The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

1864-13

Ninth Workshop on Non-linear Dynamics and Earthquake Predictions

1 - 13 October 2007

Inverse Problem of Thermal Convection: Numerical Approach & Application to Mantle Plume Restoration

Alik T. Ismail-Zadeh

Geophysical Institute
University of Karlsruhe
Karlsruhe, Germany
&
International Institute of Earthquake Prediction
Theory & Mathematical Geophysics
Moscow, Russia

United Nations Educational Scientific and Cultural Organization
and
International Atomic Energy Agency

THE ABDUS SALAM
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

Ninth Workshop
"Non-Linear Dynamics and Earthquake Prediction"
1 October to 13 October 2007

VARIATIONAL DATA ASSIMILATION

Alik T. Ismail-Zadeh

Geophysical Institute, University of Karlsruhe,
Hertzsrt. 16, Karlsruhe 76187, Germany.
E-mail: Alik.Ismail-Zadeh@gpi.uka.de

International Institute of Earthquake Prediction Theory and
Mathematical Geophysics, Russian Academy of Sciences,
84/32 Profsoyuznaya ul., Moscow 117997, Russia.

MIRAMARE-TRIESTE
October 2007

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physics of the Earth and Planetary Interiors 145 (2004) 99–114
PHYSICS OF THE EARTH AND PLANETARY INTERIORS
www.elsevier.com/locate/pepi

Inverse problem of thermal convection: numerical approach and application to mantle plume restoration

Alik Ismail-Zadeh [a,b,c,*], Gerald Schubert [c], Igor Tsepelev [d], Alexander Korotkii [d]

[a] *International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Warshavskoye shosse 79-2, Moscow 113556, Russia*
[b] *Geophysical Institute, University of Karlsruhe, Hertzstr. 16, Karlsruhe 76187, Germany*
[c] *Department of Earth and Space Sciences, Institute of Geophysics and Planetary Physics, University of California, 3806 Geology Building, 595 Charles Young Drive East, Los Angeles, CA 90095-1567, USA*
[d] *Institute of Mathematics and Mechanics, Ural Branch, Russian Academy of Sciences, ul., S. Kovalevskoy 16, Ekaterinburg 620219, Russia*

Received 7 October 2003; received in revised form 1 March 2004; accepted 1 March 2004

Abstract

Modern seismic tomographic images of the Earth's interior facilitate the inference of the complex trajectories of present-day convective flow in the upper mantle. Quantitative reconstruction of both the observed mantle structure and temperature field backwards in time requires a numerical tool for solving the inverse problem of thermal convection at infinite Prandtl number. In this paper we present a variational approach to three-dimensional numerical restoration of thermoconvective mantle flow with temperature-dependent viscosity. This approach is based on a search for the mantle temperature and flow in the geological past by minimizing differences between present-day mantle temperature derived from seismic velocities (or their anomalies) and that predicted by forward models of mantle flow for an initial temperature guess. The past mantle temperatures so obtained can be employed as constraints on forward models of mantle dynamics. To demonstrate the applicability of this technique, we restore numerically a fluid dynamic model of the evolution of upper mantle plumes and show that the initial shape of the plumes can be accurately reconstructed. We then model the evolution of the plumes forward in time (plume upbuilding) starting from the restored state to the state they were before the restoration and demonstrate the high accuracy of the model predictions. We also show that the neglect of thermal diffusion in the backward modeling of thermal plumes (in order to simplify the numerical procedure) results in erroneous restorations of the plumes.

Keywords: Mantle plume; Heat diffusion; Backward modeling; Numerical method

1. Introduction

The reconstruction of mantle plumes and lithospheric slabs to earlier stages of their evolution is a major challenge in geodynamics. High-resolution seismic tomographic studies open possibilities for detailed observations of present-day mantle structures (e.g., Grand et al., 1997; van der Voo et al., 1999;

* Corresponding author. Tel.: +49-721-608-4610; fax: +49-721-71173.
E-mail addresses: aismail@mitp.ru, alik.ismail-zadeh@gpi.uni-karlsruhe.de, aiz@ess.ucla.edu (A. Ismail-Zadeh), schubert@ucla.edu (G. Schubert), tsepelev@imm.uran.ru (I. Tsepelev), korotkii@imm.uran.ru (A. Korotkii).

0031-9201/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.pepi.2004.03.006

Ritsema and Allen, 2003) and for derivations of mantle temperature from seismic velocities or velocity anomalies (e.g., Sobolev et al., 1996; Goes et al., 2000). An accurate reconstruction would allow the test of geodynamic models by simulating the evolution of plumes or slabs starting from the restored state and comparing the derived forward state to observations.

For clarity of subsequent discussion, we introduce a few mathematical definitions used in the paper. A mathematical model for a geophysical problem has to be *well-posed* in the sense that it has to have the properties of existence, uniqueness, and stability of a solution to the problem (Hadamard, 1923). Problems for which at least one of these properties does not hold are called *ill-posed*. The requirement of stability is the most important one. If a problem lacks the property of stability then its solution is almost impossible to compute because numerical computations are polluted by unavoidable errors. If the solution of a problem does not depend continuously on the initial data, then, in general, the computed solution may have nothing to do with the true solution.

The inverse problem of thermal convection in the mantle is an ill-posed problem, since the backward heat problem, describing both heat advection and diffusion through the mantle backwards in time, possesses the properties of ill-posedness (Kirsch, 1996). In particular, the solution to the problem does not depend continuously on the initial data. This means that small changes in the present-day temperature field may result in large changes of predicted mantle temperatures in the past (see Appendix A for an explanation of this statement in the case of the one-dimensional diffusion equation).

If heat diffusion is neglected, the solution of the advection equation backwards in time does not present computational difficulties. A numerical approach to the solution of the inverse problem of the Rayleigh–Taylor (gravitational) instability was proposed by Ismail-Zadeh (1999) and was developed later for a dynamic restoration of plume (diapiric) structures to their earlier stages (Ismail-Zadeh et al., 2001a). Kaus and Podladchikov (2001) and Korotkii et al. (2002) applied the approach to study 3D Rayleigh–Taylor overturns forward and backward in time. Both direct (forward in time) and inverse (backward in time) problems of the gravitational advection are well-posed. This is because the time-dependent advection equation (for density or temperature) has the same form of characteristics for the direct and inverse velocity field (the vector velocity reverses its direction, when time is reversed). Therefore, numerical algorithms used to solve the direct problem of the gravitational instability of the geological structures can also be used in studies of the inverse problems by replacing positive timesteps with negative ones.

Steinberger and O'Connell (1997, 1998) and Conrad and Gurnis (2003) modeled the mantle flow backwards in time from present-day mantle density heterogeneities inferred from seismic observations. However, they ignored thermal diffusion in the mantle (and hence the respective term in the heat equation) and employed the advection equation in the modeling. We demonstrate that this approach (neglect of heat diffusion in backward modeling) is not valid.

There is a sizeable literature on the numerical solution of the backward heat equation (e.g., Buzbee and Carasso, 1973; Colton, 1979; Elden, 1982; Ames and Epperson, 1997; Lu, 1997; Moszynski, 2001; see also Tikhonov and Arsenin, 1977, and Kirsch, 1996, for additional references). These methods are based on a regularization of the numerical solution. Bunge et al. (2003) and Ismail-Zadeh et al. (2003a,b) have independently developed variational approaches for solving the inverse problem of mantle convection. The major differences between the two approaches are that Bunge et al. (2003) applied the variational method to a set of equations describing mantle convection, whereas Ismail-Zadeh et al. (2003a) applied the variational method to the heat equation, because time enters only into this equation and the backward heat problem is ill-posed. Ismail-Zadeh et al. (2003a) determine the temperature in the geological past and then the convective backward flow from the Stokes and continuity equations. (We will discuss other differences between these two approaches to solving the inverse problem of mantle convection later in the paper.)

In Section 1 we present a mathematical statement of the three-dimensional direct and inverse problems of thermal convection with temperature-dependent viscosity. In Section 2 we describe the variational approach to search for mantle temperature in the geological past based on estimations of its present-day temperature. The approach is based on reducing the problem to minimization of the objective functional describing the difference between the present-day

mantle temperature and that predicted by forward models of mantle flow for an initial temperature guess. The optimum solution to the minimization problem is provided by iteratively solving coupled direct and conjugate (adjoint) problems for the heat equation. The variational approach to solving the backward heat problem has been known in applied mathematics and geophysics (atmosphering modeling and oceanography, see e.g. Bennett (1992) and Kalnay (2003)), but so far has not been used in studies of mantle thermoconvective flow. In Section 3 we describe numerical techniques used in solving the inverse problem of mantle convection. We demonstrate the applicability of the numerical approach to restoration of mantle plumes and show the effect of heat diffusion on results of the backward modeling in Section 4. We discuss the physical and mathematical meaning of the time-reversible processes in Section 5 and present conclusions in Section 6.

2. Mathematical statement of the problem

We assume that the mantle behaves as a Newtonian fluid at geological time scales and consider the slow thermoconvective flow of a heterogeneous incompressible fluid at infinite Prandtl number with a temperature-dependent viscosity in a three-dimensional rectangular domain $\Omega = (0, x_1 = l_1) \times (0, x_2 = l_2) \times (0, x_3 = l_3 = h)$ heated from below; $x = (x_1, x_2, x_3)$ are the spatial coordinates; the x_3-axis is vertical and positive upward. Thermoconvective flow is described by the heat, momentum (Stokes), and continuity equations. In the Boussinesq approximation these dimensionless equations take the form (Chandrasekhar, 1961):

$$\frac{\partial T}{\partial t} + \boldsymbol{u} \cdot \nabla T - \nabla^2 T = 0, \tag{1}$$

$$-\nabla P + \nabla \cdot [\mu(T)(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{Tr}})] + Ra\, T\, \boldsymbol{e} = 0, \tag{2}$$

$$\nabla \cdot \boldsymbol{u} = 0, \tag{3}$$

for $x \in \Omega$ and $t \in (\vartheta_1, \vartheta_2)$, where T, $\boldsymbol{u}$, P, μ, and t are temperature, velocity, pressure, viscosity, and time respectively; superscript $^{\mathrm{Tr}}$ means transpose; and $\boldsymbol{e} = (0, 0, 1)$ is the unit vector. The Rayleigh number is defined as $Ra = \alpha g \rho_0 \Delta T h^3 / \mu_0 \kappa$ where α the thermal expansivity; g the acceleration due to gravity; ρ_0 and μ_0 are the reference typical density and viscosity, respectively; ΔT is the temperature contrast between the lower and upper boundaries of the model domain; and κ is the thermal diffusivity. In Eqs. (1)–(3) length, temperature, and time are normalized by h, ΔT, and h^2/κ, respectively. We do not consider the chemical convection in the mantle. The formulation of the inverse problem of thermo-chemical convection and the numerical approach to the solution of the problem are described by Ismail-Zadeh et al. (2003a).

At the boundary Γ of the model domain Ω we set the impenetrability and perfect slip conditions: $\boldsymbol{n} \cdot \nabla \boldsymbol{u}_{\mathrm{tg}} = 0$ and $\boldsymbol{n} \cdot \boldsymbol{u} = 0$, where $\boldsymbol{n}$ is the outer normal vector and $\boldsymbol{u}_{\mathrm{tg}}$ is the tangential component of velocity. We assume the heat flux through the vertical boundaries of Ω to be zero: $\boldsymbol{n} \cdot \nabla T = 0$. The upper and lower boundaries are assumed to be isothermal surfaces, and hence $T = T_{\mathrm{u}}$ at $x_3 = h$, $T = T_{\mathrm{l}}$ at $x_3 = 0$, where T_{u} and T_{l} are constant, and $\Delta T = T_{\mathrm{l}} - T_{\mathrm{u}} > 0$. To solve the direct and inverse problems of thermal convection, we assume that the temperature is known at the initial time $t = \vartheta_1$ and at the final (in terms of the direct problem) time $t = \vartheta_2$, respectively.

Thus, the direct (or inverse) problem of the thermal convection is to determine velocity, $\boldsymbol{u} = \boldsymbol{u}(t, x)$, pressure, $P = P(t, x)$, and temperature, $T = T(t, x)$, satisfying Eqs. (1)–(3) at $t \geq \vartheta_1$ (or $t \leq \vartheta_2$), the prescribed boundary conditions, and the temperature condition at $t = \vartheta_1$ (or $t = \vartheta_2$).

3. Variational approach to solving the backward heat problem

In this section, we present a variational approach to an approximate solution to the backward heat problem. Consider the following objective (quadratic) functional

$$J(\varphi) = \|T(\vartheta_2, \cdot\,; \varphi) - \chi(\cdot)\|^2 = \int_{\Omega} |T(\vartheta_2, x; \varphi) - \chi(x)|^2 \mathrm{d}x, \tag{4}$$

where $T(\vartheta_2, x; \varphi)$ is the solution of the forward heat equation (1) with the appropriate boundary and initial conditions at final time ϑ_2, which corresponds to some (unknown as yet) initial temperature distribution $\varphi = \varphi(x)$; $\chi(x) = T(\vartheta_2, x; T_0)$ is the known temper-

ature distribution at the final time for the initial temperature $T_0 = T_0(x)$; and $\|\cdot\|$ is the norm in space $L^2(\Omega)$. We seek a minimum of the objective functional with respect to the initial temperature, φ. The functional has its unique global minimum at value $\varphi = T_0$, and $J(T_0) = 0$, $\nabla J(T_0) = 0$. The uniqueness of the global minimum of the objective functional follows from the uniqueness of the solution of the relevant boundary-value problem for the heat equation and a strong convexity of the functional (Tikhonov and Samarskii, 1990).

