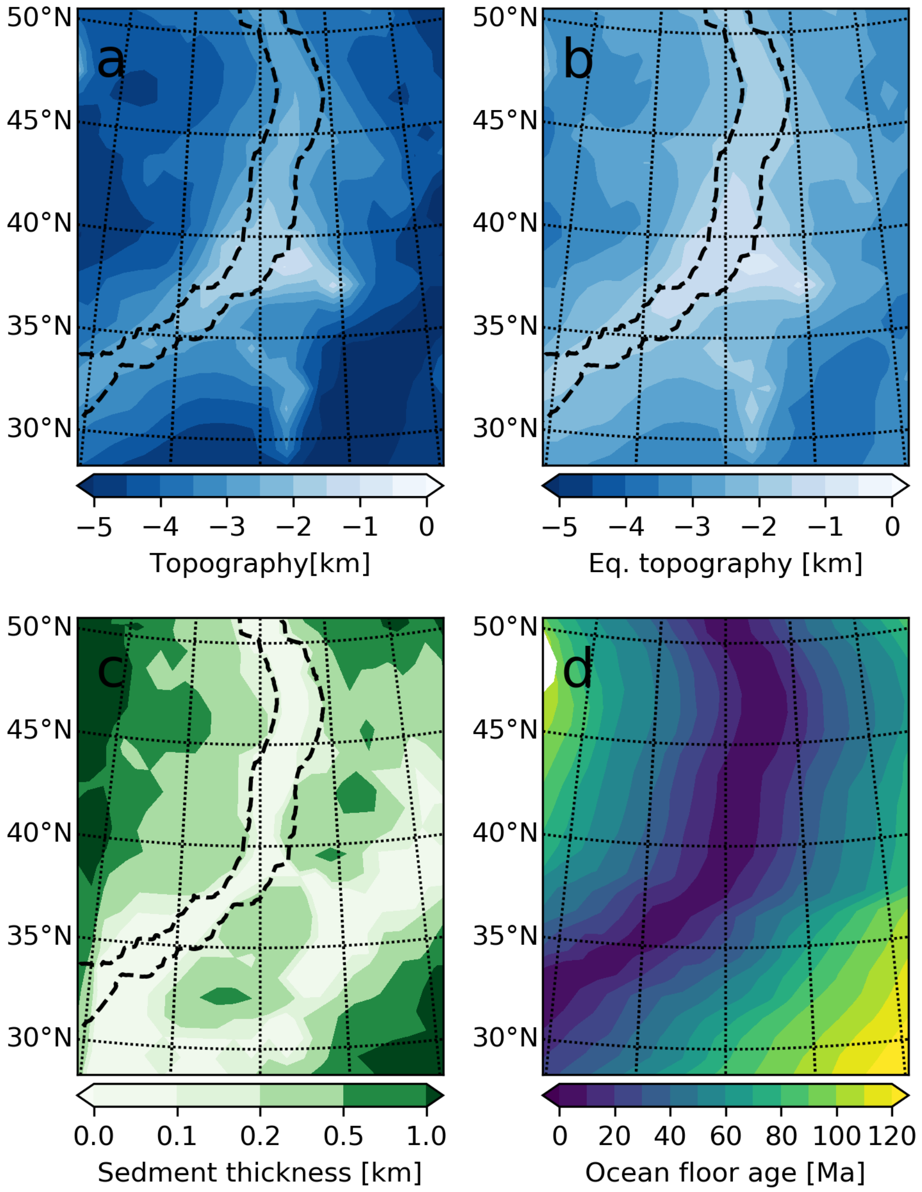

Figure 5.1: Surface data used in the inversion. All grids are sampled down to a resolution of one degree. The dashed line gives the 10 Ma isochron. a: Topography of the region. Note how the elevated ridge segments between 30 and 35 °N connects with the Azores island chain. The Great Meteor seamount lies south of the Azores. b: Equivalent topography after sediment correction. c: Sediment thickness. d: Ocean floor age (Müller et al., 2008)

in isostatic equilibrium, this approach is sufficient to account for edge and far-field effects (see chapter 2/ Szwillus et al., 2016). A topographic density of $2670\frac{\mathrm{kg}}{\mathrm{m}^3}$ is used onshore, whereas offshore a correction density of $2800 - 1030 = 1770\frac{\mathrm{kg}}{\mathrm{m}^3}$ is used. Note that we calculate the correction density relative to the crustal density used in our reference model.

The gravity effect of sediments also needs to be corrected. The gravity effect of sediments is then calculated by representing each column as a single point mass and calculating its gravity effect relative to the crustal reference density of $2800\frac{\mathrm{kg}}{\mathrm{m}^3}$. The final gradients are shown in Figure 5.2.

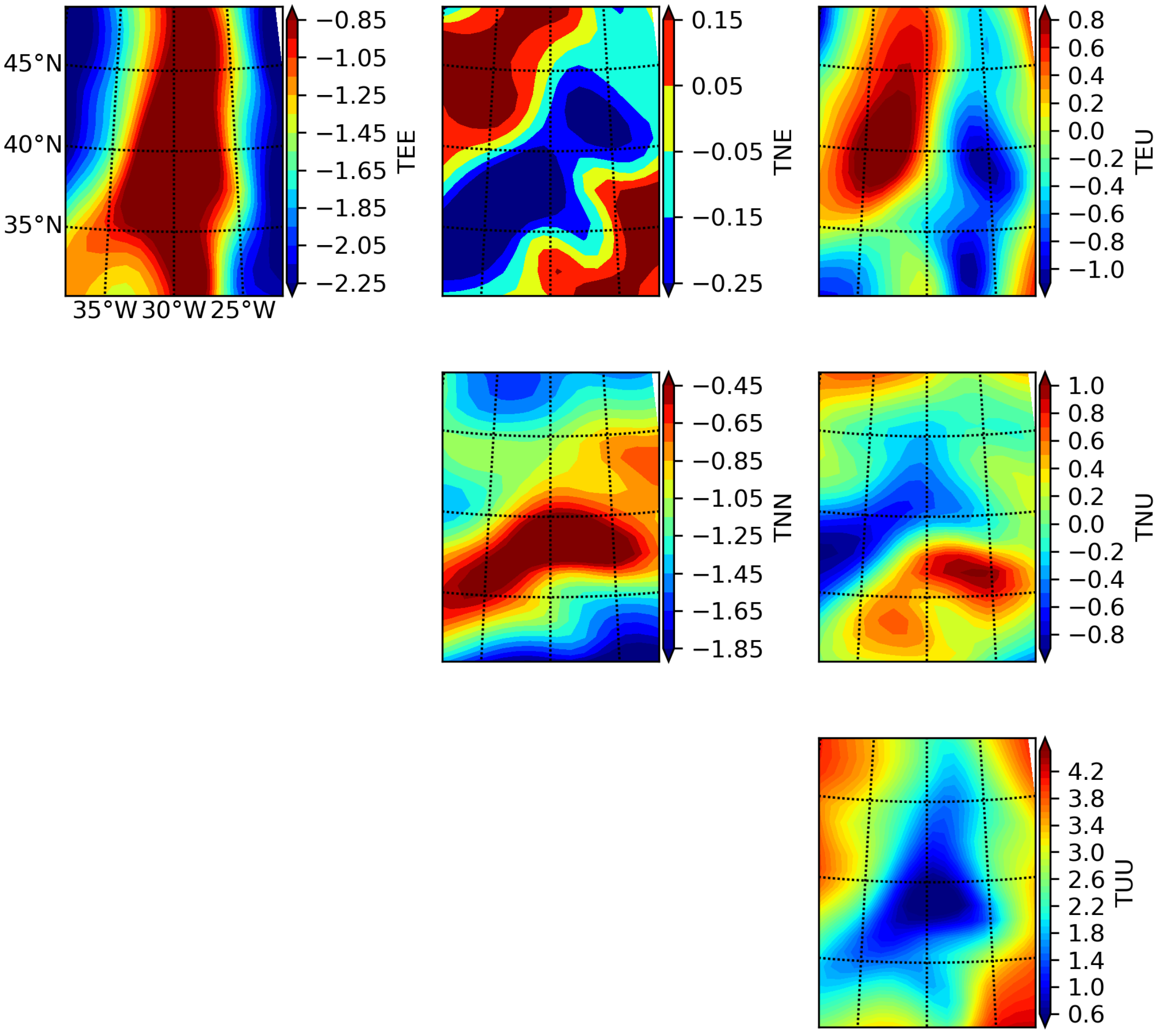

Figure 5.2: Gravity gradient components at 225 km height after topography and sediment correction. Note that the color ranges were chosen in agreement with the theoretical relation between the magnitudes of the gravity components for an isotropic potential (see next section). The East-East component is significantly stronger than expected due to the preferred orientation of geological features in the region.

5.2.3 Model parametrization

The sub-surface is separated into three layers: Crust, lithospheric mantle and sub-lithospheric mantle. We keep the top of the crust fixed to the bottom of the sediment layer (Whittaker et al., 2013) and the bottom of the model to 410 km. The unknowns are then the Moho depth as well as the density within the crust and lithospheric mantle. We chose not to invert for the sub-lithospheric density and also keep the LAB fixed at 120 km.

The starting values of Moho depth and average crustal density come from the interpolation results (chapter 4). v_p velocities were converted to density using the equations of Christensen and Mooney (1995) to obtain starting values for the crustal density. The starting LAB depth was chosen at a constant value of 120 km, while the starting density of the lithospheric mantle is 3200 $\frac{\mathrm{kg}}{\mathrm{m}^3}$, which corresponds to the mantle density of the reference model.

5.2.4 Forward calculation

Each column of the model consists of a crustal and a lithospheric mantle body. The gravity response of each body is calculated using point masses by dividing it into several sub-volumes in such a way that all sub-volumes are smaller than 25 km in every direction. The gravity effect of each sub-volume is then calculated by approximating it as a point mass at its center using the equations of Grombein et al. (2010).

Isostatic topography is calculated in each column by integrating over the density difference with respect to the reference model (Kaban et al., 1999):

$$t_{iso} = \frac{1}{2670 \frac{\mathrm{kg}}{\mathrm{m}^3}} \int\limits_0^{z_{max}} \rho_{ref}(z) - \rho(z)\,\mathrm{d}z$$

Here z_{max} is the bottom of the model at 410 km. However, since in practice the LAB is fixed at 120 km, the topography is effectively only affected by density above 120 km depth.

5.2.5 Probabilistic inversion approach

We collect all unknowns describing the sub-surface in a vector m. The measured quantities (gravity gradients and topography) are collected in two vectors t and g. Note that g and t are not vectors in a physical sense, but are simply collections of numbers. Combining g and t gives the total data vector d. The forward operator F assigns to any model m (given in terms of lithospheric structure and density) the response in terms of the observables: $d_{mod} = F(m)$. The misfit between observed and measured quantities is thus $F(m) - d$.

In addition, the models should be in agreement with *a priori* information. These two requirements can in Bayesian statistics be expressed in terms of two probability distributions: The misfit distribution $\theta(d|F(m))$ and the prior $\pi(m)$. These are combined using Bayes theorem (Gentle, 2009):

$$\mathcal{L}(m) \propto \theta(d|F(m))\pi(m)$$

Here, the $\propto$-symbol signifies proportionality, because normalizing constants that do not affect the relative likelihood of different models have been dropped from Bayes theorem.

The posteriori likelihood function $\mathcal{L}(m)$ can be interpreted in two ways: Either it is simply seen as a measure of model quality or it is seen as an actual probability distribution. In the former case the goal is optimization, i.e. finding a model with maximum $\mathcal{L}$. For example, if θ and π are Gaussian and the forward operator is linear or at least linearizable, this leads to the well-known least-squares optimization problem (Tarantola and Valette, 1982; Tarantola, 2005). In the latter case, the goal is *sampling.* i.e. finding sets of models m that are representative of the posterior probability distribution. Monte Carlo Markov Chain (MCMC) are a powerful and universal technique to generate such samples.

5.2.5.1 The misfit distribution

The misfit distribution θ encapsulates information about measurement uncertainties, potential unmodeled effects and errors of the forward calculation. It can be seen as a conditional distribution dependent on the model m, thus answering the question: How likely is it that the data d would be measured, **if** the model m was in fact the true model?

Conceptually the observed data d can be seen as the outcome of a mixed deterministic-probabilistic process:

$$d = F(m) + \varepsilon,$$

where ε is a random variable, describing all uncertainties. Often ε is assumed to be Gaussian if there is no other information available (Afonso et al., 2013a). This leads to the following misfit distribution θ:

$$\theta(d|F(m)) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} e^{-\frac{1}{2}(d-F(m))^T\Sigma^{-1}(d-F(m))}$$

A common simplification is that the uncertainties affecting d are uncorrelated, which means that Σ is simply a diagonal matrix. However, in the case of gravity gradients and topography uncertainty correlation is likely important. This correlation can occur between values of the same observable in different locations and/or between values of different observables, for example between different gravity gradient components. Under some mathematical assumptions, general statements about these correlations can be made.

5.2.5.1.1 Gravity gradient misfit function

In the following derivation we will use a Cartesian coordinate system. The gravity potential U at a location $\vec{x}$ can be represented in terms of its 2-D Fourier transform $\hat{U}(\vec{k}, z)$ as follows (Blakely, 2009):

$$U(\vec{x}) = \iint \hat{U}(\vec{k}, z) \exp(i\vec{k} \cdot \vec{x}) \, \mathrm{d}\vec{k}$$

Where $\vec{k} = (k_x, k_y)$ is the horizontal wavenumber vector. Now, the Fourier transform $\hat{U}$ can be seen as a random number (Moritz, 1976). If the potential is isotropic and stationary, it follows that the Fourier transform at different wavenumbers is uncorrelated and that the variance only depends on the magnitude of $\vec{k}$, i.e. the radial wavenumber $k_r = |\vec{k}|$.

$$\mathrm{E}[U(\vec{k})U(\vec{k}')] = \delta(\vec{k} - \vec{k}')P(k_r) \tag{5.1}$$

Here, $P(k_r)$ is called the power spectrum .

