Originally published as:

Topography caused by mantle density variations: observation-based estimates and models derived from tomography and lithosphere thickness

Bernhard Steinberger[1,2]

[1]*GFZ German Research Centre for Geosciences, D-14473 Potsdam, Germany. E-mail: bstein@gfz-potsdam.de*
[2]*Centre for Earth Evolution and Dynamics (CEED), University of Oslo, 0316 Oslo, Norway*

SUMMARY

Large-scale topography may be due to several causes, including (1) variations in crustal thickness and density structure, (2) oceanic lithosphere age differences, (3) subcrustal density variations in the continental lithosphere and (4) convective flow in the mantle beneath the lithosphere. The last contribution in particular may change with time and be responsible for continental inundations; distinguishing between these contributions is therefore important for linking Earth's history to its observed geological record. As a step towards this goal, this paper aims at such distinction for the present-day topography: the approach taken is deriving a 'model' topography due to contributions (3) and (4), along with a model geoid, using a geodynamic mantle flow model. Both lithosphere thickness and density anomalies beneath the lithosphere are inferred from seismic tomography. Density anomalies within the continental lithosphere are uncertain, because they are probably due to variations in composition and temperature, making a simple scaling from seismic to density anomalies inappropriate. Therefore, we test a number of different assumptions regarding these. As a reality check, model topography is compared, in terms of both correlation and amplitude ratio, to 'residual' topography, which follows from observed topography after subtracting contributions (1) and (2). The model geoid is compared to observations as well. Comparatively good agreement is found if there is either an excess density of ≈ 0.2 per cent in the lithosphere above ≈ 150 km depth, with anomalies below as inferred from tomography, or if the excess density is ≈ 0.4 per cent in the entire lithosphere. Further, a good fit is found for viscosity $\approx 10^{20}$ Pa s in the asthenosphere, increasing to $\approx 10^{23}$ Pa s in the lower mantle above D″. Results are quite dependent on which tomography models they are based on; for some recent ones, topography correlation is ≈ 0.6, many smaller scale features are matched, topography amplitude is less than ≈ 30 per cent too large, while geoid variance reduction exceeds 70 per cent—overall a considerable improvement compared to previous models. Correlation becomes less if smaller scale features (corresponding to spherical harmonic degrees 15 and higher), which are probably largely due to anomalies in the lithosphere, are also considered. Comparison of results with different viscosity structures, and a regional comparison of amplitude ratios, indicates that lateral viscosity variations can be quite strong, but only leading to moderate variations in model topography of a factor probably less than two.

Key words: Numerical solutions; Gravity anomalies and Earth structure; Seismic tomography; Dynamics of lithosphere and mantle; Rheology: mantle.

1 INTRODUCTION

1.1 Importance of dynamic topography throughout Earth history

Uplift and subsidence of the lithosphere, caused by both advection of density anomalies in the underlying mantle and plate motions

over it, determines which areas are below sea level, and at what depth. It is therefore of great importance for understanding of past environments and formation of natural resources. 150 m vertical movement due to mantle flow is a realistic estimate based on the Phanerozoic flooding of continental platforms (Gurnis 1990), and since a large fraction of continents is close to sea level, it makes a big difference on the amount of continent that is flooded by shallow seas.

Locally, especially under the influence of a mantle plume, probably even much larger vertical motions of several hundred metres up to ~1 km may occur, as is evidenced by ancient buried landscapes (Hartley *et al.* 2011). To understand such mantle-derived vertical movements, we should be able to distinguish which parts of topography are 'dynamic' that is, supported by density anomalies and flow in the mantle. This paper is concerned with present-day topography, but can serve as a starting point for future models that predict uplift and subsidence.

1.2 Definitions of dynamic topography

The concept of dynamic topography is in theory rather simple. Hot, less dense material tends to rise and push the overlying lithosphere upward, whereas cold, denser material tends to sink and pull the lithosphere down. In practice, though, it is not even straightforward to define what dynamic topography actually is (see e.g. Molnar *et al.* 2015): while topography isostatically supported through crustal thickness variations is clearly not included, it becomes less clear whether topography caused by density anomalies within the lithosphere should be included. As long as the lithosphere is not taking part in convection, it is conceptually appropriate to not include them. However, oceanic lithosphere is the upper limb of a mantle convection cell, therefore the increase of seafloor depth with age could be considered part of dynamic topography.

Dynamic topography can be estimated from observations by subtracting any known isostatic contributions, yielding 'residual' topography. Again, using ocean floor age grids and a depth versus age curve, a contribution due to cooling of oceanic lithosphere with age may (or may not) be subtracted, and it is less straightforward to exclude the effect of density anomalies in the subcrustal continental lithosphere. They are excluded by Le Stunff & Ricard (1995) and Panasyuk & Hager (2000). Kaban *et al.* (2003) do not correct for them and find that for spherical harmonics of degree 2–20 and density anomalies above 300 km depth, the topography kernels appropriate for dynamic flow differ at most by 10 per cent from those for a purely isostatic model.

The topography model derived here also includes the effect of density anomalies in both the convecting mantle and the subcrustal continental lithosphere, which is why the term 'model topography' rather than 'dynamic topography' is used. At the same time a lithosphere thickness model is derived, such that it is possible to distinguish both contributions in the model. While the distinction is not very important for present-day topography, it becomes important for time changes, as density anomalies within the lithosphere move with the plates and hence do not cause uplift or subsidence, whereas sublithospheric density anomalies may do so. Accordingly, time changes of dynamic topography could be typically considerably smaller than the computed amplitudes of the topography model for present day. Gurnis (1990, 1993) discusses marine inundations and is thus able to constrain uplift and subsidence, finding indeed smaller amplitudes.

1.3 Relation of density anomalies to dynamic topography and geoid

Both internal density anomalies and the boundary displacements they cause (including surface and core–mantle boundary (CMB) topography) relate to the geoid. A method of computing the geoid taking these contributions into account has been developed by Richards & Hager (1984) and Ricard *et al.* (1984). Linking these develop-

ments with the first mantle tomography models, Hager *et al.* (1985) showed a relation between lower mantle heterogeneity, dynamic topography and the geoid. Hager & Richards (1989) showed that almost 90 per cent of the observed geoid can be explained by density anomalies inferred from tomography and a model of subducted slabs. A successful fit of the geoid requires a lower mantle viscosity that is higher than upper-mantle viscosity by a factor that varies between publications from 10 (Hager *et al.* 1985) to 40 (Ricard *et al.* 1993) or 30 or greater (Thoraval & Richards 1997).

In contrast, there has always been a larger discrepancy between present-day dynamic topography predicted by such models, and inferred from observations (Le Stunff & Ricard 1995; Steinberger 2007; Flament *et al.* 2013), unless a 3-D scaling from seismic velocities to densities is invoked (Simmons *et al.* 2009, 2010). There are probably two main reasons for this discrepancy: first, topography has a large isostatic component, that is, it is largely due to crustal thickness variations and to thickness and composition variation in the lithosphere. So it is difficult to extract dynamic topography from observations, particularly in regions where crustal thickness is not well known. Secondly, topography is most sensitive to density anomalies at shallow depth. But density anomalies within the lithosphere are likely both due to temperature and compositional variations, as discussed below. So they cannot be simply converted from seismic tomography. The geoid, on the other hand, is largely due to density variations at greater depth and shallow density anomalies contribute less. So the geoid can be more reliably inferred from seismic tomography, since at these greater depths, simple 'thermal' conversion from seismic to density anomalies is probably more appropriate, as the effect of composition is smaller.

1.3.1 Mantle density anomalies and lithosphere thickness

Regarding the choice of density models, Ricard *et al.* (1993) computed dynamic topography based on subduction history, but more commonly, density anomalies have been inferred from seismic tomography, or some combination. If all seismic velocity anomalies are converted to density anomalies, assuming they are due to thermal anomalies, predicted topographies are too large; total density anomalies in the continental lithosphere are almost certainly smaller. The isopycnic hypothesis of Jordan (1988) states that the strong seismic velocity variations in the lithosphere correspond to nearly zero density anomalies. Hence, seismic velocity anomalies are often only considered below a certain depth (e.g. Lithgow-Bertelloni & Silver 1998; Steinberger *et al.* 2001).

Another approach to reduce dynamic topography predictions considers the relation between seismic, temperature and compositional variations (Karato & Karki 2001; Trampert *et al.* 2001). Non-linearities, that is, considering the temperature dependence of this relation, were considered by Cammarano *et al.* (2003). In an even more accurate approach, the pressure–temperature phase diagrams for different lithologies (pyrolite, MORB and harzburgite) are used, and densities and seismic velocities are computed for these. This approach has, for example, been taken by Nakagawa *et al.* (2009, 2010). They compute stable mineralogy and physical properties by free-energy minimization using the Perple_X software (Connolly 2005). The approach of self-consistent computation of phase equilibria and physical properties was developed for the upper mantle by Stixrude & Lithgow-Bertelloni (2005a) in order to explain the origin of the low-velocity zone. The development of the theory for the whole mantle is detailed in Stixrude & Lithgow-Bertelloni (2005b, 2011). Khan *et al.* (2009) apply such an approach

to surface wave data to infer 1-D radial profiles of Earth's mantle composition, thermal state and anisotropic structure.

It is also possible to interpret the seismically fast anomalies in the continental lithosphere together with gravity data to obtain an estimate of how much depleted (and how cold) the continental lithosphere should be (e.g. Forte & Perry 2000; Deschamps *et al.* 2002), or by accounting for petrological constraints on continents (Cammarano *et al.* 2011). Griffin *et al.* (2009) suggest that large volumes of highly depleted dunites/harzburgites, similar to the Archean orogenic massifs of western Norway, are preserved in the cores of cratons, because incorporation of such rocks in the cold upper parts of the cratonic subcontinental lithospheric mantle satisfies the seismic and gravity data. Further, they suggest a sharp dichotomy between Archean and younger tectonic regimes. Simmons *et al.* (2009, 2010) inverted seismic, mineral physical and geodynamic constraints (including dynamic topography) to infer thermal and compositional heterogeneity. The subject has been reviewed by Forte (2015). In order to distinguish thermal and compositional anomalies, and to include density anomalies at all depths, a model of lithosphere thickness is hence important. There are strong indications that 'thermal' conversion is also inappropriate for the Large Low Shear Velocity Provinces (LLSVPs) of the lowermost mantle (Masters *et al.* 2000). However, these anomalies have little effect on dynamic topography as is, for example, evident from the kernels shown in the next section.

1.3.2 Mantle viscosity

Radial viscosity variations are uncertain, although they can be constrained (e.g. Mitrovica & Forte 2004; Steinberger & Calderwood 2006; Paulson *et al.* 2007). Dynamic topography is relatively poorly known, compared to the geoid, and does not strongly depend on mantle viscosity structure, therefore does not pose a tight constraint. Flow models often only consider radial viscosity variations, but lateral viscosity variations (LVV) due to temperature variations have also been considered, for example, by Ghosh *et al.* (2013) using the Citcom-S finite-element code (Zhong *et al.* 2000; Tan *et al.* 2006), by Cammarano *et al.* (2011) using the code StagYY (Tackley 2008) or using the spectral method (Čadek & Fleitout 2003; Moucha *et al.* 2007; Petrunin *et al.* 2013). Moucha *et al.* (2007) found that present-day topography is affected relatively little by inclusion of LVV, and that their effect is significantly smaller than the variability due to uncertainties in tomography models. It is also possible to jointly derive a tomography and viscosity model that fits both seismological and geodynamic constraints, including dynamic topography and geoid (Simmons *et al.* 2009) instead of starting with a given tomography model.

1.3.3 Surface boundary conditions

Possible surface boundary conditions are free-slip (zero shear stress at the surface), no-slip (zero horizontal velocity at the surface), prescribed plate motions or free plate motions (Ricard & Vigny 1989) where forces from mantle flow on each plate are balanced. Obviously, intermediate cases, where plates are moving at some fraction of the speed of free plate motion are also possible. Forte & Peltier (1991, 1994) developed an alternative method for coupling the motion of rigid plates to buoyancy-driven mantle flow. Prescribed plate motions are not a suitable boundary condition for computing dynamic topography, if they do not correspond to the tractions acting on the plates: for example, if plates are pulled apart at a spread-

ing ridge in the model without corresponding forces, a dynamic topography low is predicted at the ridge. In all these approaches, no explicit radial displacements of the surface are computed, but instead normal stresses. These are converted to topography, using an appropriate density contrast. Assuming the crust is uplifted or subsiding as a whole without changing its thickness, the appropriate density contrast is the uppermost mantle density for dynamic topography beneath air, and the difference between uppermost mantle density and seawater density for dynamic topography beneath water. A formalism for dynamic topography and sea level consistent with partial water coverage has recently been developed (Spasojević & Gurnis 2012; Austermann & Mitrovica 2015).

1.4 A new look at dynamic topography

New developments warrant to take a renewed look at dynamic topography: recent progress in seismic tomography is illustrated in Fig. 1: already with older models, it has been possible to explain most of the geoid based on mantle density models derived from tomography. This can be understood from Fig. 1, taking Smean as an example: there is a strong positive correlation between geoid and tomography model in the degree and depth range where—for those viscosity models that allow a good fit to the geoid—kernels are most positive, and a negative correlation for degrees 2 and 3 in the lower half of the mantle, where kernels are negative (see Section 2 for a description of the kernel approach). In both regions, the tomography model has a reasonably high amplitude, hence the geoid, which is dominated by low spherical harmonic degrees can be explained well.

However, in this case geoid kernels are always strongly negative for higher degrees in the upper few hundred kilometres of the mantle. Hence, one would expect a negative correlation between geoid and tomography within that depth range, at least for depths where tomography has the highest amplitude (usually closest to the surface), and as long as seismic anomalies are correlated to density, despite the mentioned complications due to compositional variations. This negative correlation in the upper ≈200 km, where tomography amplitude is strongest, becomes much more pronounced for recent tomography models. Here, SL2013SV (Schaeffer & Lebedev 2013) is taken as an example. Hence, we expect that in the upper ≈200 km, which are most relevant for predicting dynamic topography, the new quality reached in mantle tomography can also lead to much more reliable predictions of dynamic topography, up to higher spherical harmonic degrees, and hence smaller scales. SL2013SV has already successfully been applied in a regional study of the North American lithosphere (Kaban *et al.* 2014b; Tesauro *et al.* 2014). The questions that I address here concerning density structure due to temperature and compositional anomalies are similar, but not focused on a specific region. Secondly, residual topography estimates for the oceans have also reached much higher quality recently, taking detailed information on sediment thickness and densities into account (e.g. Winterbourne *et al.* 2014), so it is also timely to provide an up-to-date geodynamic estimate of topography for comparison.

