

Chapter 3

Constraining Dynamic Properties of Mantle²

3.1 Need for Assimilation of Time-Dependent Data in Real Problems

Unlike atmospheric or oceanographic problems where many parameters within the interior of the flow can be measured and calibrated *in situ*, dynamic parameters for the mantle convection problem are obtained indirectly. A good example of such a parameter is the depth dependence of mantle viscosity, unfortunately a parameter that remains uncertain [Walcott, 1973; Hager and Clayton, 1989; Steinberger and O'Connell, 2000; Mitrovica and Forte, 2004]. This of course is problematic for the adjoint problem described in Chapter 1, because what viscosity should be used for the recovery of initial conditions? Another critical parameter for recovery is the magnitude of the temperature (density) within the anomalies. Clearly, important constraints can be placed on this problem from seismic tomography, but high-pressure, high-temperature laboratory experiments have not achieved the ability to uniquely map seismic into thermal anomalies. Thus, even for simple convection models, we should consider these basic model parameters to have uncertainty when the adjoint method is used to infer initial conditions.

As discussed in Chapter 1, a quantitative description of the time dependence of mantle flow requires time-dependent constraints. Here I will explore the surface dynamic

² This chapter is based on: 1) Liu L. and M. Gurnis (2008), *Simultaneous Inversion of Mantle Properties and Initial Conditions Using an Adjoint of Mantle Convection*, J. of Geophys. Res., 113, B08405, doi:10.1029/2008JB005594.

topography, a different type of observation from plate motions used in earlier studies [Lithgow-Bertelloni and Richards, 1998]. With the adjoint method implemented in *CitcomS*, we can compute the prior history of thermal anomalies for a given viscosity model and present-day mantle thermal structure. From the restored history, we then predict dynamic topography that can be constrained through stratigraphic constraints, such as tectonic subsidence from boreholes [Heller *et al.*, 1982; Pang and Nummedal, 1995], paleo shorelines [Bond, 1979; Sandiford, 2007; DiCaprio *et al.*, 2009], and sediment isopachs [Cross and Pilger, 1978]. Given these additional observational constraints, there is the opportunity to place limits on mantle viscosity and temperatures.

3.2 Dynamic Topography Constrains Uncertain Mantle Properties

For this study, we designed two sets of synthetic experiments, one for a simple one-layer mantle with uniform viscosity, and the other for a two-layer mantle viscosity structure.

3.2.1 One-Layer Mantle

First let us consider a mantle with a uniform viscosity throughout. On the top surface of the convection model, dynamic topography, h , is defined from

$$\sigma_{r,r} = \rho_m g h \quad (12)$$

where $\sigma_{r,r}$ is the total normal stress in the radial direction and ρ_m is the density contrast across the top surface (implicitly assuming that air overlies the solid mantle). At any instant of time, normal stress $\sigma_{r,r}$ is proportional to the temperature scaling ΔT (see Eq. 3). For an inverse problem where we use the present-day seismic tomography to interpret mantle temperature structure, ΔT is the temperature magnitude obtained by mapping seismic velocity variations to thermal anomalies. Together with Eq. (12), we relate dynamic topography with a temperature scaling via a time-dependent coefficient ζ with units, m/K . The quantity ζ describes the response of surface dynamic topography with a scaled temperature distribution and mantle rheology structure.

$$h(t) = \zeta(t) \Delta T \quad (13)$$

The rate of change of dynamic topography \dot{h} , however, is related to the absolute viscosity of the mantle when the geometry of the thermal anomaly remains fixed [Gurnis *et al.*, 2000]. Specifically, in the energy equation (Eq. 4), the rate of change of the temperature anomaly

ΔT is proportional to the product of ΔT and mantle flow speed \bar{u} (i.e., $\Delta \dot{T} \propto \bar{u} \cdot \Delta T$). In a Stokes fluid, \bar{u} is proportional to ΔT and inversely proportional to mantle viscosity η .

Considering Eq. (13), we obtain

$$\dot{h}(t) = \frac{\xi'(t)}{\eta} \Delta T^2 \quad (14)$$

For an inverse problem where ΔT and η are unknowns, and $h(t)$ and $\dot{h}(t)$ are data constraints, we simplify the problem by rewriting Eq. (14) with Eq. (13)

$$\dot{h}(t) = \frac{\xi(t)}{\eta} h_1^2 \quad (15)$$

where $h_1 = h(t_1)$, with t_1 representing present-day time; ξ (or ζ) is a kernel that describes the response of the rate of change of surface dynamic topography assuming a specific mantle viscosity; ξ has units of Pa/m . Instantaneously, when the temperature and viscosity structures remained unchanged, Eq. (15) was validated numerically for systems with temperature- and depth-dependent viscosities [Gurnis *et al.*, 2000].

Because $h(t)$ and $\dot{h}(t)$ are potentially two independent constraints, and Eq. (13) and (15) each has an independent unknown, ΔT and η , respectively, the independent unknowns might be recoverable. By using h_1 instead of $h(t)$ on the right-hand side of Eq. (14), we attempt to partially decouple this two-variable, two-constraint system. Essentially, we use the magnitude of topography $h(t)$ to constrain ΔT (Eq. 13), and use its rate of change $\dot{h}(t)$ to constrain η (Eq. 15).

The left-hand sides of equations (13) and (15) are time dependent. On the right-hand-side of the equations, the time dependence has been collapsed into the two kernels, ζ and ξ ,

which are evaluated numerically. At any moment of time, ζ and ξ can be found from the solution of Eq. (2) – (4) and are dependent on the viscosity and temperature distribution. Evaluation of ξ requires two successive solutions of Eq. (2) – (4) so that $\dot{h}(t)$ can be found.

