Bayesian inference of mantle viscosity from whole-mantle density models

Maxwell Rudolph[1], Pritwiraj Moulik[2], and Ved Lekic[2]

[1]University of California, Davis
[2]University of Maryland

November 26, 2022

Abstract

The long-wavelength geoid is sensitive to Earth's mantle density structure as well as radial variations in mantle viscosity. We present a suite of inversions for the radial viscosity profile using whole-mantle models that jointly constrain the variations in density, shear- and compressional-wavespeeds using full-spectrum tomography. We use a Bayesian approach to identify a collection of viscosity profiles compatible with the geoid, while enabling uncertainties to be quantified. Depending on tomographic model parameterization and data weighting, it is possible to obtain models with either positive- or negative-buoyancy in the large low shear velocity provinces (LLSVPs). We demonstrate that whole-mantle density models in which density and V_S variations are correlated imply an increase in viscosity below the transition zone, often near ~1000~km. Many solutions also contain a low-viscosity channel below 650~km. Alternatively, models in which density is less-correlated with V_S – which better fit normal mode data – require a reduced viscosity region in the lower mantle. This feature appears in solutions because it reduces the sensitivity of the geoid to buoyancy variations in the lowermost mantle. The variability among the viscosity profiles obtained using different density models is indicative of the strong non-linearities in modeling the geoid and the limited resolving power of the geoid kernels. We demonstrate that linearized analyses of model resolution do not adequately capture the posterior uncertainty on viscosity. Joint and iterative inversions of viscosity, wavespeeds, and density using seismic and geodynamic observations are required to reduce bias from prior assumptions on viscosity variation and scalings between material properties.

Bayesian inference of mantle viscosity from whole-mantle density models

M.L. Rudolph[1], P. Moulik[2], and V. Lekić[2]

[1]Department of Earth and Planetary Sciences, University of California, Davis CA

[2]Department of Geology, University of Maryland, College Park, MD

Key Points:

- We model mantle flow with density variations from full-spectrum tomography.
- We generate Bayesian estimates of the radial viscosity profile with uncertainties.
- There is evidence for a low-viscosity channel below the transition zone and a viscosity increase in the mid mantle.

Corresponding author: Max Rudolph, `maxrudolph@ucdavis.edu`

Abstract

The long-wavelength geoid is sensitive to Earth's mantle density structure as well as radial variations in mantle viscosity. We present a suite of inversions for the radial viscosity profile using whole-mantle models that jointly constrain the variations in density, shear- and compressional-wavespeeds using full-spectrum tomography. We use a Bayesian approach to identify a collection of viscosity profiles compatible with the geoid, while enabling uncertainties to be quantified. Depending on tomographic model parameterization and data weighting, it is possible to obtain models with either positive- or negative-buoyancy in the large low shear velocity provinces (LLSVPs). We demonstrate that whole-mantle density models in which density and V_S variations are correlated imply an increase in viscosity below the transition zone, often near 1000 km. Many solutions also contain a low-viscosity channel below 650 km. Alternatively, models in which density is less-correlated with V_S – which better fit normal mode data – require a reduced viscosity region in the lower mantle. This feature appears in solutions because it reduces the sensitivity of the geoid to buoyancy variations in the lowermost mantle. The variability among the viscosity profiles obtained using different density models is indicative of the strong non-linearities in modeling the geoid and the limited resolving power of the geoid kernels. We demonstrate that linearized analyses of model resolution do not adequately capture the posterior uncertainty on viscosity. Joint and iterative inversions of viscosity, wavespeeds, and density using seismic and geodynamic observations are required to reduce bias from prior assumptions on viscosity variation and scalings between material properties.

Plain Language Summary

The viscosity of Earth's mantle affects nearly every aspect of Earth's evolution, including its convective vigor and rate of cooling, the motion of mantle plumes and subducted oceanic lithosphere through the mantle, and the deep volatile cycle. We use the long-wavelength geoid to constrain the variation of mantle viscosity with depth. In so doing, we use newly available whole-mantle models of seismic wavespeeds and density that incorporate constraints on lowermost mantle density from the free oscillations of the Earth. We find evidence for an increase in viscosity within the mid mantle and for a low-viscosity region below the mantle transition zone. We demonstrate that depending on choices made in the data weighting and regularization of tomographic models, it is possible to obtain solutions that include a reduced viscosity region in the lower mantle as well. We argue that in joint inversions of seismic and geodynamic observations, viscosity variations must be solved for together with wavespeed and density variations, and should not be assumed *a priori*.

1 Introduction

Long-wavelength components of the geoid are precise geodetic observations that are sensitive to density and viscosity variations in the Earth's mantle, providing an observational constraint on these two dynamically significant parameters. Density variations drive convection currents that deflect the topography of the Earth's surface and internal boundaries, determining the observed geoid. Patterns and amplitudes of the geoid variations due to these flows depend on the viscosity structure of the mantle. At the longest spatial scales – those represented by the lowest degrees in spherical harmonic representations – radial variation of viscosity becomes much more important than lateral variations (e.g. Richards & Hager, 1984; Ghosh et al., 2010). If a radial viscosity profile is assumed *a priori*, geoid observations may be used to constrain density anomalies within the mantle in conjunction with seismological and other geophysical observations (e.g. Simmons et al., 2010). Alternatively, geoid observations can constrain the depth-dependence of mantle viscosity when estimates of density anomalies in the mantle are available (e.g. Hager et al., 1985). We adopt and extend the latter approach to provide new constraints on radial viscosity structure while exploring the recent seismological constraints on density variations in the mantle (cf. Moulik & Ekström, 2016).

Since the 1980s, substantial progress has been made in mapping the viscosity of the mantle by approximating density (ρ) anomalies by scaling them from shear (V_S) or compressional (V_P) velocity anomalies mapped by global seismic tomography. Due to the more uniform data coverage afforded by shear waves, global V_S tomographic models are more consistent among groups and better constrain the longest-wavelength structure (e.g. Becker & Boschi, 2002; Cottaar & Lekic, 2016; Ritsema & Lekic, 2020). The scaling factors typically used to relate wavespeed to density variations ($R_{\rho,S} = \frac{d\ln\rho}{d\ln V_s}$ and $R_{\rho,P}$) are informed by laboratory studies of mineral properties, and usually assume that density variations arise due to temperature, neglecting potential contributions from compositional variations (e.g. Karato, 1993). This assumption introduces uncertainty in inferences of mantle viscosity, compounding the limitations arising from uncertainties in the V_S tomographic models themselves.

The assumption of uniform composition may be particularly problematic in the Earth's lowermost mantle, where structure is dominated by the two large low shear-velocity provinces (LLSVPs). Multiple lines of evidence suggest that LLSVPs are compositionally distinct from the ambient lower mantle. First, LLSVPs have slower shear velocities and larger velocity gradients across their margins than is expected for uniform composition (e.g. Ni et al., 2002; Wang & Wen, 2007). Second, variations in V_S and bulk sound speed are anti-correlated in the lower mantle (e.g. Su & Dziewonski, 1997; Masters et al., 2000). Third, the ratio of $d \ln V_S / d \ln V_P$ is higher and the distribution of velocities broader in the lower mantle than is expected for purely thermal heterogeneity (e.g. Robertson & Woodhouse, 1996; Masters et al., 2000; Brodholt et al., 2007; Deschamps & Trampert, 2003). Finally, constraints from normal modes favor anti-correlation between density and velocity variations in the lowermost mantle (e.g. Ishii & Tromp, 2001; Moulik & Ekström, 2016). Depending on the relative importance of thermal and compositional contributions to density, the LLSVPs could either be positively or negatively buoyant structures. If positively buoyant, the LLSVPs would be expected to rise, precluding their stability over geologic time, which has been proposed based on multiple lines of evidence (e.g. Burke & Torsvik, 2004; Burke et al., 2008; Dziewonski et al., 2010; Torsvik et al., 2014). Further complicating matters, the LLSVPs may not be compositionally uniform and could consist of a denser sub-region surrounded by a thermal and/or compositional halo (e.g. Simmons et al., 2010; Moulik & Ekström, 2016; Ballmer et al., 2016; Lau et al., 2017). These distinct scenarios differ in their implications for the nature of LLSVPs, their relationship to geochemical signatures of primordial material, and to large-scale mantle flow.

Seismological constraints, derived primarily from normal mode splitting functions, can be used to directly map large-scale density variations in the lowermost mantle (e.g. Ishii & Tromp, 2001; Moulik & Ekström, 2016), thereby reducing the need for *a priori* $R_{\rho,S}$ and $R_{\rho,P}$ scaling factors. However, density tomographic models have not been employed yet to constrain mantle viscosity due to a lack of consensus among groups, potential dependence of the tomographic models on the starting structure (Kuo & Romanowicz, 2002), regularization (Resovsky & Ritzwoller, 1999), and lack of sensitivity to odd-degree structure (Resovsky & Ritzwoller, 1995). Improved normal mode splitting measurements based on recent large earthquakes (e.g. Deuss et al., 2011, 2013) have improved our ability to constrain lower mantle density variations. A recent joint inversion of normal mode splitting, waveforms, and travel-times reported a statistically significant argument for negative buoyancy in the bottom ∼500 km of LLSVPs (Moulik & Ekström, 2016). Additionally, new tidal constraints on lowermost mantle density structure suggest that the LLSVPs are about 0.5% denser than average (Lau et al., 2017). Further refinements on density structure from new geophysical constraints could reveal the depth extent of regions with denser than ambient mantle within LLSVPs.

In order to better understand the relationship between mantle density structure and the geoid based on currently available data, we perform Bayesian inversions for mantle viscosity structure using a similar framework to Rudolph et al. (2015). The inversions employ a suite of 18 whole-mantle density models that are created based on the approach of Moulik and Ekström (2016) while varying data weights on the seismic observations and the model regularization factor that modulates correlation between density and V_S structure. We analyze the resulting suite of viscosity structures in terms of fit to the geoid and identify persistent features and ones that depend crucially on modeling choices. Finally, we discuss the implications for making robust inferences on density, viscosity and wavespeed based on joint modeling of seismic and geodetic data.

2 Methods

We carry out a suite of inversions for the mantle viscosity profile constrained by the long-wavelength non-hydrostatic geoid. We use geoid spherical harmonic coefficients from the GRACE geoid model GGM05 (Ries et al., 2016) and the hydrostatic correction from Chambat et al. (2010). There is a rich history of inversions for the mantle viscosity structure using the long-wavelength geoid and observations related to glacial isostatic adjustment using a variety of inverse methods. In choosing an inversion methodology, it is important to consider at the outset the nature of the inverse problem at hand, and in particular its degree of non-linearity. In Figure 1 we show the variation of the misfit to the geoid for spherical harmonic degrees 2-7 for a piecewise-linear viscosity structure constrained by four control points (describing viscosity and depth). Two of the control points are fixed at the surface and the core-mantle boundary while the other two are allowed to vary in depth within the mantle. Even for this coarse parameterization of the viscosity structure using only 6 free parameters, it is evident that there are strong tradeoffs between parameters, multiple local minima or "wells" in the misfit surface (e.g. Figure 1O,Q). These basic observations suggest that the variation in the misfit may not be adequately described by linear estimates. We note that the brute-force search illustrated in Figure 1 is not a viable approach to the viscosity inversion problem because the number of forward model evaluations becomes prohibitive very quickly as the number of free parameters (control points) increases. Because of the degree of nonlinearity in the problem, we choose to use a model-space search approach. Various model space search approaches including genetic algorithms (King, 1995; Kido et al., 1998, e.g.) and Monte-Carlo approaches (e.g. Mitrovica & Peltier, 1993; Mitrovica & Forte, 1997; Paulson et al., 2007a, 2007b) have been previously applied to inversions for mantle viscosity, highlighting the viability of this class of inverse methods.