The principal goal of this paper is to assess how well geodynamic models of topography can agree with those models derived from actual topography, and under what assumptions the best fit can be obtained. The goal is further to learn, in this way, more about properties and processes in the Earth, and create a useful reference for future, improved models of dynamic topography both for the present and the geological past. Models of both lithosphere thickness and topography are derived from the same tomography
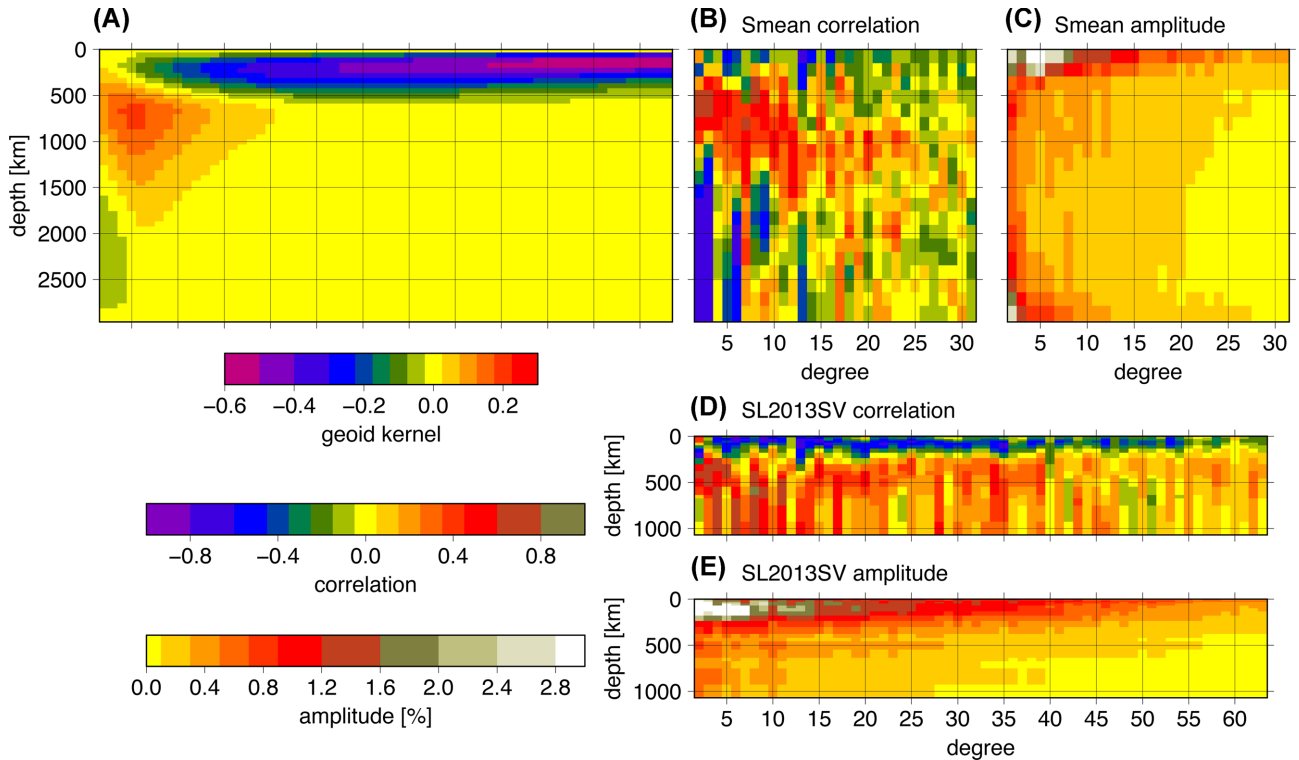
Figure 1. (A) Geoid kernels as a function of depth and spherical harmonic degree *l* for viscosity model 2b of Steinberger & Calderwood (2006), compressible flow, free-slip surface, no phase boundaries considered. (C) and (B) Amplitude of tomography model Smean (Becker & Boschi 2002) and its correlation with the geoid, as a function of depth and degree *l*. (E) and (D) Same for SL2013SV (Schaeffer & Lebedev 2013), which has smaller depth range (not the entire mantle) but higher resolution.

models. In contrast to many previous efforts, this allows including density anomalies right up to the surface in the computation of topography. This is important as shallow density anomalies are most effective in causing topography. Accordingly, together with using newer and higher resolved tomography models, this leads to residual and model topography agreeing much better and too much smaller scales than was previously possible.

2 METHODS

2.1 Observation-based estimates of dynamic topography

Topography isostatically compensated within the continental crust, and topography due to ocean floor age differences are subtracted from actual topography, and anomalous regions in the oceans are excluded. The resulting 'residual' topography is used as observation-based estimate of dynamic topography.

The integral $\int_{100\,\mathrm{km}}^{\mathrm{surface}} \rho(z)dz$ is evaluated with densities ρ including crust, oceans and ice sheets from the CRUST1.0 model (Laske *et al.* 2013) but using a constant mantle density 3350 kg m^{-3}. The lower limit of integration at depth 100 km is arbitrary as long as it is lower than the deepest crust. Although CRUST1.0 is currently probably the most accurate global crustal density model, density variations in the crust are not well constrained, relying mostly on an empirical relationship between seismic velocities and density and poor global coverage. If topography was isostatically balanced, this integral (pressure at 100 km divided by gravity acceleration) would be the same everywhere. But since it is not, it can be converted to non-isostatic topography by dividing by an appropriate density difference—either mantle density above sea level, or mantle den-

sity minus sea water density (1020 kg m^{-3}) below—and setting the global mean to zero (Fig. 2a). Apart from mid-ocean ridges, its main feature is that most continents are at lower elevation than they would be for isostatically compensated crust (e.g. Forte *et al.* 1993).

For topography due to ocean floor age differences (Fig. 2b)

$$h_{\mathrm{th}} = h_0 - 330\,\mathrm{m}\sqrt{\mathrm{age[Ma]}} \qquad (1)$$

(Korenaga & Korenaga 2008) resulting from a half-space cooling model is used for age < 100 Ma and a constant $h_{\mathrm{th}} = h_0 - 3300\,\mathrm{m}$ for age > 100 Ma. This flattening corresponds to results of litho-sphere thickness versus age obtained in the next section. The topic is further treated in Korenaga (2015). Where the age grid is undefined (essentially on continents), $h_{\mathrm{th}} = h_0 - 4365\,\mathrm{m}$ (corresponding to 175 Ma) such that residual topography has no obvious steps across the continent–ocean boundary. h_0 is chosen such that h_{th} has zero global mean. In contrast, with the older CRUST2.0 model $h_{\mathrm{th}} = h_0 - 3300\,\mathrm{m}$ on continents was found appropriate.

The oceanic crust is essentially laterally uniform in CRUST1.0 in contrast to what is suggested by the masks of thick oceanic crust excluded in Crosby *et al.* (2006) and Crosby & McKenzie (2009). Here, anomalous oceanic regions are excluded using Large Igneous Provinces (including major hotspot tracks) of Coffin *et al.* (2006) as a proxy. Residual topography in anomalous regions is interpolated by repeatedly averaging values at anomalous points from neighbouring points until convergence.

2.2 Computing mantle densities and lithospheric thickness

Neglecting compositional variations is probably adequate to first order beneath the lithosphere. Hence, in these regions mantle density
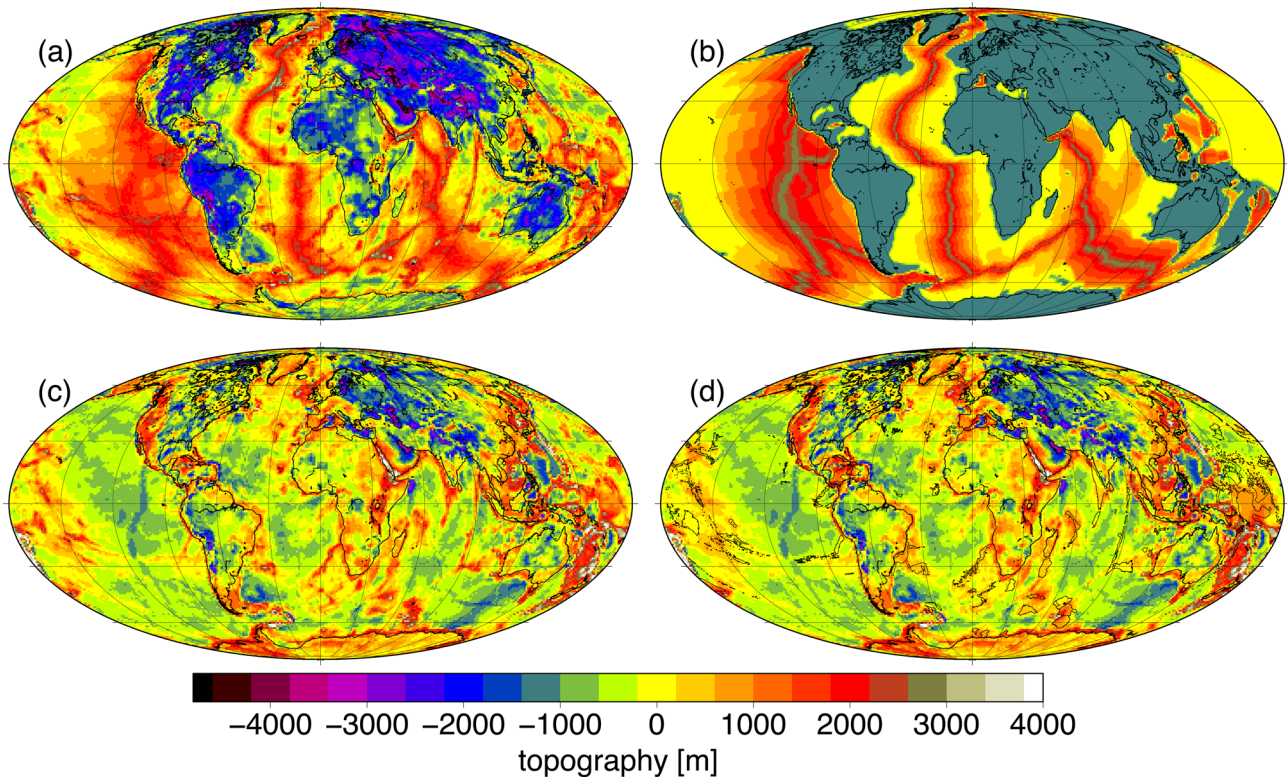
Figure 2. Steps toward an observation-based estimate of dynamic topography. (a) Non-isostatic topography derived from thickness and density of CRUST1.0 down to the Moho. (b) Topography derived from age of ocean floor according to the Müller *et al.* (2008) age grid version 3.6, available online at ftp://ftp.earthbyte.org/earthbyte/agegrid/2008/Grids. (c) Residual topography obtained by subtracting age-dependent ocean floor topography (b) from non-isostatic topography (a). (d) Oceanic Large Igneous Provinces and major hotspot tracks from Coffin *et al.* (2006), shown as black outlines, have been interpolated from their margins.

anomalies can be derived from global *S*-wave tomography models. Tomography models are expressed in a common spherical harmonic format (Becker & Boschi 2002) and formally expanded to $l_{max} = 63$ unless mentioned otherwise, although in some cases (e.g. S40RTS) the actual resolution is lower.

In order to get smoother variations, tomography models are interpolated radially, such that maximum layer spacing is 25 km above 400 km depth and 50 km between 400 and 1000 km depth. *P*-wave models are less suitable, because their ray coverage is comparatively poor in some regions of the upper mantle. Here, *S*-wave anomalies are converted to density anomalies assuming both are due to temperature variations and also considering anelasticity following Steinberger & Calderwood (2006), which is based on previous, original research cited there.

In contrast to previous work, I will not exclude density anomalies above a specified depth, but rather devise a model for the lithosphere, and consider a number of assumptions regarding density anomalies in the continental lithosphere. These are typically a lot smaller than anomalies $\delta\rho_{th}$ that would result from 'thermally' converting seismic velocity to density variations. Given that these non-thermal lithospheric density anomalies are poorly known and give a large contribution, adopting the simplified approach of Steinberger & Calderwood (2006) for thermal anomalies elsewhere should be regarded as sufficient, although adopting more accurate computations as mentioned in the introduction should be considered in future work. Distinction between continents and oceans is made here based on the Müller *et al.* (2008) age grid simply by assigning continents to regions where the age grid is undefined. The lithosphere is defined based on the tomography models themselves. The base of the

lithosphere is determined by the intersection of $\delta v_s/v_s$ and a cut-off function,

$$\delta v_s/v_s|_L = 3.1 \text{ per cent} - F_{tot} \cdot \left(1 - \text{erf}\left(\frac{z}{z_0}\right)\right) \qquad (2)$$

(black line in Fig. 4A). While this choice is somewhat ad-hoc, it is more advanced than, for example, simply using the zero anomaly as lithospheric base. Note that the value of z_0 is the same everywhere with a given tomography model, but different values may be used for different tomography models. The choice of this relatively simple procedure will be discussed in a separate manuscript (Steinberger & Becker, in preparation) and can further be justified by the fact that resulting lithosphere thicknesses for a given tomography model are always highly correlated regardless of the exact procedure. Moreover, it will turn out that the best fit between model and residual topography is obtained if minimum thickness is 70 km for continental lithosphere, and compositional variations are assumed to only occur in the upper ≈150 km of continental lithosphere. Thus, the sensitivity of our topography results to lithosphere thickness is reduced. Details on which value is adopted for lithosphere thickness in the rare cases where there is no unique intersection are described in the supplement of Steinberger *et al.* (2015).