To find a minimum of the objective functional we employ the gradient method (Vasiliev, 2002)

$$\varphi_{k+1} = \varphi_k - \alpha_k \nabla J(\varphi_k), \qquad \varphi_0 = T_*, \quad k = 0, 1, 2, \ldots, \tag{5}$$

$$\alpha_k = \min\{1/(k+1); J(\varphi_k)/\|\nabla J(\varphi_k)\|\}, \tag{6}$$

where T_* is an initial temperature guess. It can be shown that the gradient of functional J is represented as $\nabla J(\varphi) = \Psi(\vartheta_1, \cdot)$ (see Appendix B), where Ψ is the solution to the following boundary problem conjugated (adjoint) to the respective boundary problem for Eq. (1):

$$\begin{aligned} &\partial\Psi/\partial t + \boldsymbol{u}\cdot\nabla\Psi + \nabla^2\Psi = 0, \quad x \in \Omega, \quad t \in (\vartheta_1, \vartheta_2),\\ &\sigma_1\Psi + \sigma_2\partial\Psi/\partial\boldsymbol{n} = 0, \qquad x \in \Gamma, \quad t \in (\vartheta_1, \vartheta_2),\\ &\Psi(\vartheta_2, x) = 2(T(\vartheta_2, x; \varphi) - \chi(x)), \qquad x \in \Omega, \end{aligned} \tag{7}$$

where σ_1 and σ_2 are some smooth functions or constants satisfying the condition $\sigma_1^2 + \sigma_2^2 \neq 0$. Selecting σ_1 and σ_2 we can obtain corresponding boundary conditions. Problem (7) is ill-posed for positive timesteps and well-posed for negative timesteps.

The solution algorithm for the backward heat problem is based on the following three steps ($k = 0, 1, 2, \ldots, n, \ldots$):

(i) solve the forward heat equation (1) in the time interval $[\vartheta_1, \vartheta_2]$, $x \in \Omega$, with the boundary conditions defined and initial temperature $T(\vartheta_1, x) = \varphi_k(x)$ in order to find $T(\vartheta_2, x; \varphi_k)$;
(ii) solve problem (7) backwards in time and determine $\nabla J(\varphi_k) = \Psi(\vartheta_1, x; \varphi_k)$; and
(iii) determine α_k from (6) and then update the initial temperature, i.e., find φ_{k+1} from (5).

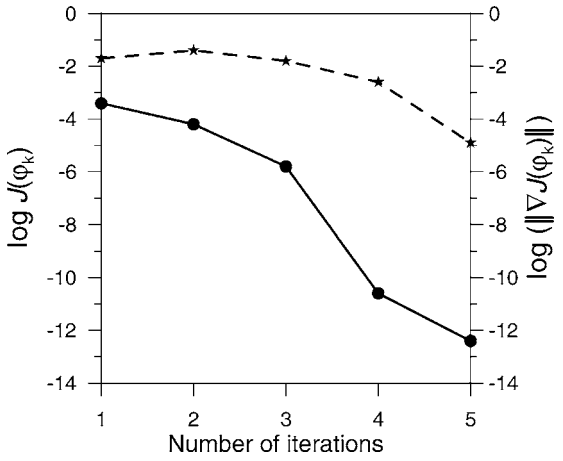

Fig. 1. Relative reductions of the objective functional J (solid line) and the norm of the gradient of the objective functional $\|\nabla J\|$ (dashed line) as functions of the number of iterations.

Computations are terminated when

$$\delta\varphi_n = J(\varphi_n) + \|\nabla J(\varphi_n)\|^2 < \varepsilon, \tag{8}$$

where ε is a small constant. The temperature φ_n is then considered to be the approximation of the target value of the initial temperature T_0. If $\delta\varphi_n \geq \varepsilon$, we return to step (i) and make the next iteration.

The performance of the algorithm is evaluated in terms of the number of iterations n required to achieve a prescribed relative reduction of $\delta\varphi_n$ (in our numerical experiments we assumed $\varepsilon = 10^{-8}$). Fig. 1 presents the evolution of the objective functional $J(\varphi_n)$ and the norm of the gradient of the objective functional $\|\nabla J(\varphi_n)\|$ versus the number of iterations at time about $(\vartheta_2 + \vartheta_1)/2$. For other time steps we observe a similar evolution of J and $\|\nabla J\|$. Numerical tests demonstrate that if the initial guess for temperature is a smooth function, then iterations converge rapidly (only 5–10 iterations); otherwise, the iterations converge very slowly (100 and more iterations).

Implementation of minimization algorithms requires the evaluation of both the objective functional (4) and its gradient ∇J. Each evaluation of the objective functional requires an integration of the model Eq. (1) with the appropriate boundary and initial conditions, whereas the gradient is obtained through the backward integration of the adjoint Eq. (7). The performance analysis shows that the CPU time required to evaluate the gradient J is about the CPU time re-

quired to evaluate the objective functional itself, and this is because the direct and adjoint heat problems are described by the same equations.

Information on the properties of the Hessian matrix ($\nabla^2 J$) is important in many aspects of minimization problems (Daescu and Navon, 2003). To obtain sufficient conditions for the existence of the minimum of the problem, the Hessian matrix must be positive definite at T_0 (optimal initial temperature). However, an explicit evaluation of the Hessian matrix in our case is prohibitive due to the number of variables.

We used the Boussinesq approximation, and hence the viscous dissipation as a heat source term in the heat equation was neglected. If viscous dissipation is included in the heat equation and viscosity depends on temperature, then the suggested method for backward modeling of the heat equation should be modified, because the adjoint problem for the heat equation becomes more complicated. However, the dissipation number, $Di = \mu_0\kappa/(C_P\rho_0\Delta T h^2)$ (where C_P is heat capacity at constant pressure) is small enough (about 10^{-7} for the upper mantle) that the viscous dissipation term can be neglected.

Thus, the solution of the backward heat problem is reduced to solutions of series of forward problems, which are known to be well-posed (Tikhonov and Samarskii, 1990). The algorithm can be used to solve the problem over any subinterval of time in $[\vartheta_1, \vartheta_2]$.

4. Numerical approach to solving the inverse problem of mantle convection

In this section, we describe briefly the numerical methods we use in the study. See Ismail-Zadeh et al. (2001b) for more detail.

4.1. Numerical method for solving the Stokes equation

To facilitate computations, Eqs. (2) and (3) are simplified by introducing a two-component representation of the vector velocity potential

$$\boldsymbol{u} = \mathrm{curl}\vec{\psi}, \quad \vec{\psi} = (\psi_1, \psi_2, 0). \tag{9}$$

We represent the vector velocity potential as a linear combination of tricubic basis splines and apply the Eulerian finite element method to Eqs. (2) and (3) with the appropriate boundary conditions. To simplify analysis, we rewrite the problem in variational form. To solve the problem numerically, the model domain Ω is discretized introducing the uniform rectangular grid

$$0 = x_i^0 < x_i^1 < \ldots < x_i^{n_i-1} < x_i^{n_i} = l_i, \quad i = 1, 2, 3,$$

with grid points $\Omega_{ijk} = (x_1^i, x_2^j, x_3^k)$, $0 \le i \le n_1$, $0 \le j \le n_2$, and $0 \le k \le n_3$. At each grid point Ω_{ijk}, we define a tricubic basis element $\omega_{ijk}^l = \omega_{ijk}^l(x_1, x_2, x_3)$, $l = 1, 2$ as the tensor product of one-dimensional cubic basis elements (Ahlberg et al., 1967). The construction of bases consisting of tricubic elements ω_{ijk}^l is described by Ismail-Zadeh et al. (1998).

The vector potential is approximated by the combinations

$$\psi_{l(t,x_1,x_2,x_3)} \approx \sum_{i,j,k} \psi_{ijk}^l(t)\omega_{ijk}^l(x_1, x_2, x_3), \quad l = 1, 2, \tag{10}$$

and viscosity is approximated by using trilinear basis elements $\phi_{ijk}(x_1, x_2, x_3)$:

$$\mu(T(t, x_1, x_2, x_3)) \approx \sum_{i,j,k} \mu_{ijk}(t)\phi_{ijk}(x_1, x_2, x_3).$$

The coefficients ψ_{ijk}^l are determined at each time step by solving a set of linear algebraic equations with a symmetric positive definite band matrix. The set is solved iteratively by conjugate gradient or Gauss–Seidel methods. The relevant software was designed for implementing the codes on parallel computers. A detailed analysis of particular implementations of iterative methods for sets of linear algebraic equations is presented by Tsepelev et al. (1999).

4.2. Numerical method for solving the heat equation

Temperature is computed by finite-difference methods. To do this, we define a regular grid in Ω (we use a grid finer by a factor of three than that employed to approximate the vector potential). The first and second order derivatives with respect to coordinates in the heat equation are approximated by central finite differences. The velocity in the heat equation is determined from (9) and (10).

We employ an implicit alternating-direction method (Marchuk, 1994) to compute temperature. Essentially, temperature T^{n+1} at time $t = t_{n+1}$ is found as

$$r^n = \tau(\nabla^2 T^n + \boldsymbol{u}\cdot\nabla T^n), \quad \left[1 - \frac{\tau}{2}\frac{\partial^2}{\partial x_3^2}\right] T^* = r^n,$$

$$\left[1 - \frac{\tau}{2}\frac{\partial^2}{\partial x_2^2}\right] T^{**} = T^*, \quad \left[1 - \frac{\tau}{2}\frac{\partial^2}{\partial x_1^2}\right] T^{***} = T^{**},$$

$$T^{n+1} = T^n + T^{***},$$

where τ is the time step. In the modeling, the parameter τ is chosen in such a way as to guarantee the stability of the finite difference method, namely:

$$\tau = \frac{1}{8}\frac{\mathrm{d}x}{u_{\max}}, \qquad \mathrm{d}x = [h_1^2 + h_2^2 + h_3^2]^{1/2},$$

$$u_{\max} = \max\{|u_i(x)| : x \in \bar{\Omega}, \quad i = 1, 2, 3\},$$

where $h_k = x_k^i - x_k^{i-1}$. To compute T^{n+1}, $n_2n_3 + n_1n_3 + n_1n_2$ tridiagonal systems are solved, and the corresponding number of independent modules can be organized to perform parallel computations of these systems by a tridiagonal method. The representation of the vector velocity potential based on cubic splines employed here makes it possible to compute both advection and diffusion of temperature simultaneously by finite-difference methods.

4.3. The algorithm for numerical solution of the inverse problem of mantle convection

We define a uniform partition of the time axis at points $t_n = \vartheta_2 - \tau n$, where τ is the time step, and n successively takes integer values from 0 to some natural number $m = (\vartheta_2 - \vartheta_1)/\tau$. At each subinterval of time $[t_{n+1}, t_n]$, the solution of the problem backwards in time consists of the following basic steps.

Step 1. Given the temperature $T = T(t_n, \cdot)$ at $t = t_n$ we solve a set of linear algebraic equations derived from Eqs. (2) and (3) and the appropriate boundary conditions to find the velocity potential $\vec{\psi} = \vec{\psi}(t_n, \cdot)$.

Step 2. Eq. (9) is used to determine the velocity $\boldsymbol{u} = \boldsymbol{u}(t_n, \cdot; T)$, corresponding to temperature $T = T(t_n, \cdot)$, from the vector potential.

Step 3. The 'advective' temperature $T_\mathrm{a} = T_\mathrm{a}(t_{n+1}, \cdot)$ is determined by solving the advection heat equation (neglecting the diffusion term) backwards in time, and steps 1 and 2 are then repeated to find the velocity $\boldsymbol{u}_\mathrm{a} = \boldsymbol{u}(t_{n+1}, \cdot; T_\mathrm{a})$, corresponding to the 'advective' temperature.

Step 4. The velocities $\boldsymbol{u}_\mathrm{a}$ and $\boldsymbol{u}$ are used in the direct problem (Eq. (1)) combined with the boundary conditions) and the conjugate problem (7), respectively, to find temperature $T = T(t_{n+1}, \cdot)$ at $t = t_{n+1}$.

Compared to the previous algorithm of Ismail-Zadeh et al. (2003a), step 3 is introduced here to accelerate the convergence of temperature iterations in solving the direct and conjugate heat problems (to satisfy inequality (8) in a few iterations at fixed ε, see Fig. 1).

After these algorithmic steps, we obtain temperature $T = T(t_n, \cdot)$, velocity potential $\vec{\psi} = \vec{\psi}(t_n, \cdot)$, and velocity $\boldsymbol{u} = \boldsymbol{u}(t_n, \cdot)$ corresponding to $t = t_n$, $n = 0, \ldots, m$. Based on the obtained results, we can use interpolation to reconstruct, when required, the entire process on the time interval $[\vartheta_1, \vartheta_2]$ in more detail. The time step is chosen automatically so that the maximal displacement of material points does not exceed a sufficiently small preset value.

Thus, at each subinterval of time we apply the variational method to the heat equation only, iterate the direct and conjugate problems for the heat equation in order to find temperature, and determine backward flow from the Stokes and continuity equations twice (for 'advective' and 'true' temperatures). Compared to the variational approach by Bunge et al. (2003), our numerical approach is computationally less expensive, because we do not involve the Stokes equation into the iterations between the direct and conjugate problems (the numerical solution of the Stokes equation is the most time consuming calculation). Moreover, our approach admits the use of temperature-dependent viscosity.

5. Restoration model of mantle plumes

In the modeling, we consider thermal plumes to be formed at the depth of 648 km, approximately the boundary between the lower mantle and upper mantle. To verify the validity of our numerical approach, we start our simulations by computing a forward model

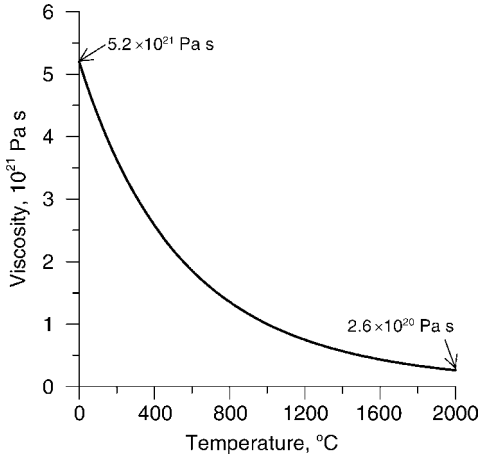

Fig. 2. Temperature-dependent viscosity used in the modeling.

of the evolution of the thermal plumes and then we restore the evolved plumes to their earlier stages.