The covariance of the potential at two different locations, is then simply:

$$cov(U(\vec{x}), U(\vec{x}')) = \mathrm{E}\left[U(\vec{x}) \cdot \overline{U(\vec{x}')}\right] = \mathrm{E}\left[\left(\iint \hat{U}(\vec{k}) \exp(i\vec{k} \cdot \vec{x}) \, \mathrm{d}\vec{k}\right) \left(\iint \hat{U}(\vec{k}') \exp(i\vec{k}' \cdot \vec{x}') \, \mathrm{d}\vec{k}'\right)\right]$$

After some rearranging, and exploiting eq. 5.1, it follows that:

$$cov(U(\vec{x}), U(\vec{x}')) = \iint P(k_r) \exp\left(i\vec{k} \cdot (\vec{x} - \vec{x}')\right) \, \mathrm{d}k_x \, \mathrm{d}k_y$$

Due to the isotropy and stationarity of the problem, a polar coordinate system is now introduced for both the spectral and spatial variables:

$$k_x = k_r \cos\phi, \quad k_y = k_r \sin\phi, \quad x - x' = r\cos\theta, \quad y - y' = r\sin\theta.$$

This leads to:

$$cov(U(\vec{x}), U(\vec{x}')) = \iint P(k_r) \exp\left(ik_r r \cos(\theta - \phi)\right) \, k_r \mathrm{d}k_r \, \mathrm{d}\phi$$

The Jacobi-Anger expansion is given by (Cantrell, 2000, eq. 12.143):

$$\exp(iz\cos(\alpha)) = \sum_{n=-\infty}^{\infty} i^n J_n(z) \exp(in\alpha),$$

where J_n is the n-th Bessel function of the first kind. Inserting this leads to:

$$cov(U(\vec{x}), U(\vec{x}')) = \int_0^{2\pi} \int_0^{\infty} P(k_r) \sum_{n=-\infty}^{\infty} i^n J_n(k_r r) \exp(in(\theta - \phi)) k_r \mathrm{d}k_r \mathrm{d}\phi$$

Exchanging summation and integration and carrying out the integration over ϕ, leads to:

$$cov(U(\vec{x}), U(\vec{x}')) = 2\pi \int_0^{\infty} P(k_r) J_0(k_r r) \, k_r \mathrm{d}k_r$$

This shows that the power spectrum is the Hankel transform of the isotropic covariance function (Debnath and Bhatta, 2015). Thus, the precise shape of the covariance function depends on the power spectrum (see for example Heller and Jordan, 1979) However, some fundamental aspects of the covariance behavior of the gravity gradients derived from the gravity potential can be found without specifying a power spectrum/covariance function for the potential.

Taking a derivative in x is equivalent to multiplying the spectrum by $ik_x = k_r \cos\phi$. Similarly the y-derivative corresponds to multiplication with $ik_y = k_r \sin\phi$. The z-derivative of a potential field is found by multiplying with k_r. As an example, the covariance function between g_{xx} and g_{yy} will be derived:

$$cov(g_{xx}(\vec{x}), g_{yy}(\vec{x}')) = \int_0^{2\pi} \cos^2\phi \sin^2\phi \int_0^{\infty} P(k_r) \sum_{n=-\infty}^{\infty} i^n J_n(k_r r) \exp(in(\theta - \phi)) k_r^5 \mathrm{d}k_r \mathrm{d}\phi$$

In general, the covariance functions between different pairs of gradient components contain different prefactors of powers of $\cos\phi$ and $\sin\phi$.

We turn now to the special case $\vec{x} = \vec{x}'$, or equivalently $r = 0$. Since only $J_0(0) = 1$ is non-zero, the expression simplifies considerably:

$$cov(g_{xx}(\vec{x}), g_{yy}(\vec{x})) = \int_0^{2\pi} \cos^2\phi \sin^2\phi \int_0^{\infty} P(k_r) k_r^5 \mathrm{d}k_r \mathrm{d}\phi \tag{5.2}$$

The integral over k_r is identical for any pairs of gradient components and determines the overall variance of the gradients, while the integral with ϕ determines the correlation between different gradient components. The variances are given in table 5.1, where the variance of g_{zz} is used as reference and set to 1.

Covariance between different components at the same locations only occurs between components on the main diagonal - as could be expected since the trace of the gradient tensor needs to be zero. Specifically, the covariance between g_{xx} and g_{yy} (normalized by dividing by 2π) is $\frac{1}{8}$, which corresponds to a correlation coefficient of $\frac{1}{3}$. The covariance between either horizontal component and g_{zz} is $-\frac{1}{2}$, corresponding to a correlation coefficient of $-\sqrt{\frac{2}{3}}$.

Table 5.1: Variance integrals for all gravity gradient components.

Component	Variance integral	Normalized Variance
g_{xx}	$\int_0^{2\pi} \cos^4 \phi \, \mathrm{d}\phi = \frac{3}{4}\pi$	$\frac{3}{8}$
g_{xy}	$\int_0^{2\pi} \sin^2 \cos^2 \phi \, \mathrm{d}\phi = \frac{1}{4}\pi$	$\frac{1}{8}$
g_{xz}	$\int_0^{2\pi} \cos^2 \phi \, \mathrm{d}\phi = \pi$	$\frac{1}{2}$
g_{yy}	$\int_0^{2\pi} \cos^4 \phi \, \mathrm{d}\phi = \frac{3}{4}\pi$	$\frac{3}{8}$
g_{yz}	$\int_0^{2\pi} \sin^2 \phi \, \mathrm{d}\phi = \pi$	$\frac{1}{2}$
g_{zz}	$\int_0^{2\pi} \mathrm{d}\phi = 2\pi$	1

These results define a misfit distribution that considers the interdependence between the components and uses consistent scaling for the standard deviation of the gravity gradient components. However, this misfit function does not consider the correlation of measurements at different locations.

Alternatively, one can consider the correlation between measurements of the same gravity gradient component at different locations. The crucial approximation here is that we only consider the *isotropic* covariance and neglect the direction dependence. Returning to equation 5.2 one can integrate over θ, effectively averaging out the direction dependence. In the example of the g_{xx} and g_{yy} component this leads to:

$$\frac{1}{2\pi} \int_0^{2\pi} cov(g_{xx}(\vec{x}), g_{yy}(\vec{x}'))\mathrm{d}\theta = \int_0^{2\pi} \cos^2 \phi \sin^2 \phi \int_0^{\infty} P(k_r) J_0(k_r r) k_r^5 \mathrm{d}k_r \mathrm{d}\phi$$

Structurally this expression is identical to the expression for $r = 0$ and the only difference is the zeroth-order Bessel function now appearing in the integral over r. Thus, the isotropic covariance functions for all the gradient components are identical except for scaling factors, which are the same as given above for the case $r = 0$.

The specific shape of the covariance function can be empirically obtained by applying semi-variogram/covariance analysis to gravity gradient data from Bouman et al. (2016). Ten sets of 2000 randomly chosen points on the globe were used to calculate ten different semi-variograms. From this set, the median semi-variance at each distance was derived (see Figure 5.3). The variability of the semi-variograms is relatively small, as given by the 33 % and 66 % percentile.

The empirical semi-variograms confirm the theoretically predicted relative variances of the different gravity components. In addition, all gravity gradients show the same isotropic correlation distance, as predicted by the preceding analysis. The semi-variogram is characterized by two regions. Up to a distance of about 500 km, the semi-variogram increases rapidly and for larger distances it continues to increase until a distance of about 2000 km. Since the overall scale of the Mid-Atlantic test region is only 2000 km, we focus on the shorter correlation distance of 500 km.

For simplicity's sake an exponential covariance function is used, which gives the covariance between two measurements at a separation h:

$$C_i(h) = \sigma_{g,i}^2 \exp(-h/\rho)$$

Where $\rho = 500\,\mathrm{km}$ and $\sigma_{g,i}^2$ is the appropriately scaled variance of gravity gradient component i.

Note that the semi-variogram analysis was applied to free-air gradients, not the topography corrected gradients. The topography introduces even more long wavelength contributions to the gravity field, leading to correlation distances of 1000 km or more. However, we chose not to use this longer correlation distance, since the long wavelength structure induced by the topography is to a large degree compensated by the isostatic or nearly isostatic compensation of topography.

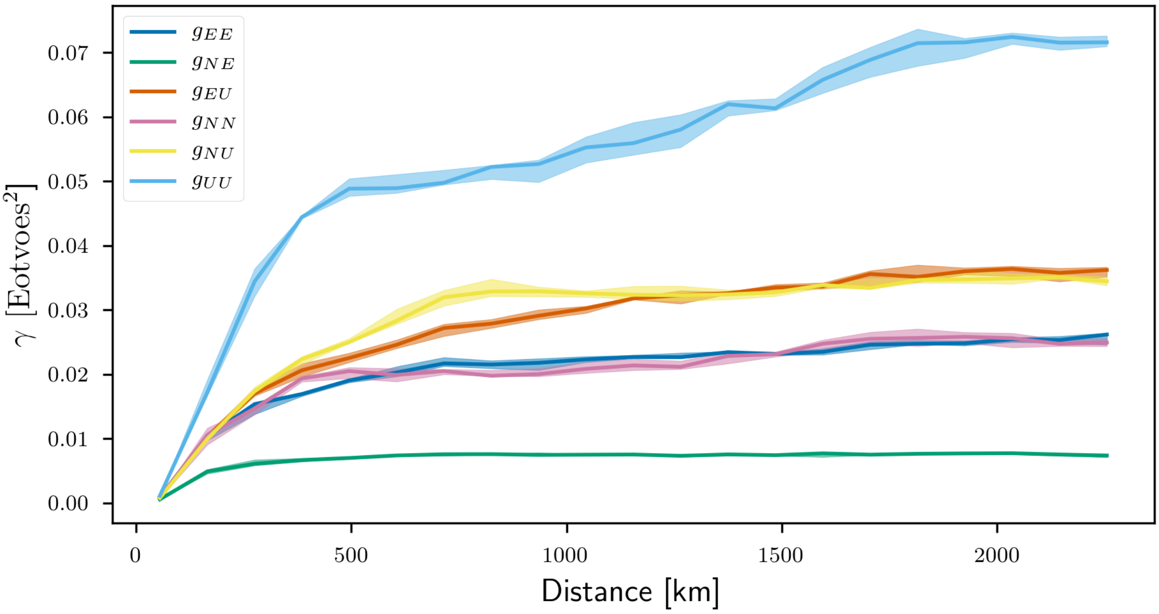

Figure 5.3: Global semi-variogram analysis of free-air gravity gradients by Bouman et al. (2016). The semi-variograms were calculated based on ten randomly generated sets of 2000 points on the globe each. The solid lines give the median over the set of ten semi-variograms, while the shaded area is limited by the 33% and 66% percentile. The observed variances for the different gravity gradient components agree with theoretical expectation. In addition, the correlation distances of all components are similar (c. 500 km).

Thus, the actual goal of the inversion should be to fit the smaller features in the free-air anomaly and not only the large scale features originating from topography.

5.2.5.1.2 Topography misfit function

Isostatic topography only depends on the sub-surface parameters inside a single vertical column. Thus, observations of isostatic topography at different locations are not intrinsically correlated. Therefore, an uncorrelated Gaussian is used for the misfit distribution of topography.

5.2.5.1.3 Data weighting using hierarchical approach

In all joint inversions, different data sets need to be weighted against each other. In the case of a Bayesian approach, this is done by determining the variances of the misfit distribution. These variances correspond to data uncertainties, which reflect both measurement and modeling uncertainties. For instance, there might be contributions from the mantle, which should not be fit by the crustal or lithospheric structure.

The Bayesian formulation can be extended to a hierarchical model that includes data uncertainties as an unknown parameter (Bodin et al., 2012). Separate standard deviations are specified for topography (σ_t) and the gravity gradients (σ_g). Thus, the misfit distribution θ is now conditional on the variances σ in addition to the model m. The posterior likelihood then becomes:

$$\mathcal{L}(m, \sigma) \propto \pi(m)\, \theta(d|m, \sigma)\, \Pi(\sigma)$$

Here $\Pi(\sigma)$ is the hyper-prior given to the standard deviations. In the absence of any preferred level of accuracy, both data uncertainties are given uniform priors in the range $[0.1, 1.0]$ (in the units Eotvoes and km respectively).

5.2.5.2 Prior distribution

Information about the crustal structure from the Kriging interpolation (chapter 4) is used into the inversion algorithm. Specifically, Moho depth and average crustal density are given multivariate Gaussian prior distributions, with the mean and covariance matrix derived from the interpolation result. The average P-wave velocity is converted into density using the following empirical velocity-density relation, which is derived by regressing the average crustal densities and velocities from Crust1.0 (Laske et al., 2013):

$$\rho \left[\frac{\mathrm{kg}}{\mathrm{m}^3}\right] = 340\, v_p \left[\frac{\mathrm{km}}{\mathrm{s}}\right] + 590$$

The resulting Moho depth and crustal density is given in Figure 5.4a and b.

The uncertainty distribution derived from Kriging is given in Figure 5.5. Note that the uncertainty distribution of the crustal density is almost constant, due to strong short-distance variation apparent in the P-wave velocity database (see chapter 4). The Moho depth uncertainty shows more structure, clearly reflecting the location of individual data points as areas of reduced uncertainty.

Kriging also gives the correlation of uncertainty at different locations. However, covariance matrices are notoriously difficult to visualize. To give an impression of the spatial correlation, isolines of the correlation coefficient for different points are used. Here, a correlation coefficient of 0.7 is used to show correlated areas. The correlation distance of the Moho depth uncertainties (Figure 5.5) is related to the uncertainty: Areas with high uncertainty are also characterized by

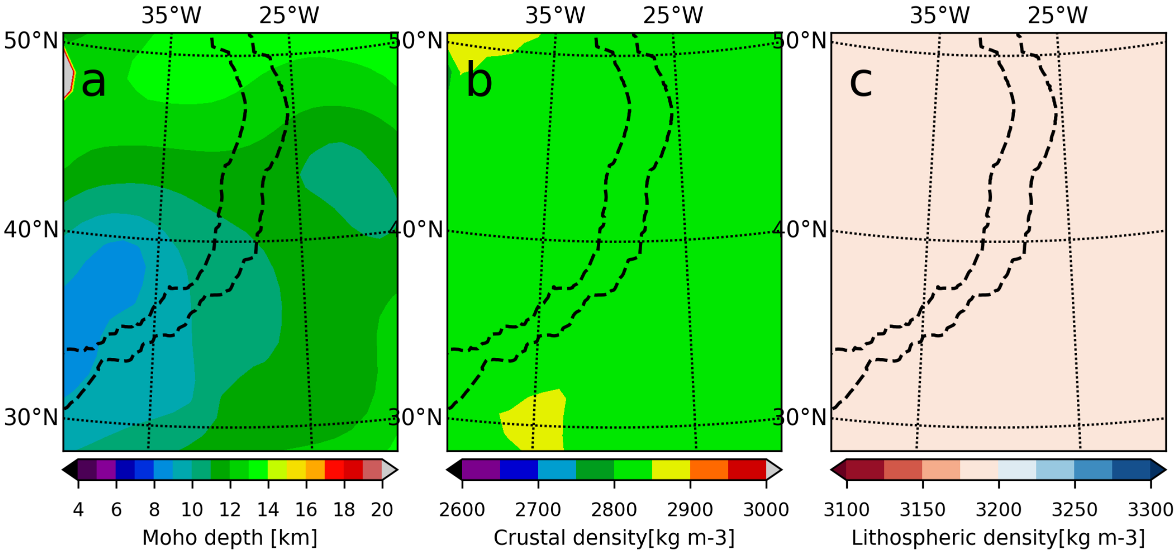

Figure 5.4: Starting values for inversion. The dashed line is the 10 Ma isochron (Müller et al., 2008) to highlight the location of the mid-oceanic ridge. a: Moho depth, b: Crustal density, c: Lithospheric density.