We use a transdimensional, hierarchical, Bayesian (THB) inversion method (Bodin et al., 2012; Sambridge et al., 2013) to estimate the mantle viscosity structure. The transdimensional approach does not specify *a priori* the number of free parameters such as viscosity values and interface depths. Rather, the number of free parameters is itself a parameter in the inversion procedure, and we explore models with varying levels of complexity, with an Ockham factor penalizing model complexity. We employ a reversible-jump Markov-Chain Monte Carlo (rjMCMC) approach to explore the model space; "reversible jump" refers to the ability of the Markov chain to jump between model spaces with different numbers of parameters. We use a parallel tempering algorithm (Sambridge, 2014), in which multiple Markov chains simultaneously explore model spaces with different "temperatures" that effectively smooth the misfit surface. Parallel tempering accelerates convergence and reduces the tendency for the Markov Chains to get stuck in local (rather than global) minima of the objective function. In contrast to our previous inferences of viscosity (Rudolph et al., 2015), we (1) use a different and more varied set of buoyancy structures from seismic tomography, (2) parameterize log-viscosity using piecewise linear functions rather than layers, (3) estimate tomographic uncertainty directly from model covariance matrices, and (4) employ parallel tempering. We describe each aspect of the modeling, starting with the generation of density models used to constrain the inversions for viscosity structure. We then describe the process used to generate the covariance matrices representing data and forward modeling uncertainty. Finally, we describe the inversion procedure used to estimate the mantle viscosity profile.

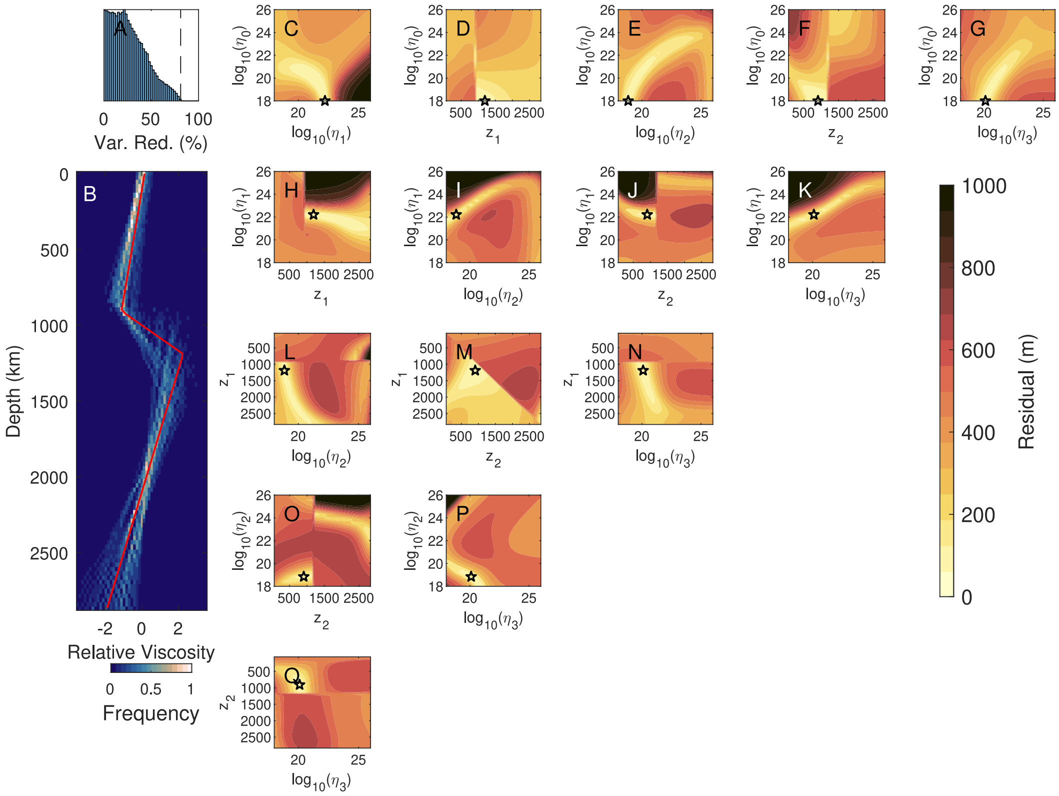

Figure 1. Results of a grid search of piecewise-linear viscosity structures with four control points. Density anomalies were derived from SEMUCB-WM1 using a uniform scaling factor $d \ln V_s / d \ln \rho \; = \; 0.2$, and the residual is computed for spherical harmonic degrees 2-7. (A) Distribution of residuals for all of the models. (B) Bivariate histogram of relative viscosity for solutions representing the 1000 smallest misfits. The solution with the smallest misfit is shown as a red curve. The dashed line in panel (A) indicates the largest residual among the solutions shown in (B). Panels (C)-(Q) show two dimensional slices through the misfit function passing through the minimum misfit value (star). The viscosity values η_0 and η_3 correspond to the core-mantle boundary and the surface. The additional control points within the mantle correspond to depth and viscosity (z_1,η_1) and (z_2,η_2).

2.1 Forward Model

We model global mantle flow and calculate the geoid anomalies using a propagator matrix technique. The forward mantle flow models are performed with HC (Becker et al., 2014), using free-slip boundary conditions at the surface and the core-mantle boundary. In the present work, we consider only depth-variation of viscosity and neglect lateral viscosity variations (LVVs). The degree to which LVVs influence the long wavelength geoid is difficult to quantify because it depends on the amplitude and wavelength of the viscosity variations (Ghosh et al., 2010). Zhong and Davies (1999) demonstrated that high-viscosity slabs in the lower mantle could have a strong effect on the geoid, even at spherical harmonic degree 2. On the other hand, lateral viscosity variations based on V_S heterogeneity from tomographic models have a less significant effect on the long wavelength geoid. Moucha et al. (2007) showed that the inclusion/omission of LVVs had a less significant impact on the $l \leq 20$ geoid than the variability among geoid predictions using density models derived from the tomographic models S20RTS (Ritsema & Van Heijst, 1999) and TX2002 (Grand, 2002). Yang and Gurnis (2016) found similar radial viscosity variations in inversions constrained by long-wavelength gravity and topography with and without LVVs, suggesting a limited sensitivity of the long-wavelength geoid to LVVs.

2.2 Joint models from full-spectrum tomography

Full-spectrum tomography (Moulik & Ekström, 2014, 2016) employs seismic waveforms and derived measurements of body waves (∼1–20s), surface waves (∼20–300s) and normal modes (∼250–3300s) to constrain physical properties – seismic velocity, anisotropy, density and the topography of discontinuities – at variable spatial resolution. We follow closely the methodology of Moulik and Ekström (2016) to generate a suite of tomographic models with different levels of structural complexity and fit to seismic data. The starting model is an anisotropic shear velocity model with topographies of transition-zone discontinuities and three (R_P^{UM}, R_P^{LM}, R_ρ^{UM}=R_ρ^{LM}) optimal scaling factors that describe the relative variations of compressional velocity (V_P) and density (ρ) in the upper (UM) and lower mantle (LM). The preferred estimate of upper (R_P^{UM}=0.7) and lower mantle velocity scaling ratio (R_P^{LM}=0.4) and a whole-mantle density scaling ratio (R_ρ=0.3) gives substantial improvements in data fits and is consistent with petrological constraints (e.g. Montagner & Anderson, 1989; Karato, 1993) and earlier tomographic studies (e.g. Robertson & Woodhouse, 1996).

The splitting of Earth's lowest-frequency observed normal mode (${}_0S_2$) is strongly sensitive to long-wavelength density structure and prefers models with negative R_ρ and anti-correlated $V_S-\rho$ structure in the lowermost mantle i.e. scenarios where the base of LLSVPs are denser and more negatively buoyant than the ambient mantle (Fig.12 in Moulik & Ekström, 2016). Motivated by these observations, we explore a range of density models by varying the weight ($w_{{}_0S_2}$) assigned to the splitting of ${}_0S_2$ and the weight assigned to the prior on $V_S-\rho$ correlation in the lowermost mantle ($\gamma_\rho^{D''}$), while keeping identical values in the other regions of the mantle. Varying $w_{{}_0S_2}$ and $\gamma_\rho^{D''}$ affects the model density structure and correlation between density and V_S below ~ 2000 km depth with the most significant changes occurring below 2500 km depth (Figure 2 and S1). The magnitude of $\gamma_\rho^{D''}$ does not have physical dimensions or significance - it is a data weighting factor in the inversion and the values reported here are consistent with Moulik and Ekström (2016). The scaling complexity is varied by modulating a modified objective function of the joint inverse problem

$$\begin{aligned}\widetilde{\chi}^2 = w_{{}_0S_2}\chi^2_{{}_0S_2} + \sum_{i=2}^{N} w_i\chi_i^2 + \gamma_\rho^{D''}\mathcal{R}^2_{R_\rho,D''} + \gamma_\rho^{\text{other}}\mathcal{R}^2_{R_\rho,\text{other}} + \\ + \sum_{k=1}^{M}\left[\gamma_{h,k}\mathcal{R}^2_{h,k} + \gamma_{v,k}\mathcal{R}^2_{v,k}\right] + \gamma_P\mathcal{R}^2_{R_P} + \gamma_a\mathcal{R}^2_{R_a},\end{aligned} \tag{1}$$

where N is the number of data in the inversion, w_i are the weights assigned to individual data. We minimize vertical ($\mathcal{R}^2_{v,k}$) and horizontal gradients ($\mathcal{R}^2_{h,k}$) for M parameters and impose scaling relationships between various pairs of parameters (v_S-v_P, $v_S-\rho$ and anisotropic a_S-a_P) with weights γ_P, γ_ρ and γ_a, respectively (eq. 8–10 of Moulik & Ekström, 2016). We construct 18 models which span the range of scenarios of ρ variations given currently-available seismic data across six values of $\gamma_\rho^{D''}$ and three of $w_{{}_0S_2}$ (Table 1). Figure 2 shows the lowermost mantle density structure for a subset of 6 of these models. Weights to the remaining (N-1) data, $v_S-\rho$ scaling in other regions ($\gamma_\rho^{\text{other}}$) and rest of the regularization parameters are kept similar to those employed in the preferred model from Moulik and Ekström (2016), which corresponds roughly to the Model 161 ($w_{{}_0S_2} = 1$X, $\gamma_\rho^{D''} = 0$) in our suite. The posterior model covariance matrix, $\widetilde{C_M}$, corresponding to each of the 18 tomographic models is also computed; our procedure for

		Prior $v_S - \rho$ covariance in the lowermost mantle ($\gamma_\rho^{D''}$)					
		0	$\mathbf{10^4}$	$\mathbf{10^6}$	$\mathbf{10^7}$	$\mathbf{10^8}$	$\mathbf{10^{11}}$
Data weight ($w_{{}_0S_2}$)	**0**	Model 155	Model 156	Model 157	Model 158	Model 159	Model 160
	1X	**Model 161**	Model 162	Model 163	Model 164	Model 165	Model 166
	10X	Model 167	Model 168	Model 169	Model 170	Model 171	Model 172

Table 1. Summary of scaling assumptions and data weights employed in the 18 whole-mantle density models. The data weight 1X represents the preferred weight for ${}_0S_2$ from Moulik and Ekström (2016).

propagating uncertainties in tomographic estimations of density variations into the potentially-correlated uncertainties on the individual geoid coefficients is described in Section 2.3.

Depending on the values of $\gamma_\rho^{D''}$ and $w_{{}_0S_2}$, we can generate models that produce either positive or negative correlation between degree-2 density structure and the geoid (Figure 2). For large values of $\gamma_\rho^{D''}$, ρ variations largely track those of V_S. As $\gamma_\rho^{D''}$ decreases, ρ variations reduce in amplitude and become de-correlated from those of V_S. Increasing $w_{{}_0S_2}$ has the effect of amplifying the strength of ρ variations and a pattern that is anti-correlated to V_S variations. It is worth noting that not all 18 scenarios are equally preferred by current seismic data; anti-correlated $V_S{-}\rho$ structure in the lowermost mantle and associated dense base of LLSVPs are preferred by most normal modes with substantial sensitivity to density at these depths. When we allow additional degrees of freedom with such independent ρ variations in the lowermost mantle, variance is reduced and the reduction is systematic, large (often exceeding 50 percent) and statistically significant for various subsets of normal-mode, surface- and body-wave data (Moulik & Ekström, 2016, Section 4.4 and Figure 12).