We assume the following dimensional model parameters: $\alpha = 3 \times 10^{-5}$ K^{-1}, $\Delta T = 2000$ K, $\rho_0 = 3.4 \times 10^3$ kg m^{-3}, and $\kappa = 0.8 \times 10^{-6}$ m^2 s^{-1} (Schubert et al., 2001); the reference mantle viscosity is $\mu_0 = 10^{21}$ Pa s (Forte and Mitrovica, 2001); $h = 720$ km, and $l_1 = l_2 = 3h$, and therefore, the Rayleigh number is $Ra = 9.5 \times 10^5$. At initial time $t = 0$ we assume that the upper mantle temperature increases linearly with depth.

We consider the mantle viscosity μ to be temperature-dependent (Busse et al., 1993):

$$\mu(T) = \exp[Q/(T+G) - Q/(0.5+G)], \qquad (11)$$

where $Q = [225/\ln(r)] - 0.25\ln(r)$, $G = [15/\ln(r)] - 0.5$, and $r = 20$ is the effective viscosity ratio between the upper and lower boundaries of the model domain. The temperature dependence of this viscosity function is shown in Fig. 2. We adopt this viscosity law for the sake of simplicity in the model and for benchmarking of our numerical codes (Busse et al., 1993), although the methodology described here is valid for more general viscosity relationships (Ismail-Zadeh et al., 2003a). The chosen temperature (and depth) dependent viscosity profile has no minimum associated with the asthenospheric layer, while an inversion of the main convection-related geophysical data (free-air gravity, plate divergence, r.m.s. topography) suggests the existence of a low-viscosity channel at depths of 100–300 km with an average viscosity of about 10^{20} Pa s (Forte and Mitrovica, 2001). A more realistic viscosity profile will influence the evolution of mantle plumes, but it will not affect results of the restoration of mantle plumes.

In order to initiate the growth of thermal plumes, we prescribe a small thermal perturbation on the horizontal plane $x_3 = 0.1$ (depth 648 km) at the initial time. The time the plumes take to develop depends on the amplitude of the initial perturbation. Hence, we computed the evolution of plumes to the stage presented in Fig. 3a and considered this stage as an initial configuration of the plumes in our forward modeling.

The model domain was divided into $37 \times 37 \times 29$ rectangular finite elements. The vector potential is approximated by tricubic splines on the elements, while temperature, velocity, and viscosity are represented on a more refined grid $112 \times 112 \times 88$. The evolution of the thermal plumes was modeled forward in time (Fig. 3a–e). We interrupted the computations at a certain time (at 75 Myr), when the plumes had developed a mushroom geometry (Fig. 3e). The final state of the plumes in the forward model was used as the initial state of the plumes in backward (or restoration) models. In the following we refer to the final state of the thermal plumes in the forward modeling as the 'present' state of the plumes.

We apply the suggested numerical approach to restore the plumes from their 'present' state to the state they were in Late Cretaceous times (75 Myr ago). To achieve the accuracy $\varepsilon = 10^{-8}$ (see Eq. (8)) we performed up to 10 iterations at each subinterval of time depending on the choice of the initial temperature guess, T_*. Despite the number of necessary iterations, a performance analysis demonstrated that the total execution time for the numerical restoration of the evolution of the plumes was only about a factor of three (depending on the number of iterations) larger than the time required for the forward modeling of the plumes. The restoration method developed by Bunge et al. (2003) is an order of magnitude more computationally expensive.

Fig. 4 (left panel) shows the restored states of the plumes and the temperature residuals δT

$$\delta T(x_1, x_2) = \left[\int_0^{l_3} \left(T(x_1, x_2, x_3) - \tilde{T}(x_1, x_2, x_3) \right)^2 \mathrm{d}x_3 \right]^{1/2}$$

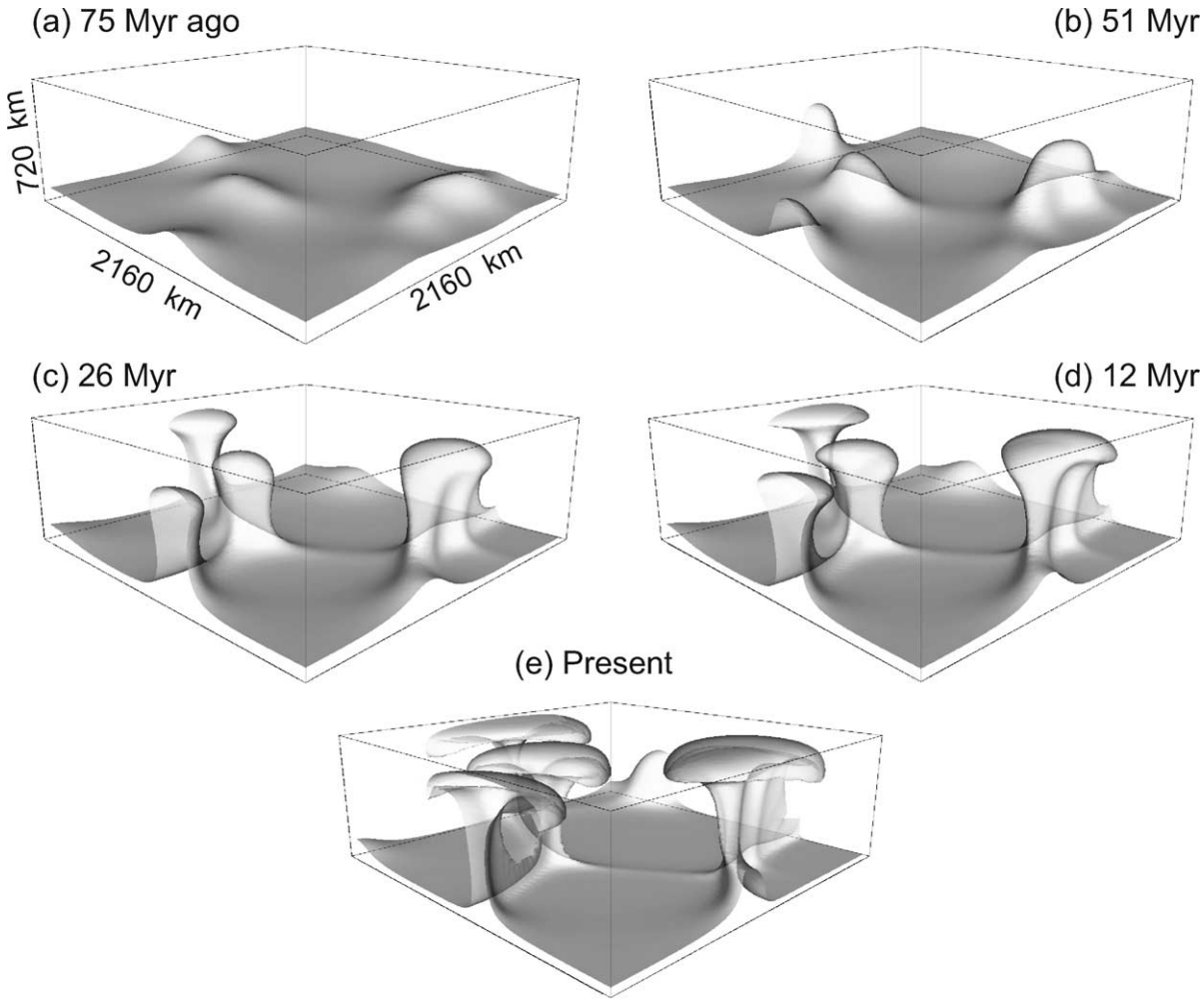

Fig. 3. Mantle plumes in the forward modeling at successive times: from 75 Myr ago (*a*) to the 'present' state of the plumes (*e*). The plumes are represented here and in Figs. 4–6 by isothermal surfaces at 1840 K.

between the temperature $\tilde{T}$ predicted by the forward model and the temperature T restored to the same age. The temperature residuals are within a thousandth of a degree for the initial restoration period (from present to about 26 Myr), and the maximum residual reaches about $\delta T = 25^{\circ}$ at the restoration time of 75 Myr. The computations show that the errors (temperature residuals) get larger the farther restorations move backwards in time. For the heat problem, it has been shown that the size of the time domain enters into the estimation of the rate of convergence, and hence this size influences the errors.

To demonstrate effects of heat diffusion (and its absence) on the temperature restoration, we computed the thermal plumes backwards in time using the heat advection equation (with no heat diffusion). The right panel of Fig. 4 presents the results of the modeling. The shapes of the restored mantle plumes become notably different from that of 'true' plumes (plumes modeled forwards in time) after 26 Myr. The temperature residuals (with no heat diffusion considered) are one to three orders of magnitude larger than those when heat diffusion is considered, and the minimum residual is about 100 K at the restoration time of 75 Myr.

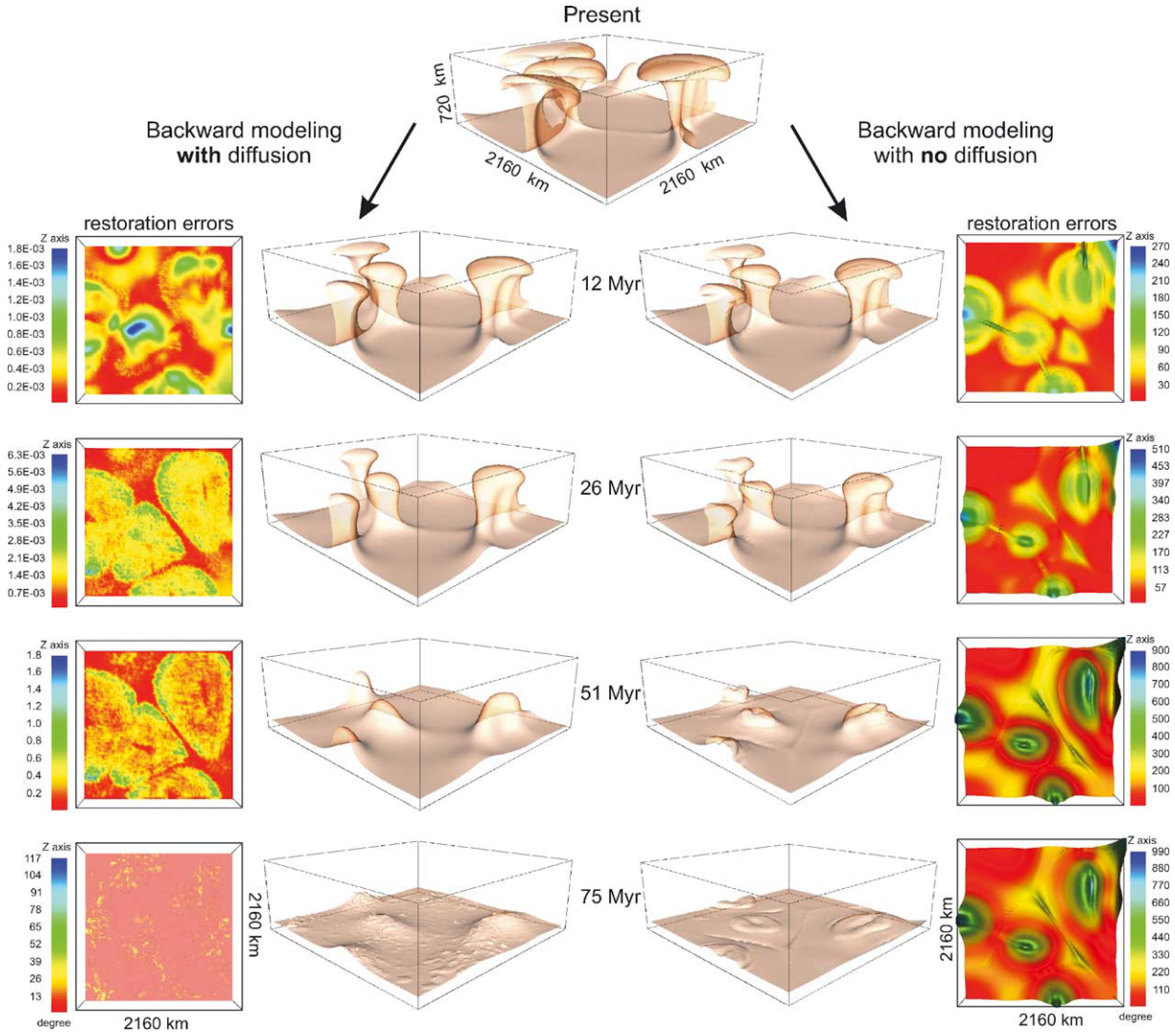

Fig. 4. Restored mantle plumes in the backward modeling and restoration errors (temperature residuals) at successive times: from the 'present' to 75 Myr ago. The left two panels present the model results in the case when diffusion is included in the heat transfer, and the right two panels are for the case in which diffusion is neglected.

Thus, we have demonstrated that the neglect of heat diffusion in the backward modeling leads to an inaccurate restoration of mantle plumes.

Even though the coefficient of heat diffusion is small, the neglect of diffusion in the heat equation results in a different solution to the heat problem because of the reduction in the order of the differential equation (Tikhonov and Samarskii, 1990). Moreover, when mantle convection is computed forwards in time using the heat diffusion equation and diffusion is ignored in the backward modeling of the same mantle convection, results are inconsistent and even unphysical.