Assuming the “structure” of the present-day temperature field is the same as the structure obtained from seismic tomography, we now show how Eq. (13) and Eq. (15) can be incorporated within an iterative scheme to solve for the unknowns ΔT and η based on observed and predicted $h(t)$ and $\dot{h}(t)$. Define j to be the index of a loop used to refine temperature and viscosity, while i remains the index over time as it was in the forward-adjoint looping (Sec. 3). At any given time i in loop j , the numerical values of the two kernels ζ and ξ are computed as $\zeta_{ij} = h_i / \Delta T^j$, $\xi_{ij} = \dot{h}_i \cdot \eta^j / (h_i)^2$, respectively. Here we treat two kernels as implicit Green’s functions. ΔT and η are updated by a method that is similar to back-projection used in seismic topography [Rowlinson and Sambridge, 2003], the difference being the use of implicit coefficients (ζ and ξ) in this case.

$$\Delta T^{j+1} = \Delta T^j + \frac{\gamma_1}{n} \sum_{i=1}^n \zeta_{ij} [h_{di} - h_i] \quad (16)$$

$$\eta^{j+1} = \eta^j + \frac{\gamma_2}{m} \sum_{i=1}^m \xi_{ij} \left[\frac{h_{d1}^2}{\dot{h}_{di}} - \frac{h_1^2}{\dot{h}_i} \right] \quad (17)$$

where m and n are the numbers of sample points within the time series and are potentially different because of the different number of constraints on topography and its rate of change; subscript d refers to data (observational constraints); γ_1 and γ_2 are two damping factors with values ≤ 0.5 .

This iteration is at a higher level than that of forward-adjoint looping and we refer to it as the outer iteration. Essentially, seismic tomography at the present day is used to constrain the geometry or depth distribution but not the precise amplitude of mantle temperature anomalies, and the forward-adjoint looping is used to find that geometry during earlier times. The outer looping is used to refine both the scaling between seismic velocity variations and temperatures (or density) and the viscosity distribution. The whole procedure is divided into two parts:

(i) Inner loop: While ΔT and η (without varying temperature dependence) remain constant, perform an adjoint calculation to recover the initial condition with the SBI first guess, and predict the dynamic topography during the final iteration.

(ii) Outer loop: Update ΔT and η via (16) and (17) through the mismatch of the predicted and target dynamic topography and its rate of change.

The whole procedure is terminated upon convergence of the two model parameters.

In a synthetic experiment, a cold spherical anomaly sinks from top to bottom of the mantle within a 3D spherical region; the system has initial (Fig. 7A) and final reference states (Fig. 7B). On the top surface, topography is sampled directly above the center of the anomaly.

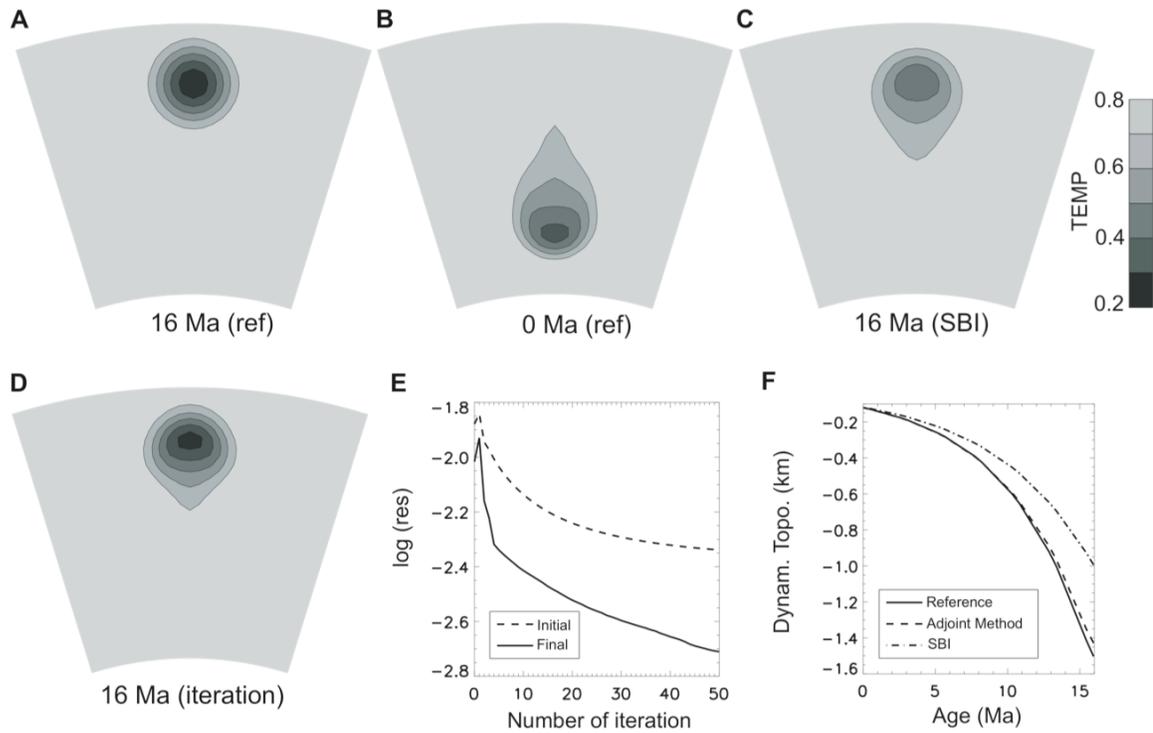


Figure 7 3D models with a single viscosity layer (modeled with a $33 \times 33 \times 33$ mesh). Reference thermal states at 16 Ma (A) and the present (B). (C) First guessed initial condition with a simple backward integration (SBI). (D) Recovered initial condition with the adjoint method after 50 iterations. (E) RMS residuals for the initial and final states based on the adjoint method. (F) The predicted dynamic topography histories based on an SBI first guess and the adjoint method, compared against the reference dynamic topography. All calculations assume a known viscosity structure.

To illustrate the effect of forward-adjoint iteration on dynamic topography, we ran the inner loop described above assuming that the temperature scaling and the absolute value of viscosity are known. The SBI initial guess (Fig. 7C) is more diffused in comparison to the finally recovered initial condition after 50 iterations (Fig. 7D). The adjoint method reduces the RMS residuals for the initial and final states by about a factor of 3 and 5, respectively

(Fig. 7E). Consequently, the associated dynamic topography curves from t_0 to t_1 are also notably different (Fig. 7F). The curve from the SBI case deviates from the reference much more than the one from the recovered solution, with a maximum deviation in magnitude by 35% vs. 5% of the reference value at 16 Ma. Although the SBI is a good method to find the best first guess for the forward-adjoint looping, the experiment demonstrates that the simple backward advection of the anomaly (SBI) does not perfectly predict the evolution of dynamic topography.