2.2.1 Model Parameterization

In previous work (Rudolph et al., 2015), we specified the viscosity profile using constant-viscosity layers. In order to permit an efficient representation of viscosity gradients, we

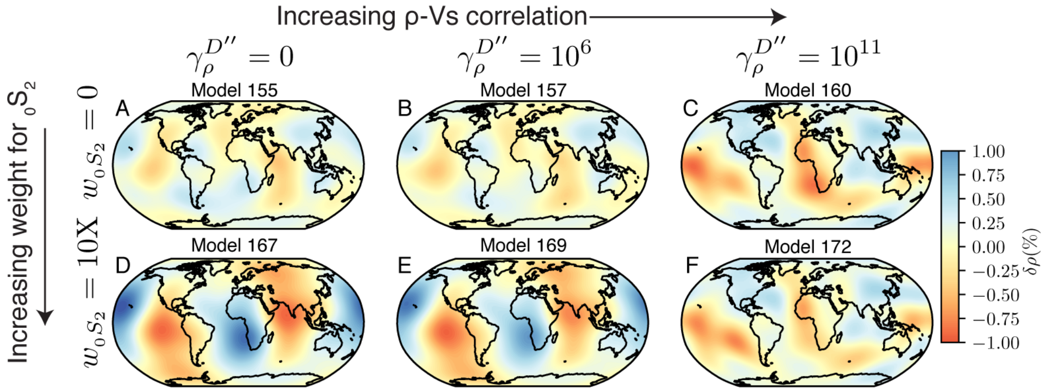

Figure 2. Here we show density variations at 2875 km depth and for spherical harmonic degrees $l = 1 - 7$ for six of the density models. Lateral variations of density ($\delta\rho$) in the lowermost mantle depend on the amount of imposed scaling ($\gamma_\rho^{D''}$) between density and V_S variation. On the other hand, the weight assigned to $_0S_2$ affects the amplitude of $\delta\rho$ variations more than their pattern. Not all scenarios are equally preferred by current seismic data (Section 2.2).

specify the viscosity profile using k control points:

$$\log_{10}(\eta(r)) = \begin{cases} \frac{\eta_2^* - \eta_1^*}{r_2 - r_1}(r - r_1) + \eta_1^*, & r_1 \le r \le r_2 \\ \frac{\eta_{i+1}^* - \eta_i^*}{r_{i+1} - r_i}(r - r_i) + \eta_i^*, & r_i \le r \le r_{i+1} \\ \dots & \\ \frac{\eta_k^* - \eta_{k-1}^*}{r_k - r_{k-1}}(r - r_{k-1}) + \eta_{k-1}^*, & r_{k-1} \le r \le r_k \end{cases} \tag{2}$$

where η_i^* and r_i specify the base-10 logarithm of the viscosity and the radial coordinate of the $i-$th control point.

2.2.2 Inversion method

We use a reversible-jump Markov chain Monte Carlo algorithm (Green, 1995),with parallel tempering (Sambridge, 2014) to find the collection of viscosity profiles compatible with the observed geoid. The parallel tempering algorithm runs several Markov chains simultaneously at different "temperatures". At increasing temperatures, the objective function is effectively smoother, and higher-temperature chains may more easily move between local minima. The chains are advanced in parallel, and after each 100 steps, pairs of chains are selected via a random permutation and given the opportunity to swap solutions at different temperatures. Chains with higher temperature therefore perform a more complete exploration of the model space, while the chains with lower temperature explore the neighborhood of the local or global minimum. Only the solutions with temperature $T = 1$ are sampled in the posterior ensemble. The benefits of parallel tempering are accelerated convergence and a reduced tendency for the chain at $T = 1$ to find a local, rather than global, minimum in the objective function.

At each step within each of the Markov Chains, we choose one of five events with equal probability: 1) Creation of a new control point; 2) Deletion of an existing control point; 3) Change of viscosity associated with a control point; 4) Change of depth of the control point; and, 5) Change in hierarchical uncertainty parameter. Given a vector $\underline{m}$ of model parameters, we calculate a vector containing the model geoid spherical harmonic coefficients $\underline{N}_{\text{model}}(\underline{m})$ for comparison with the observed geoid spherical harmonic coefficients $\underline{N}_{\text{obs}}$. We calculate a residual using the Mahalanobis distance

$$\Phi(\underline{m}) = \left(\underline{N}_{\text{model}}(\underline{m}) - \underline{N}_{\text{obs}}\right)^{\intercal} \underline{\underline{C}}_D^{-1} \left(\underline{N}_{\text{model}}(\underline{m}) - \underline{N}_{\text{obs}}\right). \tag{3}$$

We introduce a covariance matrix $\underline{\underline{C}}_D$ that represents the data and forward modeling uncertainties, discussed later. We include an additional *hierarchical parameter*, σ with which we scale the covariance matrix,

$$\underline{\underline{C}}_D = \sigma^2 \underline{\underline{\widetilde{C}}}_D. \tag{4}$$

In an inversion with a hierarchical parameter, it is assumed that the covariance matrix represents the pattern of covariance but not its amplitude, and σ^2 can be viewed as an uncertainty hyperparameter. We calculate the likelihood

$$P(\underline{N}_{obs}|\underline{m}, T_i) = \frac{1}{\sqrt{(2\pi)^{n_{lm}} (\sigma^2)^{n_{lm}} \det(\underline{\underline{C}}_D)}} \exp\left(-\frac{\Phi(\underline{m})}{2T_i}\right). \tag{5}$$

Here, T_i is the temperature of the i-th chain and n_{lm} is the length of the data vector. Given the likelihood value, we calculate an acceptance probability in log space for the proposed model (denoted with $'$ and having k' control points) given the current accepted solution (unprimed quantities)

$$\min\left(1, \frac{P(N_{\text{obs}}|m')}{P(N_{\text{obs}}|m)} \frac{k}{k'}\right). \tag{6}$$

Parallel tempering is implemented through the addition of an exchange step between chains running at different temperatures, following the approach of Sambridge (2014). At each step of the Markov chain, we propose to swap the currently-accepted solutions for *each* of the N chains having temperatures T_i, $i \in \{1, 2, \ldots, N\}$, with a randomly selected chain having temperature T_j, where $j \neq i$. The proposed swap, implemented by swapping temperatures, is accepted with probability

$$\alpha_{ij} = \min\left(1, \left[\frac{P(\underline{N}_{obs}|m_j)}{P(\underline{N}_{obs}|m_i)}\right]^{1/T_i} \left[\frac{P(\underline{N}_{obs}|m_i)}{P(\underline{N}_{obs}|m_j)}\right]^{1/T_j}\right). \tag{7}$$

We calculate the acceptance probability α_{ij} in log-space, where

$$\log(\alpha_{ij}) = \min\left(\log(1), \left(\frac{1}{T_i} - \frac{1}{T_j}\right)\left[\frac{N}{2}\log\left(\frac{\sigma_i^2}{\sigma_j^2}\right) - \frac{1}{2}\left(\frac{\Phi(m_j)}{\sigma_j^2} - \frac{\Phi(m_i)}{\sigma_i^2}\right) + \log\left(\frac{k_i}{k_j}\right)\right]\right). \tag{8}$$

For each of the inversions, we run 16 Markov chains at temperatures spaced log-uniformly from 1–50. Each chain was run for 10 million steps. During the initial phase of the MCMC procedure, we follow Kolb and Lekic (2014) and limit the dimensionality of the model space by reducing the maximum number of control points that describe the viscosity solution. During this "burn-in" period, we begin by allowing $N_{\max} = 2$ control points and wait $10,000 \times N_{\max}$ steps before increasing $N_{\max}$. The maximum number of control points (25) is accessible after 2.53×10^6 steps. We begin sampling the posterior after 5 million steps and verified that the properties of the ensemble had stabilized before sampling began. We verified that the nature of the solutions did not change as the number of control points was increased further for a test inversion. We assume a uniform prior probability density function (pdf) for $\eta(r)$ that allows for variations over six orders of magnitude. The prior on the number of control points k is proportional to $1/k$. New control points are assigned a log-viscosity sampled randomly from the prior. We propose perturbations to the currently accepted solution using gaussian pdfs with a shape parameter of 0.2 log-units for viscosity and 35 km for control point depths. The proposal distribution for changes to the hierarchical parameter σ is a gaussian pdf with shape parameter 0.05. To maintain numerical stability, we limit the distance between adjacent control points to 45 km. We verified that our MCMC procedure achieves uniform sampling of the model space, reflecting the assumption of a uniform prior on viscosity variations and a prior on number of control points that is proportional to $1/k$.

2.3 Uncertainty quantification

The THB method enables us to estimate the best-fit values of the model parameters and their uncertainty, accounting simultaneously for the effects of data uncertainty (i.e. uncertainty in the density variations from the tomographic model), and in the observational constraints (the geoid spherical harmonic coefficients). The observational uncertainties in the geoid spherical harmonic coefficients for the long wavelengths used to constrain our models are very small, on the order of parts per million to parts per thousand (Ries et al., 2016). Therefore, we ignore them in the inversion, and instead focus on quantifying uncertainties in the density variations.

First, for each of the density models, we sample the multivariate random normal distribution described by the posterior covariance matrix of the tomographic model, $\underline{\underline{\widetilde{C}}}_M$, generating an ensemble of models. Each member within this ensemble is a complete whole-mantle model of wavespeeds and density evaluated on an equispaced mesh with 2562 points laterally and at 50 km intervals in depth. At each depth, we perform a spherical harmonic expansion using routines from the Slepian software package (Simons et al., 2006) which fit the spherical harmonic basis functions to the equispaced point values using a least-squares inversion. Then, for each density model in the ensemble, we perform a forward mantle flow calculation using a reference viscosity profile (Model C from Steinberger & Holme, 2008) and obtain a set of model geoid spherical harmonic coefficients. These are used to form a sample covariance matrix that represents the data plus forward modeling uncertainty, with the caveat that we assumed a single viscosity profile when generating the ensemble. The number of samples required for accurate estimation of the covariance matrix was not known *a priori*, but we estimate the minimum number of samples required at 35 per pair of parameters. For the 60 geoid coefficients for spherical harmonic degrees 2-7, we find that the number of samples is at least $35 * \binom{60}{2} = 61,950$. We verified that the eigenvalues of the covariance matrix converge to nearly constant values as the number of samples was increased from $10^3 - 10^5$ (Figure S2). We also carried out a suite of inversions with a diagonal covariance matrix (i.e. assuming that the combined data and forward-modeling uncertainties associated with each spherical harmonic coefficient are equal and uncorrelated).

3 Results

The model viscosity profiles from our inversions are shown in Figure 3 for the end-member density models with least (panels A,C) and most (panels B, D) preferred correlation between ρ and V_S. In each panel of Figure 3 we show viscosity profiles (calculated from the mean value in the ensemble solution at each depth) for models constrained by geoid spherical harmonic degrees $l = 2-3$, $l = 2\&4$ (only even degrees), and $l = 2-7$. The shading in Figure 3 represents probability density. Figure S3 shows the posterior pdf of viscosity for each ensemble separately along with the number of control points and the distribution of variance reduction. We also show mean viscosity profiles and pseudocolor plots of probability density, model complexity, and variance reduction for inversions with a diagonal data and forward modeling covariance matrix in Figure S4-S5. All of the ensemble solutions have an increase in viscosity between the upper mantle and the lower mantle. The density models with more imposed correlation between ρ and V_S (Figure 3B,D) yield viscosity solutions that have a low-viscosity channel (roughly 1/10 the average viscosity at 660 km) below the base of the transition zone (660 km) and a rapid increase in viscosity near 1000 km depth. The overall shape of the viscosity profiles appear very consistent regardless of which subset of geoid spherical harmonics is used to constrain the inversion. However, the models constrained by more spherical harmonic coefficients in general show more complex viscosity profiles, as expected for the parsimonious inversions employed here. The models with less imposed correlation between ρ and V_S (Figure 3A,C) show a similar overall increase in viscosity between upper and lower mantle, but more overall variability among the viscosity profiles in the ensemble, as indicated by the broader shaded confidence intervals. We still find evidence for a viscosity maximum at or somewhat below 1000 km depth. The viscosity profiles shown in Figure 3A,C correspond to density models with less imposed correlation between ρ and V_S in the lowermost mantle. These density models contain lowermost mantle density heterogeneity that is uncorrelated with the geoid. The low-viscosity region centered between 2000–2500 km in panels A,C is required to minimize the sensitivity to density structure at these depths and is discussed later. The systematic development of this low-viscosity region in the lower mantle can be seen as a natural consequence of relaxing the imposed correlation between ρ and V_S (i.e. reducing $\gamma_\rho^{D''}$) and increasing the data weight assigned to the density-sensitive normal mode splitting measurements (i.e. increasing $w_{{}_0S_2}$). However, it may not be a robust feature, as will be discussed later. We show observed and modeled geoids for the four viscosity models from Figure 3 in Figure 4. For all density models, the pattern and amplitude of the synthetic geoid are in reasonably good agreement with the observations. The variance reduction for the spherical harmonic degrees included in each inversion is shown in Figure 3.