The comparison between 'true' (modeled forwards in time) and restored (modeled backwards in time) plumes is quite natural from the computational point of view, but not from the geophysical point of view, because the mantle structure in the past (initial 'true'

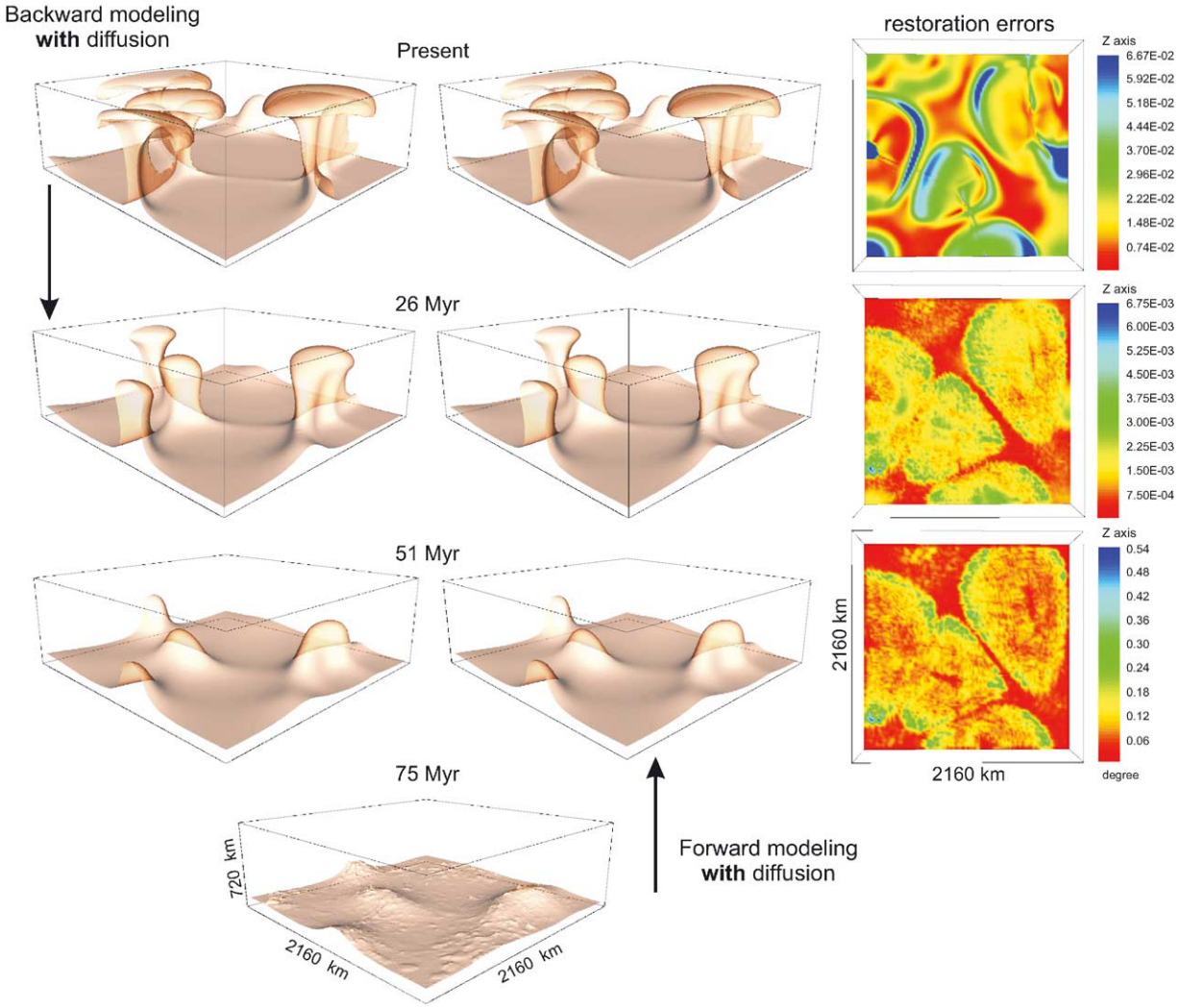

Fig. 5. Mantle plumes restored from the 'present' to 75 Myr ago (left panel), upbuilt plumes back to their 'present' state (central panel), and the restoration errors (right panel) in the case when diffusion is included.

plumes) is unknown. Hence, we perform another numerical experiment on the accuracy of the restoration technique. We start from the 'present' structure of the plumes, apply the suggested technique to restore the past structure, run a forward model of the restored plumes, and compare the 'present' structure and the one recovered after the forward modeling. Fig. 5 presents the results of this modeling which show that the restoration works quite well: temperature residuals (difference between the temperature of the restored mantle plumes and that of the plumes of the same age in the forward model) are within hundredths of a degree.

We have also performed similar computations with the heat diffusion equation replaced by the heat advection equation during the backward modeling. Fig. 6 shows the results of restoration of the 'present' state of the plumes to 75 Myr ago and upbuilding of the restored plumes to the present time. The temperature residuals are larger (by several orders of magnitude)

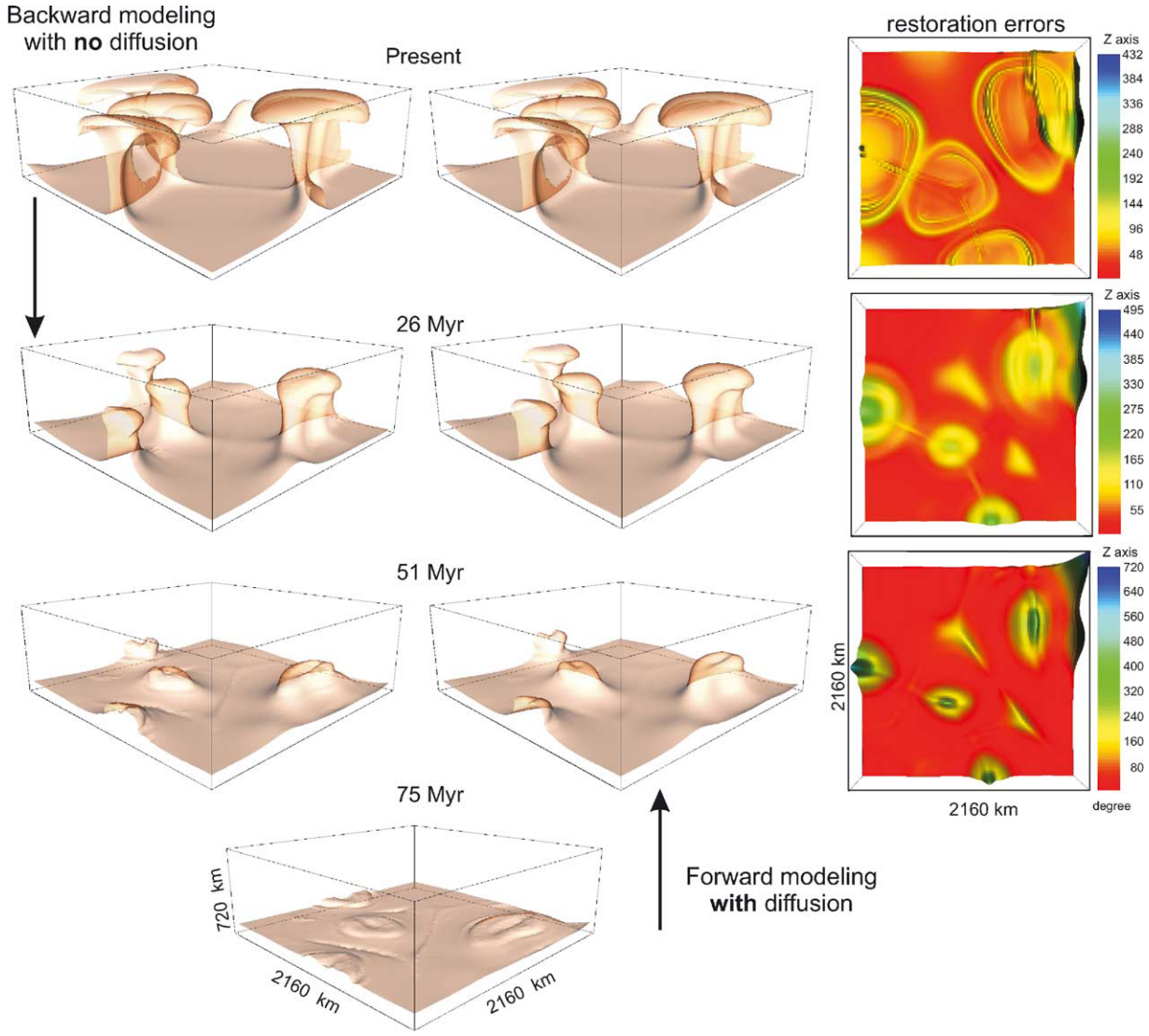

Fig. 6. Mantle plumes restored from the 'present' to 75 Myr ago (left panel), upbuilt plumes back to their 'present' state (central panel), and the restoration errors (right panel) in the case when diffusion is neglected.

than those for the case when diffusion is considered in the backward modeling. Remarkably, the upbuilt 'present' state of the plumes in these two cases (with and without diffusion in backward modeling) are very similar in appearance, giving the false impression that reconstructions are satisfactory even with zero diffusion. Our analysis demonstrates that (i) the 'present' structures restored to the past are different for these two cases and (ii) the restoration errors (temperature residuals) are large when diffusion is neglected compared to when diffusion is included in the heat transfer.

6. Discussion

Conduction and convection are two major mechanisms for the transfer of heat. Conductive heat transfer in the mantle is a diffusion process occurring due

to collisions of molecules which transmit their kinetic energies to other molecules. Convective heat transfer is associated with the mantle motion due to buoyancy and plays a dominant part in the general transport of heat from the deep interior of the Earth to the surface. In addition to transport by conduction and convection, a hot material produces blackbody radiation, and heat is diffused if the light emitted by one particle is partially scattered or absorbed by high-frequency transitions in neighboring molecules. However, according to Hofmeister (1999) the radiative contribution is relatively small across the mantle (10–15% of the total thermal conductivity).

If heat diffusion is negligible, the thermal convection in the mantle is time-reversible. "If you have a lot of particles doing something, and then you suddenly reverse the speed, they will completely undo what they did before ... If I reverse the time, the forces are not changed, and so the changes in velocity are not altered at corresponding distances. So each velocity then has a succession of alterations made in exactly the reverse of the way that they were made before, and it is easy to prove that the law of gravitation is time-reversible". With these words, the famous physicist R. Feynman introduced the time reversibility in gravity problems during the Messenger lectures on the character of physical laws he delivered at Cornell University in 1964 (Feynman, 1965).

Conductive heat transfer (heat diffusion) is a more complicated phenomenon. It is practically impossible to collect diffused heat back to the place from where it was diffused. Consider a simple example. If a 'cold' room is heated by a heater installed in the room, it becomes warmer in a few hours period. If the heater is switched off, it is ridiculous to expect that the diffused heat will return back to the heater or we could estimate the initial temperature of the heater from the current room temperature.

Similar processes occur in the Earth. The mantle is heated from the core and from inside due to decay of radioactive elements. Since mantle convection is described by heat advection and diffusion, one can ask: is it possible to tell, from the 'present' temperature estimations of the Earth, something about the Earth's temperature in the geological past?

Even though heat diffusion is irreversible in the physical sense, we can accurately predict the heat transfer backwards in time using the mathematical description of backward heat advection and diffusion without contradicting the basic thermodynamic laws. In this paper we have suggested a numerical method for modeling the backward heat equation in order to solve the inverse problem of thermal convection in the mantle. We do not solve directly the approximate backward heat equation, but rather we search for initial temperature conditions for the approximate forward heat equation.

There is a major physical limitation of the restoration of mantle plumes. If a thermal feature created, let us say, a billion years ago by a boundary layer instability has completely diffused away by the present, it is impossible to restore the feature which was more prominent in the past. The time to which a present thermal structure in the upper mantle can be restored should be restricted by *the characteristic thermal diffusion time*, the time when the temperatures of the evolved structure and the ambient mantle are nearly indistinguishable: $\tau_{\mathrm{diff}} = d^2_{\mathrm{diff}}/36\kappa$, where d_{diff} is the diffusion distance (see Turcotte and Schubert (2002); p. 155, Eq. 4–113 at $T \to T_1$, where T_1 is the ambient temperature). A maximum restoration time is therefore scale dependent, with larger structures being restorable to times further in the past. For a structure the size of the upper mantle thickness ($d_{\mathrm{diff}} = 650$ km), the time of restoration should be limited to about 470 Myr.

A part of the geophysical community may maintain a skepticism about the inverse modeling of thermal convection. This skepticism may partly have its roots in our poor knowledge of the Earth's present structure and its physical properties which cannot allow for rigorous numerical paleoreconstructions of the Earth's evolution. Even considering simplified present-day structure and thermal state of the Earth, the backward modeling of thermomechanical evolution of the Earth is a computational challenge and several numerical problems (e.g., restorations to the distant past, about 400 Myr; more realistic rheology; temperature-dependent thermal diffusivity) should be solved before the technique becomes applicable for whole mantle convection reconstruction. An increase in the accuracy of seismic tomography inversions and geodetic measurements, improvements in the knowledge of gravity and geothermal fields, and more complete experimental data on the physical and chemical properties

of mantle rocks will facilitate mantle reconstructions.

Physicists like to think that all you have to do is say: 'These are the conditions, now what happens next?' (Feynman, 1965), and hence the physicists prefer a forward modeling of phenomena. On the other hand, geologists like to predict a geological evolution based on discoveries on the Earth's surface, and therefore they prefer a modeling backwards in time. In geophysics these two approaches (forward and backward modeling) can be combined using applied mathematics as a tool in numerical modeling of the thermoconvective evolution of the Earth.

We have shown in this paper that a prominent present-day thermal feature in the mantle can be traced back into the geological past. A mathematical model of the thermal convection in the Earth's mantle is described by a set of equations, and we have demonstrated here that the set of equations can be solved numerically backwards in time. Our restoration methodology works well for the mathematical model, and we show its efficiency in the framework of this model.

We have also showed that the suggested method for backward modeling of thermal convection works well for the temperature-dependent viscosity (11). For increased values of the temperature dependence of viscosity (for more than three orders of magnitude viscosity contrast), the inversion scheme might become more sensitive to errors in backtracking the thermal state, and a more accurate inversion scheme might have to be developed.

7. Conclusions

The main motivation for this research comes from the rapid progress made by seismic tomographers in imaging deep Earth structure. Restoration of seismically imaged structures backwards in time could provide an important way to test a range of geodynamic hypotheses. We have suggested a variational approach to the numerical solution of the inverse problem of thermal convection with infinite Prandtl number. We have tested the numerical approach by restoring a model of thermal plumes. The results of the restoration models together with the error estimates demonstrate the practicality of the suggested technique. We have also demonstrated that restored 'present' structures are different when heat diffusion is neglected. The restoration errors (temperature residuals) are large when diffusion is neglect.