We then started the outer loop with two initial models (Cases AH1 and AH2) in which the temperature scaling ΔT and mantle viscosity η had “guessed values” that were different from the reference ones. The initially guessed parameters of Case AH1 (Fig. 8A, B at loop 1) were such that its effective Rayleigh number was equal to the actual Ra for the reference state while Case AH2 (Fig. 8E, F at loop 1) had an effective Rayleigh number four times smaller. In both cases, we applied the two-level looping algorithm. The inner loop was applied so that the iteration always started with the SBI first guess, and the number of forward-adjoint iterations increased as the index of the outer loop increased. We applied this simplification because the first recovered initial condition was not well known before the constraints on $h(t)$ and $\dot{h}(t)$ were applied. Due to the initially under-estimated temperature scaling ΔT in both AH1 and AH2, the first predicted temporal dynamic topography curves had small magnitudes and slopes. By applying the outer loop upon the predicted and reference dynamic topography (Fig. 8C, G), we updated model parameters ΔT and η . The difference in magnitudes of topography $h(t)$ forced ΔT to increase in both cases where AH2 increased faster due to a larger difference. Differences in slope $\dot{h}(t)$ normalized by present-

day magnitude of topography h_I updated the viscosity. The apparently smaller slope in AH1 was actually larger than that of the reference when normalized by h_I , and hence forced the viscosity to increase. The initial smaller slope for dynamic topography in AH2 forced the viscosity to decrease, and the smaller magnitude forced temperature to increase, overshooting the reference temperature. The overshoot was corrected as the viscosity also approached the true value. As a result, for both Case AH1 and AH2, the temporal (Fig. 8C, G) and spatial (Fig. 8D, H) distribution of dynamic topography converged to the target curves as the two incorrectly guessed model parameters converged to the reference values after a finite number of loops (Fig. 8A, B and E, F). Most of the model corrections occurred within the first 10 outer loops.

As discussed in Chapter 1, due to the artificially defined initial condition and low resolution of meshing, the recovered initial condition by the adjoint method is not exact, even with the same model that generates the reference states (Fig. 7D). This effect shows up in the recovered model parameters as a deviation of viscosity from the reference value by about -2% and that of the temperature scaling by about 1% . However, the final solutions in both Case AH1 and AH2 are almost identical, indicating the two-level algorithm can both recover initial conditions and unknown material properties.

Under highly controlled set of circumstances, this test shows that the history of the dynamic topography is a valuable constraint on mantle viscosity and magnitude of present day mantle thermal structures. We will then explore the limitations of this conclusion under more realistic conditions.

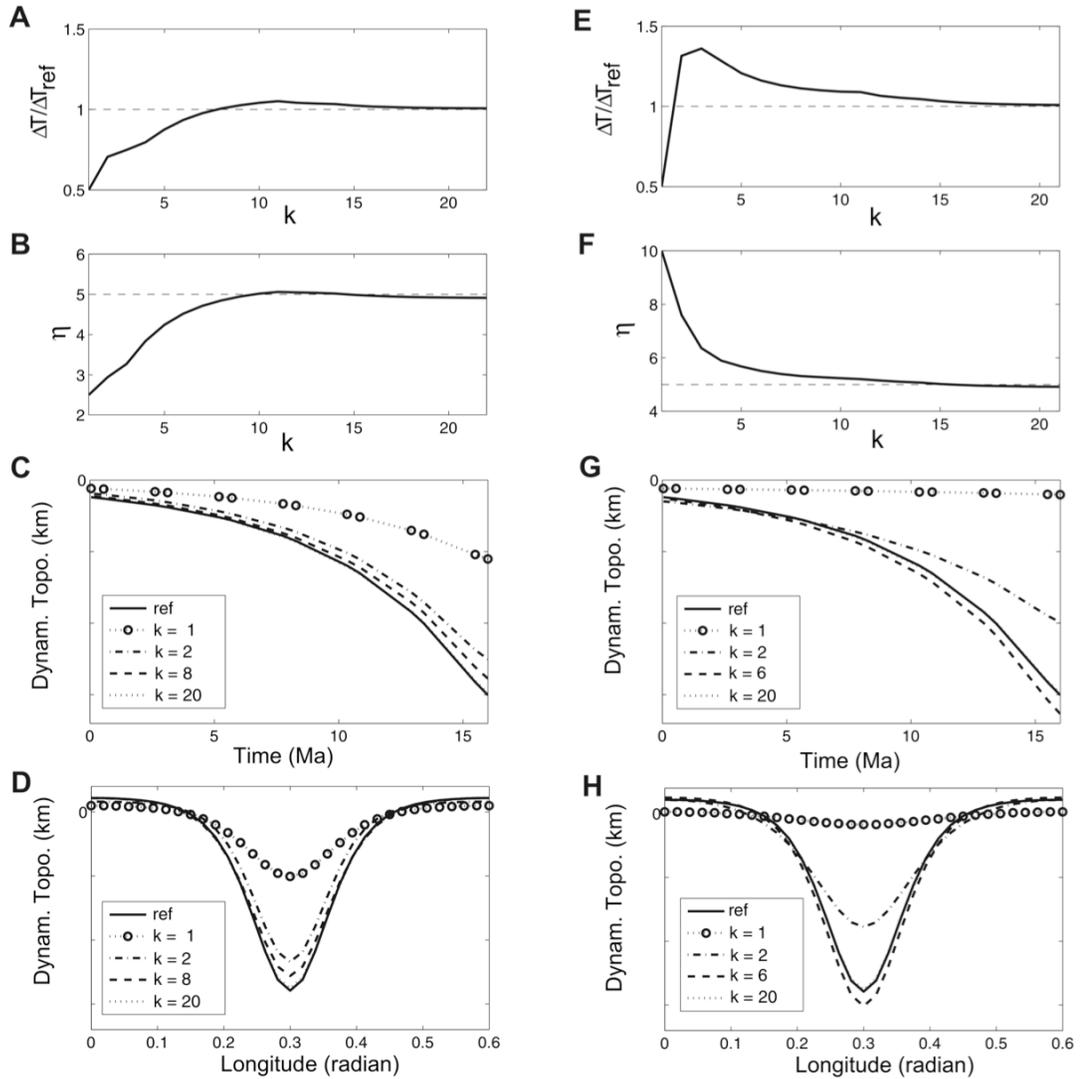


Figure 8 Recovery of model parameters using dynamic topography for models with a single layer. The starting model has either the same effective Rayleigh number (A–D, Case AH1) or a Rayleigh number four times smaller (E–H, Case AH2) than the reference value. All results plotted with respect to the outer loop index (k) and are computed from the last iteration of the inner (forward-adjoint) looping. Shown from top to bottom are the temperature scaling (A, E), viscosity (B, F), temporal evolution of dynamic topography recorded at one point on the surface (C, G), and latitudinal profile of dynamic topography across the center of the surface at 16 Ma (D, H).