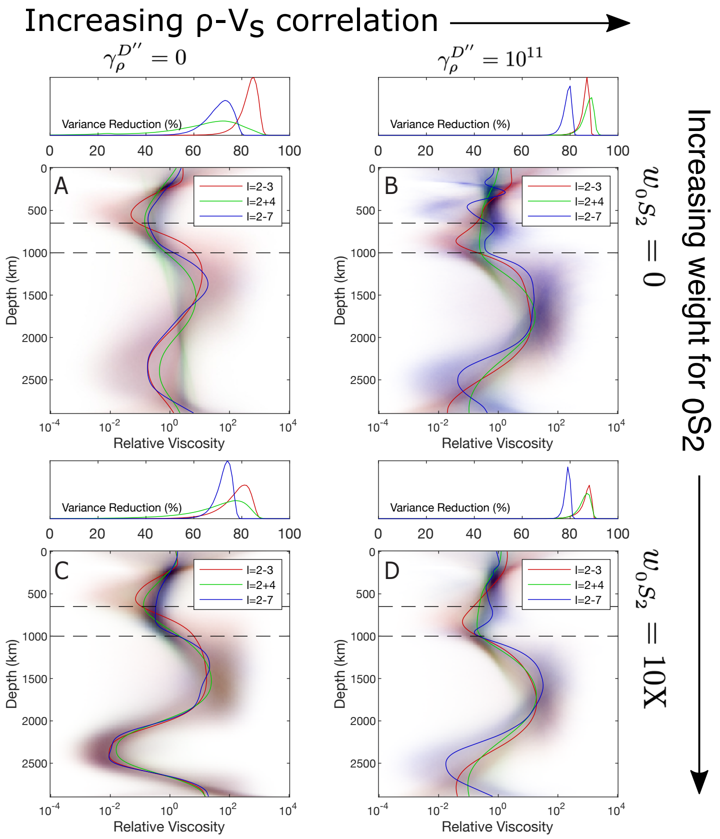

Figure 3. Results from viscosity inversions. The density models used in (A,C) have smaller misfit to the normal mode splitting measurements whereas the models in (B,D) have density variations that are strongly correlated with V_S variations, at the expense of fitting the normal mode constraints. The blue, green, and red curves in each panel correspond to models constrained by spherical harmonic degrees $l = 2-3$, $l = 2,4$, and $l = 2-7$ respectively. In each panel, the red, green, and blue colors show the probability distributions for each combination of spherical harmonic degrees. We also show the histogram of geoid variance reduction associated with each model.

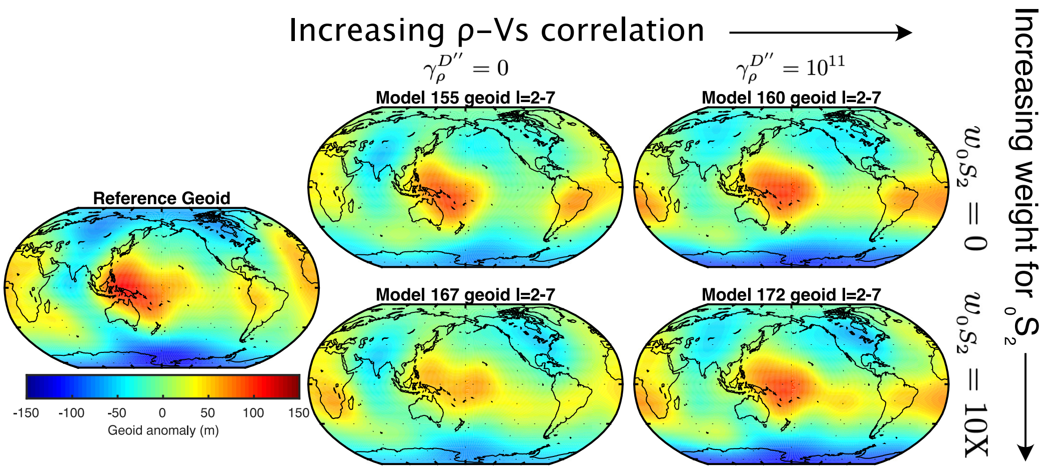

Figure 4. Model geoids for $l = 2-7$ for the four end-member density models for inversions with a hierarchical parameter. The model geoids are shown for the "median model" from each ensemble.

3.1 Interplay of radial viscosity variations and density heterogeneity

We used whole-mantle density models to invert for the mantle viscosity profile as constrained by the long-wavelength geoid. The density models are inverted with full-spectrum tomography using various data-types whose relative contributions in the inversion are set by *a priori* weights. The scenarios in the model suite differ in their choices for the weight assigned to ρ-sensitive normal-mode splitting measurements ($w_{_0S_2}$) and the degree of correlation between ρ and V_S variations in the lowermost mantle ($\gamma_\rho^{D''}$). In general, models with reduced $\gamma_\rho^{D''}$ and increased $w_{_0S_2}$ provide statistically significant improvements in data fits along with anti-correlated $V_S - \rho$ structure where the base of LLSVPs are denser than the ambient mantle. In an end-member case, illustrated in Figure 2D, the lowermost mantle density structure contains an approximately 1% density excess beneath Africa and the western Pacific, overlapping with the locations of the low-v_S LLSVPs but with somewhat discrepant patterns. The preferred density model from (Moulik & Ekström, 2016) adopted a moderate weight ($w_{_0S_2} = 1X$) to obtain a density excess of around 0.5% in the bottom 500 km of LLSVPs.

When lateral viscosity variations are ignored, lateral heterogeneity at a given spherical harmonic degree and order can only cause geoid anomalies at the same degree and order with no coupling across harmonics. It is therefore possible to represent the sensitivity of the geoid to perturbations in density structure at different depths through the geoid kernel (e.g. Hager, 1984; Richards & Hager, 1984). The misfit between modeled and observed geoid anomalies is minimized when the geoid kernel takes on positive values at depths where the pattern of mantle buoyancy variations is strongly correlated with the geoid and negative values at depths where the buoyancy structure is anti-correlated with the geoid. At depths where the mantle buoyancy structure is uncorrelated with the geoid, the residual can only be minimized by ensuring that the geoid kernel, itself a function of the radial viscosity structure, takes on a value close to zero. Approximately 70% of the power in the dynamic geoid is concentrated in the lowest spherical harmonic degrees 2-3, and power falls off rapidly with increasing degree (e.g. Kaula, 1966). Therefore, we conclude that it is possible to obtain reasonable fits to the observed geoid for all density models explored because at the long wavelengths considered here, lateral heterogeneity is either strongly correlated or anti-correlated with the geoid throughout much of the mantle.

The correlation between the long-wavelength geoid and density variations in our model suite varies across density models and with depth. In Figure 5(C,F,I,L), we show the correlation between the geoid and the density structure in four end-member density models for spherical harmonic degrees 2, 3, and 4. Models 160 and 172 both exhibit density variations that are highly correlated with V_S variations, and we can see clearly that the degree 2–4 density structure is highly correlated with the geoid in the upper mantle, and that the degree 2–3 density structure is highly anti-correlated with the geoid in the lower mantle. This pattern of correlation is a common feature of almost all published shear-wave tomographic models. On the other hand, models 155 and 167 (Figure 5C,I) that relax the imposed V_S-ρ correlation and assign more weight to the density-sensitive normal-mode data have degree-2 density structure that is anti-correlated with the geoid through the mid mantle (660–2000 km depth) but positively correlated with the geoid in the bottom ~ 500 km of the mantle. Since our normal-mode data do not constrain the odd-degree density variations (e.g. Resovsky & Ritzwoller, 1995; Deuss et al., 2011), degree-3 density variations do not depart from scaling to the V_S models, and the degree-3 density structure in all density model scenarios shows a transition from anti-correlation to correlation with the geoid, tracking the behavior seen with V_S models. Because the degree-2 and degree-3 geoid kernels are generally very similar in their overall shape for a given viscosity structure, it is more challenging to obtain a viscosity profile that matches the observed geoid. For density models 155 and 167, we recovered viscosity profiles shown in Figure 5(A,G). The geoid kernels corresponding to the "median model" from each ensemble are shown in Figure 5(B,H). Here we differentiate between the "*ensemble average*"- the mean viscosity value present in the posterior ensemble at each depth and the "*median model*" - the individual solution from the ensemble whose likelihood is closest to the peak of the likelihood distribution in the ensemble. Model 155 has degree-2 lowermost mantle structure that is overall less well-correlated with the geoid than Model 167. The viscosity profiles for inversions constrained by $l = 2 + 4$ and $l = 2 - 7$ for Model 155, shown in Figure 5B generally remove sensitivity in the lowermost mantle, where much of the power in the buoyancy field is poorly correlated with the geoid. The sensitivity is suppressed by the introduction of a viscosity reduction around 2000 km depth. Model 167 contains degree-2 structure that is more positively correlated with the observed geoid than Model 155, but other spherical harmonic degrees remain negatively correlated (e.g. degree 3) or very weakly negatively correlated (degree 4). Consequently, the viscosity profiles obtained using Model 167 contain an even more pronounced reduction in viscosity above D'', effectively suppressing geoid sensitivity in the lowermost mantle. Our results demonstrate that the relative patterns and amplitudes of long-wavelength density variations influence inferences on radial viscosity structure in a non-linear fashion. Evaluating the consistency between even- and odd-degree density variations with new seismic constraints will be critical to improved constraints on viscosity structure.

In Figure 5(A,D,G,J), we show selected models from the posterior ensemble solutions (dotted lines) as well as the ensemble mean and shaded confidence interval. It is important to note that the individual solutions in many cases possess far more overall variation in viscosity than the ensemble average. Furthermore, because the geoid kernels depend strongly on the viscosity profile, the ensemble average itself may not produce a small misfit to the geoid. Instead, the ensemble average and confidence interval should be viewed as an indicator of the common properties of accepted solutions, and the range of uncertainty in the model parameters.

3.2 Uncertainty quantification

The transdimensional, hierarchical Bayesian approach used here yields samples from the posterior that can be used to directly quantify uncertainty in the viscosity profile, regardless of the shape of the posterior distribution. This is in contrast with uncertainty quantification in linearized inversions in which the posterior in approximated by a multivariate normal distribution. When the posterior probability density is “close” to a normal distribution, the posterior covariance operator can be approximated through a linearization around the maximum likelihood solution ($\underline{m}_{ml}$) using Equation 3.56 of (Tarantola, 2005)

$$\underline{\underline{\widetilde{C}}}_M^{lin} \simeq \left(\underline{\underline{G}}^T \underline{\underline{C}}_D^{-1} \underline{\underline{G}} + \underline{\underline{C}}_M\right)^{-1} \tag{9}$$

where $\underline{\underline{C}}_D$ is the covariance matrix representing data and forward modeling uncertainty, $\underline{\underline{C}}_M$ is the prior covariance matrix, and $\underline{\underline{G}}$ contains partial derivatives of the forward model operator (geoid kernels g_i) with respect to the model parameters $\underline{m}$,

$$G_{ij} = \left.\frac{\partial g_i}{\partial m_j}\right|_{\underline{m}=\underline{m}_{ml}}. \tag{10}$$

For the uniform, very broad prior on viscosity assumed here, the prior covariance matrix in Equation 9 effectively vanishes.