Appropriate values for F_{tot} and z_0 can be found by visually matching the average thickness for a given ocean floor age determined with these values, with the thickness inferred from a half-space cooling model, which can explain the depth versus age relation (and hence presumably lithosphere thickness) for ages less than ≈80 Ma (Parsons & Sclater 1977). From the temperature
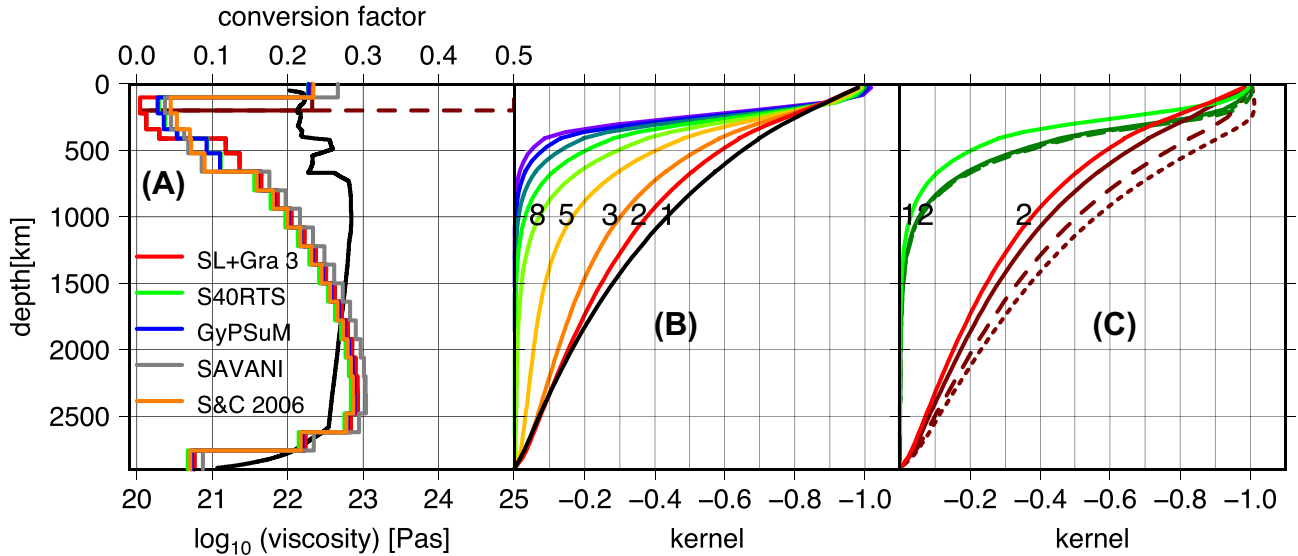
Figure 3. (A) Conversion factor from relative seismic velocity to relative density variations (black line) and viscosity structures used (other lines). The structures with 100 km lithosphere (red, green, blue and grey lines) were obtained through optimization as described in Section 3.1, for the reference case and three other tomography models as indicated. The orange line is the preferred viscosity model of Steinberger & Calderwood (2006). (B) Topography kernels for $l = 1, 2, 3, 5, 8, 12, 17, 23$ and 30 (from right to left), free-slip surface and CMB, compressible flow with self-consistent radial gravity and the viscosity structure with 100-km thick lithosphere shown as red line in the left-hand panel. Effects of phase boundaries are not shown. (C) Kernels for $l = 2$ (red) and 12 (green). Brightest colours are as in the middle panel. Darker colours are for 200 km viscous lithosphere, dashed lines for increased lithosphere viscosity 10^{25} Pa s (both also shown in left-hand panel) and the dotted line for the same 'dashed' viscosity structure, but with fixed surface instead of free-slip.

distribution $\quad T(z, t) = (T_m - T_s) \operatorname{erf}(0.5z/\sqrt{\kappa t}) + T_s \quad$ (Sandwell 2001) lithosphere thickness

$$z_L = \sqrt{\kappa t} = 10 \, \text{km} \sqrt{\text{age[Ma]}} \tag{3}$$

is obtained if the base of the lithosphere is assumed to occur at temperature $T_L = T_s + 0.843 \cdot (T_m - T_s)$ and diffusivity is $\kappa = 8 \cdot 10^{-7} \, \text{m}^2\text{s}^{-1}$.

2.3 Computation of viscous flow and model topography

Mantle flow is computed with the method of Hager & O'Connell (1979, 1981). Essentially, this method allows for the computation of flow and stresses in the mantle for given density anomalies, a viscous rheology with only radial viscosity variations, and given boundary conditions at the surface and CMB. Unless mentioned otherwise, compressibility, self-gravitation and phase boundaries at given depths in the mantle will be considered.

Assumptions about lithospheric density anomalies will be described, with each individual model, in the results section. Lithosphere thickness models and density anomalies beneath the lithosphere are inferred from seismic tomography models, as described in Section 2.2. Viscosity models are similar to the ones derived by Steinberger & Calderwood (2006) and Steinberger & Holme (2008) based on mineral physics and optimizing the fit to various observations. The optimization performed here is described in more detail in Section 2.4. As an example, Fig. 3 shows the resulting viscosity structure for four different models. They are quite similar regardless of tomography model, with sublithospheric viscosity somewhat above 10^{20} Pa s and increasing with depth to $\approx 10^{23}$ Pa s in the lower mantle above D″.

Mostly, a free-slip boundary condition is used both at the CMB and the surface, but with a lithosphere of appropriately high 'effective' viscosity. This boundary condition can give a good fit between predicted and observed geoid (Thoraval & Richards 1997), and is

consistent with neglecting LVV. However, in order to assess the influence of boundary conditions, some results for no-slip surface and prescribed plate motions will also be shown. Two phase boundaries will be normally included, mostly following Steinberger (2007) but without considering the latent heat effects discussed in that paper: phase transitions around 660 km effectively cause that density anomalies in a 89 km thick layer around that depth are 'removed', whereas the phase boundary around 410 km causes a 138 km thick layer around that depth to count twice.

Computations are performed in a fixed model domain bounded by two concentric spheres (surface and CMB), that is, topography is not an explicit part of the model. However, radial stresses τ_{rr} at the surface can be converted to topography (relative to the geoid) $h = \tau_{rr}/(\rho_s \cdot g)$ that would result if the surface was indeed free to move up and down. Here, ρ_s is an appropriate surface density contrast and g is gravity acceleration at the surface, and it is implied that the elastic strength of the lithosphere can be neglected, which is probably appropriate on large scales of 1000 km or more, given typical estimates of elastic lithosphere thickness. It will also be assumed that, in the presence of dynamic topography, the thickness of the crust and ice sheets, where present, is not affected. Therefore, the appropriate value for ρ_s is either uppermost mantle density for continents, or uppermost mantle density minus density of sea water for oceans. It is also possible to use the same density contrast everywhere, and instead modify residual topography computed in Section 2.1 by multiplying with a constant factor for either continents or oceans. This approach will be taken here with $\rho_s = 3300$ kg m^{-3} and residual topography beneath the oceans divided by $3.3/2.28 = 1.45$, corresponding to seawater density 1020 kg m^{-3}.

Mantle flow is computed in the spectral domain, that is, spherical harmonic expansion coefficients of the density field depending on depth in the mantle $\delta\rho_{lm}(z)$ (in the format introduced by Becker & Boschi 2002) are used to compute a spherical harmonic expansion of flow, stresses and model topography. For viscosity only varying in radial direction, computation can be done separately for each

spherical harmonic degree l. Specifically, this allows using a kernel formalism to express model topography coefficients as $h_{lm} = \int \delta\rho_{lm}(z)K_{t,l}(z)dz/\rho_s$. Topography kernels $K_{t,l}(z)$ depend on l, but not on order m. The middle panel of Fig. 3 shows topography kernels for the viscosity structure shown as bright red line, and for compressible flow with free-slip upper boundary. Comparison with kernels for other viscosity models (Steinberger *et al.* 2001; note the opposite sign convention) computed for stresses at the base of the lithosphere (100 km depth) shows that their shape does not strongly depend on viscosity structure or the depth level chosen (surface or 100 km). The kernels illustrate how effective density anomalies at different depth and degree l ('wavelength') are at creating topography. They all approach -1 at the surface, meaning that density anomalies near the surface are mostly isostatically compensated. They all approach zero with depth, such that density anomalies close to the surface—including those within the continental lithosphere, which are difficult to infer from tomography, as discussed before—are most important for creating topography. They approach zero faster with depth for higher degrees l. At the longest wavelength ($l = 1$ and 2), density anomalies in the lower mantle can still give a substantial contribution to dynamic topography whereas at shorter wavelengths corresponding to $l = 30$ or higher, basically only upper-mantle anomalies contribute to topography.

The right-hand panel in Fig. 3 shows results for different lithosphere thickness and viscosity, and surface boundary condition: magnitude of kernels becomes larger for thicker and more viscous lithosphere, and for no-slip instead of free-slip upper boundary. Since only radial viscosity variations are considered, the lithosphere in terms of viscosity structure has a constant thickness whereas variations in lithosphere thickness were considered when deriving the density model. Given that topography kernels are not too sensitive to lithosphere thickness (although they decrease somewhat less strongly with depth for a thicker lithosphere) it is more important to consider lithosphere thickness variations when inferring density anomalies, rather than for LVV (see also Moucha *et al.* 2007).

2.4 Optimizing the model and parameter space considered

Regarding model optimization procedure and parameter space, largely earlier work (Steinberger & Calderwood 2006) is followed, but topography is included as a quantity to be fit in the optimization. Specifically P_t, the topography misfit variance normalized to non-isostatic topography variance, for spherical harmonic degrees $l \leq 12$, is included in the misfit function. Including topography misfit in the optimization changes results merely to a somewhat lower sub-lithospheric viscosity (compare e.g. the viscosity for the reference case, bright red line in Fig. 3 to the preferred viscosity model 2b of Steinberger & Calderwood (2006)), yielding somewhat reduced topography amplitudes, but does not allow to fit individual features of residual topography. The pattern remains essentially the same. Therefore, it is really only the global rms amplitude of non-isostatic (or residual) topography acting as constraint in the optimization. Further, the misfit function includes penalties $P_{um-tz} = (q_{um} - q_{tz})^2$ if $q_{um} > q_{tz}$ and $P_{tz-lm} = (q_{tz} - q_{lm})^2$ if $q_{tz} > q_{lm}$, whereby q_{um}, q_{tz} and q_{lm} are the factors with which upper mantle, transition zone and lower mantle viscosity are multiplied relative to the unoptimized profile (Steinberger & Calderwood 2006, Fig. 4), thus suppressing solutions with transition zone viscosity higher than in the uppermost lower mantle, or lower than in the upper mantle: such features of the viscosity structure only marginally improve the fit. The Haskell constraint, that the logarithmic average of viscosity, weighted with an appropriate sensitivity kernel equals about 10^{21} Pa s (Mitrovica

1996), is required to be fit exactly. Therefore, q_{lm} is not an independent free parameter, but determined from q_{tz} and q_{um}. The misfit function to be minimized is

$$\text{MF} = 4P_{hf} + P_g + P_t + P_{tz-lm} + P_{um-tz} \qquad (4)$$

where P_t and P_g are the normalized topography and geoid misfit variance. P_t is defined above and P_g in eq. (32) of Steinberger & Calderwood (2006) with modeled and observed geoid both expanded to degree 15 and P_{hf} is heat flux misfit (eq. (31) of Steinberger & Calderwood 2006). $1 - P_g$ is referred to as geoid variance reduction. In addition to the viscosity factors, lithosphere density anomaly $\delta\rho_l$ is a further parameter varied during the optimization, in steps of 0.1 per cent. Alternatively, normalized gravity (instead of geoid) misfit variance was also used. This essentially amounts to giving more weight to higher spherical harmonic degrees and thus leads to somewhat increased variance, but with results (in particular model topography) remaining very similar.

3 RESULTS

3.1 Lithosphere thickness models

Fig. 4 shows lithosphere thickness for model number 1 of Table 1. This combination of tomography models is chosen because SL2013SV does not span the entire mantle depth, and this combination was found to give the best results (smallest misfit). With $F_{tot} = 6.2$ per cent and $z_0 = 150$ km not only the average oceanic lithosphere thickness versus age matches expectations, but also the overall pattern, with thin lithosphere close to ridges and older lithosphere being thicker. On the continents, thick lithosphere is found for many cratons (e.g. Bleeker 2003). A number of further tomography models were employed (not shown), in order to assess how robust or model-dependent our results are (Table 1). Many larger scale structures (cratons, dependence on seafloor age) generally appear independent of which model is used, whereas smaller features are often model dependent. A more thorough description of results, and comparison with other thickness estimates, will be given in a separate manuscript (Steinberger & Becker, in preparation).

3.2 Topography model reference case

In the reference case, density anomalies inferred from seismic velocity are replaced by $\delta\rho_l$ only to a maximum depth 150 km in the continental lithosphere. Furthermore, the minimum thickness of continental lithosphere is 70 km. The viscosity model shown as bright red line in the left-hand panel of Fig. 3, and $\delta\rho_l = +0.2$ per cent were found by minimizing the misfit function, as described in Section 2.4. Without applying the penalty P_{tz-lm}, transition zone viscosity would be higher than in the uppermost lower mantle, but misfit would only be marginally less. The computed geoid and the 'mantle geoid' for which variance reduction is computed are shown in Fig. 5.

Figs 6(A) and (B) show resulting model topography (effect of ocean floor cooling subtracted, as explained above) compared to residual topography. Many features in both maps match visually. If expansion is restricted to degrees $l = 1 - 12$ (not shown), correlation increases to 0.60 and ratio decreases to 1.20 (see also the analysis in the spectral domain done below). Numbers in the following will be for $l = 1 - 12$ unless mentioned otherwise.