3.2.2 Two-Layer Mantle

We now consider the geophysically more relevant possibility of a layered mantle viscosity [Hager, 1984; Mitrovica and Forte, 1997]. We used a two-layer mantle and attempted to recover three variables: ΔT (present-day temperature anomaly), η_{um} (upper mantle viscosity), and η_{lm} (lower mantle viscosity). Given this potentially underdetermined problem, we determine what we might hope to recover.

For a thermal anomaly within the upper mantle, the upper mantle viscosity controls the flow velocity, \bar{u} , and the evolving dynamic topography. Assuming that the density anomaly has not yet entered the lower mantle the system has only two variables, ΔT and η_{um} , just like the isoviscous mantle discussed above. This assumption is not entirely true since the h does depend on the ratio of upper to lower mantle viscosity [Richards and Hager, 1984; Hager, 1984]. Approximately, we still have the linear relation between $h(t)$ and ΔT , Eq. (13), and the following relation for $\dot{h}(t)$, which is similar to Eq. (14)

$$\dot{h}(t) \approx \frac{\xi'(t)}{\eta_{um}} \Delta T^2 \quad (18)$$

For a density anomaly within the lower mantle, the flow speed \bar{u} is approximately inversely proportional to η_{lm} , and the surface normal stress that defines $h(t)$ is proportional to η_{um} . So Eq. (13) becomes

$$h(t) = \zeta(t) R_\eta \Delta T \quad (19)$$

where $R_\eta = \eta_{um} / \eta_{lm}$, is the ratio of upper to lower mantle viscosity. Consider a static situation in which we neglect time dependence, Eq. (19) is nonlinear due to the fact that $\zeta =$

$\zeta(R_\eta)$, while Eq. (13) is linear since ζ is not a function of ΔT . This implies the inverse problem for a two-layer viscosity mantle is more nonlinear than for a single layer mantle.

The relation $\Delta \dot{T} \propto \bar{u} \cdot \Delta T$, together with Eq. (19), leads to the expression

$$\dot{h}(t) = \frac{\xi(t)}{\eta_{um}} h_1^2 \quad (20)$$

where h_1 is dynamic topography at the present day. Again, we use h_1 instead of $h(t)$ on the right-hand side of Eq. (20) to avoid the sharing of data constraints. In fact, Eq. (18) and (20) are equivalent: replace ΔT with h_1 in Eq. (18) and Eq. (20) is obtained. This shows that the rate of change of dynamic topography should be a good constraint on the upper mantle viscosity.

Rearranging and discretizing Eq. (20) lead to

$$\eta_{um}^{j+1} = \eta_{um}^j + \frac{\gamma_3}{m} \sum_{i=1}^m v_i \xi_{ij} \left[\frac{h_{D1}^2}{\dot{h}_{Di}} - \frac{h_1^2}{\dot{h}_i} \right] \quad (21)$$

For the other two variables, ΔT and η_{lm} , we have constraint equations (13) and (19). Ideally, we could use Eq. (13) to constrain ΔT by assimilating topographic data associated with density anomalies crossing the upper mantle through Eq. (16). Equation (19) could be used to constrain η_{lm} by topographic data with lower mantle anomalies iteratively

$$\eta_{lm}^{j+1} = \eta_{lm}^j + \frac{\gamma_4}{l} \sum_{i=1}^l w_i \frac{\Delta T^{j+1} \eta_{um}^{j+1}}{\zeta_{ij}} \left[\frac{1}{h_{Di}} - \frac{1}{h_i} \right] \quad (22)$$

In Eq. (21) and (22), m and l are the numbers of sample points within the time series; γ_3 and γ_4 are two damping factors with values $< = 0.5$; two kernels $\zeta_{ij} = R_\eta^j \Delta T^j / h_i$ and

$\xi_{ij} = \dot{h}_i \cdot \eta^j / (h_i)^2$; v_i and w_i are weighting functions that change with time. We assumed that v_i (w_i) decreases (increases) linearly from $i = 1$ to l .

However, because a thermal anomaly will move faster in the upper mantle than in the lower mantle, a topographic history would be more heavily weighted in time for the lower mantle, where ΔT largely trades off with η_{lm} when using the dynamic topography (see Eq. 19). In other words, temperature anomaly and lower mantle viscosity are coupled for most of the topographic record.

Therefore, in order to simultaneously invert for all three variables, we should avoid the trade-off between temperature scaling and lower mantle viscosity. We designed a three-level iterative scheme which solves for all three parameters while minimizing potential trade-offs between them:

(i) Inner level: While ΔT , η_{um} and η_{lm} remain fixed, perform forward-adjoint looping to recover the initial condition.

(ii) Middle level: While ΔT remains fixed, update η_{um} and η_{lm} via Eq. (21) and (22) through the mismatch of the predicted and target dynamic topography and its rate of change.

(iii) Outer level: Update ΔT according to Eq. (16).

The whole procedure is terminated upon convergence of the three model parameters.

For an explicit example, we consider a 2D model that simulates a subduction scenario, where a fragment of a cold slab sinks from the upper mantle into the low mantle over a period of 36 Myr (Fig. 9A, B). The dynamic topography is recorded at the surface at some point on the “overriding” plate. We assume that the top surface is fixed; if the plates are moving then the observational record of topography is in the plate frame [Gurnis *et al.*,

1998] and using the mismatch between observed and predicted quantities will be more involved than what the experiment given below suggests.