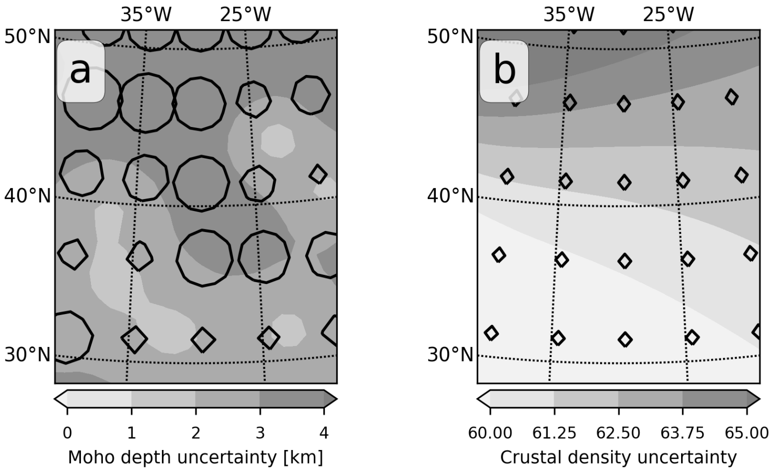

Figure 5.5: Uncertainty and uncertainty correlation of Moho depth and average crustal density from Kriging interpolation. The black lines correspond to a correlation coefficient of 0.7 to a point in the middle of the respective circle. a: Moho depth, b: Crustal density

large correlation distances. Crustal density uncertainties are spatially nearly uncorrelated due to the strong small-scale variations.

The average mantle lid density is given a uniform prior in the range $[3100, 3300]\frac{\mathrm{kg}}{\mathrm{m}^3}$, which is a variation of $\pm 100\frac{\mathrm{kg}}{\mathrm{m}^3}$ relative to the reference density. Assuming a thermal expansivity $\alpha = 3.3 \cdot 10^{-5}$, this corresponds to temperature variations $\Delta T = \frac{\Delta\rho}{\rho_0 \alpha} \approx \pm 1000\,K$, which easily allows for the required variations at a typical mid-oceanic ridge (Turcotte and Schubert, 2002).

5.2.5.3 MCMC algorithm

We mainly use a standard Metropolis-Hastings algorithm (Gentle, 2009), which is augmented by sampling the hyperparameters. The algorithm can be outlined as follows (see also Bodin et al., 2012):

- Step zero: Set the current model m and the current hyperparameters σ to its starting values.
- Step one: With equal likelihood either randomly change one model parameter **or** one hyper-parameter (i.e. one of the error standard deviations)
  - Model change: Randomly choose a single sub-surface parameter (Moho depth, LAB, crustal density, mantle lid density) at a single location and add a random perturbation to it. Recalculate the gravity and topography predicted by the new model m'.
  - Hyper parameter change: Randomly change either the gravity gradient or topography standard deviation by adding a random perturbation to it. This gives σ'.
- Step two: Evaluate the posterior likelihood function with the new parameters, giving $\mathcal{L}(m', \sigma')$.
- Step three: Calculate the ratio α of the old and new posterior likelihood $\alpha = \mathcal{L}(m', \sigma')/\mathcal{L}(m, \sigma)$.
- Step four: Accept the new values m' and σ' with probability α. (Note that if $\alpha \geq 1$ the new values are always accepted). Return to step one.

The algorithm terminates when a specified number of iterations has been reached.

The random perturbation to the variables are drawn from a normal distribution centered on zero. Each parameter has an individual standard deviation, which should be tuned to give an optimal acceptance ratio. The acceptance ratio is the fraction of accepted model changes. Roberts et al. (1997) show that an acceptance ratio of 0.234 leads to optimal efficiency for multidimensional sampling problems under quite general conditions. After numerous experiments, a proposal standard deviation of 5 km was chosen for the Moho depth, whereas crustal and lithospheric densities have proposal standard deviations of 150 $\frac{\mathrm{kg}}{\mathrm{m}^3}$ and 20 $\frac{\mathrm{kg}}{\mathrm{m}^3}$.

Since only a single model parameters is changed per iteration and only about 1 in 4 changes are accepted, we only save one in every 1000 models to reduce memory storage consumption

5.3 Synthetic test

The probabilistic inversion procedure described above contains many tunable parameters. A synthetic test gives some insight into how these parameters affect the inversion result.

5.3.1 Synthetic model generation

The synthetic model is generated by exploiting that the interpolated Moho depth and crustal density are described as multivariate Gaussian distributions with known mean vector μ and covariance matrix $\mathbf{\Sigma}$. Samples can be drawn from a multivariate Gaussian by calculating the singular value decomposition of $\mathbf{\Sigma} = \mathbf{USV^T}$. If ϕ is a vector of n independent standard normal distributed numbers, a sample $\mathbf{x}$ of the correlated distribution can be generated as:

$$\mathbf{x} = \mu + \mathbf{VS}^{\frac{1}{2}}\phi$$

The density structure in the mantle lithosphere is found by enforcing isostatic balance with respect to the real topography (Figure 5.6). The gravity gradient and topography (which is the same as real topography) responses of the synthetic model are calculated and used for further inversion. No artificial noise was added to the observables, since we found that the inherent trade-offs with such a limited set of observables was already leading to an incomplete recovery of the synthetic model.

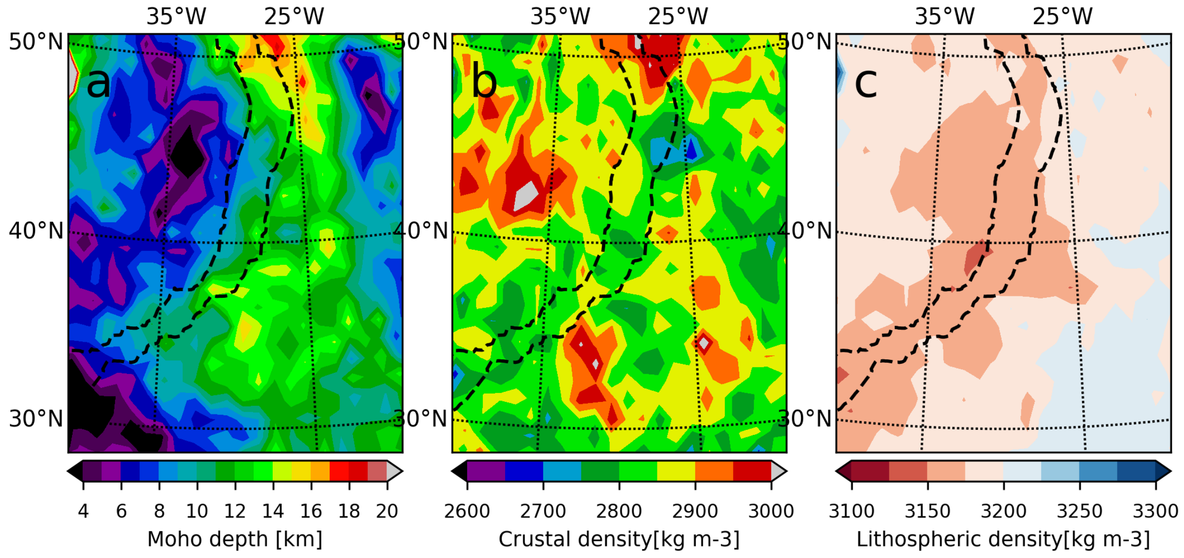

Figure 5.6: Synthetic lithosphere generated by adding a random perturbation to the Kriging starting Moho depth and crustal density. The size and spatial pattern of the perturbation agrees is based on the uncertainty covariance matrix from the Kriging interpolation. Density inside the lithospheric mantle is found by enforcing isostatic balance with respect to real topography. a: Moho depth, b: Crustal density, c: Density of the lithospheric mantle.

5.3.2 Inversion results

The inversion was run with and without a correlated misfit distribution for the gravity gradients, while all other inversion parameters were the same. The importance of properly weighting the gravity gradient components can be seen by the development of the hyperparameters (Figure 5.7). When the gravity gradients are seen as uncorrelated observations (Figure 5.7a), the inversion quickly attempts to fit the gradients well, but the topography is fit poorly (±500 m), even after 1,000,000 iterations. In contrast, when the correlation between gradient observations is explicitly considered (Figure 5.7b), both standard deviations decrease, although the gradient misfit still decreases more rapidly.

Apart from the weighting between the two data sets, the correlated misfit distribution also seems to affect the weighting between prior information and data. In the uncorrelated case, σ_g quickly reaches the minimum value of 0.1 E, but in the correlated case it converges to a value of c. 0.2 E. No artificial noise was added to the gradient data. Thus, the imperfect fit of the gravity gradients also reflects the influence of the prior, which otherwise was completely overwhelmed by the gradient data.

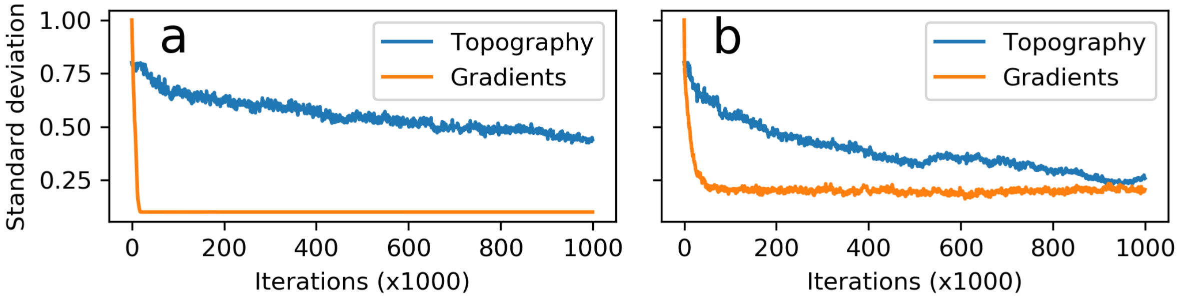

Figure 5.7: Evolution of the hyper parameters during inversion of synthetic data for different misfit distributions of the gravity gradients. a: Gradients are seen as spatially uncorrelated and all gradient components are fit up to the same accuracy. Note that the standard deviation of the topography decreases much more slowly than the gradients. b: Spatial correlation measurements of the same gravity gradient component at different locations is taken into account. The standard deviation of both topography and gradients decrease.

The two different misfit distribution also have a strong effect on the inverted model results. The mean of the 100,000 last models is used to represent the inversion result. Overall, neither inversion is capable of reproducing key features of the synthetic crustal structure (Figure 5.8a-f). In fact, the inversion results do not come closer to the synthetic input model than the starting model in terms of the RMS difference.

The lithospheric density structure is recovered more accurately with the correlated distribution (Figure 5.8g,h,i). However, with both misfit distributions density variation in the lithospheric mantle is subdued compared to the synthetic model, because the lithospheric anomalies are partially placed in the crustal density or in crustal thickness.

The synthetic tests show that the crustal structure is difficult to determine, but that the lithospheric density can be recovered reliably. Shorter wavelength gravity data might improve the recovery. For the oceans such data are available from satellite altimetry at resolution of a few arc minutes (Smith and Sandwell, 1997; Andersen et al., 2010)). Furthermore, the synthetic density variations in the crust have a very short wavelength and only have a limited impact on the observables due to the thin crust. Thus, recovering crustal density in this instance is always challenging.

5.4 Application to real data

We applied the inversion algorithm to the real data for the northern mid-Atlantic region. Based on the findings from the synthetic test, we use a spatially correlated misfit distribution. The convergence behavior of the algorithm (Figure 5.9) is similar to the synthetic case. The standard deviations converge to a minimum of about 0.2 E and km respectively after about 750,000 iterations. However, even after the hyper parameters have stabilized, the likelihood continues to