The current solution algorithm for the inverse modeling of thermal convection allows us to restore temperature for about a hundred million years into the past based on the knowledge of the present temperature distribution in the mantle. This algorithm does not allow for the thermal restoration of the upper mantle to an age of several hundred million years (within the limit of the characteristic thermal diffusion time). This is associated with a coarseness of the grid used in modeling the heat equation, and we are working on improving the algorithm to allow grid refinement.

In addition to the application of the backward modeling technique to problems of mantle plume and lithospheric slab restorations, the technique can be employed to predict paleotemperatures in sedimentary basins. The temperature estimations in the geological past can help in the forecasting of hydrocarbon generation, maturation, migration, and location in the basins.

The suggested numerical algorithm can be incorporated into many existing mantle convection codes in order to simulate the evolution of mantle structures backwards in time. The methodology opens a new possibility for restoration of mantle plumes, subducting lithosphere, plate movements, and thermoconvective mantle flow in general. Of course, real mantle plumes display more complex patterns and evolution, but our simple models represent an essential step in understanding how mantle plumes (and other mantle structures) might be reconstructed to the past.

Acknowledgements

We are grateful to P. Connolly, B. Naimark, H. Schmeling, and B. Steinberger for useful discussions and comments on the manuscript. Constructive suggestions by two anonymous reviewers are acknowledged. This research was supported by NSF (EAR 01-05945), RFBR (02-01-00354), and the program "Deep Structure of the Earth's Interior" of the Department of Earth Sciences, Russian Academy of Sciences.

Appendix A. On the stability of the solution to the one-dimensional backward diffusion equation

Consider the following boundary-value problem for the one-dimensional backward diffusion equation:

$$\partial u(t,x)/\partial t = \partial^2 u(t,x)/\partial x^2, \quad 0 \le x \le \pi, \quad t \le 0,$$

$$u(t,0) = 0 = u(t,\pi), \quad t \le 0,$$

$$u(0,x) = \phi_n(x), \quad 0 \le x \le \pi.$$

At the initial time we assume that the function $\phi_n(x)$ takes the following two forms:

$$\phi_n(x) = \frac{1}{4n+1}\sin((4n+1)x)$$

and

$$\phi_0(x) \equiv 0.$$

Note that

$$\max_{0\le x\le\pi} |\phi_n(x) - \phi_0(x)| \le \frac{1}{4n+1} \to 0 \quad \text{at } n \to \infty.$$

The following two solutions of the problem correspond to the two chosen functions of $\phi_n(x)$, respectively:

$$u_n(t,x) = \frac{1}{4n+1}\exp(-(4n+1)^2 t)\ \sin((4n+1)x) \quad \text{at } \phi_n(x) = \phi_n$$

and

$$u_0(t,x) \equiv 0 \quad \text{at} \quad \phi_n(x) = \phi_0.$$

At $t = -1$ and $x = \pi/2$ we obtain

$$u_n(-1,\pi/2) = \frac{1}{4n+1}\exp((4n+1)^2) \to \infty \quad \text{at} \quad n \to \infty.$$

At large n two closely set initial functions ϕ_n and ϕ_0 are associated with the two strongly different solutions at $t = -1$ and $x = \pi/2$. Hence, a small error in the initial data can result in very large errors in the solution to the backward problem, and therefore the solution is unstable, and the problem is ill-posed.

Appendix B. Derivation of the gradient of objective functional J

We consider the objective functional defined by (4) and determine the gradient of the functional (see Ismail-Zadeh et al. (2003a) for more details). An increment of the functional can be represented in the form:

$$\begin{aligned} J(\varphi+h) - J(\varphi) &= \int_\Omega |T(\vartheta_2,x;\varphi+h) - \chi(x)|^2 \mathrm{d}x \\ &\quad - \int_\Omega |T(\vartheta_2,x;\varphi) - \chi(x)|^2 \mathrm{d}x \\ &= 2\int_\Omega (T(\vartheta_2,x;\varphi) - \chi(x)) z(\vartheta_2,x)\mathrm{d}x \\ &\quad + \int_\Omega z(\vartheta_2,x)^2 \mathrm{d}x, \end{aligned}$$

where $h(x)$ is a small heat increment to the unknown initial temperature $\varphi(x)$, and $z = T(t,x;\varphi+h) - T(t,x;\varphi)$ is the solution to the following forward heat problem

$$\begin{aligned} &\partial z/\partial t + \boldsymbol{u}\cdot\nabla z - \nabla^2 z = 0, && x \in \Omega, t \in (\vartheta_1,\vartheta_2), \\ &\sigma_1 z + \sigma_2 \partial z/\partial \boldsymbol{n} = 0, && x \in \Gamma, t \in (\vartheta_1,\vartheta_2), \\ &z(\vartheta_1,x) = h(x), && x \in \Omega. \end{aligned} \tag{B.1}$$

We show below that

$$\begin{aligned} &2\int_\Omega (T(\vartheta_2,x;\varphi) - \chi(x)) z(\vartheta_2,x)\mathrm{d}x \\ &\quad = \int_\Omega \Psi(\vartheta_1,x) h(x)\mathrm{d}x, \end{aligned}$$

where $\Psi(t,x) = 2(T(t,x;\varphi) - \chi(x))$ is the solution to the conjugate boundary problem (7). Indeed,

$$\begin{aligned} &\int_\Omega \Psi(\vartheta_2,x) z(\vartheta_2,x)\mathrm{d}x \\ &\quad = \int_\Omega \int_{\vartheta_1}^{\vartheta_2} \frac{\partial}{\partial t}(\Psi(t,x) z(t,x))\,\mathrm{d}x\mathrm{d}t \\ &\qquad + \int_\Omega \Psi(\vartheta_1,x) h(x)\mathrm{d}x. \end{aligned}$$

Considering the fact that $\Psi = \Psi(t,x)$ and $z = z(t,x)$ are the solutions to (7) and (B.1) respectively, and the velocity $\boldsymbol{u}$ satisfies Eq. (3) and the boundary conditions specified, we obtain

$$
\begin{aligned}
&\int_\Omega \int_{\vartheta_1}^{\vartheta_2} \frac{\partial}{\partial t}(\Psi(t,x)z(t,x))\mathrm{d}t\mathrm{d}x \\
&\quad = \int_{\vartheta_1}^{\vartheta_2}\int_\Omega \left\{ \frac{\partial}{\partial t}\Psi(t,x)z(t,x) + \Psi(t,x)\frac{\partial z(t,x)}{\partial t} \right\} \mathrm{d}x\mathrm{d}t \\
&\quad = \int_{\vartheta_1}^{\vartheta_2}\int_\Omega z(t,x)\left[-\boldsymbol{u}\cdot\nabla\Psi - \nabla^2\Psi \right]\mathrm{d}x\mathrm{d}t \\
&\qquad + \int_{\vartheta_1}^{\vartheta_2}\int_\Omega \Psi(t,x)\left[-\boldsymbol{u}\cdot\nabla z + \nabla^2 z \right]\mathrm{d}x\mathrm{d}t \\
&\quad = \int_{\vartheta_1}^{\vartheta_2}\int_\Gamma \{\Psi\ \ \nabla z\cdot\boldsymbol{n} - z\ \ \nabla\Psi\cdot\boldsymbol{n}\}\, d\Gamma\mathrm{d}t \\
&\qquad + \int_{\vartheta_1}^{\vartheta_2}\int_\Omega \{\nabla\Psi\cdot\nabla z - \nabla z\cdot\nabla\Psi\}\,\mathrm{d}x\mathrm{d}t \\
&\qquad + \int_{\vartheta_1}^{\vartheta_2}\int_\Omega \{z\Psi\ \ \nabla\cdot\boldsymbol{u} + \Psi\ \ \boldsymbol{u}\cdot\nabla z \\
&\qquad - \Psi\ \ \boldsymbol{u}\cdot\nabla z\}\,\mathrm{d}x\mathrm{d}t \\
&\qquad - \int_{\vartheta_1}^{\vartheta_2}\int_\Gamma z\Psi\,\boldsymbol{u}\cdot\boldsymbol{n}\, d\Gamma\mathrm{d}t = 0.
\end{aligned}
$$

Hence, we can derive that:

$$
\begin{aligned}
J(\varphi+h) - J(\varphi) &= \int_\Omega \Psi(\vartheta_1,x)h(x)\mathrm{d}x \\
&\quad + \int_\Omega z(\vartheta_2,x)^2\mathrm{d}x \\
&= \int_\Omega \Psi(\vartheta_1,x)h(x)\mathrm{d}x + o(\|h\|).
\end{aligned}
$$

And therefore, we obtain that the gradient of the objective functional is represented as

$$\nabla J(\varphi) = \Psi(\vartheta_1,\cdot).$$

References

Ahlberg, J.H., Nilson, E.H., and Walsh, J.L., 1967. The Theory of Splines and Their Applications. Academic Press, New York, 284 pp.

Ames, K.A., Epperson, J.F., 1997. A kernel-based method for the approximate solution of backward parabolic problems. SIAM J. Numer. Anal. 34 (4), 1357–1390.

Bennett, A.F., 1992. Inverse Methods in Physical Oceanography. Cambridge University Press, Cambridge, 346 pp.

Bunge, H.-P., Hagelberg, C.R., Travis, B.J., 2003. Mantle circulation models with variational data assimilation: Inferring past mantle flow and structure from plate motion histories and seismic tomography. Geophys. J. Int. 152, 280–301.

Busse, F.H., Christensen, U., Clever, R., Cserepes, L., Gable, C., Giannandrea, E., Guillou, L., Houseman, G., Nataf, H.-C., Ogawa, M., Parmentier, M., Sotin, C., Travis, B., 1993. 3D convection at infinite Prandtl number in Cartesian geometry - a benchmark comparison. Geophys. Astrophys. Fluid Dynamics 75, 39–59.

Buzbee, B.L., Carasso, A., 1973. On the numerical computation of parabolic problems for preceding times. Math. Comput. 27, 237–266.

Chandrasekhar, S., 1961. Hydrodynamic and Hydromagnetic Stability. Oxford University Press, Oxford, 654 pp.

Colton, D., 1979. The approximation of solutions to the backwards heat equation in a nonhomogeneous medium. J. Math. Anal. Appl. 72, 418–429.

Conrad, C.P., Gurnis, M., 2003. Seismic tomography, surface uplift, and the breakup of Gondwanaland: Integrating mantle convection backwards in time. Geochem. Geophys. Geosys. 4 (3): doi:10.1029/2001GC000299.

Daescu, D.N., Navon, I.M., 2003. An analysis of a hybrid optimization method for variational data assimilation. Int. J. Comp. Fluid Dyn. 17, 299–306.

Elden, L., 1982. Time discretization in the backward solution of parabolic equations. Math. Comput. 39, 53–84.

Feynman, R., 1965. The Character of Physical Law. BBC, London, 173 pp.

Forte, A.M., Mitrovica, J.X., 2001. Deep-mantle high-viscosity flow and thermochemical structure inferred from seismic and geodynamic data. Nature 410, 1049–1056.

Goes, S., Govers, R., Vacher, P., 2000. Shallow mantle temperatures under Europe from *P* and *S* wave tomography. J. Geophys. Res. 105, 11,153–11,169.

Grand, S.P., van der Hilst, R.D., Widiyantoro, S., 1997. Global seismic tomography a snapshot of convection in the mantle. GSA Today 7, 1–7.

Hadamard, J., 1923. Lectures on the Cauchy Problem in Linear Partial Differential Equations. Yale University Press, New Haven, 316 pp.

Hofmeister, A.M., 1999. Mantle values of thermal conductivity and the geotherm from phonon lifetimes. Science 283, 1699–1706.

Ismail-Zadeh, A.T., 1999. A time-inverse problem of gravitational instability: numerical approach and applications. Geophys. Res. Abstr. 1, 61.

Ismail-Zadeh, A.T., Korotkii, A.I., Tsepelev, I.A., 2003b. Numerical approach to solving problems of slow viscous flow backwards in time. In: K.J. Bathe (Ed.), Computational Fluid and Solid Mechanics. Elsevier, Amsterdam, pp. 938–941.

Ismail-Zadeh, A.T., Talbot, C.J., Volozh, Y.A., 2001a. Dynamic restoration of profiles across diapiric salt structures: numerical approach and its applications. Tectonophysics 337, 21–36.

Ismail-Zadeh, A.T., Korotkii, A.I., Naimark, B.M., Tsepelev, I.A., 2001b. Numerical simulation of three-dimensional viscous flows with gravitational and thermal effects. Comput. Math. & Math. Phys. 41 (9), 1331–1345.

Ismail-Zadeh, A.T., Korotkii, A.I., Naimark, B.M., Tsepelev, I.A., 2003a. Three-dimensional numerical simulation of the inverse problem of thermal convection. Comput. Math. & Math. Phys. 43 (4), 587–599.

Ismail-Zadeh, A.T., Korotkii, A.I., Naimark, B.M., Suetov, A.P., Tsepelev, I.A., 1998. Implementation of a three-dimensional hydrodynamic model for evolution of sedimentary basins. Comput. Math. & Math. Phys. 38 (7), 1138–1151.

Kalnay, E., 2003. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, Cambridge, 341 p.

Kaus, B.J.P., Podladchikov, Y.Y., 2001. Forward and reverse modeling of the three-dimensional viscous Rayleigh-Taylor instability. Geophys. Res. Lett. 28, 1095–1098.

Kirsch, A., 1996. An Introduction to the Mathematical Theory of Inverse Problems. Springer-Verlag, New York, 282 pp.

Korotkii, A.I., Tsepelev, I.A., Ismail-Zadeh, A.T., Naimark, B.M., 2002. Three-dimensional backward modeling in problems of Rayleigh-Taylor instability. Izvestiya/Transactions of the Ural State University 22 (2), 94–102 (in Russian).

Lu, H., 1997. Galerkin and weighted methods for a forward-backward heat equation. Numer. Math. 75, 339–365.