If topography due to ocean floor cooling is not subtracted for both residual and model topography, their rms difference obviously
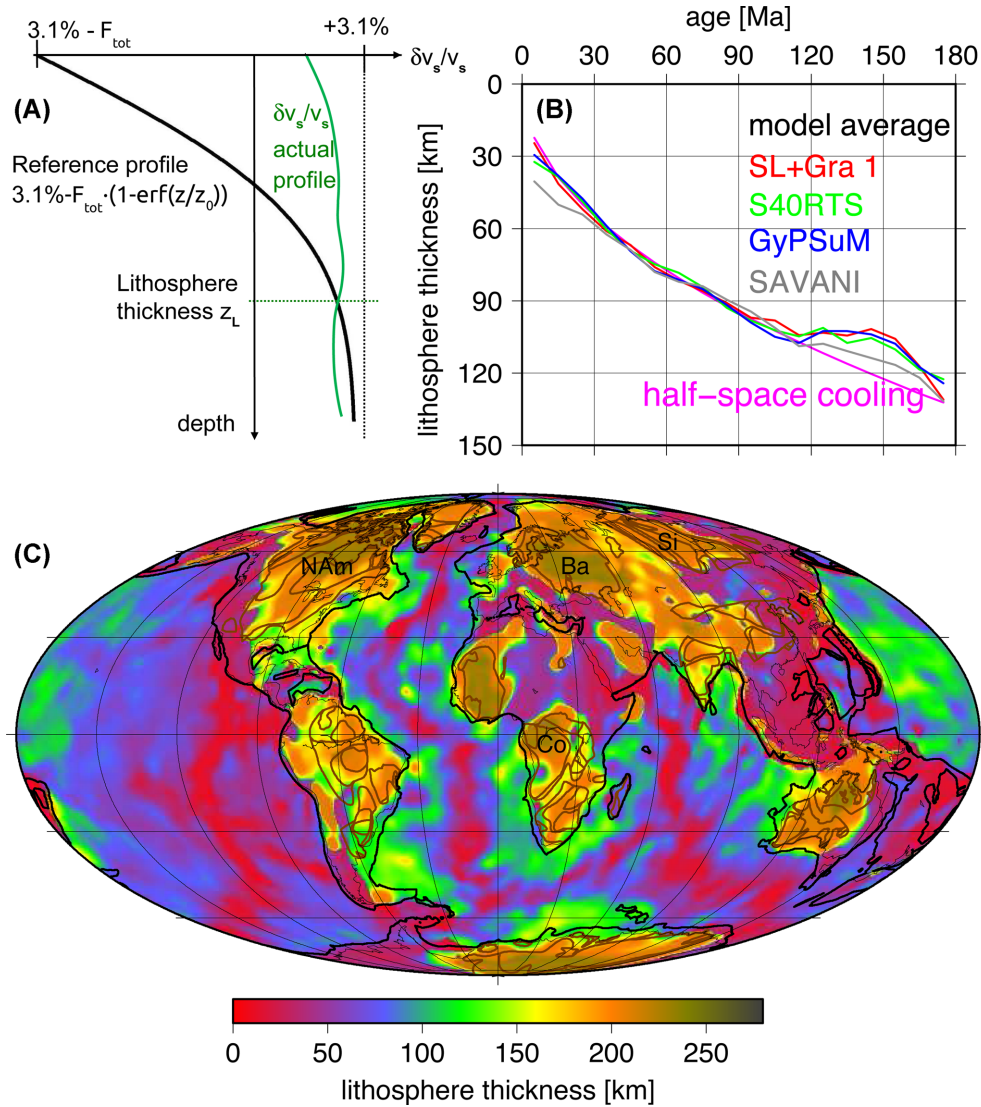
Figure 4. Lithosphere thickness determined by the method described here. (A) Sketch of seismic velocity anomaly and cut-off function as a function of depth to illustrate the method. (B) Violet line: lithosphere thickness z_L [km] $= 10\sqrt{\text{age[Ma]}}$ obtained from half-space cooling model. Red, green, blue and grey lines: average lithosphere thickness for given seafloor age determined for models numbers 1, 8, 6 and 11 of Table 1. (C) Map of lithosphere thickness for model number 1. The thick black line delimits the regions where the Müller *et al.* (2008) age grid is defined as a proxy for the continent–ocean boundary. Cratons outlines (Gubanov & Mooney 2009) are drawn in brown. NAm=North America, Ba=Baltica, Si=Siberia, Co=Congo.

remains the same, but both patterns change considerably (Figs 6C and D). Results also change on continents, since the global mean of model topography is set to be zero. Since topography due to ocean floor cooling is rather large and predicted well, including it gives a much higher correlation and the amplitude also matches better. While the interpretation of Figs 6(C) and (D) compared to parts (A) and (B) is more straightforward, as it does not involve assuming any oceanic age versus depth curve and subtracting corresponding topography, the reason why in most figures this is nevertheless subtracted is merely because in this way remaining differences between model and residual topography become more obvious.

3.3 Effect of density anomalies, phase boundaries and compressibility

While in the reference case thickness of the compositionally different part of the continental lithosphere only varies between 70 and 150 km, results do not critically depend on this depth range: they

stay similar (see Table 1) if variations between crustal thickness and 150 km are allowed or if we treat the entire continental lithosphere as compositionally distinct, but in the latter case a larger lithosphere density anomaly has to be assigned for the best fit. Thus, there is a trade-off between the depth range over which the lithosphere is assumed chemically distinct and the best-fitting density anomaly: if one still considers seismic anomalies in the lower part of the lithosphere to be due to thermal anomalies, the inferred relatively strong density anomalies pull the lithosphere sufficiently down and a rather small anomaly in the upper part of the lithosphere is required for the best fit. But if one assumes a constant density anomaly in the entire lithosphere, it has to be larger.

In the reference model, the transition from the SL2013SV tomography model to the Grand10 model is made at depth 200 km. A transition at a shallower depth would probably be inappropriate because the high correlation of SL2013SV with the geoid in the upper 200 km was the reason for using it in the first place. A transition at a larger depth leads to increased misfit, for example, from 0.679

Table 1. Summary of model parameters and results. Model: tomography model(s) used—SL+Gra combines SL2013SV (Schaeffer & Lebedev 2013) above 200 km and Grand10 (2010 model update of Grand 2002) beneath, TX2008 (Simmons *et al.* 2009), GyPSuM (Simmons *et al.* 2010), S20RTSb (Ritsema *et al.* 2004), S40RTS (Ritsema *et al.* 2011), Smean (Becker & Boschi 2002), SE+Gra combines SEMUM2 (French *et al.* 2013) above 200 km and Grand10 beneath, SAVANI (Auer *et al.* 2014). Lith: bounds of compositional lithosphere thickness (in km) on continents. c stands for crustal thickness as minimum. F_{tot} and z_0 are parameters of the cut-off function eq. (2), η_l is viscosity for 0–100 km depth, η_{as} for 100–220 km, η_{tz} for 410–520 km, η_{lm} for 660–800 km, η_0 is scaling viscosity 10^{20} Pa s. Relative variations of viscosity with depth within the upper mantle (100–410 km), transition zone (410–660 km) and lower mantle (below 660 km) are the same as for model 2b of Steinberger & Calderwood (2006)—see fig. 13 of that paper. MF is total misfit (eq. 4), VR is geoid variance reduction, Corr and MRR are correlation and ratio between model and residual topography. Model number 3 is the reference case.

| Number	Model	Lith	F_{tot} [per cent]	z_0 [km]	$\delta\rho_l$ [per cent]	$\frac{\eta_l}{\eta_0}$	$\frac{\eta_{as}}{\eta_0}$	$\frac{\eta_{tz}}{\eta_0}$	$\frac{\eta_{lm}}{\eta_0}$	MF	VR [per cent]	Corr	MRR
1	SL+Gra	c-400	6.2	150	0.4	169	1.20	14.4	40.6	0.679	71.0	0.60	1.24
2	SL+Gra	c-150	6.2	150	0.2	160	1.22	14.4	39.5	0.681	70.4	0.62	1.29
3	SL+Gra	70-150	6.2	150	0.2	209	1.10	15.1	44.1	0.667	70.4	0.60	1.20
4	Grand10	c-400	9	130	0.2	178	1.12	15.2	42.2	0.750	71.1	0.45	1.11
5	TX2008	c-400	9	130	0.3	327	2.55	5.11	45.7	0.727	71.1	0.43	0.99
6	GyPSuM	c-400	9	130	0.2	188	1.87	8.33	40.1	0.718	75.4	0.42	1.10
7	S20RTSb	c-400	7	160	0.1	187	1.90	8.66	36.9	0.969	66.4	0.42	1.39
8	S40RTS	c-400	8	140	0.1	185	1.95	8.55	36.0	1.018	71.2	0.42	1.53
9	Smean	c-400	10	120	−0.1	200	2.87	5.13	37.2	0.975	72.0	0.39	1.45
10	SE+Gra	c-400	6.6	165	0.3	170	1.15	14.9	41.5	0.687	71.6	0.54	1.15
11	SAVANI	c-400	7	160	0.5	463	2.35	4.79	56.9	1.022	44.7	0.61	1.39
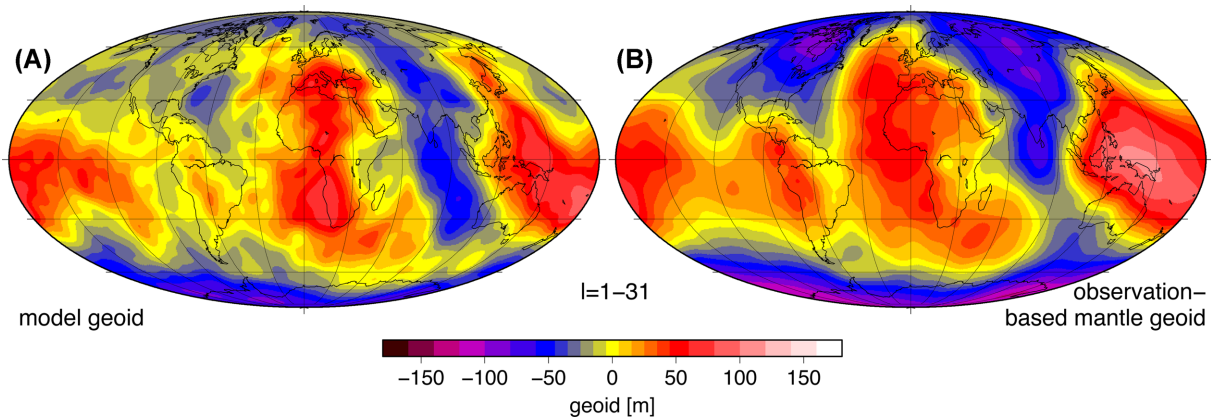
Figure 5. (A) Computed geoid in the reference case. (B) Mantle geoid = observed geoid (Pavlis *et al.* 2012) minus contribution down to the base of the crust derived from CRUST1.0 (Laske *et al.* 2013) assuming isostatic compensation and relative to the equilibrium spheroid (Nakiboglu 1982).

to 0.702 for the first model in Table 1 if it occurs at depth 250 km instead. Also it is not an obvious choice which tomography models to combine. Combination of SL2013SV with Grand10 was found to give the lowest overall misfit among several models tried (not shown). One reason for this is probably that the Grand10 model has comparatively low amplitudes except in the boundary layers below the surface and above the CMB, so this combination yields comparatively low model topography amplitudes more similar to residual topography.

A constant density anomaly in the part of the lithosphere that is assumed chemically distinct is a very simple and perhaps unrealistic assumption. Accordingly, a number of further assumptions has been tried out: first, instead of assuming a constant total density anomaly, a thermal anomaly is derived from an error function temperature profile corresponding to the lithosphere thickness model shown in Fig. 4. Then, a constant chemical anomaly is added to this in the continental lithosphere. Since the thermal anomaly alone would cause a too strong negative model topography for thick continental lithosphere, a negative (that is, buoyant) chemical density anomaly is required: if it is added to the entire lithosphere, the best fit is obtained for −1.9 per cent, if it is only added above 150 km depth, for −2.5 per cent. However, this seemingly more realistic assumption does not improve results: rather, total misfit slightly increases to 0.681 and 0.724, respectively.

Several further attempts started from converting seismic anomalies to density anomalies with a thermal conversion factor everywhere including the lithosphere and then modifying this model: model topography for the unmodified 'thermal only' model is shown in Fig. 7(A). In this case, inferred density anomalies in the continental lithosphere are much stronger than in the case of Fig. 6. Because topography kernels are largest at shallow depth (Fig. 3) this has a rather large effect on model topography: it gives very pronounced model topography lows for cratons, whereas residual topography lows tend to be much less pronounced, except for Baltica, where the residual topography low is nearly as large. Accordingly, the ratio of model topography to residual topography is increased and the overall fit gets worse. However, the fact that, for example, for Baltica computed negative residual topography is more pronounced than model topography in the reference case (Fig. 6A) and nearly as large as in Fig. 7(A) indicates that density anomalies there may on average be somewhat larger than the 0.2 per cent that were obtained for the best-fit model.