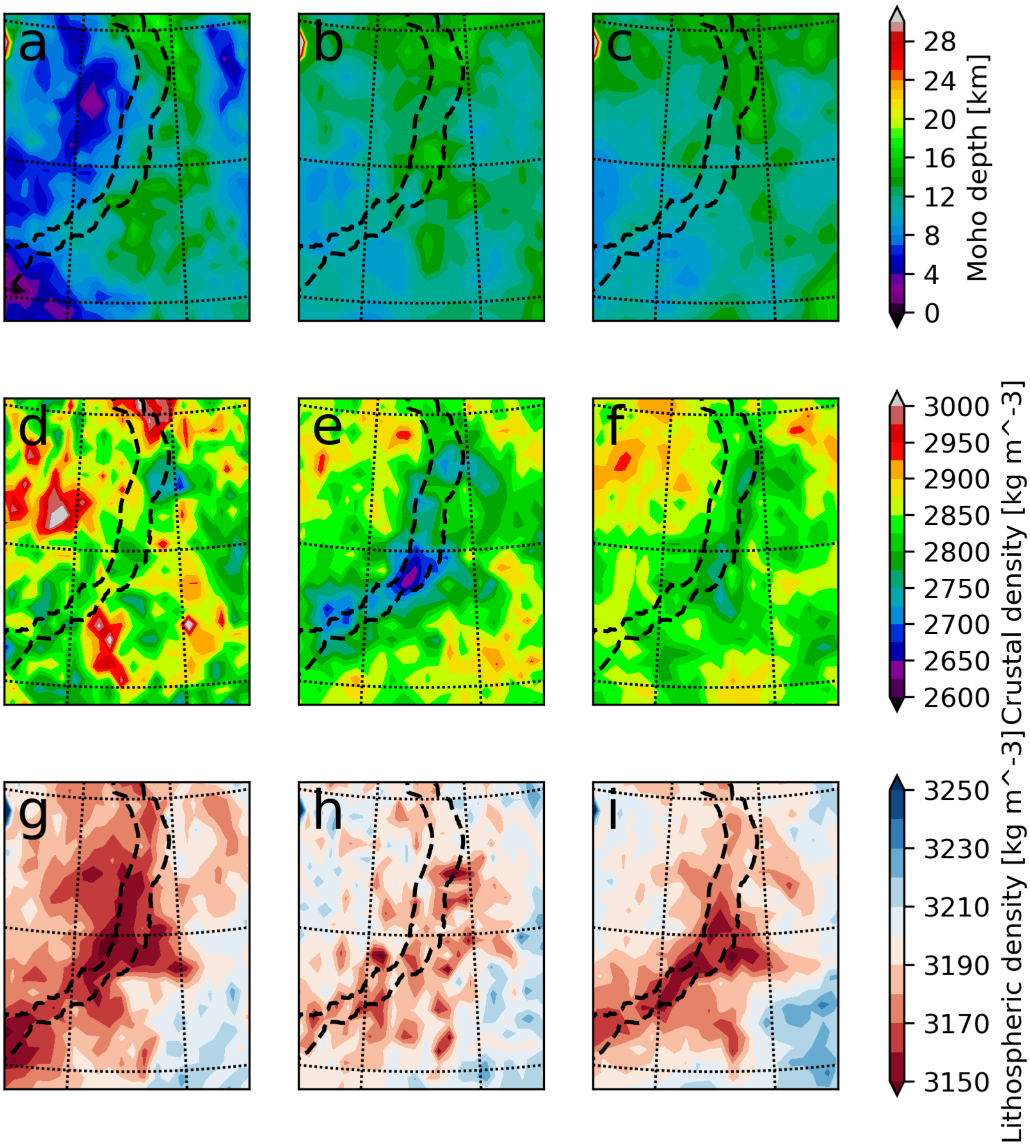

Figure 5.8: Comparison of the synthetic input model with the inversion results with and without consideration of the spatial correlation of gravity gradient components. The inversion result is the mean over the last 100,000 elements from the Markov chain. a,b,c: Moho depth: Synthetic input, uncorrelated inversion result, correlated inversion result. d,e,f: Crustal density. g,h,i: Density of the lithospheric mantle.

increase. The acceptance rate for the different sub-surface parameters decreases as the algorithm progresses, particularly for the mantle density, but stabilize at around 0.2 to 0.3.

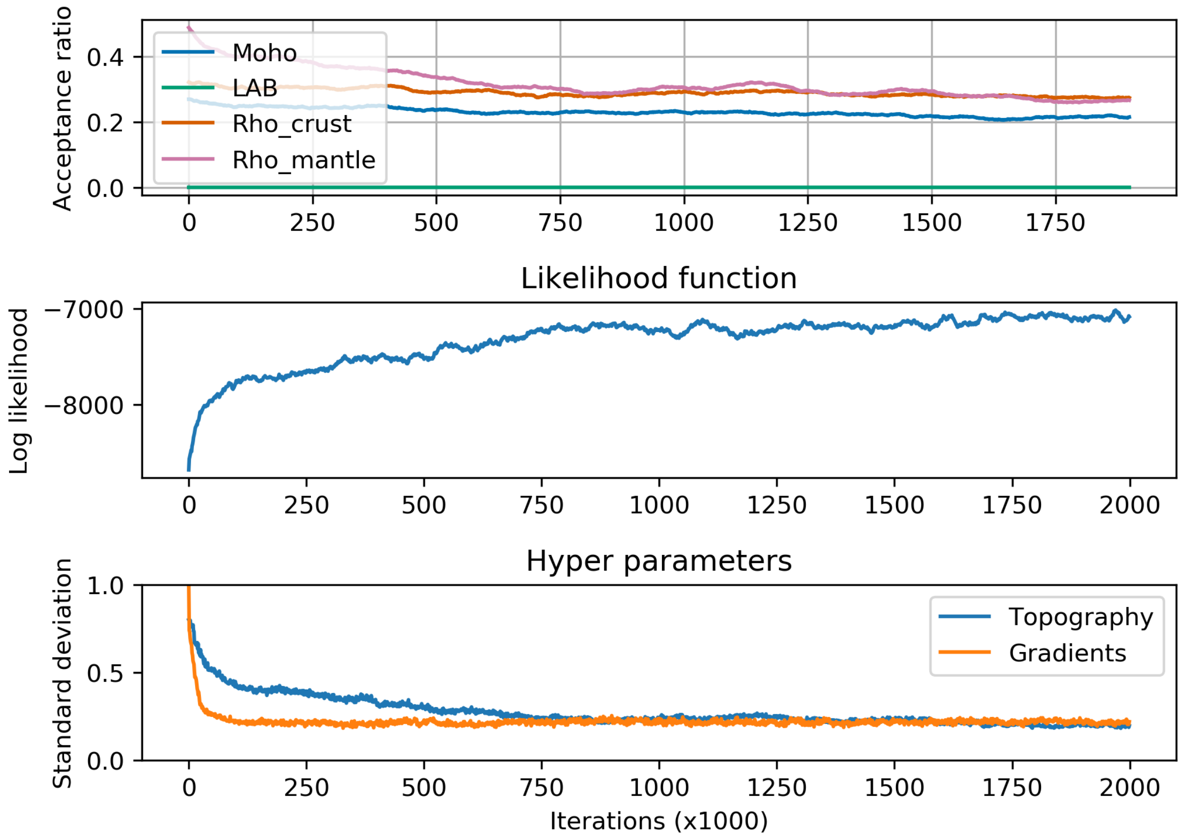

Figure 5.9: Convergence behavior of the probabilistic inversion applied to real data.

The average of the last 500,000 (Figure 5.10) illustrates the inverted model structure. Crustal thickness (Figure 5.10a) is on average around 10 km. Interestingly, thickened crust (12-15 km) is observed around the region affected by the Azores hotspot, broadly agreeing with results from receiver functions (Spieker et al., 2018). Many other variations in crustal thickness are more erratic and highly dubious, particularly spots of very thin crust appearing away from the spreading ridge. Possibly these reflect an incorrect sediment thickness or density. In addition, they are near the edge of the region covered with gravity data (30 – 50 °N, 20 – 40 °W) and thus may not be well resolved.

The recovered crustal density structure is relatively homogeneous (Figure 5.10b). This casts doubt on the strong small-scale variability obtained from Kriging. Likely these small-scale variations are largely due to P-wave velocity measurement errors. Furthermore, the velocity-density conversion might be inaccurate, since the inverted densities are significantly higher than the starting values (see also Brocher, 2005; Godfrey et al., 1997). Overall, the inversion provides no convincing case for anything but a nearly constant crustal density.

Secular cooling of the oceanic lithosphere is clearly resolved by the inversion. The mid-Atlantic ridge is characterized by a negative anomaly of about 50 $\frac{\mathrm{kg}}{\mathrm{m}^3}$ (equivalent to a temperature anomaly of $\approx +500\,K$). As the age of the plate increases, the lithospheric density increases to a maximum positive anomaly of around 50 $\frac{\mathrm{kg}}{\mathrm{m}^3}$.

The area of the Azores and the Great Meteor seamount chain display clear negative density anomalies in the lithospheric mantle. In addition, the density anomaly in the segment of the mid-Atlantic ridge between ~35 – 40 °N has stronger negative density anomalies than the rest of the mid-Atlantic ridge, which would be consistent with ridge-plume interaction (Gente et al., 2003).

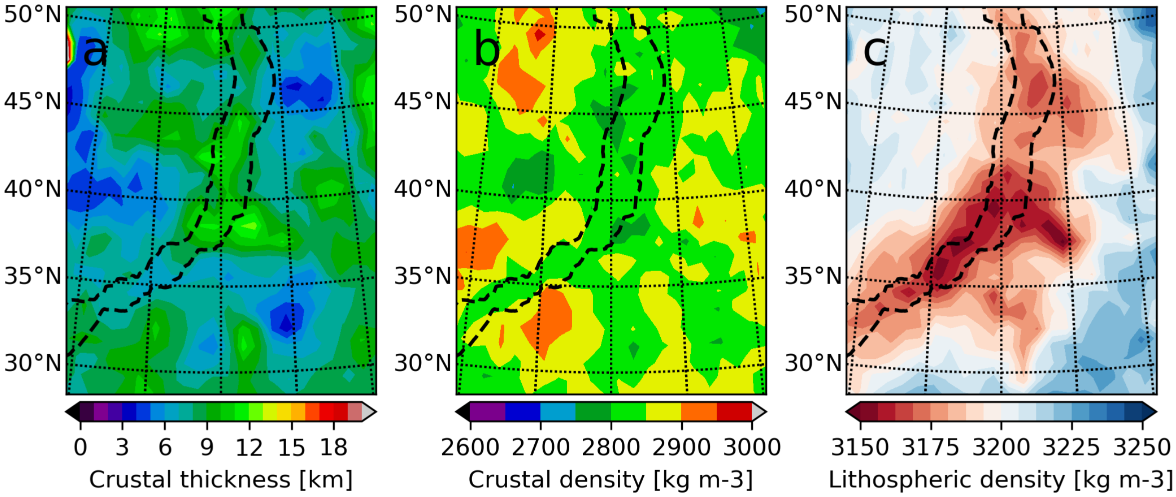

Figure 5.10: Inversion result (mean of the last 500,000 models). The dashed lines represent the 10 Ma isochron. a: Inverted **Crustal thickness** b: Average crustal density. c: Density inside the lithospheric mantle.

5.5 Discussion

The hierarchical approach, where the data uncertainty is itself treated as unknown, proved to be promising for joint inversion, as found previously (Bodin et al., 2012). Joint inversions with more observables might benefit particularly from this technique. Especially, since the hierarchical approach can also expose issues with individual observables, such as the importance of including the spatial correlation of the gravity gradient components.

The benefit of using the complete gravity gradient tensor might be doubtful, as all entries of the tensor are simply derivatives of the potential. However, when the derivatives are converted into their representation in terms of a covariance function, only the entries of the main diagonal are found to be intrinsically correlated at the same measurement location. This suggests that using the full gravity gradient tensor might benefit the inversion.

The behavior of the inversion is strongly influenced by the spatial correlation due to the long wavelength nature of satellite gravity gradients. This analysis could be extended by including anisotropy and the lower order gravity functionals (gravity field and geoid) in order to find an optimal way to combine different gravity data.

Using Kriging interpolation to derive a prior distribution in the crust might lead to a better recovery of lithospheric structure. However, this depends on the accuracy of the starting model and also the accuracy of the uncertainty estimates. In other words: What is the error of the error? In chapter 4 cross-validation showed that uncertainty estimates are overall reliable, but this does not preclude unrealistic error estimates in a particular area. Thus, the interpolation results should not be trusted blindly, but they still represent an improvement compared to ad-hoc error estimates. The interpolation could be extended to pointwise estimates of LAB depth from receiver functions (e.g. Rychert et al., 2010). However, it is not always clear that LAB definitions of different methods agree (see Eaton et al., 2009; Fischer et al., 2010; Steinberger, 2016).

In principle, the reasonable recovery of the lithospheric density structure with a limited data set is promising. However, it should be noted that the central role of the lithospheric mantle has long been known. Famously the “square root of age law” was deduced from the half-space cooling model based on heat flow measurements (Turcotte and Schubert, 2002). Indeed, the topography alone would suggest a strong mantle contribution, since crustal structure is mainly dependent on mantle potential temperature at the ridge axis (McKenzie and Bickle, 1988) and hence cannot be invoked to explain the sea floor subsidence with age.

Furthermore, the mainly thermal origin of the density anomalies also shows a weakness of the chosen sub-surface parametrization. At the ridge axis the entire lithospheric column might be approximated as a negative density anomaly with respect to the reference density. However, as the lithosphere cools, the top eventually becomes **heavier** than the reference, while the lower parts are still **lighter**. Even if the LAB is allowed to vary during the inversion, this distribution cannot be adequately described. A possible indication, that this is an issue, is that the inversion results show some tendency towards crustal thinning with greater age (Figure 5.10a). Effectively the thin crust acts as a positive density anomaly at shallow depth. Possibly the inversion is thus trying to emulate the cooling by altering crustal structure. Note that in integrated thermo-chemical modeling, the thermal field is assumed to be stationary (Zeyen and Fernàndez, 1994; Fullea et al., 2009), which is a valid assumption for stable continents (Artemieva and Mooney, 2001) but certainly not for the oceans (McKenzie et al., 2005).

5.6 Conclusions and outlook

In this chapter a proof of concept for a probabilistic inversion algorithm for gravity gradients and topography was developed. The main features of the inversion are (i) that a hierarchical approach is used to co-estimate the data uncertainty, (ii) that the specific spatial characteristics of gravity gradients are taken into account and (iii) that the prior distribution is based on Kriging interpolation.

Application of the algorithm to the northern mid-Atlantic demonstrated that the crustal thickness and density are difficult to resolve with this data set. The lithospheric density distribution related to the cooling of the lithosphere was resolved more reliably.

From a more technical point of view, it was found that extending the classical Monte-Carlo-Markov-Chain technique with variable data uncertainty and an estimate of data covariance improved the behavior of the algorithm. These findings could be transferred to integrated thermo-chemical inversion (i.e. Afonso et al., 2013b), particularly where many different observables are used and weighing these observables is non-trivial.

A possibly promising extension of the hierarchical approach would allow representing deep-earth contributions without explicitly modeling them. Deep-earth contributions have long wavelengths and are characterized by a different ratio of gravity to topography than isostatic contributions (see for example the kernels in Colli et al., 2016; Molnar et al., 2015). Hence, a probabilistic term representing deep-earth contributions could be added to the data-generation process:

$$d = F(m) + \varepsilon + D,$$

where D represents unmodelled effects coming from below the model. The simplest way to parametrize D is to use a correlation distance to describes the spatial structure. Two different standard deviations are used for the magnitude of gravity and topography effects. The three parameters describing D are then treated as hyper-parameters and altered during the runs of the MCMC algorithm.

The range of these hyper-parameters can be limited based on geostatistical and spectral analysis of seismic tomography models (i.e. Becker and Boschi, 2002), because this describes the spatial structure of the source. When combined with the Green's functions of the topography and gravity response to density anomalies in the mantle (Richards and Hager, 1984), reasonable bounds on the deep-earth contribution to gravity and topography can be derived. Note that the spectral representation in terms of spherical harmonics and the statistical representation in terms of a covariance function are directly related (Jones, 1963). Thus, one is free to choose the more practical representation for the specific task.