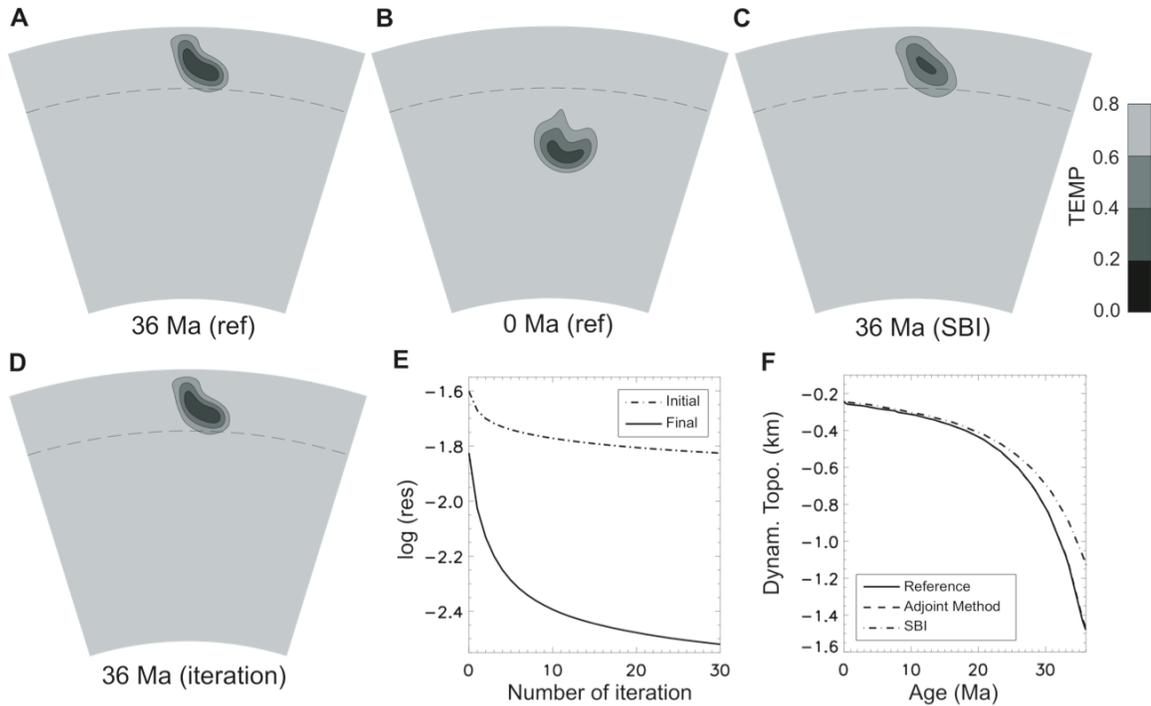


Figure 9 Same as Fig. 7 except for 2D models (on a 129×129 mesh) with a two-layer viscosity. The dashed lines (A–D) indicate the upper and lower mantle interface.

To avoid numerical artifacts, we generated the initial condition by first defining a smooth slab on the surface and then allowing the slab to sink to the position shown in Fig. 7a. A fine resolution mesh with a 129×129 grid is used, to mimic the trench-normal cross-section. We compared the SBI first guess (Fig. 9C) and the recovered solution (Fig. 9D) through the same model that generates the reference states, with residual curves shown in Fig. 9E. The SBI first guess is visually diffused while the recovered solution is nearly

identical to the reference initial state. Moreover, the dynamic topography associated with the SBI estimate deviates from the reference by about 20% at 36 Ma while that with the recovered solution is less than 1%. This indicates that the recovered initial condition with simple forward-adjoint looping is almost perfect if the viscosity and temperature scaling are known *a priori*.

Since the inner level involving the forward-adjoint looping has been described in Section 2.2.1, we focus our discussion on the middle and outer levels. For the middle level, we show several cases with different ΔT values, where upper and lower mantle viscosities are recovered from several initial guesses.

In a set of experiments, we chose ΔT at its reference value but incorrectly guessed both viscosities. We tried two starting viscosity models, AH3 and AH4, that were both guessed to be isoviscous with $(\eta_{um}, \eta_{lm}) = (5, 5)$ and $(20, 20)$, respectively, while the target had a layered viscosity, $(\eta_{um}, \eta_{lm}) = (1, 10)$ (Fig. 10A, C). Because the initial upper/lower mantle viscosity ratio was overestimated in both models, Eq. (19) implies that the present-day dynamic topography should be overpredicted, as verified as loop 1 in Fig. 10B, D. Since η_{lm} was controlled by the magnitude of topography during the later part of its evolution, the over predicted magnitude of h caused η_{lm} to increase (Fig. 10A, C). Since the upper mantle viscosity η_{um} was over-estimated in both AH3 and AH4, the rate of the change of topography was small during the early stages of evolution (Fig. 10B, D). This difference forced η_{um} to decrease quickly in both cases. Changes in both η_{um} and η_{lm} likewise reduced their ratio. As a result, in AH3, both viscosities changed monotonically (Fig. 10A), while in

AH4, η_{lm} first overshoot the target (Fig. 10C). This overshoot happened because η_{lm} was forced to increase at the beginning due to an initially overpredicted h , but as η_{um} decreased the viscosity ratio went below the reference, h became underpredicted, which led to the final decrease of η_{lm} . As the viscosities converged, the topographic evolution conformed to the target in both cases after a finite number of loops (Fig. 10B, D). We conclude that the solution is potentially robust as it does not depend on the initial models. Additional experiments demonstrate that solution errors of both upper and lower mantle viscosities are within 1%.