Since such a thermal scaling is obviously not applicable to continental lithosphere where compositional effects are known to play an important role (see references in the Introduction) this model was modified in a couple of ways: in one attempt, a constant density anomaly is again added within the continental lithosphere above 150 km. Again, because cratons tend to be too low without it, the
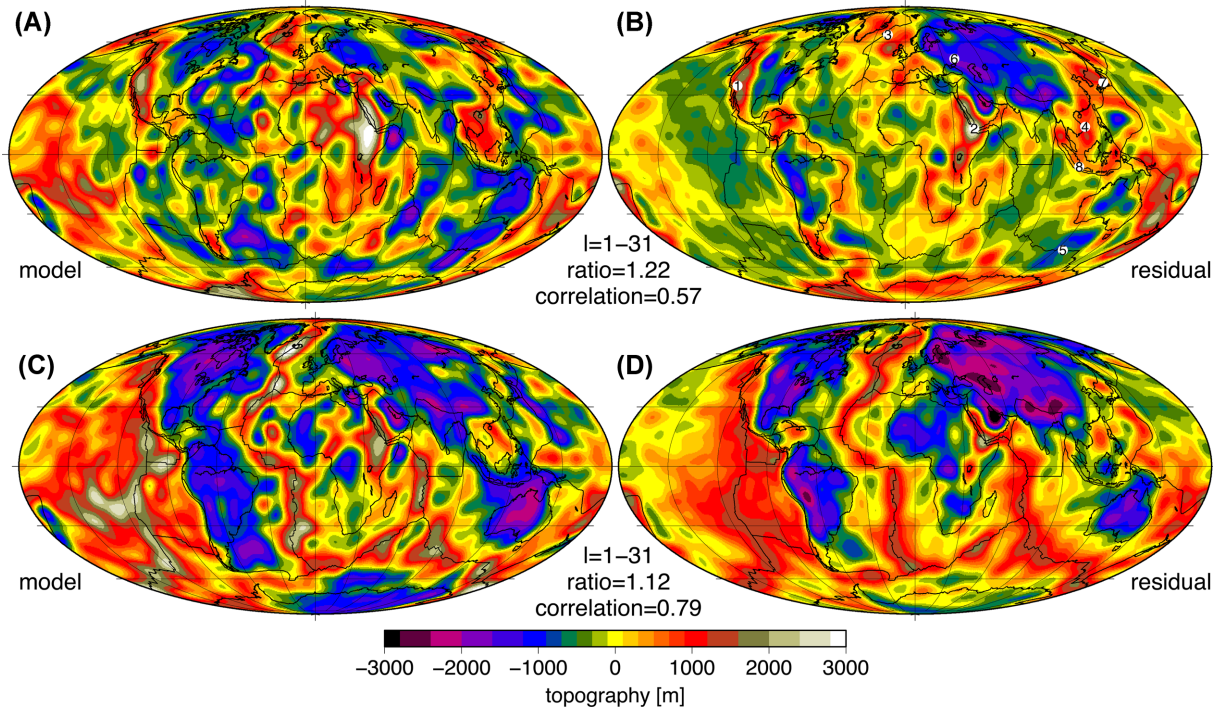
Figure 6. (A) and (C) Computed model topography in the reference case computed 'beneath air' with uppermost mantle density 3300 kg m^{-3}. (B) and (D) Residual topography, divided by 1.45 in the oceans to account for water coverage. In (A) and (B) the effect of ocean floor cooling is subtracted whereas in (C) and (D) it is included. Locations mentioned in the text shown in panel (B) 1=western U.S., 2=East Africa, 3=Iceland, 4=South China Sea, 5=Australia-Antarctic Discordance, 6=Eurasian part of Russia, 7=Japan, 8=Indonesia.
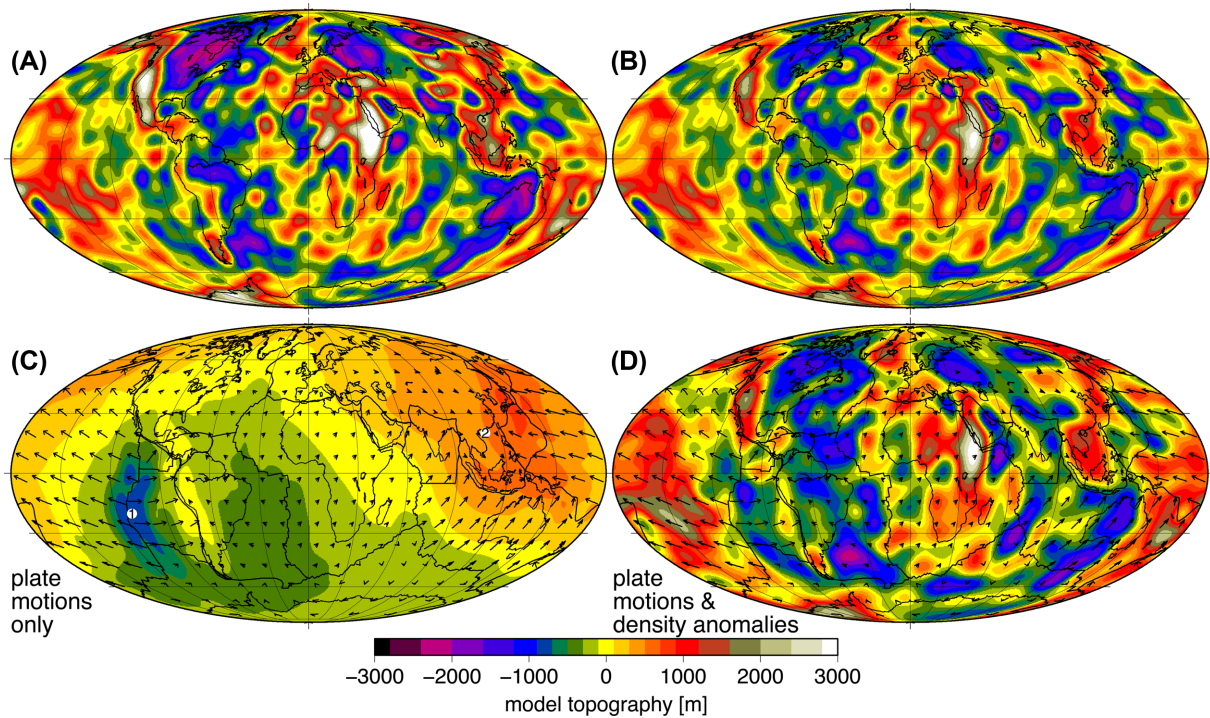
Figure 7. Computed model topography (A) with density anomaly in continental lithosphere converted from seismic tomography as in other regions, (B) with variable lithosphere thickness: for each point, the model topography is computed for only radial viscosity variations, but with lithosphere thickness modified to match Fig. 4, (C) and (D) with 0–10 Ma plate motion boundary condition (Seton *et al.* 2012), expanded to $l = 31$ and with a cosine taper for $l = 16-31$, shown with arrows (1° of arc = 1 cm yr^{-1}) and plate boundaries. (C) Topography due to prescribed plate motions only, with no internal density anomalies. 1=East Pacific (Rise), 2=East Asia. (D) Topography with density anomalies and prescribed plate motions. Other modeling assumptions as in Fig. 6(A).
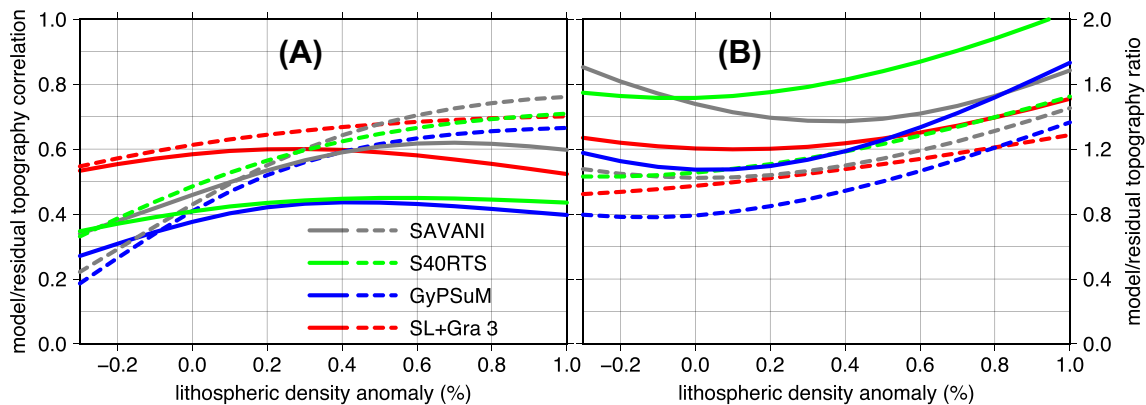
Figure 8. Correlation (A) and ratio (B) between residual topography and model topography, for four cases as indicated, with corresponding viscosity structures (Fig. 3) as a function of $\delta\rho_1$. Surface boundary conditions are free-slip. Dashed lines are for continents only.

fit improves for a negative anomaly, and the minimum misfit MF = 0.845 occurs for −0.3 per cent (compared to 0.879 for 0 per cent). In a further attempt, the difference between the density model inferred from scaling the tomography model thermally, and the theoretical model with an error function profile corresponding to the lithosphere thickness model is scaled with a different factor. The idea behind this assumption is that seismic anomalies are due to both temperature and compositional anomalies, but the scaling from velocity to density anomalies is different for compositional anomalies. In this case, the optimum fit (MF = 0.866) occurs if the conversion for the difference (thought to correspond to compositional anomalies) is scaled down by a factor 0.09. But in both cases, misfit is only slightly reduced, compared to the much larger reduction in the reference model where seismic anomalies in the lithosphere are ignored altogether. Hence, results seem to indicate that it is not straightforward to relate seismic velocity and density anomalies in the continental lithosphere.

In order to assess model variability depending on tomography model, several of them are used. Lithosphere parameters F_{tot} and z_0 are determined as explained in Section 2.2. Parameters are listed and results are summarized in Table 1. Among them, SEMUM2 is again not defined for the entire mantle and therefore again combined with Grand10, giving a very similar misfit MF. All other tomography models are used for the whole mantle. For them, overall misfit is higher and correlation between model and residual topography lower, justifying our choice of reference model. A more extensive comparison for many different tomography models, although with somewhat different modeling assumptions, for example, regarding definition of the lithosphere, can be found in Steinberger *et al.* (2010). Results shown and discussed here are fairly representative for this larger set of results.

Features that are repeatedly found and that can also be seen in Figs 6(A) and (B) include the large topographic highs in the western U.S. and East Africa, which approximately match corresponding features in the residual topography map. Also, models generally predict topography highs around Iceland and in the South China Sea, and lows along the Australia-Antarctic discordance. For all of these features, similar ones can be found in residual topography. Typical mismatches include that negative model topography in the European part of Russia is much less pronounced than the residual topography low. On the other hand, pronounced negative topography is predicted near subduction zones, such as around Japan and Indonesia, whereas residual topography can be even positive there. A possible reason for such discrepancy is that in the upper mantle subducted slabs are attached to the plate ('slab pull'), whereas

our model only considers viscous coupling ('slab suction'; Conrad & Lithgow-Bertelloni 2004). In addition, the results in Figs 6(A) and (B) agree better in several smaller scale features than previous results in Steinberger *et al.* (2010).

To further illustrate the effect of varying density anomalies in the continental lithosphere, Fig. 8 shows how correlation and ratio changes as a function of lithospheric density anomalies $\delta\rho_1$. There is generally a trade-off: while the model versus residual topography correlation mostly increases with $\delta\rho_1$ their ratio also increases to values much larger than the optimum value 1.0 for $\delta\rho_1 \gtrsim 0.4$. The best overall fit occurs for $\delta\rho_1$ around 0.1–0.5, where both the correlation is close to its maximum, and the ratio is close to 1.0.

Another region in the Earth where compositional density anomalies possibly play an important role are the LLSVPs (e.g. Masters *et al.* 2000; not shown). Following Steinberger & Holme (2008), in a further model, a (compositional) density anomaly of 1.2 per cent has been added wherever in the lowermost 300 km of the mantle the *s*-wave anomaly is more than 1 per cent negative. As expected from kernel magnitude for the lowermost mantle, the effect of compositional anomalies was found to be small. Model-to-residual ratio is reduced from 1.20 to 1.17, whereas correlation remains 0.60. Also, whether or not phase boundaries or compressibility are considered has only a minor effect on model topography. If phase boundaries are excluded correlation is 0.58 instead of 0.60 and amplitude ratio 1.18 instead of 1.20. Latent heat release at phase boundaries, which is not considered here, causes about 5 per cent decrease in topography amplitude (Steinberger 2007). If compressibility is not considered, correlation remains 0.60 and ratio again decreases from 1.20 to 1.18.

3.4 Effect of viscosity structure and surface boundary condition

Variation in amplitude for different viscosity structures and surface boundary conditions are shown in Fig. 9. If further lithosphere viscosity increases to 10^{25} Pa s for thickness 200 km (dashed brown line in Fig. 3 left; otherwise as reference model), amplitude ratio becomes 1.47, and if furthermore a no-slip instead of free-slip surface boundary condition is applied, it reaches 1.68—a trend that would be expected from the shape of the kernels in Fig. 3. On the other hand, the pattern of predicted topography does not change so much. For the cases corresponding to the red line in the right-hand panel of Fig. 9 (like the reference case, apart from lithosphere thickness) correlation only varies between 0.55 and 0.62. For the case
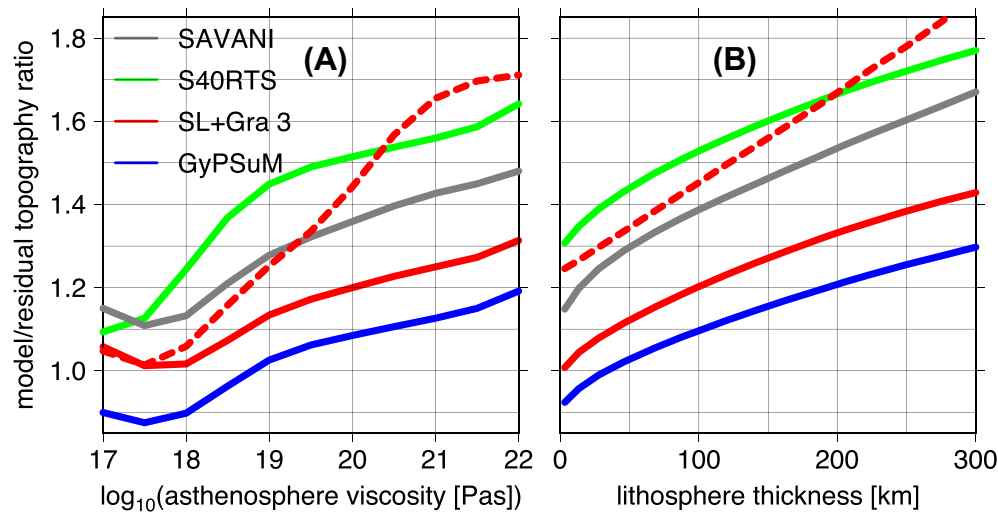
Figure 9. Model topography amplitude (normalized to residual topography) (A) as a function of viscosity in the depth range 100–220 km, for free-slip (continuous lines) or no-slip (dashed line) surface boundary condition; (B) as a function of lithosphere thickness in the viscosity model. Elsewhere, viscosity is identical to the respective case in Fig. 3(A).

with lithosphere viscosity 10^{25} Pa s correlation is also 0.55. In the case with no-slip boundary condition, it is slightly reduced to 0.53, mainly due to a large degree-one component in model topography (not shown), negative in East Asia and positive in the East Pacific, which is not present in the residual topography pattern. This discrepancy can be attributed to the fact that at the global scale, a free-slip boundary condition represents the actual plate motions more appropriately than no-slip.

Increasing lithosphere thickness in the viscosity structure means decreasing thickness of the low-viscosity upper mantle, hence the lithosphere is more strongly coupled to the underlying mantle convection, and model topography amplitude is increased. In the same way, amplitude also varies as a function of viscosity in a certain depth layer: Fig. 9 shows that model topography amplitude increases by not more than ≈ 50 per cent, if asthenosphere viscosity in the depth range 100–220 km is increased from 10^{17} to 10^{22} Pa s with a free-slip surface boundary condition. With no-slip, amplitude increases by about a factor 1.7. Again, the pattern remains rather similar over a wide viscosity range. Correlation between model and residual topography remains between 0.58 and 0.61 for asthenosphere viscosity 10^{18} to 10^{22} Pa s, and everything else like the reference model. However, for 10^{17} Pa s it drops to 0.45. Also, for the other tomography models, correlation remains similar from 10^{18} to 10^{22} Pa s but is lower for 10^{17} Pa s.

The similarity of resulting model topographies for various viscosity models shows again that adjusting the viscosity structure to optimize the fit between model and residual topography does not change results in a major way. As another example, if we use viscosity model 2b of Steinberger & Calderwood (2006) instead of the 'optimized' model in the reference case, correlation between model and residual topography remains 0.60, and their ratio only slightly increases to 1.24.