Ultimately, a true joint model consisting of the entire crust and mantle would be preferable. However, the parametrized representation of the deep-earth contribution might still be useful as an intermediate step towards an integrated model.

6 Summary and concluding remarks

In this thesis, I investigated the use of gravity, magnetic and seismic data to constrain lithospheric structure. The three relatively divergent methods are linked through their common sensitivity to lithospheric properties, particularly density.

The primary outcomes of each chapter are:

- Analysis of gravity forward calculation based on a topographic-isostatic model leads to the recommendation that, in order to minimize far-field effects in regional density modeling, topographic and isostatic masses should always be considered together. This can be easily achieved in practice by extending the model area a few 100 km around the area covered by gravity measurements. This approach works for ground and satellite data as well as for gravity and gravity gradients without increasing the computational complexity. It does mean however, that the topographic effect needs to be recalculated for each study region. Thus, a specific Bouguer anomaly map is tailored to one study region and global Bouguer anomaly maps (such as Balmino et al., 2012) should not be used directly for regional density modeling, if accurate long wavelength results are desired.

- The comparison of susceptibility inversion results obtained using different model geometries suggests that satellite magnetic data are not on their own capable of constraining the thermal structure inside the lithosphere. This limited sensitivity is due to the strong trade-offs between susceptibility and thickness variations and the relatively wide range of possible susceptibility values in the crust. In addition, the longest wavelengths (>~2500 km) of the lithospheric magnetic field are hidden by the core field. As a result, the susceptibility and thickness structure at that scale is likewise inaccessible using magnetic data alone.

- The crustal model obtained using kriging interpolation demonstrates that a reasonable global crustal model can be derived using minimal manual input. While the resulting crustal model is not necessarily geologically plausible in all regions, it follows the existing data base very closely. In addition, uncertainty and uncertainty covariance estimates can be obtained using this technique. Statistical analysis of the uncertainty estimates suggests that the estimated uncertainties are overall reliable, i.e. neither overly optimistic nor pessimistic.

- Probabilistic joint inversion of satellite gravity gradients and topography (assuming isostatic balance) shows that simple lithospheric models can be derived using this limited data set. The Bayesian formulation that underlies the probabilistic inversion technique affords sufficient flexibility to consider many types of additional information. This way, uncertainty information about the crustal structure, fundamental mathematical relations between gravity gradient measurements and variable uncertainties can be included.

6.1 Isostatic and dynamic topography

There is no reason to assume that Earth's topography is compensated purely isostatically or dynamically. After all, isostasy requires a dynamic Earth, since otherwise there would be no *mechanism* to explain how isostatic balance is attained. However, it remains contentious how much the two mechanisms contribute to topography at different wavelengths.

The results of the joint inversion (chapter 5) would limit the dynamic contribution in the mid-Atlantic region to about 200 m, which corresponds to about 5 % of total topographic variability. However, the investigated region only has a size of 2000 km, which might not be sufficient to detect the larger scale dynamic topography signals, which concentrate in the first 20 spherical harmonic degrees (Colli et al., 2016). In addition, the lithospheric mantle density was effectively unconstrained, so most topography could be accommodated using the lithospheric density. However, if these density variations were completely incorrect and most topography was actually supported dynamically, simultaneously fitting the gravity field and topography would not have been possible. Thus, the findings from chapter 5 agree with the "few hundreds of meters" predicted by Molnar et al. (2015).

6.2 Possibilities for integrated lithospheric modeling with satellite data

The benefit of integrated modeling of several data sets has been amply demonstrated (c.f. Afonso et al., 2016a). How could the methodologies developed in this thesis be linked to existing and future efforts of integrated lithospheric modeling? There are at least three possible aspects to consider: (i) the parametrization (i.e. description) of the crust and mantle, (ii) the observables used during the inversion and (iii) the spatial extent of the model volume.

When mainly using satellite data, it might be sufficient to only explicitly consider the temperature structure (Fullea.2007; Afonso et al., subm.) instead of the full thermochemical structure. Temperature exerts first order control on all rock properties and is strongly reflected in surface heat flow (Jaupart and Mareschal, 1999). In addition, global data coverage might be insufficient to reliably resolve chemical trends in the mantle. Instead, the impact of composition on properties could be included into the prior range of the properties (e.g. seismic velocity, density). In this way, the effect of composition is taken into account, without requiring explicit thermodynamic modeling.

In principle, including satellite magnetic data into thermochemical inversion might improve both the understanding of the lithospheric magnetic field and the temperature structure of the lithosphere. The magnetic field inversion might benefit from a better separation of thickness from susceptibility variations, because the temperature structure is better constrained. The thermochemical inversion would gain an additional observable, that is not related to the typical observables, such as the gravity field.

However, the specific sensitivities of the magnetic field might limit joint inversion with other observables. Rock magnetism depends on the presence of trace amounts of magnetizable phases like magnetite (Reynolds et al., 1990). These phases hardly affect the bulk properties of the rock, since they only make up a few vol % of the rock. While the stability of these magnetic phases does depend on temperature, pressure and iron content, it also depends on oxidation conditions (Pilchin and Eppelbaum, 1997) and titanium content (Reynolds et al., 1990), which do not have a major impact on the main phase assemblage. Thus, at least two new subsurface parameters need to be introduced (e.g. oxygen fugacity and Ti-content), to predict one observable (e.g. the magnetic field). These added degrees of freedom might lead to stronger trade-offs in an already non-unique system (Afonso et al., 2013a). Subsequently, the joint inversion might fragment into several – effectively independent – single inversions, because the link between the different observables is too weak.

In a wider sense, integrated modeling also concerns the vertical extent of the modeled volume, because mantle convection and lithospheric structure are often treated separately. In mantle convection studies, at least the upper 150 km of the Earth are typically ignored (e.g. Steinberger, 2007). In lithospheric studies the contributions from the deep earth are typically removed using wavelength filtering (e.g. Afonso et al., 2013b). Often the work of Bowin (2000) is cited to justify a specific cut-off wavelength in relation to a target source depth. However, as was often pointed out (e.g. Root et al., 2015), sources from all depths overlap in spectral domain. Thus, a more comprehensive approach would be to jointly model the complete crust-mantle system.

Bibliography

Afonso, J. C., M. Fernàndez, G. Ranalli, W. L. Griffin, and J. A. D. Connolly (2008). "Integrated geophysical-petrological modeling of the lithosphere and sublithospheric upper mantle: Methodology and applications". In: *Geochemistry, Geophysics, Geosystems* 9.5. ISSN: 15252027. DOI: 10.1029/2007GC001834.

Afonso, J. C., J. Fullea, W. L. Griffin, Y. Yang, A. G. Jones, J. A. D. Connolly, and S. Y. O'Reilly (2013a). "3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. I: A priori petrological information and geophysical observables". In: *Journal of Geophysical Research: Solid Earth* 118.5, pp. 2586–2617. ISSN: 21699313. DOI: 10.1002/jgrb.50124.

Afonso, J. C., J. Fullea, Y. Yang, J. Connolly, and A. G. Jones (2013b). "3-D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. II: General methodology and resolution analysis". In: *Journal of Geophysical Research: Solid Earth* 118.4, pp. 1650–1676. ISSN: 21699313. DOI: 10.1002/jgrb.50123.

Afonso, J. C., M. Moorkamp, and J. Fullea (2016a). "Imaging the Lithosphere and Upper Mantle". In: *Integrated Imaging of the Earth*. Ed. by M. Moorkamp, P. G. Lelièvre, N. Linde, and A. Khan. Hoboken, NJ: John Wiley & Sons, Inc. ISBN: 9781118929063.

Afonso, J. C., N. Rawlinson, Y. Yang, D. L. Schutt, A. G. Jones, J. Fullea, and W. L. Griffin (2016b). "3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central U.S". In: *Journal of Geophysical Research: Solid Earth* 121.10, pp. 7337–7370. ISSN: 21699313. DOI: 10.1002/2016JB013049.

Afonso, J. C., F. Salajegheh, W. Szwillus, J. Ebbing, and C. Gaina (subm.). "A global reference model of the lithosphere and upper mantle from joint inversion and analysis of multiple data sets". In: *Journal of Geophysical Research: Solid Earth*. ISSN: 2169-9313.

Airy, G. B. (1855). "On the Computation of the Effect of the Attraction of Mountain-Masses, as Disturbing the Apparent Astronomical Latitude of Stations in Geodetic Surveys". In: *Philosophical Transactions of the Royal Society of London* 145.0, pp. 101–104. ISSN: 0261-0523. DOI: 10.1098/rstl.1855.0003.

Amante, C. and B. W. Eakins (2009). *ETOPO1 Global Relief Model converted to PanMap layer format*. DOI: 10.1594/PANGAEA.769615.

Ambikasaran, S., J. Y. Li, P. K. Kitanidis, and E. Darve (2013). "Large-scale stochastic linear inversion using hierarchical matrices". In: *Computational Geosciences* 17.6, pp. 913–927. ISSN: 1420-0597. DOI: 10.1007/s10596-013-9364-0.

Andersen, O. B., P. Knudsen, and P. A. M. Berry (2010). "The DNSC08GRA global marine gravity field from double retracked satellite altimetry". In: *Journal of Geodesy* 84.3, pp. 191–199. ISSN: 0949-7714. DOI: 10.1007/s00190-009-0355-9.

Artemieva, I. M. and W. D. Mooney (2001). "Thermal thickness and evolution of Precambrian lithosphere: A global study". In: *Journal of Geophysical Research: Solid Earth* 106.B8, pp. 16387–16414. ISSN: 01480227. DOI: 10.1029/2000JB900439.

Assumpção, M., M. Feng, A. Tassara, and J. Julià (2013). "Models of crustal thickness for South America from seismic refraction, receiver functions and surface wave tomography". In: *Tectonophysics* 609, pp. 82–96. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2012.11.014.

Balmino, G., N. Vales, S. Bonvalot, and A. Briais (2012). “Spherical harmonic modelling to ultra-high degree of Bouguer and isostatic anomalies”. In: *Journal of Geodesy* 86.7, pp. 499–520. ISSN: 0949-7714. DOI: 10.1007/s00190-011-0533-4.

Baranov, A. and A. Morelli (2013). “The Moho depth map of the Antarctica region”. In: *Tectonophysics* 609, pp. 299–313. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2012.12.023.

Baratchart, L. and C. Gerhards (2017). “On the Recovery of Core and Crustal Components of Geomagnetic Potential Fields”. In: *SIAM Journal on Applied Mathematics* 77.5, pp. 1756–1780. ISSN: 0036-1399. DOI: 10.1137/17M1121640.

Bassin, C., G. Laske, and T. G. Masters (2000). “The Current Limits of Resolution for Surface Wave Tomography in North America”. In: *Transactions of the AGU*.

Baykiev, E. (2017). “Analysis of modelled satellite magnetic data to explore lithospheric magnetization”. PhD Thesis. Trondheim: Norwegian University of Science and Technology.

Baykiev, E., J. Ebbing, M. Brönner, and K. Fabian (2016). “Forward modeling magnetic fields of induced and remanent magnetization in the lithosphere using tesseroids”. In: *Computers & Geosciences* 96, pp. 124–135. ISSN: 00983004. DOI: 10.1016/j.cageo.2016.08.004.

Becker, T. W. and L. Boschi (2002). “A comparison of tomographic and geodynamic mantle models”. In: *Geochemistry, Geophysics, Geosystems* 3.1. ISSN: 15252027. DOI: 10.1029/2001GC000168.

Blakely, R. J. (1988). “Curie temperature isotherm analysis and tectonic implications of aeromagnetic data from Nevada”. In: *Journal of Geophysical Research* 93.B10, p. 11817. ISSN: 0148-0227. DOI: 10.1029/JB093iB10p11817.

– (2009). *Potential theory in gravity and magnetic applications*. Transferred to digital print. Cambridge: Univ. Press. ISBN: 0-521-41508-X.

Bleil, U. and N. Petersen (1983). “Variations in magnetization intensity and low-temperature titanomagnetite oxidation of ocean floor basalts”. In: *Nature* 301.5899, pp. 384–388. DOI: 10.1038/301384a0.

Bodin, T., M. Sambridge, H. Tkalčić, P. Arroucau, K. Gallagher, and N. Rawlinson (2012). “Transdimensional inversion of receiver functions and surface wave dispersion”. In: *Journal of Geophysical Research: Solid Earth* 117.B2, n/a–n/a. ISSN: 01480227. DOI: 10.1029/2011JB008560.

Bouman, J., J. Ebbing, M. Fuchs, J. Sebera, V. Lieb, W. Szwillus, R. Haagmans, and P. Novak (2016). “Satellite gravity gradient grids for geophysics”. In: *Scientific reports* 6, p. 21050. DOI: 10.1038/srep21050.

Bowin, C. (2000). “Mass anomalies and the structure of the Earth”. In: *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy* 25.4, pp. 343–353. URL: http://www.sciencedirect.com/science/article/pii/S1464189500000569.

Braitenberg, C. (2015). “Exploration of tectonic structures with GOCE in Africa and across-continents”. In: *International Journal of Applied Earth Observation and Geoinformation* 35, pp. 88–95. DOI: 10.1016/j.jag.2014.01.013.

Braitenberg, C., P. Mariani, and A. de Min (2013). “The European Alps and nearby orogenic belts sensed by GOCE”. In: *Bolletino di Geofisica Teorica e Applicata* 54.4, pp. 321–334.

Braitenberg, C., P. Mariani, M. Reguzzoni, and N. Ussami (2010). “GOCE observations for detecting unknown tectonic features”. In: *Proccedings of ESA Living Planet Symposium*. Ed. by ESA Communication Office. URL: http://www.lithoflex.org/bib/ESA_Bergen_10.pdf (visited on 11/27/2015).

Brocher, T. M. (2005). “Empirical relations between elastic wavespeeds and density in the Earth’s crust”. In: *Bulletin of the Seismological Society of America* 95.6, pp. 2081–2092.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu (1995). “A Limited Memory Algorithm for Bound Constrained Optimization”. In: *SIAM Journal on Scientific Computing* 16.5, pp. 1190–1208. ISSN: 1064-8275. DOI: 10.1137/0916069.

Cande, S. C. and D. V. Kent (1995). “Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic”. In: *Journal of Geophysical Research: Solid Earth* 100.B4, pp. 6093–6095. ISSN: 01480227. DOI: 10.1029/94JB03098.