To better understand the uncertainty estimates recovered by our method and to compare these estimates with those that could be obtained using other, more established, methods in geophysics, we compute a few illustrative results. First, we compute the estimate of the posterior covariance matrix around the maximum likelihood point. We identify the “median model” from the posterior ensemble, interpolate it onto 20 regularly-spaced points in depth, and calculate the partial derivatives of the geoid coefficients with respect to viscosity using numerical differentiation. We compute $\underline{\underline{\widetilde{C}}}_M^{lin}$ using the Moore-Penrose pseudoinverse because in some cases (especially without a prior on viscosity) the covariance matrices are nearly singular. The posterior covariance matrix $\underline{\underline{\widetilde{C}}}_M^{lin}$ is shown in Figure 6A and selected rows of the covariance matrix are shown Figure 6B-C. We also computed an estimate of the posterior covariance matrix including prior uncertainties on $\log_{10}\eta(r)$ of 0.5 (log-units), shown in Figure 6D-F. Regardless of whether the prior on η is included, the covariance matrices shown here represent very strong tradeoffs between parameters and depths at which the model has little sensitivity.

Figure 7 illustrates the uncertainty on the viscosity profiles implied by the actual posterior ensemble, a re-sampling of the same ensemble, and by the covariance around the maximum likelihood point (Equation 9). The results shown correspond to density Model 155, constrained by spherical harmonic degrees $l = 2{-}7$. In Figure 7B, we generated a sample covariance matrix from the posterior ensemble, and then sampled viscosity profiles from this covariance matrix. In Figure 7, we show the pdf corresponding to the linear estimate of $\widetilde{C}_M^{lin}$ from Equation 9 with a flat, broad prior on η ($\underline{\underline{C}}_M^{-1} = \underline{\underline{0}}$). Three key observations emerge from this analysis. First, the posterior pdf in our ensemble solutions is not well-described by a multivariate normal distribution. Second, the posterior is multimodal at some depths. Third, the linearized covariance operator grossly underestimates the true variability among solutions accepted in the posterior ensemble. All three are a direct consequence of the non-linearity inherent in the inversion of geoid data for viscosity, as illustrated in Figure 1.

3.3 Mechanisms for radial viscosity variations

In a mantle with pyrolitic composition, phase transitions of olivine to its high-pressure polymorphs have important implications for the inferences of viscosity in the transition zone. We recover no large variations in viscosity at 410 km and within the mantle transition zone (410–650 km), consistent with recent global inversions of shear attenuation (Moulik, 2016). Large changes in viscosity or the potentially related shear attenuation in the transition zone can be disfavored based on mineralogical considerations; phase transition from olivine to wadsleyite ($\sim$410 km) and ringwoodite ($\sim$550 km) do not involve a wholesale reordering of the unit cell structure as in the ringwoodite to perovskite transition ($\sim$650 km). Other complicating effects such as grain size reduction could influence the nature of viscosity variations across the 650-km discontinuity (e.g. Panasyuk & Hager, 1998; Solomatov & Reese, 2008; Dannberg et al., 2017).

An increase in viscosity below the base of the transition zone is a persistent feature among the viscosity profiles from our inversions (Figure 3). The origin of this feature cannot be easily attributed to a single physical mechanism since its depth is not coincident with known mantle phase transitions; nevertheless, several plausible explanations exist, many of which are not mutually exclusive. The inversions using models 160 and 172, both of which have density structures that closely resemble scaled V_S tomography, generally favor the presence of a reduction in viscosity at 660 km depth and a subsequent increase in viscosity close to 1000 km depth. Similar low-viscosity channels were recovered in viscosity inversions constrained by the global long-wavelength geoid (Forte et al., 1993) and by shorter-wavelength ($l = 12 - 25$) variations in the oceanic geoid (Kido et al., 1998).

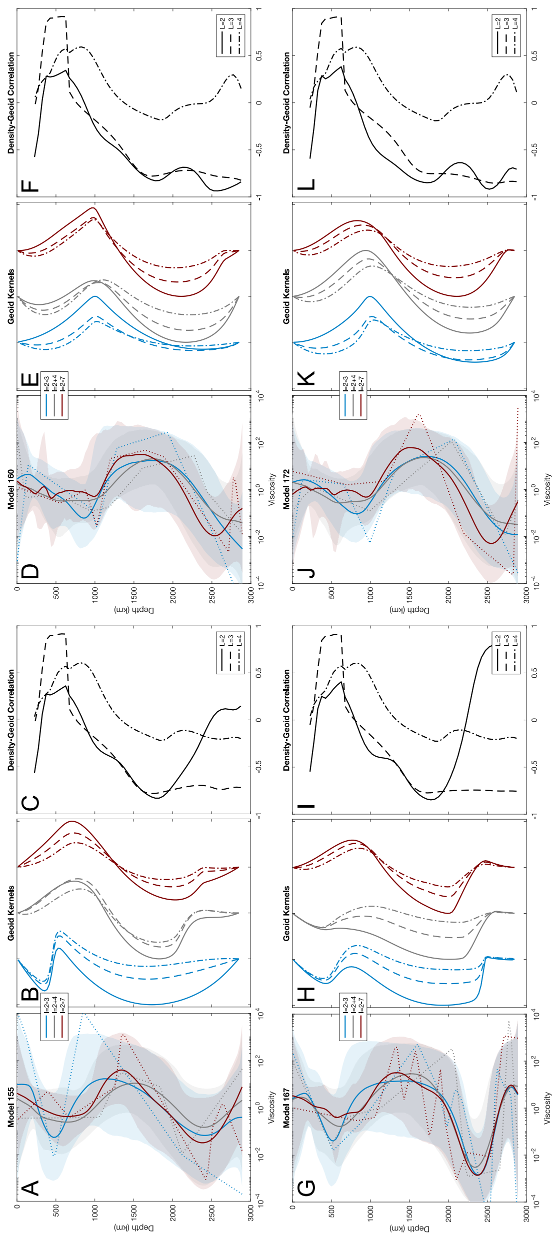

Figure 5. Viscosity profiles, geoid kernels, and the correlation between ρ and geoid anomaly for Models 155 (A-C), 160 (D-F), 167 (G-I), and 172 (J-L). In panels (A,D,G,J), we show the ensemble mean (solid) and the median model (dotted) for the inversions with hierarchical parameter. The shaded region indicates a 90% confidence interval. (B,E,H,K) show geoid kernels for the median model for each inversion. Blue, grey, and red curves correspond to inversions constrained by spherical harmonic degrees l=2-3, l=2,4, and l=2-7, respectively and the kernels for degrees 2-4 are shown using solid, dashed, and dash-dotted lines. The kernels are normalized to unit amplitude and shifted for clarity. (C,F,I,L) Here we show the correlation between ρ and the geoid for spherical harmonic degrees 2, 3, and 4 for each of the density models.

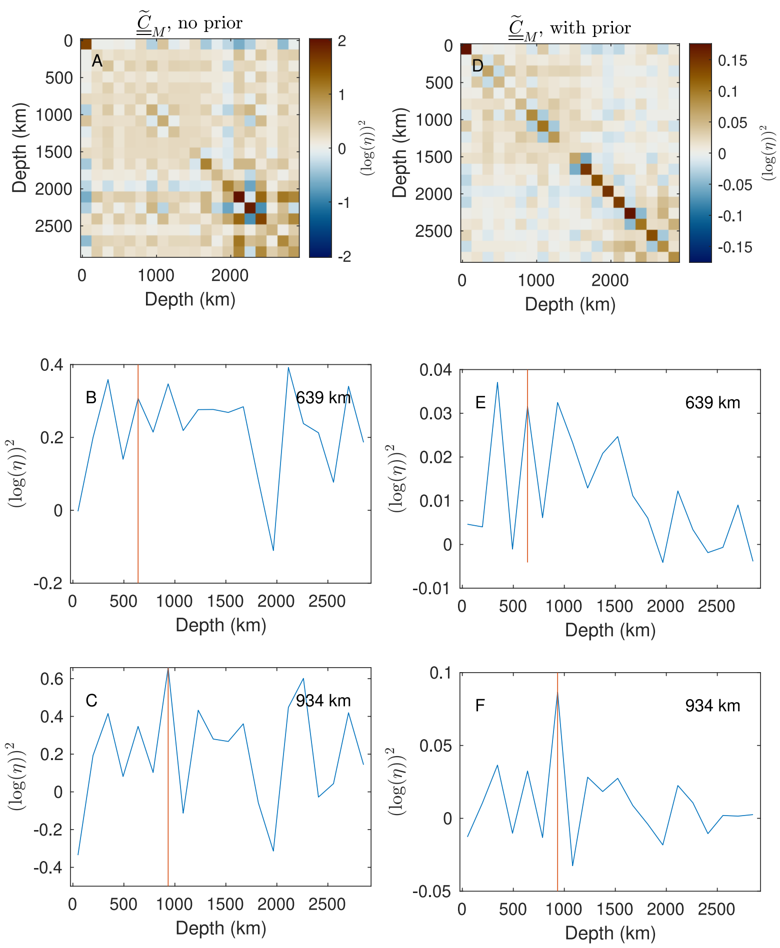

Figure 6. Linearized estimates of the posterior covariance matrix $\underset{=}{\widetilde{C}}_M^{lin}$ for viscosity for the inversion using density Model 155. Panels (A) and (D) illustrate the pattern of covariance without and with a prior on viscosity. In (D), the prior assumes an uncorrelated uncertainty in viscosity of 0.5 log-units. In panels (B-C) we show rows of the covariance operator for depths closest to 600 km and 1000 km corresponding to (A). Panels E-F show the same information as (B-C) for the covariance operator in panel D.

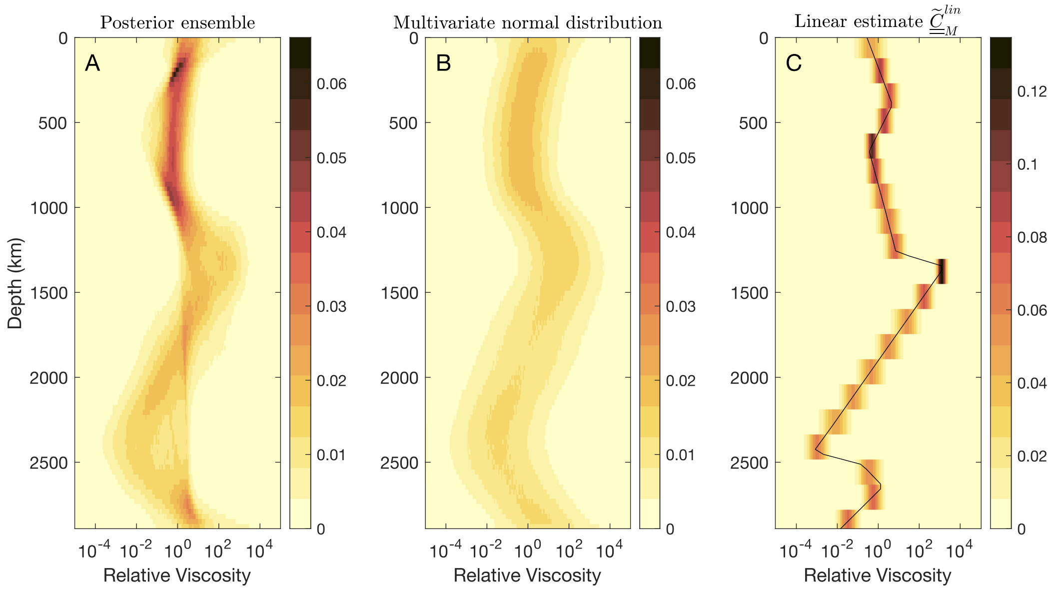

Figure 7. Posterior probability distribution (pdf) for viscosity. (A) shows the true sampling of the posterior ensemble. (B) shows a multivariate normal distribution approximating the pdf in (A). (C) shows a pdf centered around the maximum likelihood solution (black curve) described by the linearization in Equation 9. We note that the pdf in (B) significantly overestimates the uncertainty of model at most depths while the pdf in (C) underestimates the uncertainty represented by the true posterior.