Since results for different lithosphere thicknesses are only mildly different, a series of calculations was carried out for different values, and the results were combined in Fig. 7(B), using at each point the lithosphere thickness as in the bottom panel of Fig. 4. Model topography most strongly depends on density anomalies at shallow depth beneath, therefore it can be expected that this pattern is more realistic than for any constant lithosphere thickness, and can be compared

to results of full 3-D computations, which should be the topic of future work. The pattern matches residual topography slightly better than for constant lithosphere thickness—correlation increases from 0.60 to 0.61 and ratio from 1.20 to 1.21. Also, visual agreement is slightly better. For example, there are more pronounced negative anomalies for the Congo craton and Russia (Baltica and Siberia Cratons, Fig. 4), and modeled topography maxima in the Pacific are less pronounced. Because of the relatively weak dependence of results on lithosphere thickness (Fig. 9B), effects are small and any discontinuities introduced by this procedure are not obvious in Fig. 7(B).

Results so far were with free-slip or no-slip boundary conditions, but the most realistic boundary conditions are plate-like. However, if plate motions are prescribed and do not correspond to forces acting on plates, computed topography will show features that may not be realistic. Accordingly, free-slip boundary conditions result in the best fits to the geoid (Thoraval & Richards 1997). This is illustrated in Fig. 7(C): with no underlying density anomalies driving mantle flow that would be the cause of plate motions, just kinematically prescribing these plate motions causes almost 1 km negative topography above the fastest spreading centre, the East Pacific Rise and conversely, almost 1 km positive topography in East Asia, a region where plates converge. One can, for example, imagine that diverging flow caused by forcing plates apart creates a void, that is, negative topography.

In the Hager & O'Connell (1979, 1981) approach, total model topography with prescribed plate motions and underlying density anomalies (Fig. 7D) is computed as the sum of topography with prescribed plate motions only (Fig. 7C) and topography with no-slip boundary condition (not shown). As mentioned, in the latter case, predicted topography has a large degree-one component—negative in East Asia and positive in the East Pacific. But this is largely compensated by the topography shown in Fig. 7(C), such that the resulting topography with prescribed plate motions and internal anomalies is again very similar to the one for a free-slip boundary condition (compare Figs 6A and 7D): both show a very similar pattern, but topography amplitude is somewhat higher for prescribed plate motions (amplitude ratio 1.43 instead of 1.20). In this case, also correlation with residual topography is slightly higher (0.61 instead of 0.60).
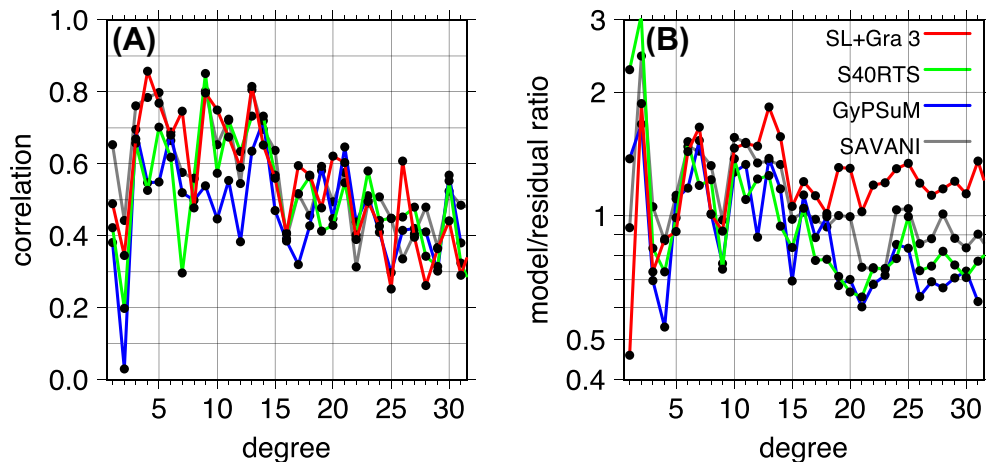
Figure 10. Correlation and ratio of model and residual topography for four models as indicated, with corresponding viscosity structures as in Fig. 3, as a function of degree *l*.

3.5 Comparison in the spectral domain

To further analyse the causes of differences between model and residual topography, correlation and ratio of model and residual topography are also computed separately for each degree *l*. Results shown in Fig. 10 for four tomography models primarily show two trends: both correlation and ratio tend to be higher for lower degrees. In particular, there is a peak in model-to-residual ratio at degree two, because model topography contains a large degree-two signal with highs above the LLSVPs, and lows above the ring of subduction zones surrounding the Pacific, which is much less prominent for residual topography. These results are a common feature for a large number of tomography models (Steinberger *et al.* 2010). The case based on GyPSuM does not overpredict topography as much as other cases at *l* = 2. Fig. 10 shows particularly high correlations in the range *l* = 3−14.

Generally higher correlation for lower degrees is probably a consequence of the topography model and/or residual topography being better constrained at larger scales. The trend in model-to-residual ratio could be due to too little model topography computed at high degrees, for example, if tomography models have limited resolution, or too much residual topography, for example, if there are artefacts resulting from the crustal model, or a combination. The fact that the second trend mostly disappears for the model based on Schaeffer & Lebedev (2014), which is presumably better resolved at shallow depths than earlier models, points toward the first cause.

3.6 Regional differences

If the analysis for the reference model is done separately for continental and oceanic (Müller *et al.* 2008) regions, correlation is higher (0.64 versus 0.55) and ratio is lower (1.02 versus 1.57) for continents. These differences in correlation and amplitude ratio are also typical for a larger number of models (Steinberger *et al.* 2010). Higher correlation could indicate that model topography or residual topography, or both, are better known for continents, due to better resolution of tomography models and/or more accurate crustal thickness models, despite the seemingly simpler crustal structure in the oceans. The difference in correlation is even larger for older tomography models—that is, the increase of correlation with newer models is stronger in the oceans, where tomography is becoming better resolved and more appropriate for mapping density anomalies.

Higher amplitude ratio in the oceans reflects lower residual topography amplitude in the oceans. Rms residual topography amplitude in the oceans (computed for seawater coverage) is 473 m for expansion to degree *l* = 12 and 534 m for *l* = 31, ≈20–30 per cent larger than the 400 m found by Korenaga & Korenaga (2008). Lower amplitude ratio for continents is probably due to neglecting LVV. Since the low-viscosity asthenosphere is probably more pronounced beneath oceans, the consideration of LVV would reduce amplitudes for oceans, thus making amplitude ratios more similar to continents. For the case shown in Fig. 7(B), which considers lithosphere thickness variations in a simplified fashion, the difference in ratio for continents versus oceans is reduced from 54 to 43 per cent. This rather small drop indicates that differences between continents and oceans are probably even larger than assumed here—in particular, it is likely that the asthenosphere also has lower viscosity beneath the oceans, which is not considered here. For comparison, Ghosh *et al.* (2010) find that dynamic topography magnitudes are affected by as much as 20 per cent when LVV were introduced.

To further analyse regional differences, Fig. 11 shows how correlation and ratio varies regionally. Correlation is often good in regions with strong features, where the pattern agrees in model and residual topography—including the prominent highs in eastern Africa and the western U.S. The pattern for ratio can be partly explained in a similar way as the difference between continents and oceans: the smallest ratios occur in northern Eurasia (Russia), where there are also two large cratons (Baltica and Siberia; Fig. 4) hence a good coupling of the lithosphere with mantle convection, such that model topography is underpredicted. The largest ratio occurs around the East Pacific Rise, the fastest spreading centre and therefore presumably the largest area of thin lithosphere and pronounced low-viscosity asthenosphere beneath, such that model topography is overpredicted. Again, this general pattern of correlation and ratio is a robust feature found for a large number of models (Steinberger *et al.* 2010). However, there is also a large craton in North America, where the ratio is not particularly low. Hence there are probably also other factors influencing this ratio.

3.7 Comparison with and among different residual topography models

The observation-based estimates to which model topographies are compared are also subject to uncertainties. In particular, the
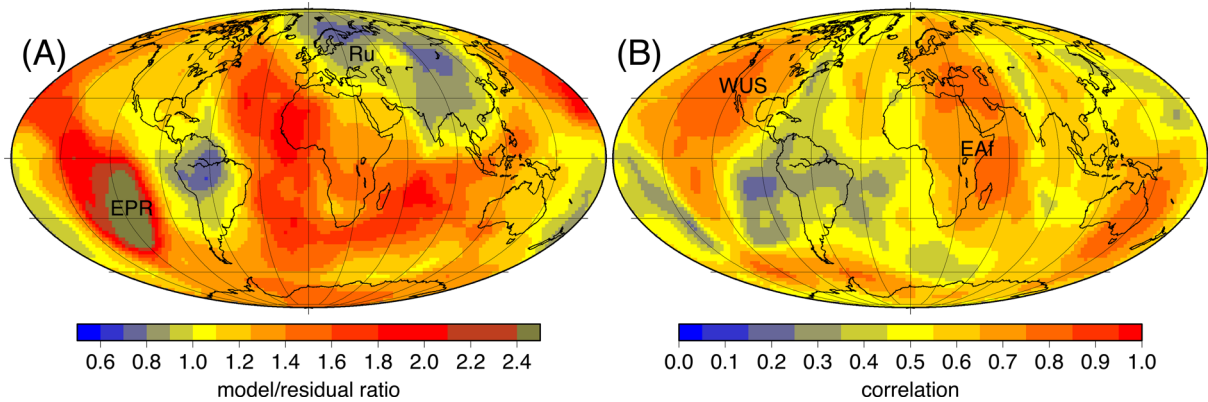
Figure 11. Ratio (A) and correlation (B) of model and residual topography within a circle with radius 30° of arc around each point, for the reference model. EPR=East Pacific Rise, Ru=Russia, WUS=western U.S., EAf=eastern Africa.

residual topography model of Fig. 2 is based on a global model of crustal thickness and densities and does not account for subcrustal lithospheric density variations, and comparison to more detailed regional studies or compilations, both on continents (Tesauro *et al.* 2008) and in the oceans (Wheeler & White 2000; Czarnota *et al.* 2013; Winterbourne *et al.* 2014) is more appropriate.

In order to get an estimate of the possible errors introduced by the approximations made in CRUST1.0, we also include a comparison with the recent results of Winterbourne *et al.* (2014), Table 1, who treat crustal thickness and sediment cover in the oceans in greater detail, as the link between them is not straightforward. Differences in Fig. 6 can be due to errors in both the model and residual topography estimate, and this comparison further illustrates the considerable uncertainties. Since Winterbourne *et al.* (2014) only list residual topographies for certain locations, a different representation is used in Fig. 12. For comparison, the model versus residual topography estimates of Figs 6(A) and (B), and the two residual topography estimates are plotted against each other in the same fashion: there are considerable differences not only between model and residual topography, but also between the two residual topography estimates (Fig. 12C).

Correlations between model and residual topography are generally lower than in Table 1, however this is largely because correlations in oceans tend to be lower than on continents, and they also tend to be lower if topographies are not filtered to maintain only longer wavelengths. So absolute values of correlations cannot be directly compared. Correlation is highest in the reference case, which could again indicate that, among those models and including smaller scale features, the SL2013SV tomography model gives the best representation of density heterogeneities in the upper ≈200 km. As was found for oceans in general, model topography rms amplitudes exceed those of residual topography by up to a factor ≈2—model topography mostly ranges between ±2 km, whereas the residual topography estimates mostly range between ±1 km.

4 DISCUSSION

The main purpose of this work is to assess what topography signal one can expect based on what is known about density anomalies and rheology of the Earth mantle, and how this compares with what can be inferred from observed topography. Often, seismic velocity anomalies are taken as a proxy for density and temperature anomalies (Steinberger & Calderwood 2006; Simmons *et al.* 2009), but in the lithosphere, variable composition also influences

seismic velocities and densities (Deschamps *et al.* 2002; Griffin *et al.* 2009; Cammarano *et al.* 2011) so a simple velocity-to-density anomaly conversion does not work there. Hence, density anomalies at shallow depth, which are most effective in causing topography, are poorly constrained. Lithosphere thickness models (Shapiro *et al.* 2004; Artemieva 2006; Conrad & Lithgow-Bertelloni 2006; Li *et al.* 2007; Rychert *et al.* 2010; Priestley & McKenzie 2013; Pasyanos *et al.* 2014) help to infer where the 'thermal' velocity-to-density conversion probably would not work. For the sake of consistency, thickness models are derived here based on the same tomography models used to infer mantle density.

The best fit between models and observations is found for very small and slightly positive density anomalies in the continental lithosphere, approximately consistent with the isopycnic hypothesis of Jordan (1988). However, results do not deteriorate if a thermal scaling of tomography is assumed at any depth greater than 150 km regardless of whether the anomaly is located within or beneath the lithosphere, roughly consistent with the finding of a mid-lithospheric discontinuity (Rader *et al.* 2015), and the idea that this separates a compositionally distinct lithospheric layer above from a thermal boundary layer beneath (Yuan & Romanowicz 2010). In this case, lithosphere thickness >150 km inferred here does not affect results. In the reference model, the total density anomaly is $\delta \rho_l = +0.2$ per cent. From eqs (1) and (3), an average thermal anomaly of 2.28 per cent can be inferred. Therefore, this corresponds to a 'chemically buoyant' (i.e. chemical density anomaly negative) continental lithosphere, where the relative importance of chemical to thermal buoyancy is ≈0.9, similar to the value ≈0.8 found by Forte & Perry (2000). Below the lithosphere, the scaling factor from seismic velocity to density anomalies was not varied. Model topography amplitudes, which are somewhat larger than for residual topography, could be reduced by using a lower scaling factor.