Cantrell, C. D. (2000). *Modern mathematical methods for physicists and engineers.* Cambridge, UK and New York: Cambridge University Press. ISBN: 0521591805.

Cathles, L. M. (1975). *The viscosity of the earth’s mantle.* Princeton Legacy Library. Princeton, New Jersey: Princeton University Press. ISBN: 9780691617831. DOI: 10.2307/j.ctt13x0t47. URL: http://www.jstor.org/stable/10.2307/j.ctt13x0t47.

Chiles, J.-P. and P. Delfiner (1999). *Geostatistics: Modeling spatial uncertainty.* Wiley series in probability and statistics. Applied probability and statistics section. New York: Wiley. ISBN: 0470317833.

Christensen, N. I. and W. D. Mooney (1995). “Seismic velocity structure and composition of the continental crust: A global view”. In: *Journal of Geophysical Research: Solid Earth* 100.B6, pp. 9761–9788. ISSN: 01480227. DOI: 10.1029/95JB00259.

Chulick, G. S., S. Detweiler, and W. D. Mooney (2013). “Seismic structure of the crust and uppermost mantle of South America and surrounding oceanic basins”. In: *Journal of South American Earth Sciences* 42, pp. 260–276. ISSN: 08959811. DOI: 10.1016/j.jsames.2012.06.002.

Colli, L., S. Ghelichkhan, and H.-P. Bunge (2016). “On the ratio of dynamic topography and gravity anomalies in a dynamic Earth”. In: *Geophysical Research Letters* 43.6, pp. 2510–2516. ISSN: 00948276. DOI: 10.1002/2016GL067929.

Connolly, J. A. D. (2009). “The geodynamic equation of state: What and how”. In: *Geochemistry, Geophysics, Geosystems* 10.10. ISSN: 15252027. DOI: 10.1029/2009GC002540.

Cressie, N. A. C. (2015). *Statistics for spatial data.* Revised edition. Hoboken, NJ: John Wiley & Sons Inc. ISBN: 978-1-119-11461-1.

Davies, J. H. (2013). “Global map of solid Earth surface heat flow”. In: *Geochemistry, Geophysics, Geosystems* 14.10, pp. 4608–4622. ISSN: 15252027. DOI: 10.1002/ggge.20271.

Debnath, L. and D. Bhatta (2015). *Integral transforms and their applications.* Third edition. Boca Raton, London, and New York: CRC Press/Taylor & Francis Group. ISBN: 9781482223576.

Eaton, D. W., F. Darbyshire, R. L. Evans, H. Grütter, A. G. Jones, and X. Yuan (2009). “The elusive lithosphere-asthenosphere boundary (LAB) beneath cratons”. In: *Lithos* 109.1-2, pp. 1–22. ISSN: 0024-4937. DOI: 10.1016/j.lithos.2008.05.009. URL: http://dx.doi.org/10.1016/j.lithos.2008.05.009.

Ebbing, J., J. Bouman, M. Fuchs, S. Gradmann, and R. Haagmans (2014). “Sensitivity of GOCE Gravity Gradients to Crustal Thickness and Density Variations: Case Study for the Northeast Atlantic Region”. In: *Gravity, Geoid and Height Systems.* Ed. by U. Marti. Vol. 141. International Association of Geodesy Symposia. Cham, s.l.: Springer International Publishing, pp. 291–298. ISBN: 978-3-319-10836-0. DOI: 10.1007/978-3-319-10837-7{\textunderscore}37.

Ferré, E. C., S. A. Friedman, F. Martín-Hernández, J. M. Feinberg, J. L. Till, D. A. Ionov, and J. A. Conder (2014). “Eight good reasons why the uppermost mantle could be magnetic”. In: *Tectonophysics* 624-625, pp. 3–14. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2014.01.004. URL: http://dx.doi.org/10.1016/j.tecto.2014.01.004.

Fischer, K. M., H. A. Ford, D. L. Abt, and C. A. Rychert (2010). “The Lithosphere-Asthenosphere Boundary”. In: *Annual Review of Earth and Planetary Sciences* 38.1, pp. 551–575. ISSN: 0084-6597. DOI: 10.1146/annurev-earth-040809-152438.

Fishwick, S. and I. D. Bastow (2011). “Towards a better understanding of African topography: A review of passive-source seismic studies of the African crust and upper mantle”. In: *Geological Society, London, Special Publications* 357.1, pp. 343–371. ISSN: 0305-8719. DOI: 10.1144/SP357.19.

Flament, N., M. Gurnis, and R. D. Müller (2013). “A review of observations and models of dynamic topography”. In: *Lithosphere* 5.2, pp. 189–210.

Floberghagen, R., M. Fehringer, D. Lamarre, D. Muzi, B. Frommknecht, C. Steiger, J. Piñeiro, and A. da Costa (2011). “Mission design, operation and exploitation of the gravity field and steady-state ocean circulation explorer mission”. In: *Journal of Geodesy* 85.11, pp. 749–758. ISSN: 0949-7714. DOI: 10.1007/s00190-011-0498-3.

French, S. W. and B. Romanowicz (2015). “Broad plumes rooted at the base of the Earth’s mantle beneath major hotspots”. In: *Nature* 525.7567, pp. 95–99. DOI: 10.1038/nature14876.

Friis-Christensen, E., H. Lühr, D. Knudsen, and R. Haagmans (2008). “Swarm – An Earth Observation Mission investigating Geospace”. In: *Advances in Space Research* 41.1, pp. 210–216. ISSN: 02731177. DOI: 10.1016/j.asr.2006.10.008.

Fullea, J., J. C. Afonso, J. A. D. Connolly, M. Fernàndez, D. García-Castellanos, and H. Zeyen (2009). “LitMod3D: An interactive 3-D software to model the thermal, compositional, density, seismological, and rheological structure of the lithosphere and sublithospheric upper mantle”. In: *Geochemistry, Geophysics, Geosystems* 10.8, n/a–n/a. ISSN: 15252027. DOI: 10.1029/2009GC002391.

Fullea, J., M. Fernàndez, H. Zeyen, and J. Vergés (2007). “A rapid method to map the crustal and lithospheric thickness using elevation, geoid anomaly and thermal analysis. Application to the Gibraltar Arc System, Atlas Mountains and adjacent zones”. In: *Tectonophysics* 430.1-4, pp. 97–117. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2006.11.003.

Gente, P., J. Dyment, M. Maia, and J. Goslin (2003). “Interaction between the Mid-Atlantic Ridge and the Azores hot spot during the last 85 Myr: Emplacement and rifting of the hot spot-derived plateaus”. In: *Geochemistry, Geophysics, Geosystems* 4.10, p. 391. ISSN: 15252027. DOI: 10.1029/2003GC000527.

Gentle, J. E. (2009). *Computational Statistics.* New York, NY: Springer New York. ISBN: 978-0-387-98143-7. DOI: 10.1007/978-0-387-98144-4.

Globig, J., M. Fernàndez, M. Torne, J. Vergés, A. Robert, and C. Faccenna (2016). “New insights into the crust and lithospheric mantle structure of Africa from elevation, geoid, and thermal analysis”. In: *Journal of Geophysical Research: Solid Earth* 121.7, pp. 5389–5424. ISSN: 21699313. DOI: 10.1002/2016JB012972.

Godfrey, N. J., B. C. Beaudoin, and S. L. Klemperer (1997). “Ophiolitic basement to the Great Valley forearc basin, California, from seismic and gravity data: Implications for crustal growth at the North American continental margin”. In: *Geological Society of America Bulletin* 109.12, pp. 1536–1562. ISSN: 00167606. DOI: 10.1130/0016-7606(1997)109{\textless}1536:OBTTGV{\textgreater}2.3.CO;2.

Götze, H. .-.-.-.-J. and B. Lahmeyer (1988). “Application of three–dimensional interactive modeling in gravity and magnetics”. In: *Geophysics* 53.8, pp. 1096–1108. DOI: 10.1190/1.1442546.

Goutorbe, B., J. Poort, F. Lucazeau, and S. Raillard (2011). “Global heat flow trends resolved from multiple geological and geophysical proxies”. In: *Geophysical Journal International* 187.3, pp. 1405–1419. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.2011.05228.x.

Grombein, T., X. Luo, K. Seitz, and B. Heck (2014). “A Wavelet-Based Assessment of Topographic-Isostatic Reductions for GOCE Gravity Gradients”. In: *Surveys in Geophysics* 35.4, pp. 959–982. ISSN: 0169-3298. DOI: 10.1007/s10712-014-9283-1.

Grombein, T., K. Seitz, and B. Heck (2010). *Untersuchungen zur effizienten Berechnung topographischer Effekte auf den Gradiententensor am Fallbeispiel der Satellitengradiometriemission GOCE.* Vol. 7547. KIT Scientific Publishing.

– (2013). “Optimized formulas for the gravitational field of a tesseroid”. In: *Journal of Geodesy* 87.7, pp. 645–660. ISSN: 0949-7714. DOI: 10.1007/s00190-013-0636-1.

Gubbins, D., D. Ivers, S. M. Masterton, and D. E. Winch (2011). “Analysis of lithospheric magnetization in vector spherical harmonics”. In: *Geophysical Journal International* 187.1, pp. 99–117. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.2011.05153.x.

Guinness, J. and M. Fuentes (2016). “Isotropic covariance functions on spheres: Some properties and modeling considerations”. In: *Journal of Multivariate Analysis* 143, pp. 143–152. ISSN: 0047259X. DOI: 10.1016/j.jmva.2015.08.018.

Hammer, S. (1939). “Terrain corrections for gravimeter stations”. In: *Geophysics* 4.3, pp. 184–194.

Hamoudi, M., Y. Cohen, and J. Achache (1998). “Can the thermal thickness of the continental lithosphere be estimated from Magsat data?” In: *Tectonophysics* 284.1-2, pp. 19–29. ISSN: 0040-1951. DOI: 10.1016/S0040-1951(97)00166-2.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The elements of statistical learning: Data mining, inference, and prediction.* Second edition. Springer eBook collection Mathematics and statistics. New York, NY: Springer. ISBN: 978-0-387-84858-7. DOI: 10.1007/978-0-387-84858-7. URL: http://dx.doi.org/10.1007/978-0-387-84858-7.

Hayford, J. and W. Bowie (1912). *Geodesy: Effect of topography and isostatic compensation upon the intensity of gravity. (Second paper).* URL: http://docs.lib.noaa.gov/rescue/cgs_specpubs/QB275U35no121912.pdf (visited on 11/16/2015).

Heck, B. and K. Seitz (2007). “A comparison of the tesseroid, prism and point-mass approaches for mass reductions in gravity field modelling”. In: *Journal of Geodesy* 81.2, pp. 121–136. ISSN: 0949-7714. DOI: 10.1007/s00190-006-0094-0.

Heller, W. G. and S. K. Jordan (1979). “Attenuated white noise statistical gravity model”. In: *Journal of Geophysical Research* 84.B9, p. 4680. ISSN: 0148-0227. DOI: 10.1029/JB084iB09p04680.

Hemant, K. and S. Maus (2005). “Geological modeling of the new CHAMP magnetic anomaly maps using a geographical information system technique”. In: *Journal of Geophysical Research* 110.B12. ISSN: 0148-0227. DOI: 10.1029/2005JB003837.

Herceg, M., I. M. Artemieva, and H. Thybo (2015). “Sensitivity analysis of crustal correction for calculation of lithospheric mantle density from gravity data”. In: *Geophysical Journal International* 204.2, pp. 687–696. ISSN: 0956-540X. DOI: 10.1093/gji/ggv431.

Hinze, W. J., C. Aiken, J. Brozena, B. Coakley, D. Dater, G. Flanagan, R. Forsberg, T. Hildenbrand, G. R. Keller, J. Kellogg, R. Kucks, X. Li, A. Mainville, R. Morin, M. Pilkington, D. Plouff, D. Ravat, D. Roman, J. Urrutia-Fucugauchi, M. Véronneau, M. Webring, and D. Winester (2005). “New standards for reducing gravity data: The North American gravity database”. In: *Geophysics* 70.4, J25–J32. DOI: 10.1190/1.1988183.

Hirt, C., M. Kuhn, W. E. Featherstone, and F. Göttl (2012). “Topographic/isostatic evaluation of new-generation GOCE gravity field models”. In: *Journal of Geophysical Research: Solid Earth (1978–2012)* 117.B5.

Hoggard, M. J., N. White, and D. Al-Attar (2016). “Global dynamic topography observations reveal limited influence of large-scale mantle flow”. In: *Nature Geoscience* 9.6, pp. 456–463. ISSN: 1752-0894. DOI: 10.1038/ngeo2709.

Holschneider, M., V. Lesur, S. Mauerberger, and J. Baerenzung (2016). “Correlation-based modeling and separation of geomagnetic field components”. In: *Journal of Geophysical Research: Solid Earth* 121.5, pp. 3142–3160. ISSN: 21699313. DOI: 10.1002/2015JB012629.

Holzrichter, N. and J. Ebbing (2016). “A regional background model for the Arabian Peninsula from modeling satellite gravity gradients and their invariants”. In: *Tectonophysics* 692, pp. 86–94. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2016.06.002.

Hunt, C. P., B. M. Moskowitz, and S. K. Banerjee (1995). “Magnetic Properties of Rocks and Minerals”. In: *Rock Physics & Phase Relations: A Handbook of Physical Constant.* Ed. by T. Ahrens, C. P. Hunt, B. M. Moskowitz, and S. K. Banerjee, p. 189.

Jaupart, C. and J. C. Mareschal (1999). “The thermal structure and thickness of continental roots”. In: *Composition, Deep Structure and Evolution of Continents.* Vol. 24. Developments in Geotectonics. Elsevier, pp. 93–114. ISBN: 9780444503091. DOI: 10.1016/S0419-0254(99)80007-X.

Jones, E., Oliphant Travis, P. Peterson, et al. (2001). *SciPy: Open source scientific tools for Python.* URL: http://www.scipy.org (visited on 09/20/2017).

Jones, R. H. (1963). "Stochastic Processes on a Sphere". In: *The Annals of Mathematical Statistics* 34.1, pp. 213–218. ISSN: 0003-4851. DOI: 10.1214/aoms/1177704257.

Jordan, T. H. (1978). "Composition and development of the continental tectosphere". In: *Nature* 274.5671, pp. 544–548. DOI: 10.1038/274544a0.