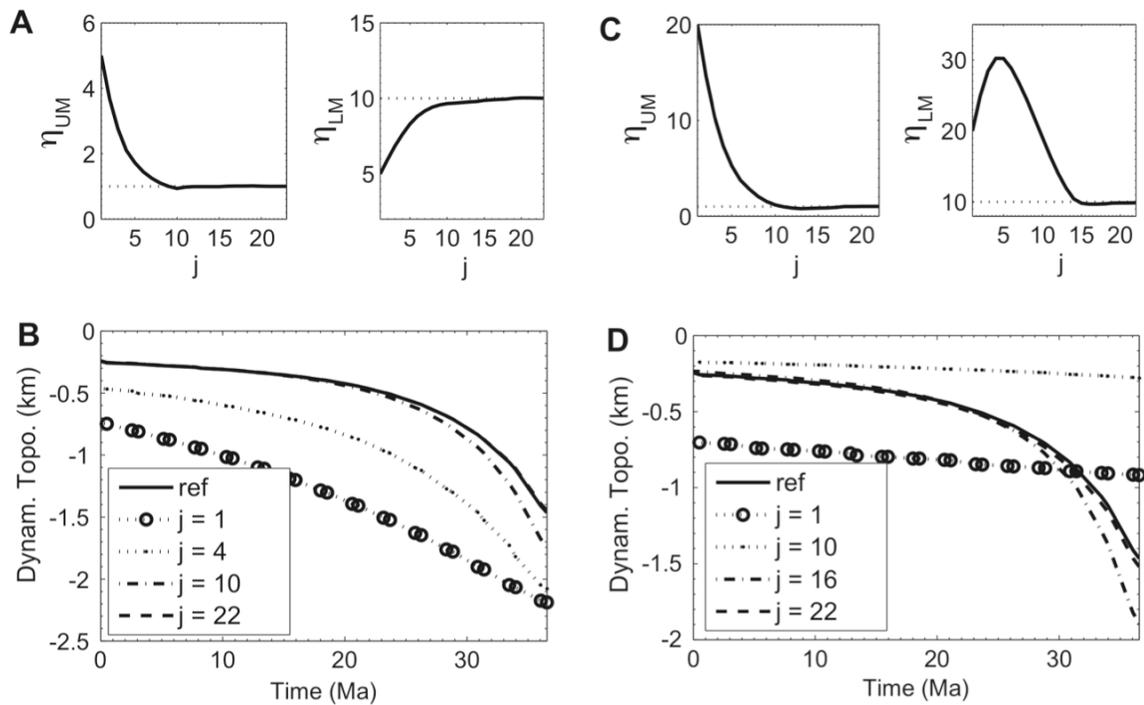


Figure 10 A two-level looping for recovery of both viscosities and initial condition, with temperature scaling at its reference value. Evolution of upper and lower mantle viscosities with respect to middle loop index (j) for Case AH3 (A) and AH4 (C), with dotted lines indicating the

reference values. (B, D) Convergence of temporal dynamic topography recorded at one point for Case AH3 and AH4, respectively.

With another set of experiments with all target values as those just described (AH3 and AH4), we incorrectly guessed ΔT so that it was either smaller (AH5) or larger (AH6) than the true value by 50%. AH5 started with an initially isoviscous state, $(\eta_{um}, \eta_{lm}) = (5, 5)$ (see Fig. 11A, loop 1); and AH6 started with a higher viscosity, $(\eta_{um}, \eta_{lm}) = (20, 20)$ (see Fig. 11C, loop 1). The initial models were chosen such that their effective Rayleigh numbers were not too far from the target values. Parameter recovery in these two cases was similar to what we observed above. Although the viscosity ratio R_η was the same in both AH5 and AH6, the present-day dynamic topographies were different in loop 1, in proportion to the different temperature scaling ΔT (Eq. 19). Consequently, lower mantle viscosities η_{lm} evolved very differently when the temperature was incorrectly guessed. In both Case AH5 and AH6, converged solutions for both viscosities and dynamic topography were obtained. However, although the recovered upper mantle viscosities were always close, there was a tradeoff between lower mantle viscosity and the temperature scaling, as expected from Eq. (19). With more tests on different initial viscosity models, we found that the solutions were robust in that the converged viscosities oscillated around some mean values by no more than 5%. Deviations of the topographic evolutions from the target are instructive (Fig. 11B, D): Due to the tradeoff between ΔT and η_{lm} , the later portion of the predicted curve (closer to present day) always matched the reference curve; however the early portion of the curves were never well predicted since ΔT did not trade off with η_{um} .

Specifically, the earliest portion of the curve was flatter than the reference when ΔT was smaller, and steeper when ΔT larger.

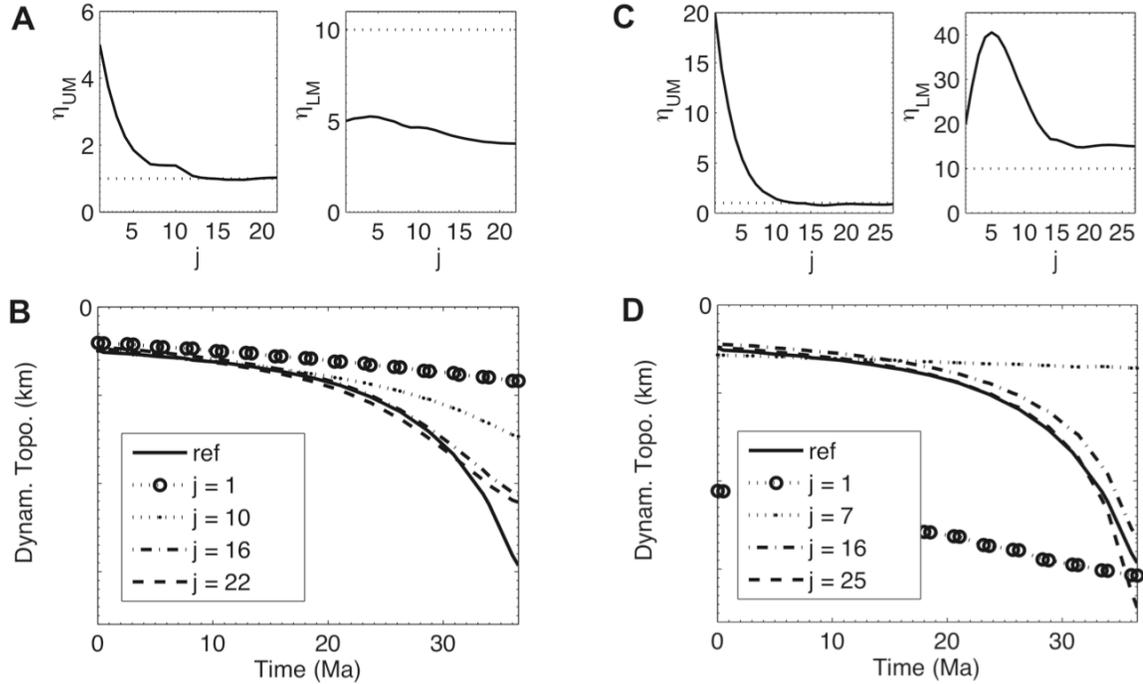


Figure 11 Same as Figure 10, except that the temperature scaling is either smaller (A, B, Case AH5) or larger (C, D, Case AH6) than the reference value by 50%.