The best fit found between model and residual topography is a correlation of 0.6 and amplitude ratio 1.2. As also found by Steinberger *et al.* (2010) for a larger set of models, correlation is somewhat higher for lower resolution. Large-scale patterns are generally matched better than smaller scale structure. But compared to Steinberger *et al.* (2010), it is now possible to obtain a good match even for many regional-scale features in both location and amplitude. Using independent estimates of lithosphere thickness for computing model topography generally gives a worse fit (results not shown), presumably because in this case more often strong positive seismic anomalies are interpreted to be located beneath the lithosphere and hence converted to density anomalies, giving features in model
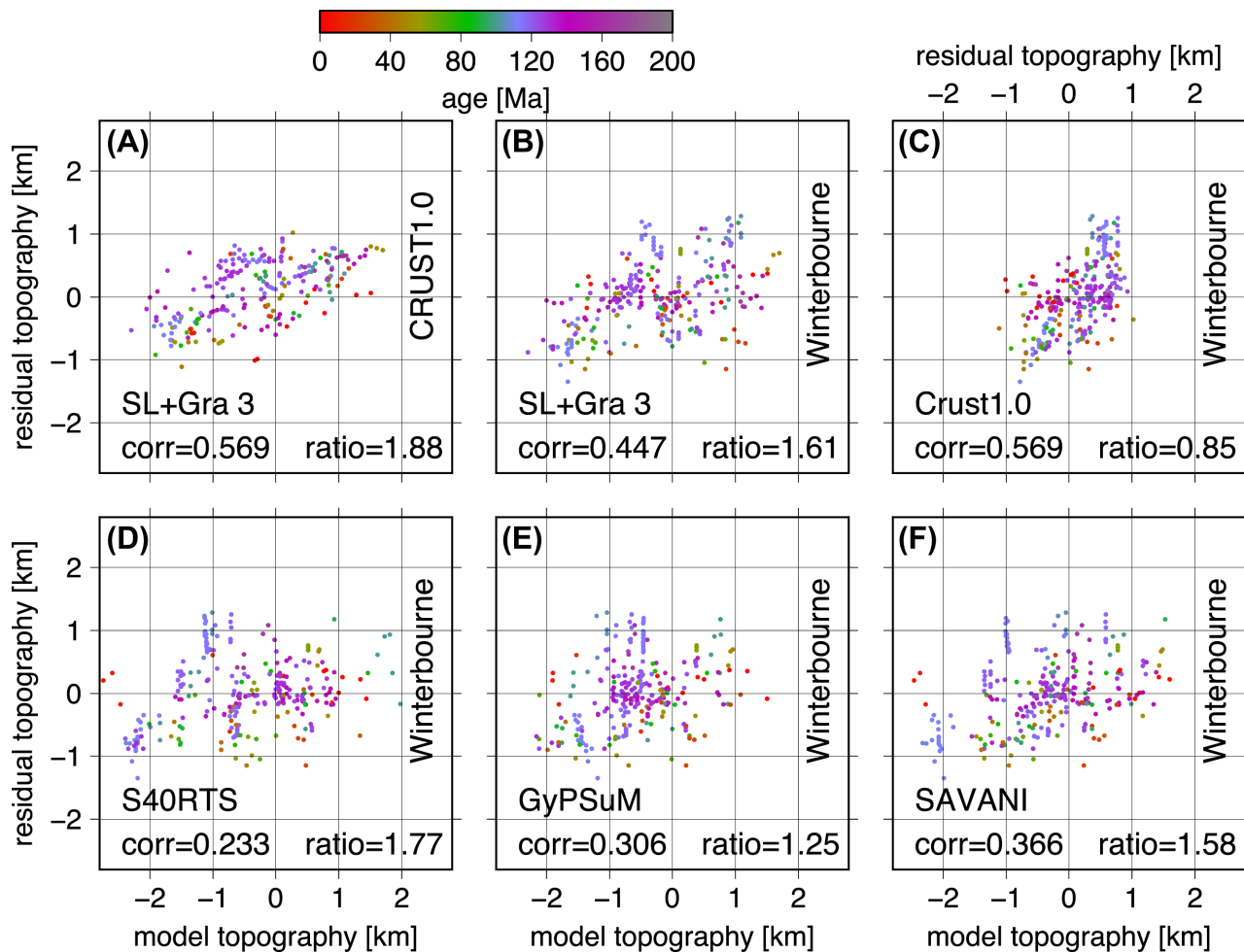
Figure 12. Comparison of different model topographies and residual topography estimates. For model topography cases see Table 1; SL+Gra 3 is reference case. 'CRUST1.0' is residual topography as in Fig. 6(B), 'Winterbourne' refers to Winterbourne *et al.* (2014), but correcting with the same ocean floor age versus depth relation eq. (1) as elsewhere in the paper, and appropriate depth shift. Colours correspond to seafloor age at the location corresponding to a given data point. ratio=rms amplitude ratio; corr=correlation.

topography for which there is no corresponding feature in residual topography. Conversely, strong negative anomalies may be interpreted to be inside the lithosphere and removed.

Viscosity structures with a minimum of $\approx 10^{20}$ Pa s below the lithosphere, rising to a maximum of $\approx 10^{23}$ Pa s in the lower part of the lower mantle and dropping again to somewhere between 10^{20} and 10^{21} Pa s in the lowermost mantle yield the best fit. Including topography as a constraint in the optimization leads to infer somewhat lower viscosities in the asthenosphere than otherwise. Mantle viscosity remains poorly known, although further constraints exist (e.g. Mitrovica & Forte 2004; Paulson *et al.* 2007). Model topography does not place strong constraints on mantle viscosity, as it does not strongly dependent on viscosity structure, and because residual topography is poorly known, although it regionally reached much higher quality recently (e.g. Winterbourne *et al.* 2014).

Although LVV have less effect on present-day model topography than uncertainties in density models (Moucha *et al.* 2007), a laterally variable viscoelastoplastic lithosphere (Popov & Sobolev 2008) should be included in future studies, in particular when considering the very large differences between strong plate interiors and weak plate boundaries. It is expected that results are affected by the decoupling of lithospheric plates (Kaban *et al.* 2014a) and, more generally, lithosphere rheology (Burov *et al.* 2014).

Computations of how topography changes with time have been performed (e.g. Gurnis *et al.* 2000; Conrad & Gurnis 2003; Liu *et al.* 2008; Moucha & Forte 2011; Spasojević & Gurnis 2012; Steinberger *et al.* 2015), and observations exist to infer palaeoelevations (e.g. Garzione *et al.* 2008) and uplift rates (e.g. Roberts & White 2010; Japsen *et al.* 2014). We expect that the improved present-day topography models presented here can help to further improve models of past topography. Also, in regions where tomography models with higher resolution are available, such as the western U.S., a more detailed comparison down to smaller scales is possible (Boschi *et al.* 2010; Becker *et al.* 2014).

5 CONCLUSIONS

A topography model can be inferred from mantle density anomalies based on seismic tomography. But in order to obtain correlations ≈ 0.6 with residual topography and rms amplitudes that are only moderately ($\lesssim 30$ per cent) larger, it is necessary to assume that continental lithosphere contains compositional density variations. This is in accord with previous findings that the continental lithosphere has a different composition from the rest of the mantle, rather than merely being colder. Therefore,

for the purpose of deriving a topography model, also a model of lithospheric thickness has been derived. The best agreement between model and observation is obtained with a viscosity profile that rises from about 10^{20} Pa s beneath the lithosphere to $\approx 10^{23}$ Pa s in the lower part of the lower mantle and drops again in the D'' layer. Compared to viscosity profiles that were derived based on fitting the geoid without consideration of topography, this profile gives a fit to the geoid that is almost as good, while the rather low asthenosphere viscosity yields a model topography amplitude only moderately larger than residual topography. Lateral variations in asthenosphere thickness and viscosity could be substantial, and only lead to rather small changes in model topography amplitude. Nevertheless considering these variations should lead to model improvements, in particular when calculating the evolution of mantle flow and therefore the time changes of topography, because anomalies inside the lithosphere move with the plates, therefore only anomalies beneath the lithosphere cause uplift or subsidence.

ACKNOWLEDGEMENTS

REFERENCES

Artemieva, I.M., 2006. Global $1° \times 1°$ thermal model TC1 for the continental lithosphere: implications for lithosphere secular evolution, *Tectonophysics,* **416,** 245–277.

Auer, L., Boschi, L., Becker, T.W., Nissen-Meyer, T. & Giardini, D., 2014. Savani: a variable resolution whole-mantle model of anisotropic shear velocity variations based on multiple data sets, *J. geophys. Res.,* **119,** 3006–3034.

Austermann, J. & Mitrovica, J.X., 2015. Calculating gravitationally self-consistent sea level changes driven by dynamic topography, *Geophys. J. Int.,* **203,** 1909–1922.

Becker, T.W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.,* **3,** doi:10.1029/2001GC000168.

Becker, T.W., Faccenna, C., Humphreys, E.D., Lowry, A.R. & Miller, M.S., 2014. Static and dynamic support of western United States topography, *Earth planet. Sci. Lett.,* **402,** 234–246.

Bleeker, W., 2003. The late Archean record: a puzzle in ca. 35 pieces, *Lithos,* **71,** 99–134.

Boschi, L., Faccenna, C. & Becker, T.W., 2010. Mantle structure and dynamic topography in the Mediterranean Basin, *Geophys. Res. Lett.,* **37,** L20303, doi:10.1029/2010GL045001.

Burov, E., Gerya, T. & Koptev, A., 2014. Impact of lithosphere rheology on the dynamic topography, *Geophys. Res. Abstr.,* **16,** Abstract EGU2014-9186.

Čadek, O. & Fleitout, L., 2003. Effect of lateral viscosity variations in the top 300 km on the geoid and dynamic topography, *Geophys. J. Int.,* **152,** 566–580.

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

Cammarano, F., Tackley, P. & Boschi, L., 2011. Seismic, petrological and geodynamical constraints on thermal and compositional structure of the upper mantle: global thermo-chemical models, *Geophys. J. Int.,* **187,** 1301–1318.

Coffin, M.F. *et al.*, 2006. Large igneous provinces and scientific ocean drilling: Status quo and a look ahead, *Oceanography,* **19**(4), 150–160.

Connolly, J.A.D., 2005. Computation of phase equilibria by linear programming: a tool for geodynamic modeling and its application to subduction zone decarbonation, *Earth planet. Sci. Lett.,* **236,** 524–541.

Conrad, C. & Gurnis, M., 2003. Seismic tomography, surface uplift, and the breakup of Gondwanaland: integrating mantle convection backwards in time, *Geochem. Geophys. Geosyst.,* **4,** 1031, doi:10.1029/2001GC000299.

Conrad, C.P. & Lithgow-Bertelloni, C., 2004. The temporal evolution of plate driving forces: importance of "slab suction" versus "slab pull" during the Cenozoic, *J. geophys. Res.,* **109,** B10407, doi:10.1029/2004JB002991.

Conrad, C.P. & Lithgow-Bertelloni, C., 2006. Influence of continental roots and asthenosphere on plate-mantle coupling, *Geophys. Res. Lett.,* **33,** L05312, doi:10.1029/2005GL025621.

Crosby, A.G. & McKenzie, D., 2009. An analysis of young ocean depth, gravity and global residual topography, *Geophys. J. Int.,* **178,** 1198–1219.

Crosby, A.G., McKenzie, D. & Sclater, J.G., 2006. The relationship between depth, age and gravity in the oceans, *Geophys. J. Int.,* **166,** 553–573.

Czarnota, K., Hoggard, M.J., White, N. & Winterbourne, J., 2013. Spatial and temporal patterns of Cenozoic dynamic topography around Australia, *Geochem. Geophys. Geosyst.,* **14,** 634–658.

Deschamps, F., Trampert, J. & Snieder, R., 2002. Anomalies of temperature and iron in the uppermost mantle inferred from gravity data and tomographic models, *Phys. Earth planet. Inter.,* **129,** 245–264.

Flament, N., Gurnis, M. & Müller, R.D., 2013. A review of observations and models of dynamic topography, *Lithosphere,* **5,** 189–210.

Forte, A.M., 2015. Constraints on 3-D seismic models from global geodynamic observables: implications for the global mantle convective flow, in *Treatise on Geophysics,* 2nd edn, pp. 805–858, ed. Schubert, G., Elsevier.

Forte, A.M. & Peltier, W.R., 1991. Viscous flow models of global geophysical observables, 1. Forward problems, *J. geophys. Res.,* **96,** 20 131–20 159.

Forte, A.M. & Peltier, W.R., 1994. The kinematics and dynamics of poloidal-toroidal coupling in mantle flow: the importance of surface plates and lateral viscosity variations, *Adv. Geophys.,* **36,** 1–119.

Forte, A.M. & Perry, H.K.C., 2000. Geodynamic evidence for a chemically depleted continental tectosphere, *Science,* **290,** 1940–1944.

Forte, A.M., Peltier, W.R., Dziewonski, A.M. & Woodward, R.L., 1993. Dynamic surface topography: a new interpretation based upon mantle flow models derived from seismic tomography, *Geophys. Res. Lett.,* **20,** 225–228.

French, S., Lekić, V. & Romanowicz, B., 2013. Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere, *Science,* **342,** 227–230.

Garzione, C.N., Hoke, G.D., Libarkin, J.C., Withers, S., MacFadden, B., Eiler, J., Ghosh, P. & Mulch, A., 2008. Rise of the Andes, *Science,* **320,** 1304–1307.

Ghosh, A., Becker, T.W. & Zhong, S.J., 2010. Effects of lateral viscosity variations on the geoid, *Geophys. Res. Lett.,* **37,** L01301, doi:10.1029/2009GL040426.

Ghosh, A., Becker, T.W. & Humphreys, E.D., 2013. Dynamics of the North American continent, *Geophys. J. Int.,* **194,** 651–669.

Grand, S.P., 2002. Mantle shear-wave tomography and the fate of subducted slabs, *Phil. Trans. R. Soc. Lond., A,* **360,** 2475–2492.

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

Gubanov, A.P. & Mooney, W.D., 2009. New global maps of crustal basement age, *EOS, Trans. Am. geophys. Un.,* **90,** Fall Meet. Suppl., Abstract T53B-1583.

Gurnis, M., 1990. Bounds on global dynamic topography from Phanerozoic flooding of continental platforms, *Nature,* **344,** 754–756.

Gurnis, M., 1993. Comment on 'Dynamic Surface Topography' by A.M. Forte, *et al.*, *Geophys. Res. Lett.* **20,** 1663–1664.

Gurnis, M., Mitrovica, J.X., Ritsema, J. & van Heijst, H.-J., 2000. Constraining mantle density structure using geological evidence of surface uplift rates: the case of the African Superplume, *Geochem. Geophys. Geosyst.,* **1,** 1020, doi:10.1029/1999GC000035.

Hager, B.H. & O'Connell, R.J., 1979. Kinematic models of large-scale mantle flow, *J. geophys. Res.,* **84,** 1031–1048.