Kaban, M. and W. D. Mooney (2001). "Density structure of the lithosphere in the southwestern United States and its tectonic significance". In: *Journal of Geophysical Research: Solid Earth* 106.B1, pp. 721–739. ISSN: 01480227. DOI: 10.1029/2000JB900235.

Kaban, M., P. Schwintzer, and S. A. Tikhotsky (1999). "A global isostatic gravity model of the Earth". In: *Geophysical Journal International* 136.3, pp. 519–536. ISSN: 1365-246X. DOI: 10.1046/j.1365-246x.1999.00731.x.

Kaban, M., P. Schwintzer, I. M. Artemieva, and W. D. Mooney (2003). "Density of the continental roots: compositional and thermal contributions". In: *Earth and Planetary Science Letters* 209.1-2, pp. 53–69. ISSN: 0012-821X. DOI: 10.1016/s0012-821x(03)00072-4. URL: http://dx.doi.org/10.1016/S0012-821X(03)00072-4.

Karato, S.-i. (1993). "Importance of anelasticity in the interpretation of seismic tomography". In: *Geophysical Research Letters* 20.15, pp. 1623–1626. ISSN: 00948276. DOI: 10.1029/93GL01767.

Khan, A., A. Zunino, and F. Deschamps (2013). "Upper mantle compositional variations and discontinuity topography imaged beneath Australia from Bayesian inversion of surface-wave phase velocities and thermochemical modeling". In: *Journal of Geophysical Research: Solid Earth* 118.10, pp. 5285–5306. ISSN: 21699313. DOI: 10.1002/jgrb.50304.

Kind, R., X. Yuan, and P. Kumar (2012). "Seismic receiver functions and the lithosphere–asthenosphere boundary". In: *Tectonophysics* 536, pp. 25–43. ISSN: 0040-1951.

King, S. D. (2016). "An evolving view of transition zone and midmantle viscosity". In: *Geochemistry, Geophysics, Geosystems* 17.3, pp. 1234–1237. ISSN: 15252027. DOI: 10.1002/2016GC006279.

Köther, N., H.-J. Götze, B. D. Gutknecht, T. Jahr, G. Jentzsch, O. H. Lücke, R. Mahatsente, R. Sharma, and S. Zeumann (2012). "The seismically active Andean and Central American margins: Can satellite gravity map lithospheric structures?" In: *Journal of Geodynamics* 59-60, pp. 207–218. ISSN: 02643707. DOI: 10.1016/j.jog.2011.11.004.

Lachenbruch, A. H. and P. Morgan (1990). "Continental extension, magmatism and elevation; formal relations and rules of thumb". In: *Tectonophysics* 174.1-2, pp. 39–62. ISSN: 0040-1951. DOI: 10.1016/0040-1951(90)90383-J.

Langel, R. A. and W. J. Hinze (1998). *The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective.* Cambridge University Press.

Laske, G., T. G. Masters, Z. Ma, and M. E. Pasyanos (2013). "Update on CRUST1.0 - A 1-degree Global Model of Earth's Crust". In: *Geophysical Research Abstracts* 15. URL: https://igppweb.ucsd.edu/~gabi/crust1.html (visited on 11/02/2015).

Leone, F. C., L. S. Nelson, and R. B. Nottingham (1961). "The Folded Normal Distribution". In: *Technometrics* 3.4, p. 543. ISSN: 00401706. DOI: 10.2307/1266560.

Lesur, V., M. Hamoudi, Y. Choi, J. Dyment, and E. Thébault (2016). "Building the second version of the World Digital Magnetic Anomaly Map (WDMAM)". In: *Earth, Planets and Space* 68.1, p. 6093. ISSN: 1880-5981. DOI: 10.1186/s40623-016-0404-6.

Liu, L. and M. Gurnis (2010). "Dynamic subsidence and uplift of the Colorado Plateau". In: *Geology* 38.7, pp. 663–666. ISSN: 0091-7613. DOI: 10.1130/G30624.1.

Martos, Y. M., M. Catalan, T. A. Jordan, A. Golynsky, D. Golynsky, G. Eagles, and D. G. Vaughan (2017). "Heat flux distribution of Antarctica unveiled". In: *Geophysical Research Letters*. ISSN: 00948276. DOI: 10.1002/2017GL075609.

Matzka, J., D. Krása, T. Kunzmann, A. Schult, and N. Petersen (2003). "Magnetic state of 10–40 Ma old ocean basalts and its implications for natural remanent magnetization". In: *Earth and Planetary Science Letters* 206.3-4, pp. 541–553. ISSN: 0012-821X. DOI: 10.1016/S0012-821X(02)01094-4.

Maule, C. F., M. E. Purucker, N. Olsen, and K. Mosegaard (2005). “Heat flux anomalies in Antarctica revealed by satellite magnetic data”. In: *Science (New York, N.Y.)* 309.5733, pp. 464–467. ISSN: 0036-8075. DOI: 10.1126/science.1106888.

Maus, S. (2008). “The geomagnetic power spectrum”. In: *Geophysical Journal International* 174.1, pp. 135–142. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.2008.03820.x.

– (2010). *Magnetic field model MF7*. URL: www.geomag.us/models/MF7.html (visited on 09/09/2017).

Maus, S. and V. Haak (2003). “Magnetic field annihilators: Invisible magnetization at the magnetic equator”. In: *Geophysical Journal International* 155.2, pp. 509–513. ISSN: 1365-246X. DOI: 10.1046/j.1365-246X.2003.02053.x.

Maus, S., F. Yin, H. Lühr, C. Manoj, M. Rother, J. Rauberg, I. Michaelis, C. Stolle, and R. D. Müller (2008). “Resolution of direction of oceanic magnetic lineations by the sixth-generation lithospheric magnetic field model from CHAMP satellite magnetic measurements”. In: *Geochemistry, Geophysics, Geosystems* 9.7, n/a–n/a. ISSN: 15252027. DOI: 10.1029/2008GC001949.

Mayer-Guerr, T. (2015). “The combined satellite gravity field model GOCO05s”. In: *Geophysical Research Abstracts* 17. URL: http://adsabs.harvard.edu/abs/2015EGUGA..1712364M.

McKenzie, D. and M. J. Bickle (1988). “The volume and composition of melt generated by extension of the lithosphere”. In: *Journal of petrology* 29.3, pp. 625–679.

McKenzie, D. (2016). “A note on estimating $$T_e$$ T e from Bouguer coherence”. In: *GEM - International Journal on Geomathematics* 7.1, pp. 103–116. ISSN: 1869-2672. DOI: 10.1007/s13137-015-0078-4.

McKenzie, D., J. Jackson, and K. Priestley (2005). “Thermal structure of oceanic and continental lithosphere”. In: *Earth and Planetary Science Letters* 233.3, pp. 337–349. ISSN: 0012-821X.

Mikuška, J., R. Pašteka, and I. Marušiak (2006). “Estimation of distant relief effect in gravimetry”. In: *Geophysics* 71.6, J59–J69. DOI: 10.1190/1.2338333.

Molnar, P., P. C. England, and C. H. Jones (2015). “Mantle dynamics, isostasy, and the support of high terrain”. In: *Journal of Geophysical Research: Solid Earth* 120.3, pp. 1932–1957. ISSN: 2169-9313. DOI: 10.1002/2014jb011724. URL: http://dx.doi.org/10.1002/2014JB011724.

Mooney, W. D. (2015). “Global crustal structure”. In: *Treatise on geophysics*. Ed. by G. Schubert. Amsterdam, Netherlands: Elsevier, pp. 361–417. ISBN: 978-0-444-53802-4.

Mooney, W. D. and M. K. Kaban (2010). “The North American upper mantle: Density, composition, and evolution”. In: *Journal of Geophysical Research* 115.B12. ISSN: 0148-0227. DOI: 10.1029/2010jb000866. URL: http://dx.doi.org/10.1029/2010JB000866.

Mooney, W. D., G. Laske, and T. G. Masters (1998). “CRUST 5.1: A global crustal model at 5° x 5°”. In: *Journal of Geophysical Research: Solid Earth* 103.B1, pp. 727–747. ISSN: 01480227. DOI: 10.1029/97JB02122.

Moorkamp, M., P. G. Lelièvre, N. Linde, and A. Khan, eds. (2016). *Integrated Imaging of the Earth*. Hoboken, NJ: John Wiley & Sons, Inc. ISBN: 9781118929063. DOI: 10.1002/9781118929063.

Moritz, H. (1976). *Covariance Functions in Least-Squares Collocation*.

Müller, R. D., M. Sdrolias, C. Gaina, and W. R. Roest (2008). “Age, spreading rates, and spreading asymmetry of the world’s ocean crust”. In: *Geochemistry, Geophysics, Geosystems* 9.4, n/a–n/a. ISSN: 15252027. DOI: 10.1029/2007GC001743.

Nabighian, M. N., M. E. Ander, V. J. S. Grauch, R. O. Hansen, T. R. LaFehr, Y. Li, W. C. Pearson, J. W. Peirce, J. D. Phillips, and M. E. Ruder (2005). “Historical development of the gravity method in exploration”. In: *Geophysics* 70.6, 63ND–89ND. DOI: 10.1190/1.2133785.

Olsen, N., C. C. Finlay, S. Kotsiaros, and L. Tøffner-Clausen (2016). “A model of Earth’s magnetic field derived from 2 years of Swarm satellite constellation data”. In: *Earth, Planets and Space* 68.1, p. 6339. ISSN: 1880-5981. DOI: 10.1186/s40623-016-0488-z.

Olsen, N., D. Ravat, C. C. Finlay, and L. K. Kother (2017). “LCS-1: A high-resolution global model of the lithospheric magnetic field derived from CHAMP and Swarm satellite observations”. In: *Geophysical Journal International* 211.3, pp. 1461–1477. ISSN: 1365-246X. DOI: 10.1093/gji/ggx381.

Paciorek, C. and Schervish M. J. (2004). “Nonstationary Covariance Functions for Gaussian Process Regression”. In: *Advances in Neural Information Processing Systems 16*. Ed. by Thrun S., L. Saul, and B. Schölkopf. MIT Press.

Pail, R., H. Goiginger, W.-D. Schuh, E. Höck, J. M. Brockmann, T. Fecher, T. Gruber, T. Mayer-Gürr, J. Kusche, A. Jäggi, and D. Rieser (2010). “Combined satellite gravity field model GOCO01S derived from GOCE and GRACE”. In: *Geophysical Research Letters* 37.20, n/a–n/a. ISSN: 00948276. DOI: 10.1029/2010GL044906.

Panet, I., G. Pajot-Métivier, M. Greff-Lefftz, L. Métivier, M. Diament, and M. Mandea (2014). “Mapping the mass distribution of Earth/’s mantle using satellite-derived gravity gradients”. In: *Nature Geoscience* 7.2, pp. 131–135.

Paoletti, V., M. Fedi, F. Italiano, G. Florio, and S. Ialongo (2016). “Inversion of gravity gradient tensor data: Does it provide better resolution?” In: *Geophysical Journal International* 205.1, pp. 192–202. ISSN: 0956-540X. DOI: 10.1093/gji/ggw003.

Pasyanos, M. E., T. G. Masters, G. Laske, and Z. Ma (2014). “LITHO1.0: An updated crust and lithospheric model of the Earth”. In: *Journal of Geophysical Research: Solid Earth* 119.3, pp. 2153–2173. ISSN: 2169-9313. DOI: 10.1002/2013jb010626. URL: http://dx.doi.org/10.1002/2013JB010626.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011). “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Pilchin, A. and L. Eppelbaum (1997). “Determination of the lower edges of magnetized bodies by using geothermal data”. In: *Geophysical Journal International* 128.1, pp. 167–174. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.1997.tb04077.x.

Pratt, J. H. (1855). “On the Attraction of the Himalaya Mountains, and of the Elevated Regions beyond Them, upon the Plumb-Line in India”. In: *Philosophical Transactions of the Royal Society of London* 145.0, pp. 53–100. ISSN: 0261-0523. DOI: 10.1098/rstl.1855.0002.

Pratt, J. H. (1864). *On the degree of uncertainty which Local Attraction, if not allowed for, occasions in the Map of a Country, and in the Mean Figure of the Earth as determined by Geodesy.* Vol. no. 64. Proc. Roy. Soc. London: [publisher not identified].

Prezzi, C. B., H.-J. Götze, and S. Schmidt (2009). “3D density model of the Central Andes”. In: *Physics of the Earth and Planetary Interiors* 177.3-4, pp. 217–234. ISSN: 0031-9201. DOI: 10.1016/j.pepi.2009.09.004.

Priestly, K. and D. McKenzie (2006). “The thermal structure of the lithosphere from shear wave velocities”. In: *Earth and Planetary Science Letters* 244.1-2, pp. 285–301. ISSN: 0012-821X. DOI: 10.1016/j.epsl.2006.01.008. URL: http://dx.doi.org/10.1016/j.epsl.2006.01.008.

Prodehl, C. and W. D. Mooney (2012). *Exploring the earth’s crust: History and results of controlled-source seismology.* Vol. 208. Memoir / Geological Society of America. Boulder, Colo.: Geological Soc. of America. ISBN: 9780813712086.

Purucker, M. E., R. A. Langel, M. Rajaram, and C. Raymond (1998). “Global magnetization models with a priori information”. In: *Journal of Geophysical Research: Solid Earth* 103.B2, pp. 2563–2584. ISSN: 01480227. DOI: 10.1029/97JB02935.

Reguzzoni, M., D. Sampietro, and F. Sanso (2013). “Global Moho from the combination of the CRUST2.0 model and GOCE data”. In: *Geophysical Journal International* 195.1, pp. 222–237. ISSN: 1365-246X. DOI: 10.1093/gji/ggt247.

Reigber, C., H. Lühr, P. Schwintzer, and J. Wickert (2005). *Earth Observation with CHAMP*. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-540-22804-2. DOI: 10.1007/b138105.

Reynolds, R., J. G. Rosenbaum, M. R. Hudson, and N. S. Fishman (1990). “Rock Magnetism, The distribution of Magnetic Minerals in the Earth’s Crust, and Aeromagnetic Anomalies”. In: *Geologic applications of modern aeromagnetic surveys*. Ed. by W. F. Hanna.

Ricard, Y., C. Froidevaux, and L. Fleitout (1988). “Global plate motion and the geoid: A physical model”. In: *Geophysical Journal International* 93.3, pp. 477–484. ISSN: 1365-246X. DOI: 10.1111/j.1365-246X.1988.tb03875.x.

Richards, M. A. and B. H. Hager (1984). “Geoid anomalies in a dynamic Earth”. In: *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5987–6002. ISSN: 01480227. DOI: 10.1029/JB089iB07p05987.