A reduction in viscosity below 650 km depth might be associated with grain size reduction as downwelling material crosses the perovskite-forming phase transition. The radial extent of the low-viscosity region below the 660 km phase transition could be much smaller than the feature recovered in our inversions, with a thickness of about 1 km estimated on theoretical grounds (Panasyuk & Hager, 1998). Alternatively, upwelling plumes could be blocked partially by an endothermic phase transition, resulting in ponding of warm, low-viscosity material below the transition zone. The presence of a reduced-viscosity channel below the 650 km phase transition may have important dynamical implications. Sinking slabs could move laterally with relative ease through a low viscosity region, promoting stagnation in the transition zone beneath the northwest Pacific and eastern China as observed in tomographic models (e.g. Fukao et al., 2009; Moulik & Ekström, 2014; French & Romanowicz, 2014) and confirmed using geodynamic models that include a low-viscosity channel and an endothermic phase transition (Mao & Zhong, 2018; Lourenç & Rudolph, 2020).

An increase in viscosity around 1000 km depth has been supported by multiple studies of the mantle viscosity profile constrained by glacial isostatic adjustment and geoid (King & Masters, 1992; Mitrovica & Forte, 1997). For highly simplified two-layer viscosity structures in which the depth and magnitude of a viscosity contrast are the only parameters, preferred depth of the viscosity increase between the shallow and deep mantle depends on the definition of misfit (e.g. correlation, L2-norm) and on the assumptions used to generate buoyancy structures from mantle tomography, but the results uniformly favor a viscosity increase below 660 km, with preferred depths c. 800-1200 km (Forte, 1989; Rudolph et al., 2015). The reason that geoid inversions favor a deeper viscosity increase can be understood from an inspection of the geoid kernels and the correlation between the geoid and tomography models. The degree-2 geoid is positively correlated with buoyancy structure throughout the upper mantle, and while there is a decrease in the correlation coefficient between geoid and buoyancy at 650 km depth, the correlation remains positive for several models down to almost 1000 km. In the presence of an increase in viscosity, the geoid kernels for layered viscosity structure change sign from positive in the upper layer to negative in the lower layer, and the depth of the change in sign increases with the depth of the viscosity increase.

Our previous work found that the positive V_S-geoid correlation persists to $\sim$1000 km depth in models that employ continuous and smooth parameterization thereby recovering smooth changes in the heterogeneity spectrum across the 650-km phase transition (French & Romanowicz, 2014). Such a paramaterization can lead to smearing of heterogeneity from the transition zone to the uppermost lower mantle (Gu et al., 2001), especially when employing data sensitive to these depths such as normal modes (Moulik & Ekström, 2014). Nevertheless, all of the density models used here contain an abrupt reduction in correlation between density variations and the degree 2-3 geoid at 650-km depth (Figure 5(C,F,I,L)), as well as a decrease in the RMS amplitude of both overall power and power at low degrees (Moulik & Ekström, 2014). However, at degrees 2-4, the correlation between density heterogeneity and the geoid remains positive below 650 km, persisting to 800-1200 km (e.g. Figure 5F). The dramatic reduction in density-geoid correlation at 650 km depth coincides with an allowed discontinuity between the 16 cubic splines in the radial direction (Moulik & Ekström, 2016). This choice, common across a subset of recent tomographic models (Kustowski et al., 2008; Moulik & Ekström, 2014), represents the *a priori* information that deviations in large-scale pattern of mantle heterogeneity could coincide with the 650-km phase transition and is substantiated by the improved fits to precursors of the body-wave phase SS that reflect off this discontinuity (Gu et al., 2003). Boschi and Becker (2011) reported improved fit to body wave travel times with a greater depth of decorrelation ($\sim$ 800 km) but did not include additional data such as SS precursors that directly constrain transition zone topography, which trades off with volumetric wavespeed variations at shallow lower mantle depths (Moulik & Ekström, 2014). In our previous inferences of viscosity based on SEMUCB-WM1 (French & Romanowicz, 2014), which has a continuous, smooth, parameterization in the radial direction, we recovered solutions favoring an increase in viscosity near 1000 km depth (Rudolph et al., 2015). It is noteworthy that in spite of the allowed decorrelation at 650 km depth in the density models used here, we still recover viscosity structures that prefer a viscosity increase somewhat deeper than the base of the transition zone, which we attribute to the fact that for spherical harmonic degrees 2 and 4, the correlation between density structure and geoid remains positive below the base of the transition zone, and only becomes negative deeper within the lower mantle. The deeper viscosity increase is driven by the dominant degree-2 variation that has a positive correlation with the geoid which persists to the uppermost lower mantle.

Several potential mechanisms exist that could explain an increase in viscosity below the 650-km discontinuity. First, Marquardt and Miyagi (2015) measured the strength of ferropericlase and found a threefold increase in strength over the pressure range 20-65 GPa. Though ferropericlase constitutes the minority of the lower mantle, Marquardt and Miyagi (2015) argue that it could form interconnected layers/sheets especially in high strain-rate regions, controlling the lower mantle rheology, an idea that is supported by two-phase deformation experiments on analogue materials (Kaercher et al., 2016) and on mixtures of Bridgmanite and Magnesiowüstite (Girard et al., 2016). Second, changes in the proportionation of iron between bridgmanite and ferropericlase in the depth range of 1200-1600 km could produce a mid-mantle viscosity hill (Shim et al., 2017). If a greater proportion of iron is incorporated in ferropericlase, the melting temperature of bridgmanite increases, and viscosity is expected to increase based on homologous temperature scaling. Shim et al. (2017) predict a viscosity hill with a maximum viscosity at $\sim$ 1200 km depth and a value approximately two orders of magnitude larger than the viscosity at 660-km depth. The magnitude of the predicted viscosity variation and the depth of the viscosity maximum are in reasonably good agreement with the viscosity profiles shown here. Third, Deng and Lee (2017) measured the solidus and liquidus temperatures of ferropericlase and found a maximum at pressures corresponding to 1000 km depth, again implying an increase in viscosity on the basis of homologous temperature scaling. These various mechanisms do not appear to be mutually incompatible.

3.4 Implications for joint modeling of seismic and geodynamic data

Some of the earliest tomographic models (e.g. Woodhouse & Dziewoński, 1984) exhibited large-scale structures that were fairly well correlated with the major surface manifestations of mantle convection, such as the long-wavelength non-hydrostatic geoid (Hager et al., 1985) and the large-scale tectonic plate motions (Forte & Peltier, 1987). Recent tomographic models, such as the ones used here based on Moulik and Ekström (2016), have afforded refined images and better fits to diverse measurements from broadband seismograms. The analysis presented above has focused on how seismically-inferred mantle density structure may be used to model mantle flow in order to predict surface observables. However, our results also have major implications for the converse approach, i.e. inferring mantle structure from either geodynamic observations in isolation (e.g. Hager, 1984; Ricard et al., 1989; Forte, 1989) or jointly with seismic data (e.g. Simmons et al., 2010).

The earliest inferences of radial viscosity variation (e.g. Hager et al., 1985) assumed constant scaling throughout the mantle to convert velocity variations to those of density that drive mantle flow. This is only appropriate for a purely thermal contribution to seismic velocity heterogeneity throughout the mantle, and is contrary to multiple lines of seismic evidence (e.g. Ritsema & Lekic, 2020) including normal mode (Moulik & Ekström, 2016) and tidal constraints (Lau et al., 2017). Such scaling assumptions are also employed in the construction of recent tomographic models (e.g. Ritsema et al., 2011; French & Romanowicz, 2014), which assume fixed scalings between $d\ln V_S$, $d\ln V_P$, and $d\ln\rho$, in order to account for data sensitivity to parameters that are not directly inverted for. When attempts are made to jointly model seismic and geodynamic data (e.g. Simmons et al., 2009), observations such as normal modes that can uniquely disentangle the density contributions from those of other elastic parameters are often excluded. Our results demonstrate that inferences of viscosity variations and thereby mantle flow depend strongly on the density models. Inferred jumps and gradients in viscosity from a correlated V_S-ρ model can differ from an independent ρ inversion by up to 2 orders of magnitude and contain features such as low-viscosity channels. Therefore, dynamical inferences based on constant V_S-ρ scaling assumptions or tomographic models constructed therewith may be biased in the dynamically important boundary regions of the Earth e.g. transition zone and lowermost mantle.

Several recent studies have attempted to relate dynamical observations (e.g. geoid) to structural heterogeneity (e.g. temperature, velocity, density) to jointly constrain mantle flow dynamics. For example, Simmons et al. (2009) inverted some geodynamic and a small subset of available seismic constraints, primarily body wave arrival times, for lateral heterogeneity assuming a fixed viscosity profile (Mitrovica & Forte, 2004). We demonstrate that sensitivity kernels that relate plate motions and geoid to density variations (Figure 5) depend strongly on viscosity variations. Radial viscosity changes can amplify sensitivity in depth ranges where density variations are consistent with the geoid anomalies, and even nullify sensitivity in regions where the two are dissimilar. The use of a constant radial viscosity profile in earlier studies implicitly introduces a strong *a priori* assumption about the relative contribution of heterogeneity at various depths to surface geodynamic observations. An iterative procedure of recalculating sensitivity kernels may help converge towards a self-consistent solution of radial viscosity variations and structural heterogeneity.

Our results on radial viscosity variations could potentially inform the parameterization and regularization choices in seismic tomography. Moulik and Ekström (2016) employed a parameterization that allowed various spherical harmonic degrees in density structure to deviate from a scaled V_S structure as dictated by seismic observations. Since the self-coupled normal-mode splitting observations constrain only even-degree density variations, all inversions strongly disfavored even-degree V_S-ρ correlation ($R_2 \sim$ -0.46 to -0.25) in the lowermost mantle while retaining the starting assumptions on positive V_S-ρ correlation in the remaining regions and for odd degree variations. The opposing sign of the correlation of the longest wavelength even- vs. odd-degree structure with the geoid maps into a region of reduced viscosity in the lower mantle in our viscosity inversions. Since this viscosity feature is a product of current limitations in data, it needs further evaluation with odd-degree sensitive observations (e.g. Resovsky & Ritzwoller, 1995).

An alternative approach to suppress this even-odd degree dichotomy in V_S-ρ correlation is to parameterize structural heterogeneity in terms of a radial V_S-ρ scaling ratio and three-dimensional velocity heterogeneity, as used in previous inversion (e.g. Robertson & Woodhouse, 1996; Simmons et al., 2009) and forward modeling schemes (e.g. Lau et al., 2017; Koelemeijer et al., 2017). In contrast to the methods employed in our modeling, this approach enforces perfect V_S-ρ (anti)correlation with a radially varying scaling factor that is consistent across spherical harmonic degrees, and would likely disfavor a low viscosity channel in the lower mantle. However, it is not immediately clear if such a strong prior assumption on V_S-ρ scaling is either compatible with data or is physically reasonable in a strongly heterogeneous boundary region. Various mechanisms (e.g. partial melt, iron enrichment, primordial material, grain size variations) may manifest more strongly at different spatial scales in the lowermost mantle and get expressed as spatially varying correlations and amplitudes of V_S-ρ scaling. For instance, small-scale structures such as Ultra Low Velocity Zones (e.g. Thorne & Garnero, 2004; Rost et al., 2005; Cottaar & Romanowicz, 2012) may have a different physical origin (and associated $V_S-\rho$ scaling) than the larger-scale LLSVPs. Joint and iterative inversions of structural heterogeneity and radial viscosity variations with new and improved measurements may help disentangle such effects in the Earth's deep interior.

4 Conclusions

We used recently-developed whole-mantle V_P, V_S and density models from full-spectrum tomography (Moulik & Ekström, 2016), together with their associated covariance matrices, to infer the mantle viscosity profile as constrained by the long-wavelength geoid. The resulting inferences of depth-variation in viscosity contain several persistent features, including an increase in viscosity below the base of the mantle transition zone (often near 1000 km depth) and a maximum in mantle viscosity at mid-mantle depths. Our ensemble solutions permit the quantification of uncertainties in the inferences of viscosity. We found that uncertainty estimates based on linearized inversions are likely to woefully understate true uncertainty, and therefore the robustness of specific complexities in viscosity profiles that were justified using an uncertainty analysis based on covariance around the maximum likelihood point. This might explain why so many different viscosity profiles have been proposed based on inversions constrained by similar gravity data and using similar forward modeling assumptions.