Hager, B.H. & O'Connell, R.J., 1981. A simple global model of plate dynamics and mantle convection, *J. geophys. Res.,* **86,** 4843–4867.

Hager, B.H. & Richards, M.A., 1989. Long-wavelength variations in Earth's geoid: physical models and dynamical implications, *Phil. Trans. R. Soc. Lond., A,* **328,** 309–327.

Hager, B.H., Clayton, R.W., Richards, M.A., Comer, R.P. & Dziewonski, A.M., 1985. Lower mantle heterogeneity, dynamic topography and the geoid, *Nature,* **313,** 541–545.

Hartley, R.A., Roberts, G.G., White, N. & Richardson, C., 2011. Transient convective uplift of an ancient buried landscape, *Nat. Geosci.,* **4,** 562–565.

Japsen, P., Green, P.F., Bonow, J.M., Nielsen, T.F.D. & Chalmers, J.A., 2014. From volcanic plains to glaciated peaks: burial and exhumation history of southern East Greenland after opening of the NE Atlantic, *Glob. Planet. Change,* **116,** 91–114.

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

Kaban, M.K., Schwintzer, P., Artemieva, I.M. & Mooney, W.D., 2003. Density of the continental roots: compositional and thermal contributions, *Earth planet. Sci. Lett.,* **209,** 53–69.

Kaban, M.K., Petrunin, A.G., Schmeling, H. & Shahraki, M., 2014a. Effect of decoupling of lithospheric plates on the observed geoid, *Surv. Geophys.,* **35,** 1361–1373.

Kaban, M.K., Tesauro, M., Mooney, W.D. & Cloetingh, S.A.P.L., 2014b. Density, temperature, and composition of the North American lithosphere—new insights from a joint analysis of seismic, gravity, and mineral physics data: 1. Density structure of the crust and upper mantle, *Geochem. Geophys. Geosyst.,* **15,** 4781–4807.

Karato, S. & Karki, B.B., 2001. Origin of lateral variation of seismic wave velocities and density in the deep mantle, *J. geophys. Res.,* **106,** 21 771–21 783.

Khan, A., Boschi, L. & Connolly, J.A.D. 2009. On mantle chemical and thermal heterogeneities and anisotropy as mapped by inversion of global surface wave data, *J. geophys. Res.,* **114,** B09305, doi:10.1029/2009JB006399.

Korenaga, J., 2015. *Seafloor Topography and the Thermal Budget of Earth,* Geological Society of America Special Paper SPE514.

Korenaga, T. & Korenaga, J., 2008. Subsidence of normal oceanic lithosphere, apparent thermal expansivity, and seafloor flattening, *Earth planet. Sci. Lett.,* **268,** 41–51.

Laske, G., Masters, G., Ma, Z. & Pasyanos, M., 2013. Update on CRUST1.0—a 1-degree global model of Earth's crust, *Geophys. Res. Abstr.,* **15,** Abstract EGU2013-2658.

Le Stunff, Y. & Ricard, Y., 1995. Topography and geoid due to lithospheric mass anomalies, *Geophys. J. Int.,* **122,** 982–990.

Li, X., Yuan, X. & Kind, R., 2007. The lithosphere-asthenosphere boundary beneath the western United States, *Geophys. J. Int.,* **170,** 700–710.

Lithgow-Bertelloni, C. & Silver, P.G., 1998. Dynamic topography, plate driving forces and the African superswell, *Nature,* **395,** 269–272.

Liu, L., Spasojević, S. & Gurnis, M., 2008. Reconstructing Farallon plate subduction beneath North America back to the Late Cretaceous, *Science,* **322,** 934–938.

Masters, G., Laske, G., Bolton, H. & Dziewonski, A., 2000. The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure, *AGU Geophys. Monogr. Ser.,* **117,** 63–87.

Mitrovica, J.X., 1996. Haskell [1935] revisited, *J. geophys. Res.,* **101,** 555–569.

Mitrovica, J.X. & Forte, A.M., 2004. A new inference of mantle viscosity based upon a joint inversion of convection and glacial isostatic adjustment data, *Earth planet. Sci. Lett.,* **225,** 177–189.

Molnar, P., England, P.C. & Jones, C.H., 2015. Mantle dynamics, isostasy, and the support of high terrain, *J. geophys. Res.,* **120,** 1932–1957.

Moucha, R. & Forte, A.M., 2011. Changes in African topography driven by mantle convection, *Nat. Geosci.,* **4,** 707–712.

Moucha, R., Forte, A.M., Mitrovica, J.X. & Daradich, A., 2007. Lateral variations in mantle rheology: implications for convection related surface observables and inferred viscosity models, *Geophys. J. Int.,* **169,** 113–135.

Müller, R.D., Sdrolias, M., Gaina, C. & Roest, W.R., 2008. Age, spreading rates, and spreading asymmetry of the world's ocean crust, *Geochem. Geophys. Geosyst.,* **9,** Q04006, doi:10.1029/2007GC001743.

Nakagawa, T., Tackley, P.J., Deschamps, F. & Connolly, J.A.D., 2009. Incorporating self-consistently calculated mineral physics into thermochemical mantle convection simulations in a 3-D spherical shell and its influence on seismic anomalies in Earth's mantle, *Geochem. Geophys. Geosyst.,* **10,** Q03004, doi:10.1029/2008GC002280.

Nakagawa, T., Tackley, P.J., Deschamps, F. & Connolly, J.A.D., 2010. The influence of MORB and harzburgite composition on thermo-chemical mantle convection in a 3-D spherical shell with self-consistently calculated mineral physics, *Earth planet. Sci. Lett.,* **296,** 403–412.

Nakiboglu, S.M., 1982. Hydrostatic theory of the Earth and its mechanical implications, *Phys. Earth planet. Inter.,* **28,** 302–311.

Panasyuk, S.V. & Hager, B.H., 2000. Models of isostatic and dynamic topography, geoid anomalies, and their uncertainties, *J. geophys. Res.,* **105**(B12), 28 199–28 209.

Parsons, B. & Sclater, J.G., 1977. An analysis of the variation of ocean floor bathymetry and heat flow with age, *J. geophys. Res.,* **82,** 803–827.

Pasyanos, M.E., Masters, T.G., Laske, G. & Ma, Z., 2014. LITHO1.0: an updated crust and lithospheric model of the Earth, *J. geophys. Res.,* **119,** 2153–2173.

Paulson, A., Zhong, S. & Wahr, J., 2007. Inference of mantle viscosity from GRACE and relative sea level data, *Geophys. J. Int.,* **171,** 497–508.

Pavlis, N.K., Holmes, S.A., Kenyon, S.C. & Factor, J.K., 2012. The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), *J. geophys. Res.,* **117,** B04406, doi:10.1029/2011JB008916.

Petrunin, A., Kaban, M., Rogozhina, I. & Trubitsyn, V., 2013. Revising the spectral method as applied to modeling mantle dynamics, *Geochem. Geophys. Geosyst.,* **14,** 3691–3702.

Popov, A. & Sobolev, S.V., 2008. SLIM3D: a tool for the three-dimensional thermomechanical modeling of the lithospheric deformation with elasto-visco-plastic rheology, *Phys. Earth planet. Inter.,* **171,** 55–75.

Priestley, K. & McKenzie, D., 2013. The relationship between shear wave velocity, temperature, attenuation and viscosity in the shallow part of the mantle, *Earth planet. Sci. Lett.,* **381,** 78–91.

Rader, E., Emry, E., Schmerr, N., Frost, D., Cheng, C., Menard, J., Yu, C.-Q. & Geist, D., 2015. Characterization and petrological constraints of the midlithospheric discontinuity, *Geochem. Geophys. Geosyst.,* **16,** 3484–3504.

Richards, M.A. & Hager, B.H., 1984. Geoid anomalies in a dynamic Earth, *J. geophys. Res.,* **89,** 5987–6002.

Ricard, Y. & Vigny, C., 1989. Mantle dynamics with induced plate tectonics, *J. geophys. Res.,* **94,** 17 543–17 560.

Ricard, Y., Fleitout, L. & Froidevaux, C., 1984. Geoid heights and lithospheric stresses for a dynamic Earth, *Ann. Geophys.,* **2,** 267–286.

Ricard, Y., Richards, M., Lithgow-Bertelloni, C. & Le Stunff, Y., 1993. A geodynamic model of mantle density heterogeneity, *J. geophys. Res.,* **98,** 21 895–21 909.

Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 2004. Global transition zone tomography, *J. geophys. Res.,* **109,** B02302, doi:10.1029/2003JB002610.

Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.,* **184,** 1223–1236.

Roberts, G.G. & White, N., 2010. Estimating uplift rate histories from river profiles using African examples, *J. geophys. Res.,* **115,** B02406, doi:10.1029/2009JB006692.

Rychert, C.A., Shearer, P.M. & Fischer, K.M., 2010. Scattered wave imaging of the lithosphere-asthenosphere boundary, *Lithos,* **120,** 173–185.

Sandwell, D., 2001. Cooling of oceanic lithosphere and ocean floor topography. Available at: http://topex.ucsd.edu/geodynamics/07cooling.pdf, last accessed 8 February 2016.

Schaeffer, A.J. & Lebedev, S., 2013. Global shear speed structure of the upper mantle and transition zone, *Geophys. J. Int.,* **194,** 417–449.

Seton, M. *et al.*, 2012. Global continental and ocean basin reconstructions since 200 Ma, *Earth-Sci. Rev.,* **113,** 212–270.

Shapiro, N.M., Ritzwoller, M. H., Mareschal, J.C. & Jaupart, C., 2004. Lithospheric structure of the Canadian Shield inferred from inversion of surface-wave dispersion with thermodynamic a priori constraints, *Geol. Soc. Lond. Spec. Publ.,* **239,** 175–194.

Simmons, N.A., Forte, A.M. & Grand, S.P., 2009. Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity: implications for the relative importance of thermal versus compositional heterogeneity, *Geophys. J. Int.,* **177,** 1284–1304.

Simmons, N.A., Forte, A.M., Boschi, L. & Grand, S.P., 2010. GyPSuM: A joint tomographic model of mantle density and seismic wave speeds, *J. geophys. Res.,* **115,** B12310, doi:10.1029/2010JB007631.

Spasojević, S. & Gurnis, M., 2012. Sea level and vertical motion of continents from dynamic Earth models since the Late Cretaceous, *AAPG Bull.,* **96,** 2037–2064.

Steinberger, B., 2007. Effect 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, *Phys. Earth planet. Inter.,* **164,** 2–20.

Steinberger, B. & Calderwood, A., 2006. Models of large-scale viscous flow in the Earth's mantle with constraints from mineral physics and surface observations, *Geophys. J. Int.,* **167,** 1461–1481.

Steinberger, B. & Holme, R., 2008. Mantle flow models with core-mantle boundary constraints and chemical heterogeneities in the lowermost mantle, *J. geophys. Res.,* **113,** B05403, doi:10.1029/2007JB005080.

Steinberger, B., Schmeling, H. & Marquart, G., 2001. Large-scale lithospheric stress field and topography induced by global mantle circulation, *Earth planet. Sci. Lett.,* **186,** 75–91.

Steinberger, B., Buiter, S., Medvedev, S. & Tetreault, J., 2010. The mantle under Africa: an overview of global and regional tomography models, lithosphere thickness models and their effect on present-day and past dynamic topography. NGU report no. 2010.058, http://www.ngu.no/upload/Publikasjoner/Rapporter/2010/2010_058.pdf, last accessed 8 February 2016.

Steinberger, B., Spakman, W., Japsen, P. & Torsvik, T.H., 2015. The key role of global solid Earth processes in the late Cenozoic intensification of Greenland glaciation, *Terra Nova,* **27,** 1–8.

Stixrude, L. & Lithgow-Bertelloni, C., 2005a. Mineralogy and elasticity of the oceanic upper mantle: origin of the low-velocity zone, *J. geophys. Res.,* **110,** B03204, doi:10.1029/2004JB002965.

Stixrude, L. & Lithgow-Bertelloni, C., 2005b. Thermodynamics of mantle minerals—I. Physical properties, *Geophys. J. Int.,* **162,** 610–632.

Stixrude, L. & Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals—II. Phase equilibria, *Geophys. J. Int.,* **184,** 1180–1213.

Tackley, P.J., 2008. Modelling compressible mantle convection with large viscosity contrasts in a three-dimensional spherical shell using the yin-yang grid, *Phys. Earth planet. Inter.,* **171** (1–4), 7–18.

Tan, E., Choi, E., Thoutireddy, P., Gurnis, M. & Aivazis, M., 2006. GeoFramework: coupling multiple models of mantle convection within a computational framework, *Geochem. Geophys. Geosyst.,* **7,** Q06001, doi:10.1029/2005GC001155.

Tesauro, M., Kaban, M.K. & Cloetingh, S.A.P.L., 2008. EuCRUST-07: a new reference model for the European crust, *Geophys. Res. Lett.,* **35,** L05313, doi:10.1029/2007GL032244.

Tesauro, M., Kaban, M.K., Mooney, W.D. & Cloetingh, S.A.P.L., 2014. Density, temperature, and composition of the North American lithosphere—new insights from a joint analysis of seismic, gravity, and mineral physics data: 2. Thermal and compositional model of the upper mantle, *Geochem. Geophys. Geosyst.,* **15,** 4808–4830.

Thoraval, C. & Richards, M.A., 1997. The geoid constraint in global geodynamics: viscosity structure, mantle heterogeneity models and boundary conditions, *Geophys. J. Int.,* **131,** 1–8.

Trampert, J., Vacher, P. & Vlaar, N., 2001. Sensitivities of seismic velocities to temperature, pressure and composition in the lower mantle, *Phys. Earth planet. Int.,* **124,** 255–267.

Wessel, P. & Smith, W.H.F., 1998. New, improved version of generic mapping tools released, *EOS, Trans. Am. geophys. Un.,* **79,** 579, doi:10.1029/98EO00426.

Wheeler, P. & White, N., 2000. Quest for dynamic topography: observations from Southeast Asia, *Geology,* **28,** 963–966.

Winterbourne, J., White, N.J. & Crosby, A.G., 2014. Accurate measurements of residual topography from the oceanic realm, *Tectonics,* **33,** 982–1015.

Yuan, H. & Romanowicz, B., 2010. Lithospheric layering in the North American craton, *Nature,* **466,** 1063–1068.

Zhong, S. J., Zuber, M.T., Moresi, L. & Gurnis, M., 2000. Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection, *J. geophys. Res.,* **105,** 11 063–11 082.