Risser, M. D. and C. A. Calder (2017). “Local likelihood estimation for covariance functions with spatially-varying parameters: the convoSPAT package for R”. In: *arXiv.org*. URL: https://arxiv.org/abs/1507.08613.

Roberts, G. O., A. Gelman, and W. R. Gilks (1997). “Weak convergence and optimal scaling of random walk Metropolis algorithms”. In: *The Annals of Applied Probability* 7.1, pp. 110–120. DOI: 10.1214/aoap/1034625254.

Root, B. C., W. Wal, P. Novák, J. Ebbing, and L. L. Vermeersen (2015). “Glacial isostatic adjustment in the static gravity field of Fennoscandia”. In: *Journal of Geophysical Research: Solid Earth* 120.1, pp. 503–518. ISSN: 2169-9313.

Royden, L. H. (1993). “The tectonic expression slab pull at continental convergent boundaries”. In: *Tectonics* 12.2, pp. 303–325. DOI: 10.1029/92TC02248.

Rychert, C. A., P. M. Shearer, and K. M. Fischer (2010). “Scattered wave imaging of the lithosphere–asthenosphere boundary”. In: *Lithos* 120.1-2, pp. 173–185. ISSN: 0024-4937. DOI: 10.1016/j.lithos.2009.12.006.

Schaeffer, A. J. and S. Lebedev (2015). “Global Heterogeneity of the Lithosphere and Underlying Mantle: A Seismological Appraisal Based on Multimode Surface-Wave Dispersion Analysis, Shear-Velocity Tomography, and Tectonic Regionalization”. In: *The Earth’s Heterogeneous Mantle*. Ed. by A. Khan and F. Deschamps. Cham: Springer International Publishing, pp. 3–46. ISBN: 978-3-319-15626-2. DOI: 10.1007/978-3-319-15627-9{\textunderscore}1.

Shapiro, N. M. and M. Campillo (2004). “Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise”. In: *Geophysical Research Letters* 31.7, n/a–n/a. ISSN: 00948276. DOI: 10.1029/2004GL019491.

Shaw, H. and E. L. Jones (1923). “The Eötvös Torsion Balance and its Use in the Field”. In: *Nature* 111.2799, pp. 849–851. DOI: 10.1038/111849a0.

Shive, P. N. (1989). “Can remanent magnetization in the deep crust contribute to long wavelength magnetic anomalies?” In: *Geophysical Research Letters*. ISSN: 00948276.

Smith, W. H. F. and D. T. Sandwell (1997). “Global sea floor topography from satellite altimetry and ship depth soundings”. In: *Science* 277.5334, pp. 1956–1962. ISSN: 1095-9203.

Spector, A. and F. S. Grant (1970). “Statistical models for interpreting aeromagnetic data”. In: *Geophysics* 35.2, pp. 293–302. DOI: 10.1190/1.1440092.

Spieker, K., S. Rondenay, R. Ramalho, C. Thomas, and G. Helffrich (2018). “Constraints on the structure of the crust and lithosphere beneath the Azores Islands from teleseismic receiver functions”. In: *Geophysical Journal International* 213.2, pp. 824–835. ISSN: 1365-246X. DOI: 10.1093/gji/ggy022.

Steinberger, B. (2000). “Plumes in a convecting mantle: Models and observations for individual hotspots”. In: *Journal of Geophysical Research: Solid Earth* 105.B5, pp. 11127–11152. ISSN: 01480227. DOI: 10.1029/1999JB900398.

– (2007). “Effects of latent heat release at phase boundaries on flow in the Earth’s mantle, phase boundary topography and dynamic topography at the Earth’s surface”. In: *Physics of the Earth and Planetary Interiors* 164.1-2, pp. 2–20. ISSN: 0031-9201. DOI: 10.1016/j.pepi.2007.04.021.

Steinberger, B. (2016). “Topography caused by mantle density variations: Observation-based estimates and models derived from tomography and lithosphere thickness”. In: *Geophysical Journal International* 205.1, pp. 604–621. ISSN: 0956-540X. DOI: 10.1093/gji/ggw040.

Stolk, W., M. Kaban, F. Beekman, M. Tesauro, W. D. Mooney, and S. Cloetingh (2013). “High resolution regional crustal models from irregularly distributed data: Application to Asia and adjacent areas”. In: *Tectonophysics* 602, pp. 55–68. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2013.01.022.

Szwillus, W., J. Ebbing, and N. Holzrichter (2016). “Importance of far-field topographic and isostatic corrections for regional density modelling”. In: *Geophysical Journal International* 207.1, pp. 274–287. ISSN: 0956-540X. DOI: 10.1093/gji/ggw270.

Tanaka, A. (2017). “Global centroid distribution of magnetized layer from World Digital Magnetic Anomaly Map (WDMAM)”. In: *Tectonics*. DOI: 10.1002/2017TC004770.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004). “The gravity recovery and climate experiment: Mission overview and early results”. In: *Geophysical Research Letters* 31.9. ISSN: 00948276. DOI: 10.1029/2004GL019920.

Tapley, B. D., J. Ries, S. Bettadpur, D. Chambers, M. Cheng, F. Condi, B. Gunter, Z. Kang, P. Nagel, R. Pastor, T. Pekker, S. Poole, and F. Wang (2005). “GGM02 – An improved Earth gravity field model from GRACE”. In: *Journal of Geodesy* 79.8, pp. 467–478. ISSN: 0949-7714. DOI: 10.1007/s00190-005-0480-z.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Philadelphia, Pa.: SIAM - Soc. for Industrial and Applied Math. ISBN: 978-0-89871-572-9. URL: http://www.ipgp.fr/~tarantola/Files/Professional/SIAM/InverseProblemTheory.pdf.

Tarantola, A. and B. Valette (1982). “Generalized nonlinear inverse problems solved using the least squares criterion”. In: *Reviews of Geophysics* 20.2, p. 219. ISSN: 8755-1209. DOI: 10.1029/RG020i002p00219.

Tesauro, M., M. K. Kaban, W. D. Mooney, and S. Cloetingh (2014). “NACr14: A 3D model for the crustal structure of the North American Continent”. In: *Tectonophysics* 631, pp. 65–86. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2014.04.016.

Thébault, E., C. C. Finlay, C. D. Beggan, P. Alken, J. Aubert, O. Barrois, F. Bertrand, T. Bondar, A. Boness, L. Brocco, E. Canet, A. Chambodut, A. Chulliat, P. Coïsson, F. Civet, A. Du, A. Fournier, I. Fratter, N. Gillet, B. Hamilton, M. Hamoudi, G. Hulot, T. Jager, M. Korte, W. Kuang, X. Lalanne, B. Langlais, J.-M. Léger, V. Lesur, F. J. Lowes, S. Macmillan, M. Mandea, C. Manoj, S. Maus, N. Olsen, V. Petrov, V. Ridley, M. Rother, T. J. Sabaka, D. Saturnino, R. Schachtschneider, O. Sirol, A. Tangborn, A. Thomson, L. Tøffner-Clausen, P. Vigneron, I. Wardinski, and T. Zvereva (2015). “International Geomagnetic Reference Field: The 12th generation”. In: *Earth, Planets and Space* 67.1, p. 68. ISSN: 1880-5981. DOI: 10.1186/s40623-015-0228-9.

Thébault, E., M. Purucker, K. A. Whaler, B. Langlais, and T. J. Sabaka (2010). “The Magnetic Field of the Earth’s Lithosphere”. In: *Space Science Reviews* 155.1-4, pp. 95–127. ISSN: 0038-6308. DOI: 10.1007/s11214-010-9667-6.

Tibshirani, R. and T. Hastie (1987). “Local Likelihood Estimation”. In: *Journal of the American Statistical Association* 82.398, p. 559. ISSN: 0162-1459. DOI: 10.2307/2289465.

Tork Qashqai, M., J. Carlos Afonso, and Y. Yang (2016). “The crustal structure of the Arizona Transition Zone and southern Colorado Plateau from multiobservable probabilistic inversion”. In: *Geochemistry, Geophysics, Geosystems* 17.11, pp. 4308–4332. ISSN: 15252027. DOI: 10.1002/2016GC006463.

Turcotte, D. L. and G. Schubert (2002). *Geodynamics*. Cambridge: Cambridge University Press. ISBN: 9780511807442. DOI: 10.1017/cbo9780511807442.

Uieda, L. and V. C. Barbosa (2016). “Fast nonlinear gravity inversion in spherical coordinates with application to the South American Moho”. In: *Geophysical Journal International* 208.1, pp. 162–176. ISSN: 0956-540X. DOI: 10.1093/gji/ggw390.

Uieda, L., E. P. Bomfim, C. Braitenberg, and E. Molina (2011). “Optimal forward calculation method of the Marussi tensor due to a geologic structure at GOCE height”. In: *Proceedings of the 4th International GOCE User Workshop*. Ed. by L. Ouwehand. Vol. 696. ESA SP. Noordwijk: ESA Publications Div. ISBN: 978-92-9092-260-5. URL: http://fatiando.org/papers/Uieda,etal._2011.pdf.

van der Meijde, M., J. Julià, and M. Assumpção (2013). “Gravity derived Moho for South America”. In: *Tectonophysics* 609, pp. 456–467. ISSN: 0040-1951. DOI: 10.1016/j.tecto.2013.03.023.

Vervelidou, F. and E. Thébault (2015). “Global maps of the magnetic thickness and magnetization of the Earth’s lithosphere”. In: *Earth, Planets and Space* 67.1, p. 774. ISSN: 1880-5981. DOI: 10.1186/s40623-015-0329-5.

Wasilewski, P. J. and M. A. Mayhew (1992). “The moho as a magnetic boundary revisited”. In: *Geophysical Research Letters* 19.22, pp. 2259–2262. ISSN: 00948276. DOI: 10.1029/92GL01997.

Watts, A. B. (2001). *Isostasy and flexure of the lithosphere*. Cambridge: Cambridge Univ. Press. ISBN: 0521006007.

Watts, A. B. and E. Burov (2003). “Lithospheric strength and its relationship to the elastic and seismogenic layer thickness”. In: *Earth and Planetary Science Letters* 213.1-2, pp. 113–131. ISSN: 0012-821X. DOI: 10.1016/S0012-821X(03)00289-9.

Wessel, P., W. H. F. Smith, R. Scharroo, J. Luis, and F. Wobbe (2013). “Generic Mapping Tools: Improved Version Released”. In: *Eos, Transactions American Geophysical Union* 94.45, pp. 409–410. ISSN: 00963941. DOI: 10.1002/2013EO450001.

White, R. S., D. McKenzie, and R. K. O’Nions (1992). “Oceanic crustal thickness from seismic measurements and rare earth element inversions”. In: *Journal of Geophysical Research* 97.B13, p. 19683. ISSN: 0148-0227. DOI: 10.1029/92JB01749.

Whittaker, J. M., A. Goncharov, S. E. Williams, R. D. Müller, and G. Leitchenkov (2013). “Global sediment thickness data set updated for the Australian-Antarctic Southern Ocean”. In: *Geochemistry, Geophysics, Geosystems* 14.8, pp. 3297–3305. ISSN: 15252027. DOI: 10.1002/ggge.20181.

Wild-Pfeiffer, F. (2008). “A comparison of different mass elements for use in gravity gradiometry”. In: *Journal of Geodesy* 82.10, pp. 637–653. DOI: 10.1007/s00190-008-0219-8.

Worm, H.-U. (1989). “Comment on “Can remanent magnetization in the deep crust contribute to long wavelength magnetic anomalies? ” by Peter N. Shive”. In: *Geophysical Research Letters* 16.6, pp. 595–597. ISSN: 00948276. DOI: 10.1029/GL016i006p00595.

Yang, T. and M. Gurnis (2016). “Dynamic topography, gravity and the role of lateral viscosity variations from inversion of global mantle flow”. In: *Geophysical Journal International* 207.2, pp. 1186–1202. ISSN: 0956-540X. DOI: 10.1093/gji/ggw335.

Zeyen, H. and M. Fernàndez (1994). “Integrated lithospheric modeling combining thermal, gravity, and local isostasy analysis: Application to the NE Spanish Geotransect”. In: *Journal of Geophysical Research: Solid Earth* 99.B9, pp. 18089–18102. ISSN: 01480227. DOI: 10.1029/94JB00898.

Zhu, C., R. H. Byrd, P. Lu, and J. Nocedal (1997). “Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization”. In: *ACM Transactions on Mathematical Software* 23.4, pp. 550–560. ISSN: 00983500. DOI: 10.1145/279232.279236.

Zoback, M. L. and W. D. Mooney (2003). “Lithospheric Buoyancy and Continental Intraplate Stresses”. In: *International Geology Review* 45.2, pp. 95–118. ISSN: 0020-6814. DOI: 10.2747/0020-6814.45.2.95.

Acknowledgments

This work would not have been possible without the support of many people. Especially, I thank:

- Prof. Dr. Jörg Ebbing for supervising this PhD thesis. Without his support, ideas and patience this thesis would not have been possible. He really struck just the right balance between letting me explore my own ideas and helping me, when I got stuck or sidetracked. For this I am deeply grateful.
- Professor Richard England for acting as co-referee.
- Nils Holzrichter for discussion and support on anything from computer troubles to the intricate workings of university administration.
- All other past and current members of the AG Ebbing a.k.a. AG "Geoinformation and Geophysics" a.k.a. AG "Satellite and Airborne Geophysics". The climate in the team was always wonderful.
- Eldar Baykiev for a great and productive cooperation.
- Juan Carlos Afonso and all members of the group in Sydney for welcoming me and giving me the opportunity to learn about Litmod and all the things behind it.
- All the developers of open source software that I used in this work: The Generic Mapping Toolbox, NumPy, SciPy, Matplotlib, Project Jupyter, scikit-learn.
- Deutsche Forschungsgemeinschaft for funding our project in the SPP 1788 "Dynamic Earth".
- Johanna for her support and for always believing in me.

Erklärung

Diese Arbeit ist, abgesehen von der Beratung durch meinen Betreuer und die Zuhilfenahme der angegebenen Mittel, nach Inhalt und Form meine eigene. Die Arbeit hat weder ganz noch zum Teil bereits an anderer Stelle im Rahmen eines Prüfungsverfahren vorgelegen. Die Arbeit wurde nicht veröffentlicht noch zur Veröffentlichung eingereicht. Die Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft enstanden.

(Ort) (Datum) (Wolfgang Szwillus)