It is noteworthy that the mid-mantle viscosity increase persists in the suite of inversions presented here despite the use of a radial parameterization in the tomographic model suite that includes a discontinuity at 650 km depth as opposed to the smooth radial parameterization used in SEMUCB-WM1, the basis for our previous inferences that favored a viscosity increase at $\sim$ 1000 km depth (Rudolph et al., 2015). The detailed features of our inversions are refined in the upper mantle, within and below the transition zone. The robustness of our inferences is evaluated across a suite of density models. Density variations that most-closely resemble the V_S structure predict viscosity structures that contain a low-viscosity channel below the 660 km discontinuity. While models with anti-correlated $V_S-\rho$ structure in the lowermost mantle with the base of LLSVPs denser than the ambient mantle fit the seismological constraints significantly better, they currently provide discrepant even and odd-degree correlations with the geoid and poorer overall correlation. The best-fitting viscosity profiles for these seismically-preferred models tend to contain a low-viscosity channel between $\sim$2000-2500 km depth that acts to reduce the geoid sensitivity to buoyancy variations in the lowermost mantle. For all of the density models considered, it is possible to find a viscosity profile that accurately predicts the geoid. Iteratively solving for radial viscosity variations and density heterogeneity is likely to account for their strong nonlinear relationship in joint inversions of seismological and geophysical observations. New seismological constraints on density structure and joint modeling with the geoid may provide improved insights on the thermochemical nature of the lowermost mantle.

Acknowledgments

All authors received support from the National Science Foundation through NSF Geophysics grant EAR-1825104. Computational resources were provided through NSF Major Research Instrumentation grant DMS-1624776 to Portland State University. The geoid data used to constrain the models are freely available from the GRACE project. The propagator matrix code HC used to carry out the viscosity inversions is available through the Computational Infrastructure for Geodynamics `https://github.com/geodynamics/hc`. The Slepian software for spherical harmonic expansions was downloaded from `https://github.com/csdms-contrib/slepian_alpha`. Codes to carry out the inversions for radial viscosity structure are on GitHub `https://github.com/maxrudolph/hc/tree/thb_hc` and will be archived on Zenodo at the time of article acceptance. The figures in this paper use Fabio Crameri's scientific colormaps (Crameri, 2019). We thank the editor, Thorsten Becker, and two anonymous reviewers and Scott King for helpful comments that improved the quality of the manuscript.

References

Ballmer, M. D., Schumacher, L., Lekic, V., Thomas, C., & Ito, G. (2016). Compositional layering within the large low shear-wave velocity provinces in the lower mantle. *Geochemistry, Geophysics, Geosystems*, *17*(12), 5056–5077. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GC006605` doi: 10.1002/2016GC006605

Becker, T. W., & Boschi, L. (2002, January). A comparison of tomographic and geodynamic mantle models. *Geochemistry, Geophysics, Geosystems*, *3*(1), 1003–n/a. doi: 10.1029/2001GC000168

Becker, T. W., O'Neill, C., & Steinberger, B. (2014). *HC, a global mantle circulation solver.* Retrieved from `https://github.com/geodynamics/hc`

Bodin, T., Sambridge, M., Rawlinson, N., & Arroucau, P. (2012, June). Transdimensional tomography with unknown data noise. *Geophysical Journal International*, *189*(3), 1536–1556. Retrieved 2019-12-20, from `https://academic.oup.com/gji/article/189/3/1536/611337` doi: 10.1111/j.1365-246X.2012.05414

.x

Boschi, L., & Becker, T. W. (2011, January). Vertical coherence in mantle heterogeneity from global seismic data. *Geophysical Research Letters*, *38*(20), L20306.

Brodholt, J. P., Helffrich, G., & Trampert, J. (2007, October). Chemical versus thermal heterogeneity in the lower mantle: The most likely role of anelasticity. *Earth and Planetary Science Letters*, *262*(3), 429–437. Retrieved 2019-12-23, from `http://www.sciencedirect.com/science/article/pii/S0012821X07004992` doi: 10.1016/j.epsl.2007.07.054

Burke, K., Steinberger, B., Torsvik, T. H., & Smethurst, M. A. (2008, January). Plume Generation Zones at the margins of Large Low Shear Velocity Provinces on the coremantle boundary. *Earth and Planetary Science Letters*, *265*(1-2), 49–60. doi: 10.1016/j.epsl.2007.09.042

Burke, K., & Torsvik, T. H. (2004). Derivation of Large Igneous Provinces of the past 200 million years from long-term heterogeneities in the deep mantle. *Earth and Planetary Science Letters*, *227*(34), 531–538.

Chambat, F., Ricard, Y., & Valette, B. (2010, November). Flattening of the Earth: further from hydrostaticity than previously estimated. *Geophysical Journal International*, *183*(2), 727–732. doi: 10.1111/j.1365-246X.2010.04771.x

Cottaar, S., & Lekic, V. (2016). Morphology of seismically slow lower-mantle structures. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *207*(2), 1122–1136.

Cottaar, S., & Romanowicz, B. (2012, November). An unusually large ULVZ at the base of the mantle near Hawaii. *Earth Planet. Sci. Lett.*, *355-356*, 213–222.

Crameri, F. (2019, May). *Scientific Colour Maps.* Zenodo. Retrieved 2020-07-27, from `https://zenodo.org/record/3596401#.Xx9BifhKhqc` (Language: eng) doi: 10.5281/zenodo.3596401

Dannberg, J., Eilon, Z., Faul, U., Gassmller, R., Moulik, P., & Myhill, R. (2017). The importance of grain size to mantle dynamics and seismological observations. *Geochemistry, Geophysics, Geosystems*, *18*(8), 3034–3061. Retrieved 2019-10-02, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GC006944` doi: 10.1002/2017GC006944

Deng, J., & Lee, K. K. M. (2017, December). Viscosity jump in the lower mantle inferred from melting curves of ferropericlase. *Nature Communications*, *8*(1), 1997. Retrieved 2019-07-27, from `https://www.nature.com/articles/s41467-017-02263-z` doi: 10.1038/s41467-017-02263-z

Deschamps, F., & Trampert, J. (2003, December). Mantle tomography and its relation to temperature and composition. *Physics of the Earth and Planetary Interiors*, *140*(4), 277–291. Retrieved 2019-12-23, from `http://www.sciencedirect.com/science/article/pii/S0031920103001894` doi: 10.1016/j.pepi.2003.09.004

Deuss, A., Ritsema, J., & Heijst, H. v. (2011). Splitting function measurements for Earth's longest period normal modes using recent large earthquakes. *Geophysical Research Letters*, *38*(4). Retrieved 2019-07-12, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010GL046115` doi: 10.1029/2010GL046115

Deuss, A., Ritsema, J., & van Heijst, H. (2013, May). A new catalogue of normal-mode splitting function measurements up to 10 mHz. *Geophysical Journal International*, *193*(2), 920–937. Retrieved 2019-07-12, from `https://academic.oup.com/gji/article/193/2/920/633769` doi: 10.1093/gji/ggt010

Dziewonski, A. M., Lekic, V., & Romanowicz, B. A. (2010, October). Mantle Anchor Structure: An argument for bottom up tectonics. *Earth and Planetary Science Letters*, *299*(1-2), 69–79. doi: 10.1016/j.epsl.2010.08.013

Forte, A. M. (1989). *Mantle Convection and Global Geophysical Observables* (Doctoral Thesis). University of Toronto, Toronto, Canada.

Forte, A. M., Dziewonski, A. M., & Woodward, R. L. (1993). Aspherical Structure of the Mantle, Tectonic Plate Motions, Nonhydrostatic Geoid, and Topography of the Core-Mantle Boundary. In *Dynamics of Earth's Deep Interior and Earth Rotation* (pp. 135–166). American Geophysical Union (AGU). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GM072p0135` doi: 10.1029/GM072p0135

Forte, A. M., & Peltier, W. R. (1987, April). Plate tectonics and aspherical earth structure: The Importance of poloidal-toroidal coupling. *J. Geophys. Res.*, *92*(B5), 3645–3679. doi: 10.1029/JB092iB05p03645

French, S. W., & Romanowicz, B. A. (2014, December). Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophysical Journal International*, *199*(3), 1303–1327. doi: 10.1093/gji/ggu334

Fukao, Y., Obayashi, M., & Nakakuki, T. (2009, May). Stagnant Slab: A Review. *Annual Review of Earth and Planetary Sciences*, *37*(1), 19–46.

Ghosh, A., Becker, T. W., & Zhong, S. J. (2010, January). Effects of lateral viscosity variations on the geoid. *Geophysical Research Letters*, *37*(1), 01301. doi: 10.1029/2009GL040426

Girard, J., Amulele, G., Farla, R., Mohiuddin, A., & Karato, S.-i. (2016, January). Shear deformation of bridgmanite and magnesiowstite aggregates at lower mantle conditions. *Science*, *351*(6269), 144–147. Retrieved 2020-01-13, from https://science.sciencemag.org/content/351/6269/144 doi: 10.1126/science.aad3113

Grand, S. P. (2002, November). Mantle shearwave tomography and the fate of subducted slabs. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *360*(1800), 2475–2491. doi: 10.1098/rsta.2002.1077

Green, P. J. (1995, December). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, *82*(4), 711–732. doi: 10.1093/biomet/82.4.711

Gu, Y. J., Dziewonski, A. M., Su, W., & Ekstrm, G. (2001, June). Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities. *J. Geophys. Res.*, *106*(B6), 11169–11199. doi: 10.1029/2001JB000340

Gu, Y. J., Dziewoski, A. M., & Ekstrm, G. (2003, August). Simultaneous inversion for mantle shear velocity and topography of transition zone discontinuities. *Geophysical Journal International*, *154*(2), 559–583. Retrieved 2019-07-27, from https://academic.oup.com/gji/article/154/2/559/610947 doi: 10.1046/j.1365-246X.2003.01967.x

Hager, B. H. (1984, July). Subducted slabs and the geoid: Constraints on mantle rheology and flow. *J. Geophys. Res.*, *89*(B7), 6003–6015. doi: 10.1029/JB089iB07p06003

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

Ishii, M., & Tromp, J. (2001, April). Even-degree lateral variations in the Earth's mantle constrained by free oscillations and the free-air gravity anomaly. *Geophysical Journal of the Royal Astronomical Society*, *145*(1), 77–96. doi: 10.1111/j.1365-246X.2001.00385.x

Kaercher, P., Miyagi, L., Kanitpanyacharoen, W., Zepeda-Alarcon, E., Wang, Y., Parkinson, D., … Wenk, H.-R. (2016, December). Two-phase deformation of lower mantle mineral analogs. *Earth and Planetary Science Letters*, *456*, 134–145. doi: 10.1016/j.epsl.2016.09.030

Karato, S.-i. (1993). Importance of anelasticity in the interpretation of seismic tomography. *Geophysical Research Letters*, *20*, 1623–1626.

Kaula, W. M. (1966). Tests and combination of satellite determinations of the gravity field with gravimetry. *Journal of Geophysical Research (1896-1977)*, *71*(22), 5303–5314. Retrieved 2019-07-25, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ071i022p05303` doi: 10.1029/JZ071i022p05303

Kido, M., Yuen, D. A., Čadek, O., & Nakakuki, T. (1998). Mantle viscosity derived by genetic algorithm using oceanic geoid and seismic tomography for whole-mantle versus blocked-flow situations. *Physics of the Earth and Planetary Interiors*, *107*(4), 307–326.