This deviation in topographies during the early part of evolution is the basis of an outer level iteration for the update of ΔT . When ΔT is incorrect, lower mantle viscosity trades off with temperature, upper mantle viscosity does not; in theory, dynamic topography can never be predicted exactly if ΔT is incorrect. Eq. (13) and (18) imply that different ΔT , lead to different early topographic evolutions. In practice, we used the very simple relation described by Eq. (13) to update ΔT , constrained from the deviation described above. The iterative relation is given by Eq. (16), where n is the number of data points within the time period when this deviation occurred. Instead of using the absolute magnitude of dynamic

topography, we used the amount of change of topography from the initial time to the n^{th} point. Essentially, we use the difference in the slope at the initial stage of subduction.

As an example, we used the values of ΔT in AH5 and AH6 as two starting guesses for the temperature scaling, and then applied an additional outer loop (calling these new cases, AHT1 and AHT2). The procedure for the outer loop is described above. Note that with different values of ΔT , the converged dynamic topography had different slopes at initial times. We calculated the mismatch between the predicted and reference dynamic topography over the early part of topographic evolution and applied Eq. (16) to update ΔT . Consequently, the deviated topographic curves in both AHT1 (Fig. 12A) and AHT2 (Fig. 12E) moved toward the reference as the number of outer loops increased. Convergence of ΔT 's were shown in Fig. 12B and F with respect to outer loop, where the symbol size was proportional to the residual between predicted and reference dynamic topography. Both the evolution of topography and that of ΔT indicated a correct convergence. To show the interior process of this three-level looping scheme, we picked some value of ΔT during the evolution as an example. For this ΔT , we plotted the updating mantle viscosities, i.e., the middle level loop (Fig. 12C, G). When the two viscosities converged, the corresponding RMS residuals between predicted and target mantle thermal structure at the initial and final (present-day) time also converged (the innermost loop, Fig. 12D, H). These experiments illustrate well that when ΔT is incorrect, recovered η_{lm} is also incorrect; recovered mantle initial conditions based on these model parameters are far from the reference initial since the RMS residuals remain large (compare Fig. 12D, H with Fig. 9E). By fitting the slope of the topography (Fig. 12A, E), ΔT was constrained to converge. The recovered temperature

scaling in both cases approximated its target value within 1%. How closely the final solution fits the reference values will be affected by the discretization of data and the form of weighting functions in Eq. (16), (20), and (21). In the final solution, all recovered model parameters have errors less than 1%, where the lower mantle viscosity linearly trades off with temperature scaling.

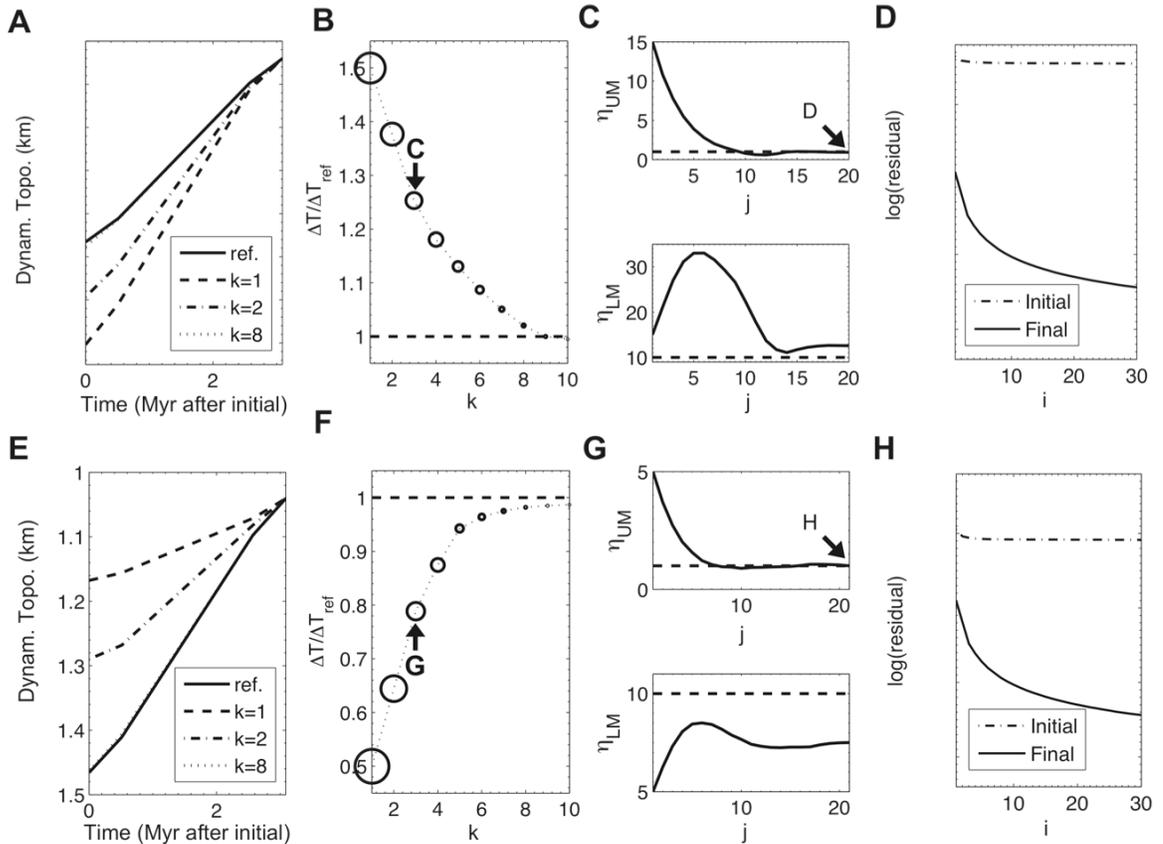


Figure 12 The three-level looping algorithm shown for Case AHT1 (A–D) and AHT2 (E–H), with i , j and k denoting the index of inner, middle, and outer loops, respectively. Shown are evolution of topography at the earliest time (A, E) and temperature scaling (B, F) with respect to outer loop, convergence of upper and lower mantle viscosities versus middle loop, and RMS residuals for both initial and final states of temperature fields for chosen temperature scaling and viscosities. In B and

F, the size (area) of the open circles correspond to the mismatch between magnitudes of predicted and reference dynamic topography in A and E, respectively. All dashed lines indicate the target values.

In summary, our experiments show that, given a temporal record of surface dynamic topography and the present-day mantle seismic tomography showing the geometry of anomalies, this three-level looping scheme allows the recovery of all three mantle dynamic parameters, including upper mantle viscosity, lower mantle viscosity, and the magnitude of the temperature anomaly scaled from seismic perturbations.