Marchuk, G.I., 1994. Numerical Methods and Applications. CRC Press, Boca Raton, 272 pp.

Moszynski, K., 2001. Approximation with frequency filter for backward parabolic equations. Numer. Math. 88, 159–183.

Ritsema, J., Allen, R.M., 2003. The elusive mantle plume. Earth Planet. Sci. Lett. 207, 1–12.

Schubert, G., Turcotte, D.L., Olson, P., 2001. Mantle Convection in the Earth and Planets. Cambridge University Press, Cambridge, 940 pp.

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

Steinberger, B., O'Connell, R.J., 1997. Changes of the earth's rotation axis owing to advection of mantle density heterogeneities. Nature 387, 169–173.

Steinberger, B., O'Connell, R.J., 1998. Advection of plumes in mantle flow: implications for hotspot motion, mantle viscosity and plume distribution. Geophys. J. Int. 132, 412–434.

Tikhonov, A.N., Arsenin, V.Y., 1977. Solution of Ill-Posed Problems. Winston, Washington, DC, 258 pp.

Tikhonov, A.N., Samarskii, A.A., 1990. Equations of Mathematical Physics. Dover Publications, New York, 765 pp.

Tsepelev, I.A., Korotkii, A.I., Ismail-Zadeh, A.T., Naimark, B.M., 1999. Parallel algorithms for modelling of thermal convection in viscous fluid (in Russian). In: Ulianov, O.N. (Ed.), Algorithms and Software for Parallel Computations. Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Issue 4, pp. 252–279.

Turcotte, D.L., Schubert, G., 2002. Geodynamics. 2nd ed. Cambridge University Press, Cambridge, 456 pp.

Vasiliev, F.P., 2002. Methody optimizatsii (in Russian). Factorial Press, Moscow, 824 pp.

van der Voo, R., Spakman, W., Bijwaard, H., 1999. Mesozoic subducted slabs under Siberia. Nature 397, 246–249.

Three-dimensional forward and backward numerical modeling of mantle plume evolution: Effects of thermal diffusion

Alik Ismail-Zadeh,[1,2,3] Gerald Schubert,[4] Igor Tsepelev,[5] and Alexander Korotkii[5]

Received 14 April 2005; revised 6 February 2006; accepted 22 February 2006; published 8 June 2006.

[1] We investigate the effects of thermal diffusion on the evolution of mantle plumes by means of three-dimensional numerical modeling forward and backward in time. Mantle plumes are fed by a hot, low-viscous material from the thermal boundary layer. The material of the plumes is mainly advected toward the Earth's surface with some effects of thermal diffusion. However, the feeding can become weaker with time, and then thermal diffusion can take over and control the evolution of the plumes. Numerical experiments forward in time show that a week feeding of mantle plumes by the hot material from the boundary layer results in the diffusive disappearance of plume tails first and plume heads later. This is the most likely explanation for the seismically detected low-velocity mantle structures (mantle plumes) with prominent heads and almost invisible tails at midmantle depths. We develop restoration models (backward in time) to recover strong features of mantle plumes in the geological past after they have dissipated due to thermal diffusion and analyze effects of thermal diffusion and temperature-dependent viscosity on the reconstruction of the mantle plumes. We investigate the impact of thermal diffusion on the performance of our restoration (variational data assimilation) algorithm. For a given range of Rayleigh number *Ra* and two values of the viscosity ratio *r* (between the upper and lower boundaries of the model domain) we show that (1) the residuals between the temperature predicted by the forward model and that reconstructed by the backward modeling become larger and (2) the restoration process becomes poorer as *Ra* decreases and *r* increases. We assimilate temperature obtained from high-resolution seismic tomography data for the southeastern Carpathians and show that present diffused mantle structures can be restored to their prominent state in the Miocene times. We discuss the problems of smoothness of model input and output data, errors associated with the modeling, and some other challenges in the data assimilation for thermoconvective flow in the mantle.

Citation: Ismail-Zadeh, A., G. Schubert, I. Tsepelev, and A. Korotkii (2006), Three-dimensional forward and backward numerical modeling of mantle plume evolution: Effects of thermal diffusion, *J. Geophys. Res.*, *111*, B06401, doi:10.1029/2005JB003782.

1. Introduction

[2] Mantle plumes are among the most spectacular features of mass and heat transport from the mantle to the Earth's surface. Thermal plumes in the mantle plausibly originate near either the core-mantle boundary or the upper mantle–lower mantle transition due to instabilities in the hot thermal boundary layers that could exist at these locations. Although some mantle plumes appear to last for more than 150 Myr, they are nonetheless transient features: no tracks older than the Mesozoic are well established [e.g., *Condie*, 2001; *Jellinek and Manga*, 2002]. Direct observational evidence of mantle plumes comes from seismic tomography, which provides constraints on temperature and composition of present mantle structures [e.g., *Ritsema et al.*, 1999; *Montelli et al.*, 2004]. Our understanding of mantle plume dynamics comes from numerical (*Schubert et al.* [2001] provide an overview) and laboratory [e.g., *Davaille*, 1999] experiments.

[3] Numerical models of mantle plume evolution have been mainly carried out forward in time, i.e., from the onset of plumes to late stages of maturity. The main drawback of these models is that the initial conditions (conditions in the geological past) for the models are unknown. However, temperature and flow at the time of plume onset can be inferred from the present mantle temperature and flow using data assimilation based on combined forward and backward numerical modeling of plume evolution. The main motivation for the data assimilation comes from the rapid progress

[1] International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Moscow, Russia.
[2] Geophysikalishes Institut, Universität Karlsruhe, Karlsruhe, Germany.
[3] Institut de Physique du Globe de Paris, Paris, France.
[4] Department of Earth and Space Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California, USA.
[5] Institute of Mathematics and Mechanics, Ural Branch, Russian Academy of Sciences, Yekaterinburg, Russia.

0148-0227/06/2005JB003782$09.00

made by seismic tomographers in imaging deep Earth structure. Restoration of seismically imaged structures backward in time could provide an important way to test a range of geodynamics hypotheses.

[4] Data assimilation is defined as the incorporation of present (observations) and past (initial conditions) data in an explicit dynamical model to provide time continuity and coupling among the physical fields. The basic principle of data assimilation is to consider the initial condition as a control variable and to optimize the initial condition in order to minimize the discrepancy between the observations and the solution of the model. Data related to a thermoconvective mantle flow can be assimilated by using sequential filtering, variational technique, and some others methods [e.g., *Ghil and Malanotte-Rizzoli*, 1991; *Wunsch*, 1996; *Talagrand*, 1997]. In sequential filtering a numerical model is computed forward in time for the interval for which observations have been made, updating the model each time where observations are available. *Bunge et al.* [1998, 2002] used this approach to compute mantle circulation models. Despite sequential data assimilation well adapted to mantle circulation studies, each individual observation influences the model state at later times. Information propagates from the geological past into the future, although our knowledge of the Earth's mantle at earlier times is much poor than that at present.

[5] The use of variational data assimilation in solid Earth dynamics (to estimate initial mantle temperature and flow in the geological past) has been put forward by *Bunge et al.* [2003] and *Ismail-Zadeh et al.* [2003a, 2003b]. This idea is based on a variational technique applied to solve the coupled heat, momentum and continuity equations in order to find the model representation that is most consistent with the observations. That best estimate can then be used to analyze geodynamic processes or initialize a model setup more accurately. *Ismail-Zadeh et al.* [2004] presented a data assimilation algorithm for numerical restoration of a three-dimensional model of present prominent mantle plumes to past stages and showed a high-accuracy in recovering the initial configurations of these plumes. The two major objectives of this study are (1) to estimate effects of thermal diffusion and temperature-dependent viscosity on the evolution of mantle plumes and (2) to recover the structure of mantle plumes prominent in the past from that of present plumes weakened by thermal diffusion.

[6] Conduction and convection are two major mechanisms for the transfer of heat. Conductive heat transfer in the mantle is a diffusion process occurring due to collisions of molecules, which transmit their kinetic energies to other molecules. Convective heat transfer is associated with the mantle motion due to buoyancy and plays a dominant part in the general transport of heat from the deep interior of the Earth to the surface. The thermal conductivity of mantle material depends on pressure and temperature. A model for thermal conductivity in the sublithospheric mantle, based on the experimental study (photon lifetimes obtained from infrared reflectivity) by *Hofmeister* [1999], shows that the thermal conductivity increases with depth from about 2 to 7 W m^{-1} K^{-1}. In addition to transport by conduction and convection, a hot material produces blackbody radiation, and heat is diffused if the light emitted by one particle is partially scattered or absorbed by high-frequency transitions in neighboring molecules. *Badro et al.* [2004] showed experimentally a substantial increase in radiative thermal conductivity in the lower mantle. The change in the radiative conductivity of lower mantle minerals will influence the lower mantle dynamics and plume evolution, because the increase in thermal conductivity results in a decrease of the Rayleigh number and hence in an increase of thermal diffusion.

[7] We briefly describe the model setup and numerical method in section 2 and the variational data assimilation approach to the reconstruction of mantle plumes in section 3 (details of this approach are given by *Ismail-Zadeh et al.* [2004]). We present three-dimensional forward numerical models of mantle plume weakening due to thermal diffusion and analyze the influence of thermal diffusion and viscosity ratio on the evolution of mantle plumes in section 4. These diffused plume structures are then restored to their prominent state in the past, and we analyze the effects of thermal diffusion and viscosity on the reconstruction of mantle plumes in section 5. The efficiency of the data assimilation technique is illustrated in section 6 in terms of the number of iterations required to obtain the target temperature and flow velocity in the past. In section 7.1 we discuss how the numerical results on fading mantle plumes can explain the recent seismic tomography observations of low-velocity anomalies extending down to midmantle depths. We show in section 7.2 the applicability of the numerical reconstruction method (data assimilation approach) to "real" (that is, imaged by seismic tomography) mantle structures, and present conclusions in section 8.

2. Model Problem and Numerical Approach

[8] We study the problem of mantle plume evolution in the three-dimensional model domain $\Omega = [0, x_1 = 3h] \times [0, x_2 = 3h] \times [0, x_3 = h]$, where $\mathbf{x} = (x_1, x_2, x_3)$ are the Cartesian coordinates and h is the depth of the domain. We assume that the mantle behaves as a Newtonian incompressible fluid with a temperature-dependent viscosity and infinite Prandtl number. Rising mantle plumes are modeled as hot fluid jets ascending into the relatively cold ambient fluid heated from below. The mantle flow is described by heat, motion, and continuity equations [*Chandrasekhar*, 1961]. To simplify the governing equations, we make the Boussinesq approximation [*Boussinesq*, 1903] keeping the density constant everywhere except for buoyancy term in the equation of motion. We note that a variable (temperature-dependent) density [*Ismail-Zadeh et al.*, 2003a] and an internal heating [*Bunge et al.*, 2003] can be also used in the forward and backward modeling of thermoconvective mantle circulation. In the Boussinesq approximation the dimensionless equations take the form

$$\partial T/\partial t + \mathbf{u} \cdot \nabla T = \nabla^2 T, \quad t \in (0, \vartheta), \quad \mathbf{x} \in \Omega, \tag{1}$$

$$\nabla P = \mathrm{div}[\eta(T)\mathbf{E}] + RaT\mathbf{e}, \quad \mathbf{E} = \{\partial u_i/\partial x_j + \partial u_j/\partial x_i\}, \quad \mathbf{e} = (0, 0, 1), \tag{2}$$

$$\mathrm{div}\mathbf{u} = 0, \quad t \in (0, \vartheta), \quad \mathbf{x} \in \Omega. \tag{3}$$

Here T, t, $\mathbf{u} = (u_1, u_2, u_3)$, P, and η are dimensionless temperature, time, velocity, pressure, and viscosity, respectively. The Rayleigh number is defined as $Ra = \alpha g \rho_{ref} \Delta T h^3 \eta_{ref}^{-1} \kappa^{-1}$, where α is the thermal expansivity, g is the acceleration due to gravity, ρ_{ref} and η_{ref} are the reference typical density and viscosity, respectively; ΔT is the temperature contrast between the lower and upper boundaries of the model domain; and κ is the thermal diffusivity. In equations (1)–(3), length, temperature, and time are normalized by h, ΔT, and $h^2\kappa^{-1}$, respectively.

[9] At the boundary of the model domain we set the impenetrability condition with perfect slip conditions: $\partial \mathbf{u}_\tau / \partial \mathbf{n} = 0$, $\mathbf{u} \cdot \mathbf{n} = 0$, where $\mathbf{n}$ is the outward unit normal vector at a point on the model boundary, and $\mathbf{u}_\tau$ is the projection of the velocity vector onto the tangent plane at the same point on the model boundary. We assume zero heat flux through the vertical boundaries of the box. The upper and lower boundaries are isothermal surfaces, and we set $T = 0$ and $T = 1$ at these boundaries, respectively.

[10] Equations (1)–(3) together with the boundary conditions describe a thermoconvective mantle flow. To solve the problem forward or backward in time we assume the temperature to be known at the time of plume onset ($t = 0$) or at the present time ($t = \vartheta$).

[11] Temperature in the heat equation (1) is approximated by finite differences and determined by the semi-Lagrangian method, which allows for relatively large time steps, high accuracy, and low numerical diffusion [*McDonald*, 1984]. A numerical solution to the Stokes equations (2) is based on the introduction of a two-component vector velocity potential and on the application of the Eulerian finite element method with a tricubic-spline basis for computing the potential [*Ismail-Zadeh et al.*, 2001]. Such a procedure results in a set of linear algebraic equations with a symmetric positive-definite banded matrix. We solve the set of equations by the conjugate gradient method [*Fletcher and Reeves*, 1964]. The numerical algorithm was designed to be implemented on parallel computers. The reader is referred to *Ismail-Zadeh et al.* [2001, 2004] for more detail.

3. Variational Data Assimilation

[12] Data assimilation techniques has been pioneered by meteorologists and used very successfully to improve operational weather forecasts [e.g., *Kalnay*, 2003]. Data assimilation has also been widely used in oceanography [e.g., *Bennett*, 1992] and in hydrological studies [e.g., *McLaughlin*, 2002]. However, the application of the method to problems of mantle dynamics is still in its infancy.