King, S. D. (1995, December). Radial models of mantle viscosity: results from a genetic algorithm. *Geophysical Journal International*, *122*(3), 725–734. Retrieved 2020-07-14, from `https://academic.oup.com/gji/article/122/3/725/716918` (Publisher: Oxford Academic) doi: 10.1111/j.1365-246X.1995.tb06831.x

King, S. D., & Masters, G. (1992, August). An inversion for radial viscosity structure using seismic tomography. *Geophysical Research Letters*, *19*(15), 1551–1554. doi: 10.1029/92GL01700

Koelemeijer, P., Deuss, A., & Ritsema, J. (2017). Density structure of Earths lowermost mantle from Stoneley mode splitting observations. *Nature Communications*, *8*, 15241. doi: 10.1038/ncomms15241https://www.nature.com/articles/ncomms15241#supplementary-information

Kolb, J. M., & Lekic, V. (2014, June). Receiver function deconvolution using trans-dimensional hierarchical Bayesian inference. *Geophysical Journal International*, *197*(3), 1719–1735. doi: 10.1093/gji/ggu079

Kuo, C., & Romanowicz, B. (2002, July). On the resolution of density anomalies in the Earth's mantle using spectral fitting of normal-mode data. *Geophysical Journal International*, *150*(1), 162–179. Retrieved 2019-07-11, from `https://academic.oup.com/gji/article/150/1/162/591652` doi: 10.1046/j.1365-246X.2002.01698.x

Kustowski, B., Ekstrm, G., & Dziewoski, A. M. (2008). Anisotropic shear-wave velocity structure of the Earth's mantle: A global model. *Journal of Geophysical Research: Solid Earth*, *113*(B6). Retrieved 2019-07-27, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JB005169` doi: 10.1029/2007JB005169

Lau, H. C. P., Mitrovica, J. X., Davis, J. L., Tromp, J., Yang, H.-Y., & Al-Attar, D. (2017). Tidal tomography constrains Earths deep-mantle buoyancy. *Nature*, *551*, 321. doi: 10.1038/nature24452

Lourenç, D., & Rudolph, M. (2020). Shallow lower mantle viscosity modulates the pattern of mantle structure. *Geochemistry, Geophysics, Geosystems*, *21*(9), e2020GC008934.

Mao, W., & Zhong, S. (2018, November). Slab stagnation due to a reduced viscosity layer beneath the mantle transition zone. *Nature Geoscience*, *11*(11), 876. Retrieved 2019-06-30, from `https://www.nature.com/articles/s41561-018-0225-2` doi: 10.1038/s41561-018-0225-2

Marquardt, H., & Miyagi, L. (2015, March). Slab stagnation in the shallow lower mantle linked to an increase in mantle viscosity. *Nature Geoscience*, *8*(4), 311–314. doi: 10.1038/ngeo2393

Masters, G., Laske, G., Bolton, H., & Dziewonski, A. (2000). The Relative Behavior of Shear Velocity, Bulk Sound Speed, and Compressional Velocity in the Mantle: Implications for Chemical and Thermal Structure. In S.-I. Karato, A. M. Forte, R. Lieberman, G. Masters, & L. Stixrude (Eds.), (Vol. 117, pp. 63–87). Washington, D. C.: American Geophysical Union.

Mitrovica, J. X., & Forte, A. M. (1997). Radial profile of mantle viscosity: Results from the joint inversion of convection and postglacial rebound observables. *Journal of Geophysical Research: Solid Earth*, *102*(B2), 2751–2769. Retrieved 2019-07-27, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/`

10.1029/96JB03175 doi: 10.1029/96JB03175

Mitrovica, J. X., & Forte, A. M. (2004, August). A new inference of mantle viscosity based upon joint inversion of convection and glacial isostatic adjustment data. *Earth and Planetary Science Letters*, *225*(12), 177–189. doi: 10.1016/j.epsl.2004.06.005

Mitrovica, J. X., & Peltier, W. R. (1993). A new formalism for inferring mantle viscosity based on estimates of post glacial decay Times: Application to RSL variations in N.E. Hudson Bay. *Geophysical Research Letters*, *20*(20), 2183–2186. Retrieved 2020-07-14, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/93GL02136 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/93GL02136) doi: 10.1029/93GL02136

Montagner, J., & Anderson, D. L. (1989). Petrological constraints on seismic anisotropy. *Physics of the Earth and Planetary Interiors*, *54*, 82–105.

Moucha, R., Forte, A. M., Mitrovica, J. X., & Daradich, A. (2007, April). Lateral variations in mantle rheology: implications for convection related surface observables and inferred viscosity models. *Geophysical Journal of the Royal Astronomical Society*, *169*(1), 113–135. doi: 10.1111/j.1365-246X.2006.03225.x

Moulik, P. (2016). *Earth's elastic and density structure from diverse seismological observations* (Unpublished doctoral dissertation). Columbia Univ., New York, NY

Moulik, P., & Ekström, G. (2014, December). An anisotropic shear velocity model of the earth's mantle using normal modes, body waves, surface waves and long-period waveforms. *Geophysical Journal International*, *199*(3), 1713–1738. doi: 10.1093/gji/ggu356

Moulik, P., & Ekström, G. (2016, April). The relationships between largescale variations in shear velocity, density, and compressional velocity in the earth's mantle. *Journal of Geophysical Research (Solid Earth)*, *121*(4), 2737–2771. doi: 10.1002/2015JB012679

Ni, S., Tan, E., Gurnis, M., & Helmberger, D. (2002, January). Sharp Sides to the African Superplume. *Science*, *296*(5574), 1850–1852. doi: 10.1126/science.1070698

Panasyuk, S. V., & Hager, B. H. (1998, June). A model of transformational superplasticity in the upper mantle. *Geophysical Journal International*, *133*(3), 741–755. Retrieved 2019-07-25, from https://academic.oup.com/gji/article/133/3/741/601224 doi: 10.1046/j.1365-246X.1998.00539.x

Paulson, A., Zhong, S., & Wahr, J. (2007a, November). Inference of mantle viscosity from GRACE and relative sea level data. *Geophysical Journal International*, *171*(2), 497–508. doi: 10.1111/j.1365-246X.2007.03556.x

Paulson, A., Zhong, S., & Wahr, J. (2007b, March). Limitations on the inversion for mantle viscosity from postglacial rebound. *Geophysical Journal International*, *168*(3), 1195–1209. doi: 10.1111/j.1365-246X.2006.03222.x

Resovsky, J. S., & Ritzwoller, M. H. (1995). Constraining odd-degree earth structure with coupled free-oscillations. *Geophysical Research Letters*, 2301–2304. Retrieved 2019-07-25, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95GL01996%4010.1002/%28ISSN%291944-8007.BOLIVEQ1 doi: 10.1029/95GL01996@10.1002/(ISSN)1944-8007.BOLIVEQ1

Resovsky, J. S., & Ritzwoller, M. H. (1999). Regularization uncertainty in density models estimated from normal mode data. *Geophysical Research Letters*, *26*(15), 2319–2322. Retrieved 2019-07-25, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999GL900540 doi: 10.1029/1999GL900540

Ricard, Y., Vigny, C., & Froidevaux, C. (1989, October). Mantle heterogeneities, geoid, and plate motion: A Monte Carlo inversion. *J. Geophys. Res.*, *94*(B10), 13739–13754. doi: 10.1029/JB094iB10p13739

Richards, M. A., & Hager, B. H. (1984). Geoid anomalies in a dynamic Earth. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5987–6002. Retrieved 2019-07-19, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB089iB07p05987 doi: 10.1029/JB089iB07p05987

Ries, J., Bettadpur, S., Eanes, R., Kang, Z., Ko, U., McCullough, C., ... Tapley, B. (2016). *Development and Evaluation of the Global Gravity Model GGM05* (Tech. Rep. No. CSR-16-02). Center for Space Research: The University of Texas at Austin.

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

Ritsema, J., & Lekic, V. (2020). Heterogeneity of Seismic Wave Velocity in Earth's Mantle. *Annual Review of Earth and Planetary Sciences*, *48*(1), 377–401. Retrieved 2020-07-30, from https://doi.org/10.1146/annurev-earth-082119-065909 (_eprint: https://doi.org/10.1146/annurev-earth-082119-065909) doi: 10.1146/annurev-earth-082119-065909

Ritsema, J., & Van Heijst, H. (1999, December). Seismic imaging of structural heterogeneity in Earth's mantle: evidence for large-scale mantle flow. *Science progress*, *83 (3)*, 243–259.

Robertson, G. S., & Woodhouse, J. H. (1996). Ratio of relative S to P velocity heterogeneity in the lower mantle. *Journal of Geophysical Research*, *101*(B9), 20041–20052.

Rost, S., Garnero, E. J., Williams, Q., & Manga, M. (2005, June). Seismological constraints on a possible plume root at the core–mantle boundary. *Nature*, *435*(7042), 666–669.

Rudolph, M. L., Lekic, V., & Lithgow-Bertelloni, C. (2015, December). Viscosity jump in Earths mid-mantle. *Science*, *350*(6266), 1349–1352. doi: 10.1126/science.aad1929

Sambridge, M. (2014, January). A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization. *Geophysical Journal International*, *196*(1), 357–374. Retrieved 2019-07-05, from https://academic.oup.com/gji/article/196/1/357/585739 doi: 10.1093/gji/ggt342

Sambridge, M., Bodin, T., Gallagher, K., & Tkalčić, H. (2013). Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1984), 20110547.

Shim, S.-H., Grocholski, B., Ye, Y., Alp, E. E., Xu, S., Morgan, D., … Prakapenka, V. B. (2017, June). Stability of ferrous-iron-rich bridgmanite under reducing midmantle conditions. *Proceedings of the National Academy of Sciences*, *114*(25), 6468–6473. Retrieved 2019-07-27, from https://www.pnas.org/content/114/25/6468 doi: 10.1073/pnas.1614036114

Simmons, N. A., Forte, A. M., Boschi, L., & Grand, S. P. (2010, December). GyPSuM: A joint tomographic model of mantle density and seismic wave speeds. *J. Geophys. Res.*, *115*(B12), B12310. doi: 10.1029/2010JB007631

Simmons, N. A., Forte, A. M., & Grand, S. P. (2009, June). Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity: Implications for the relative importance of thermal versus compositional heterogeneity. *Geophysical Journal International*, *177*(3), 1284–1304.

Simons, F., Dahlen, F., & Wieczorek, M. (2006, January). Spatiospectral Concentration on a Sphere. *SIAM Review*, *48*(3), 504–536. Retrieved 2019-09-08, from `https://epubs.siam.org/doi/10.1137/S0036144504445765` doi: 10.1137/S0036144504445765

Solomatov, V. S., & Reese, C. C. (2008). Grain size variations in the Earth's mantle and the evolution of primordial chemical heterogeneities. *Journal of Geophysical Research: Solid Earth*, *113*(B7). Retrieved 2019-10-02, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JB005319` doi: 10.1029/2007JB005319

Steinberger, B., & Holme, R. (2008). Mantle flow models with core-mantle boundary constraints and chemical heterogeneities in the lowermost mantle. *Journal of Geophysical Research: Solid Earth*, *113*(B5). Retrieved 2019-07-10, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JB005080` doi: 10.1029/2007JB005080

Su, W.-j., & Dziewonski, A. M. (1997, March). Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle. *Physics of the Earth and Planetary Interiors*, *100*(14), 135–156.

Tarantola, A. (2005). *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation.* Philadelphia: Society for Industrial and Applied Mathematics.

Thorne, M. S., & Garnero, E. J. (2004, August). Inferences on ultralow-velocity zone structure from a global analysis of SPdKS waves. *Journal of Geophysical Research: Solid Earth*, *109*(B8), 421.

Torsvik, T. H., Van der Voo, R., Doubrovine, P. V., Burke, K., Steinberger, B., Ashwal, L. D., ... Bull, A. L. (2014, June). Deep mantle structure as a reference frame for movements in and on the Earth. *Proceedings of the National Academy of Sciences*.

Wang, Y., & Wen, L. (2007, January). Geometry and P and S velocity structure of the African Anomaly. *J. Geophys. Res.*, *112*(B5), B05313. doi: 10.1029/

2006JB004483

Woodhouse, J. H., & Dziewoński, A. M. (1984). Mapping the Upper Mantle: Three-Dimensional Modeling of Earth Structure by Inversion of Seismic Waveforms. *Journal of Geophysical Research*, *89*, 5953–5986.

Yang, T., & Gurnis, M. (2016, November). Dynamic topography, gravity and the role of lateral viscosity variations from inversion of global mantle flow. *Geophysical Journal International*, *207*(2), 1186–1202. Retrieved 2020-07-20, from https://academic.oup.com/gji/article/207/2/1186/2583789 (Publisher: Oxford Academic) doi: 10.1093/gji/ggw335

Zhong, S., & Davies, G. F. (1999). Effects of plate and slab viscosities on the geoid. *Earth and Planetary Science Letters*, *170*(4), 487–496.