3.2.3 Discussion

Combined with dynamic topography observations, the application of the adjoint method can be expanded so that not only can past mantle structures be recovered but also constraints placed on mantle properties. Based on the governing equations, we developed multi-level iteration schemes that constrain both mantle thermal anomalies (the scaling between seismic velocity and temperature or density) and absolute values of upper and lower mantle viscosities. With synthetic experiments, we show that our algorithm is stable and robust. It is worthwhile to note that although this algorithm allows all three model parameters to vary while the final solution remains unique (the uniqueness depends on the recovering power of the adjoint method). In practice, however, we should take advantage of *a priori* knowledge of these quantities, which will reduce the calculation time substantially. This will be essential for large 3D models, which are always computationally expensive.

Synthetic experiments are ideal, and realistic modeling is limited by other factors, including the availability and quality of data. Dynamic topography can be spatially and temporally incomplete, but a complete record is not required for convergence. For example, the dynamic topography constraints on Australian vertical motion since the Cretaceous are limited to specific intervals of time [Gurnis *et al.*, 1998; DiCaprio *et al.*, 2009]. Rate of change of dynamic topography associated with upper mantle thermal structures is especially useful and requires better data coverage both in time and in space. Given these features, our method may work well when applied to realistic modeling problems, for example, reconstructing the subduction history of the Farallon plate underneath North America, using the stratigraphic record on the continent of North America and tomographic images under the continent. Errors associated with recovered dynamic topography from stratigraphy will propagate into the solution of lower mantle viscosity and temperature scaling, mainly due to the fact that these two quantities strongly trade off.

The physical significance of this study is two-fold. First, it is a new way to calculate mantle viscosities, which is almost perfect in synthetic tests as discussed in this paper. This method has an equally good recovery of both upper and lower mantle viscosity, not like glacial rebound models sampling mostly the upper mantle. Furthermore, our method puts constraints on the absolute viscosities of the mantle, while previous geoid studies constrain the upper-lower mantle viscosity ratio only. Second, when using real data including seismic tomography, the constraint on the temperature to seismic velocity scaling by this method can be important for understanding the relation between seismic velocities and temperature and pressure under mantle conditions. As discussed later in the recovery of past Farallon subduction processes, we can put constraints on the flat slab subduction mechanism by

tracking an initial buoyant oceanic plateau whose extra buoyancy has slowed the sinking of the slab, causing flattening, and that the relation of the plateau composition and seismic properties can be further measured via extensive seismic ray tracing and waveform fitting.

The inherent power of our inverse method is that a single density anomaly rises (sinks) through the whole mantle and at each depth samples the viscosity through the rate at which dynamic topography changes. Inherent in this idea is that the magnitude of the density anomaly remains fixed. Of course, the assumption of a constant density difference at each depth is unlikely to be true for mantle convection because the mantle is compressible and different materials will compress at different rates as they ascend (descend) through the mantle (e.g., Tan and Gurnis [2007]). Another complexity is the possible temporal variation of mantle viscosity structures before and after certain geophysical processes, say subduction, as is not included in our model either. To further constrain these complexities, the adjoint model should be expanded to allow for assimilation of more independent data constraints, as is one of the future goals this kind of model should move into.

The recovery of initial conditions, two viscosities, and the temperature scaling, from only dynamic topography and the shape of the present day temperature structure (i.e., “seismic tomography”) appears to be underdetermined. However, for the synthetic experiments this is not the case. The reason is that we have a set of dynamic topography values: One value of dynamic topography places some constraints on the temperature scaling while two dynamic topographies early in the evolution constrain the shallow viscosity (for example), while two dynamic topographies later in the evolution constrain the deeper viscosity. Together, the set of data leads to the recovery of the viscosities, temperature scaling, and initial condition. In reality, however, the problem might be

underdetermined because the number of unique dynamic topography values will be limited and the mantle presumably has a more complex viscosity structure than simulated by a two-layer model. On the other hand, we also realize that for the synthetic experiments carried out earlier, only one single record of the dynamic topography is used. But in reality, many more geographic sites with extracted dynamic subsidence signals are available, although they may not span the whole time sequence; there are other types of constraints besides signals from isolated points, for example, the distribution, wavelength, and magnitude of subsidence will provide extra constraints.

Another issue for problems tailored to the observational record is plate motions, which are an important constraint on mantle flow [e.g., Hager and O'Connell, 1979]. When dynamic topography is used in the inverse procedure, plate motions complicate the problem since the stratigraphic record moves with respect to the rising and sinking mantle anomalies [Gurnis *et al.*, 1998]. Therefore, a transfer function between the two reference frames is needed. In solving the real geophysical problem, say the Farallon subduction, a software package dealing with the transformation of reference frames has been developed, and the moving plates no longer represent a barrier for the inversion.

We want to emphasize that the use of tomography-based adjoint methods is not the only way to infer the internal state of mantle at earlier times and we realize that there are two other means to arrive at such “initial conditions”. In plate-motion constrained mantle flow models, mantle structures can be estimated by a “forward approach” in which a flow model is started earlier in time and then integrated forward to the time of interest [Bunge *et al.*, 1998; Bunge *et al.*, 2002]. Initial conditions can also be inferred via a tectonic approach

based on generating the internal state at a previous time from some combination of surface observations (such as proxies for the configuration of plates, the position of subduction, or hot-spot volcanism). For example, based on the orientation of Mesozoic subduction, Gurnis *et al.* [1998] developed an initial condition at 130 Ma and then forward simulated convection beneath Australia. Bunge and Grand [2000] used the geological arguments that the Farallon slab was flat lying at ~80 Ma and then forward modeled flow beneath North America. These methods complement the inverse models and have different sources of errors. The forward and inverse approaches shared two sources of error: subduction parameterization and radial viscosity structure. However, forward models were sensitive to uncertainties in the plate reconstruction further back in time while inverse models were sensitive to error in reconstructions from the time of interest to the present. We suspect that uncertainty (e.g., the range of structural models permissible at any given time) can be better estimated by comparing the results from these three different methods of generating paleo mantle structures. Consequently, we believe that the adjoint and inverse methods we have developed here should be used in conjunction with the forward and tectonic approaches. Together, these methods will allow a new generation of global and dynamic models to be developed that have well constrained initial conditions.