[13] The variational data assimilation is based on a search of the best fit between the forecast model state and the observations by minimizing an objective functional (a normalized residual between the target model and observed variables) over space and time. To minimize the objective functional over time, an assimilation time interval is defined and an adjoint model is typically used to find the derivatives of the objective functional with respect to the model states. The variational data assimilation is well suited for smoothing problems (we discuss the problem of smoothness of the initial data and solution in Appendix A).

[14] The method for variational data assimilation can be formulated with a weak constraint where errors in the model formulation are taken into account as control parameters (generalized inverse) [*Bunge et al.*, 2003] or with a strong constraint where the model is assumed to be perfect except for the errors associated with the initial conditions [*Bunge et al.*, 2003; *Ismail-Zadeh et al.*, 2003a]. There are several sources of errors in forward and backward modeling of thermoconvective mantle flow, which we discuss in Appendix B. The generalized inverse of mantle convection considers model errors, data misfit and the misfit of parameters as control variables. Unfortunately the generalized inverse presents a tremendous computational challenge and is difficult to solve in practice. Hence *Bunge et al.* [2003] considered a simplified generalized inverse imposing a strong constraint on errors (ignoring all errors except for the initial condition errors). Therefore the strong constraint makes the problem computationally tractable.

[15] We consider the following objective functional $J(\varphi) = \|T(\vartheta, \cdot\ ; \varphi) - \chi(\cdot)\|^2$, where parallels denote the norm in the space L_2 (Ω) (the Hilbert space with the norm defined as $\|y\| = [\int_\Omega y^2(\mathbf{x}) d\mathbf{x}]^{1/2}$). Since in what follows the dependence of solutions of the thermal boundary value problems on initial data is important, we introduce these data explicitly into the mathematical representation of temperature. Here $T(\vartheta, \cdot\ ; \varphi)$ is the solution of the thermal boundary value problem (1) at the final time ϑ, which corresponds to some (unknown as yet) initial temperature distribution $\varphi(\mathbf{x})$; $\chi(\mathbf{x}) = T(\vartheta, \mathbf{x}; T_0)$ is the known temperature distribution at the final time, which corresponds to the initial temperature $T_0(\cdot)$. The functional has its unique global minimum at value $\varphi \equiv T_0$ and $J(T_0) \equiv 0$, $\nabla J(T_0) \equiv 0$. To find the minimum of the functional we employ the gradient method ($k = 0, \ldots, j, \ldots$):

$$\varphi_{k+1} = \varphi_k - \beta_k \nabla J(\varphi_k), \tag{4}$$

$$\beta_k = \min\{1/(k+1), J(\varphi_k)/\|\nabla J(\varphi_k)\|\}, \quad \varphi_0 = T_*, \tag{5}$$

where T_* is an initial temperature guess. The minimization method belongs to a class of limited-memory quasi-Newton methods [*Zou et al.*, 1993], where approximations to the inverse Hessian matrices are chosen to be the identity matrix. The gradient of the objective functional $\nabla J(\varphi_k)$ decreases steadily with the number of iterations, and it provides the convergence of the method. Meanwhile the absolute value of β_k increases with the number of iterations, and it can result in instability of the iteration process [*Samarskii and Vabischevich*, 2004]. To avoid the instability, we use equation (5) to minimize the parameter β_k.

[16] The minimization algorithm requires the calculation of the gradient of the objective functional, ∇J. This can be done through the use of the adjoint problem for the model equations (1)–(3) with the relevant boundary and initial conditions. In the case of the heat problem, the adjoint problem can be represented in the following form:

$$\begin{aligned} &\partial Z/\partial \tau - \mathbf{u} \cdot \nabla Z = \nabla^2 Z, \qquad \tau = \vartheta - t \in (-\vartheta, 0), \\ &Z(0, \mathbf{x}) = 2(T(\vartheta, \mathbf{x}; \varphi) - \chi(\mathbf{x})), \qquad \mathbf{x} \in \Omega, \end{aligned} \tag{6}$$

with uniform boundary conditions. The solution $Z(\vartheta, \cdot)$ to this adjoint problem is the gradient of the objective functional, and

the gradient is derived by using the Fréchet derivative of the functional [see *Ismail-Zadeh et al.*, 2004, Appendix B]. The correctness of the solution has been verified by the gradient accuracy test [*Navon et al.*, 1992, equation 2.20].

[17] We define a uniform partition of the time axis at points $t_n = \vartheta - n\delta t$, where δt is the time step, and n successively takes integer values from 0 to some natural number $m = \vartheta/\delta t$. At each subinterval of time $[t_{n+1}, t_n]$, the search of the temperature T and flow velocity $\mathbf{u}$ at $t = t_{n+1}$ consists of the following basic steps.

[18] 1. Given the temperature $T = T(t_n, \mathbf{x})$ at $t = t_n$ we solve a set of linear algebraic equations derived from equations (2) and (3) with the appropriate boundary conditions in order to determine the velocity $\mathbf{u}$.

[19] 2. The "advective" temperature $T_{adv} = T_{adv}(t_{n+1}, \mathbf{x})$ is determined by solving the advection heat equation backward in time, neglecting the diffusion term in equation (1). This can be done by replacing positive time steps by negative ones [see *Ismail-Zadeh et al.*, 2003b].

[20] Given the temperature $T = T_{adv}$ at $t = t_{n+1}$ steps 1 and 2 are then repeated to find the velocity $\mathbf{u}_{adv} = \mathbf{u}(t_{n+1}, \mathbf{x}; T_{adv})$.

[21] 3. The heat equation (1) is solved with appropriate boundary conditions and initial condition $\varphi_0(\mathbf{x}) = T_{adv}(t_{n+1}, \mathbf{x})$ forward in time using velocity $\mathbf{u}_{adv}$ in order to find $T(t_n, \mathbf{x}; \varphi_0)$.

[22] 4. The adjoint equation (6) is then solved backward in time with appropriate boundary conditions and initial condition $T(t_n, \mathbf{x}) = T(t_n, \mathbf{x}; \varphi_0)$ using velocity $\mathbf{u}$ in order to determine $\nabla J(\varphi_0)$.

[23] 5. The coefficient β_0 is determined from equation (5), and the temperature is updated (i.e., φ_1 is determined) from equation (4).

[24] Steps 3 to 5 are repeated for φ_j and β_j ($j = 1, 2, 3, \ldots$) until $\delta\varphi_j = J(\varphi_j) + \|\nabla J(\varphi_j)\|^2 < \varepsilon$, where ε is a small constant. Temperature φ_j is then considered to be the approximation to the target value of the initial temperature $T(t_{n+1}, \mathbf{x})$. Finally, step 1 is used to determine the flow velocity $\mathbf{u}(t_{n+1}, \mathbf{x}; T(t_{n+1}, \mathbf{x}))$.

[25] Step 2 introduces a preconditioner to accelerate the convergence of temperature iterations in steps 3 to 5 at high Rayleigh number. At low Ra, step 2 is omitted and $\mathbf{u}_{\mathrm{adv}}$ is replaced by $\mathbf{u}$.

4. Forward Modeling of Mantle Plume Diffusion

[26] Mantle plumes evolve in three distinguishing stages: (1) immature, i.e., an origin and initial rise of the plumes; (2) mature, i.e., plume-lithosphere interaction, gravity spreading of plume head and development of overhangs beneath the bottom of the lithosphere, and partial melting of the plume material [e.g., *Ribe and Christensen*, 1994; *Moore et al.*, 1998]; and (3) overmature, i.e., slowing down of the plume rise and fading of the mantle plumes due to thermal diffusion [*Davaille and Vatteville*, 2005]. The ascent and evolution of mantle plumes depend on the properties of the source region (that is, the thermal boundary layer) and the viscosity and thermal diffusivity of the ambient mantle. The properties of the source region determine temperature and viscosity of the mantle plumes. Structure, flow rate, and heat flux of the plumes are controlled by the properties of the mantle through which the plumes rise. While properties of the lower mantle (e.g., viscosity, thermal conductivity) are relatively constant during about 150 Myr lifetime of most plumes, source region properties can vary substantially with time as the thermal basal boundary layer feeding the plume is depleted of hot material. Complete local depletion of this boundary layer cuts the plume off from its source. It is the subsequent evolution of the plume that interests us here.

[27] We study only the late stage of the mantle plume evolution associated with the fading of the plume due to thermal diffusion and model the evolution of mantle plumes deprived of source material through numerical experiments of three-dimensional thermal convection in a bottom heated box. The mantle behaves as a Newtonian fluid on geological timescales, and a dimensionless temperature-dependent viscosity law [*Busse et al.*, 1993] given by

$$\eta(T) = \exp\left(\frac{M}{T+G} - \frac{M}{0.5+G}\right)$$

is used in the modeling, where $M = [225/\ln(r)] - 0.25\ln(r)$, $G = 15/\ln(r) - 0.5$ and r is the viscosity ratio between the upper and lower boundaries of the model domain. We model the evolution of mantle plumes for two viscosity profiles: $r = 20$ and $r = 200$. The temperature-dependent viscosity profile has its minimum at the core-mantle boundary. A more realistic viscosity profile [e.g., *Forte and Mitrovica*, 2001] will influence the evolution of mantle plumes, though it will not influence the restoration of the plumes. The model domain is divided into $37 \times 37 \times 29$ rectangular finite elements to approximate the vector velocity potential by tricubic splines, and a uniform grid $112 \times 112 \times 88$ is employed for approximation of temperature, velocity, and viscosity.

[28] Initially, we model the evolution of mature mantle plumes to obtain initial temperature data for models of mantle plume diffusion. With $\alpha = 3 \times 10^{-5}$ K^{-1}, $\rho_{ref} = 4000$ kg m^{-3}, $\Delta T = 3000$ K, $h = 2800$ km, $\eta_{ref} = 8 \times 10^{22}$ Pa s, and $\kappa = 10^{-6}$ m^{-2} s^{-1}, the initial Rayleigh number is $\mathrm{Ra} = 9.5 \times 10^5$. While plumes evolve in the convecting heterogeneous mantle, at the initial time we assume that the plumes develop in a laterally homogeneous temperature field and hence consider that the mantle temperature in the model increases linearly with depth.

[29] Mantle plumes are generated by random temperature perturbations at the top of the thermal source layer associated with the core-mantle boundary (Figure 1a). The mantle material in the basal source layer flows horizontally toward the plumes. The reduced viscosity in this basal layer promotes the flow of the material to the plumes. Vertical upwelling of hot mantle material is concentrated in low-viscosity conduits near the centerlines of the emerging plumes (Figures 1b and 1c). The plumes move upward through the model domain, gradually forming structures with well-developed heads and tails. The plumes diminish in size with time (Figure 1d), and the plume tails disappear before the plume heads (Figures 1e and 1f). We note that the figures present a hot isothermal surface of the plumes. If colder isotherms are considered, the disappearance of the isotherms will occur later. However, anyhow, hot or cold isotherms are plotted, plume tails will vanish before their heads. Results of recent laboratory experiments [*Davaille and Vatteville*, 2005] support strongly our numerical find-

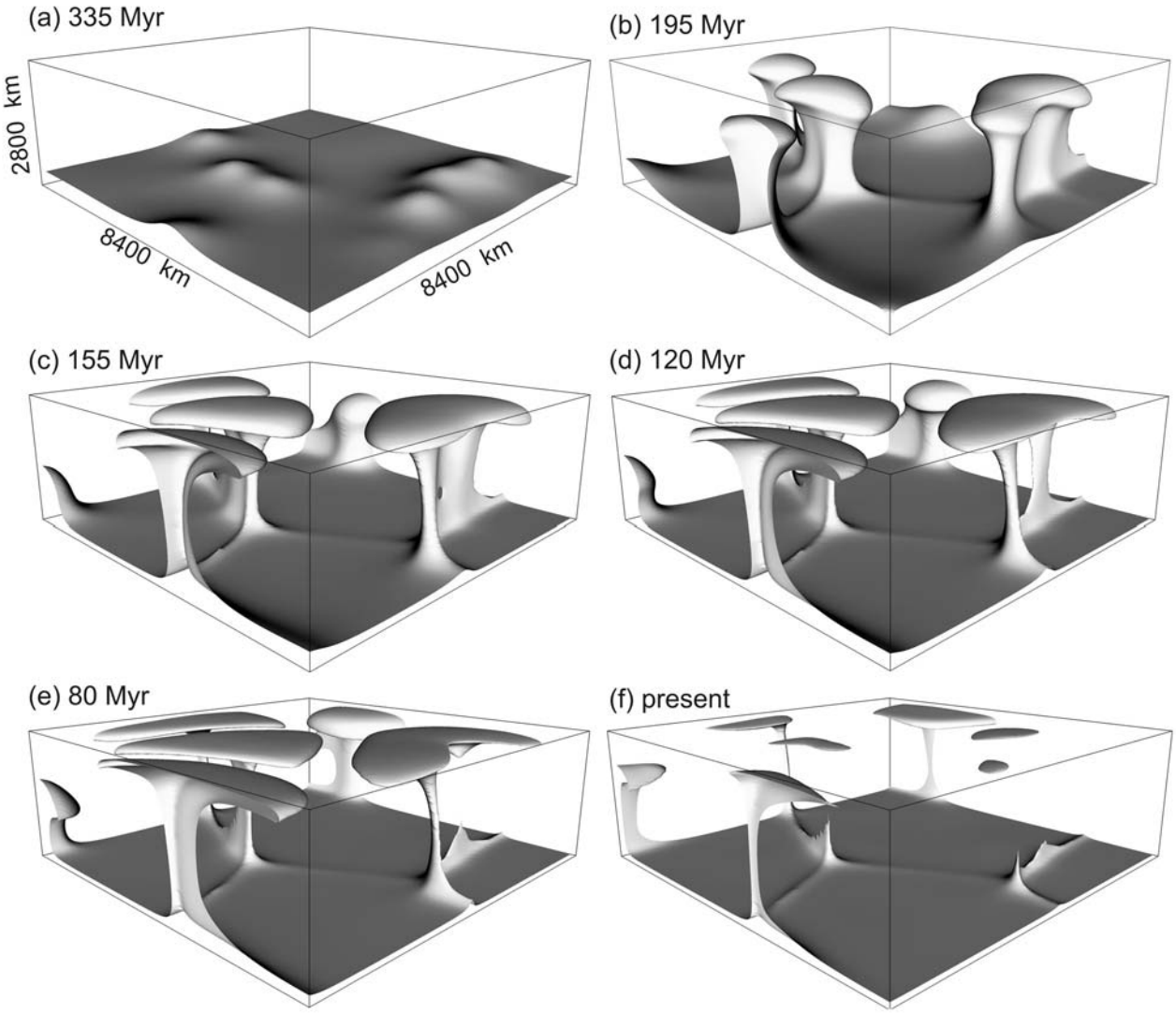

Figure 1. Mantle plumes in the forward modeling at successive diffusion times: from (a) 335 Myr ago to (f) the "present" state of the plumes. The plumes are represented here and in Figures 2 and 3 by isothermal surfaces at 3000 K.

ings that plumes start disappearing from bottom up and fade away by thermal diffusion.

[30] At different stages in the plume decay one sees quite isolated plume heads, plume heads with short tails, and plumes with nearly pinched off tails. Different amounts of time are required for different mantle plumes to vanish into the ambient mantle, the required time depending on the geometry of the plume tails. Temperature loss is greater for sheet-like tails than for cylindrical tails. The tails of the cylindrical plumes (e.g., Figure 1c, in the left part of the model domain) are still detectable after about 155 Myr. However, at this time the sheet-like tail of the large plume in the right part of the model domain (Figure 1c) is already invisible and only its head is preserved in the uppermost mantle (Figure 1f). Two-dimensional numerical experiments of steady state convection [*Leitch et al.*, 1996] reveal a significant change in the centerline temperature of sheet-like plume tails compared to the cylindrical plume tail due to heat conduction in the horizontal direction.

5. Recovering Prominent Mantle Plumes From Their Weakened Present Stage

[31] We use the numerical approach described in section 3 to reconstruct the prominent state of the plumes (Figure 1d) in the past from their "present" weak state (Figure 1f). Figure 2 illustrates the reconstructed states of the plumes (Figures 2e–2g) and the temperature residuals δT (Figures 2h–2j) between the temperature $T(\mathbf{x})$ predicted by the forward model and the temperature $\widetilde{T}(\mathbf{x})$ reconstructed to the same age:

$$\delta T(x_1, x_2) = \left[\int_0^h \left(T(x_1, x_2, x_3) - \widetilde{T}(x_1, x_2, x_3) \right)^2 dx_3 \right]^{1/2}.$$

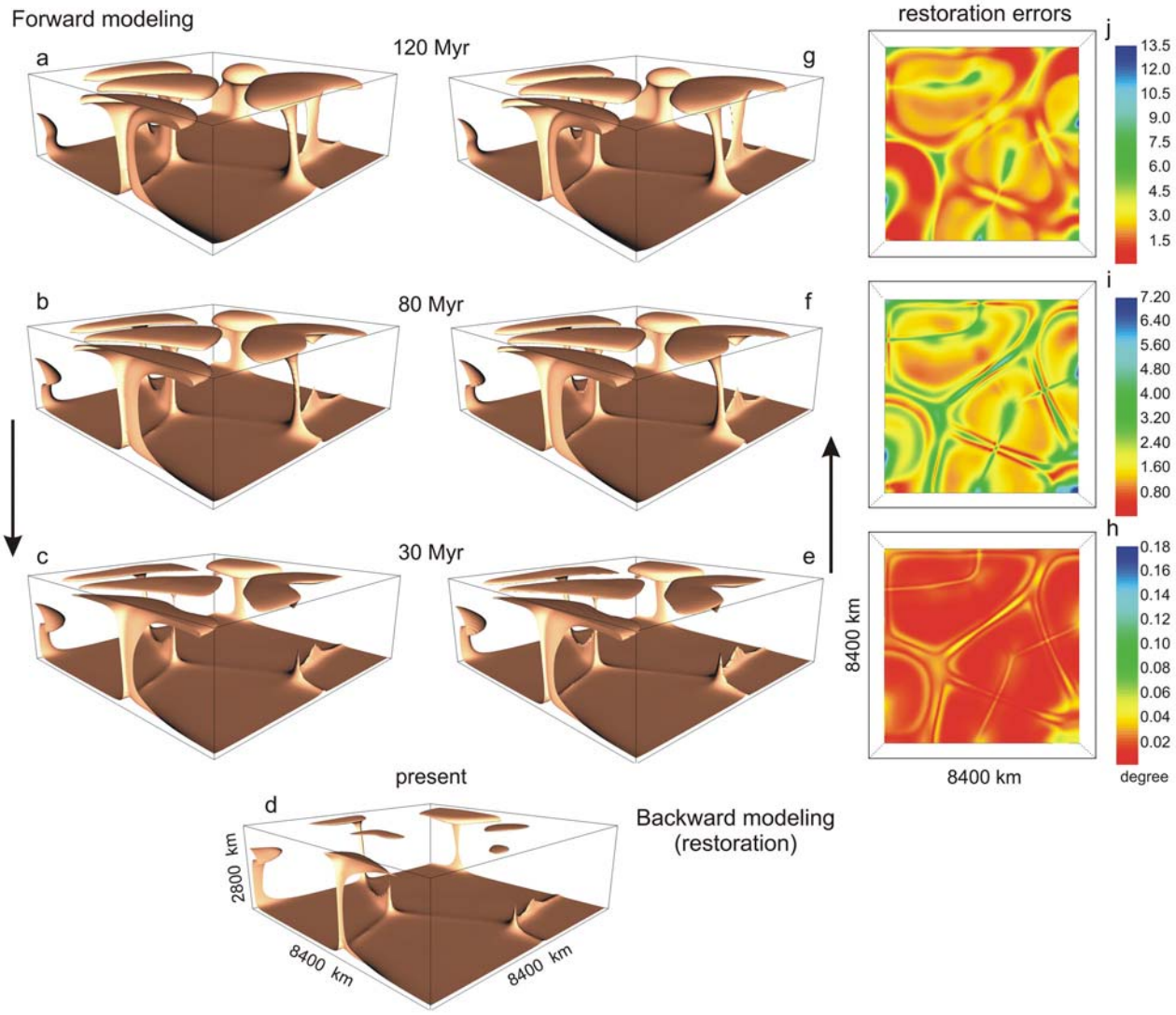

Figure 2. Mantle plume diffusion (r = 20 and $Ra = 9.5 \times 10^5$) in the forward modeling at successive diffusion times: (a–d) from 120 Myr ago to the "present" state of the plumes. (e–g) Restored mantle plumes in the backward modeling and (h–j) restoration errors.

[32] To study the effect of thermal diffusion on the restoration of mantle plumes, we develop several independent experiments on mantle plume restoration assigning several Ra values less than the initial Ra by one to 3 orders of magnitude at two values of viscosity ratio r. Figure 3 presents the case of r = 200 and $Ra = 9.5 \times 10^3$ and shows several stages in the diffusive decay of the mantle plumes.

[33] The dimensional temperature residuals are within a few degrees for the initial restoration period (Figures 2i and 3h). The computations show that the errors (temperature residuals) get larger the farther the restorations move backward in time (e.g., $\delta T \approx 300$ K at the restoration time of more than 300 Myr, $r = 200$, and $Ra = 9.5 \times 10^3$). Compared to the case of $Ra = 9.5 \times 10^5$, one can see that the residuals become larger as the Rayleigh number decreases or thermal diffusion increases and viscosity ratio increases.

[34] We introduce the critical temperature residual $\delta T_{cr} = 0.2\ \Delta T$ such that the quality of mantle structure recovery is estimated to be bad if $\delta T > \delta T_{cr}$. The quality of the restoration depends on the dimensionless Peclet number $Pe = hu_{\max}\kappa^{-1}$, where $u_{\max}$ is the maximum flow velocity. According to the numerical experiments, the Peclet number corresponding to the critical temperature residual δT_{cr} = 600 K is Pe = 10; Pe should not be less than about 10 for a high-quality plume restoration.

[35] In numerical experiments backward in time we observe an increase in the noise of the restored temperatures with time. *Samarskii et al.* [1997] studied a one-dimensional backward heat diffusion problem and showed that the solution to this problem becomes noisy if the initial temperature guess is slightly perturbed, and the amplitude of this noise increases with the initial perturbations of the temperature guess. They suggest using a special filter to reduce the noise and illustrate the efficiency of the filter. This filter is based on the replacement of iterations (4) by the following iterative scheme:

$$\mathbf{B}(\varphi_{k+1} - \varphi_k) = -\beta_k \nabla J(\varphi_k), \tag{7}$$

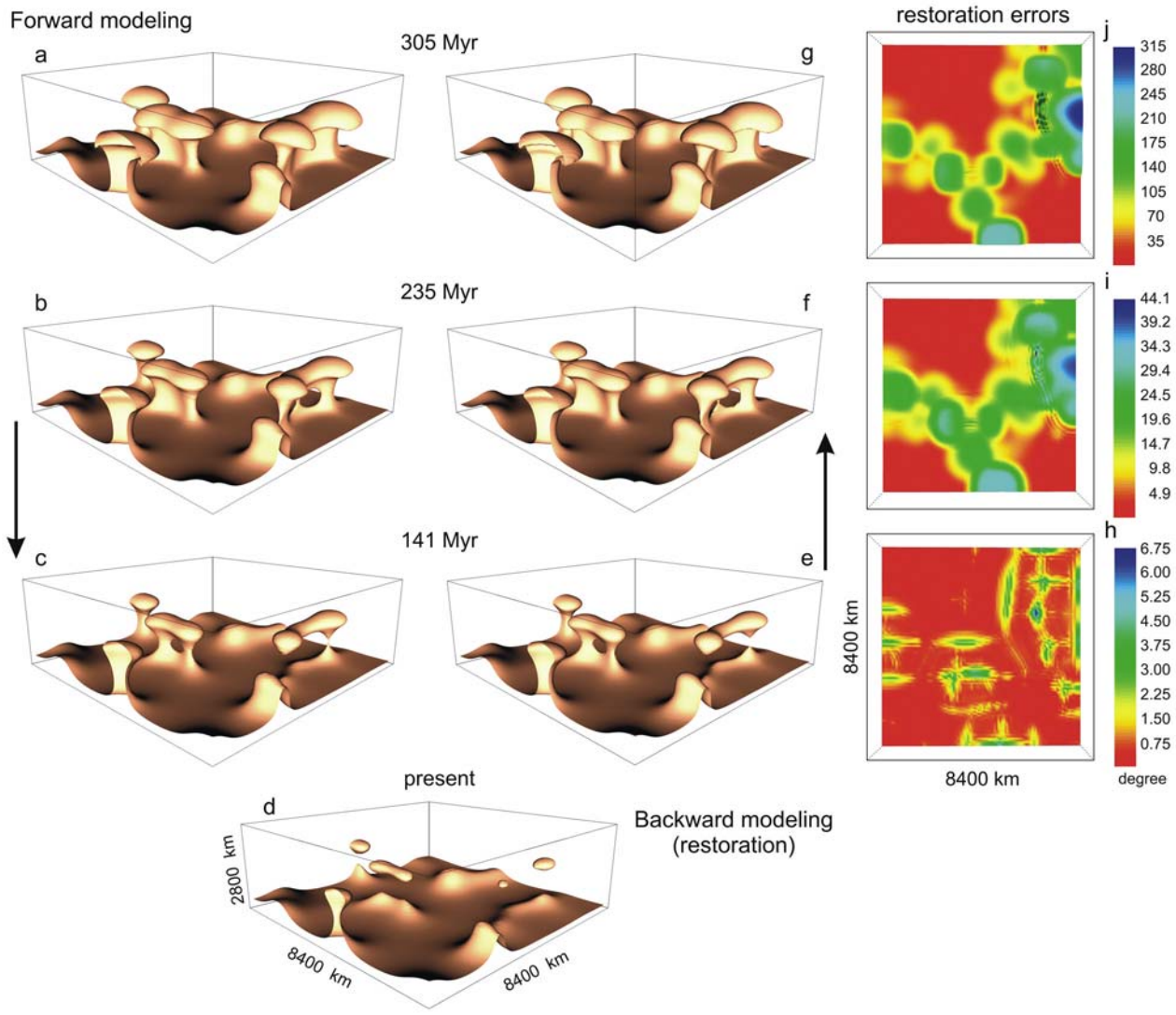

Figure 3. Mantle plume diffusion ($r = 200$ and $Ra = 9.5 \times 10^3$) in the forward modeling at successive diffusion times: (a–d) from 305 Myr ago to the "present" state of the plumes. (e–g) Restored mantle plumes in the backward modeling and (h–j) restoration errors.

where $\mathbf{B}y = y - \nabla^2 y$. Unfortunately, employment of this filter increases the number of iterations to obtain the target temperature and it becomes quite expensive computationally, especially when the model is three-dimensional. Therefore our approach to this problem is to run the model backward to the point of time when the noise becomes relatively large.

6. Performance of the Numerical Algorithm

[36] Here we investigate the impact of diffusion on the performance of our restoration algorithm for various Ra and r. The performance of the algorithm is evaluated in terms of the number of iterations n required to achieve a prescribed relative reduction of $\delta\varphi_n$. Figure 4 presents the evolution of the objective functional $J(\varphi_n)$ and the norm of the gradient of the objective functional $\|\nabla J(\varphi_n)\|$ versus the number of iterations at time about 0.5θ. For other time steps we observe a similar evolution of J and $\|\nabla J\|$.

[37] Both the objective functional and the norm of its gradient show a quite rapid decrease after about 7 iterations for $Ra = 9.5 \times 10^5$ and $r = 20$ (curves 1). The same rapid convergence as a function of adjoint iterations is observed in the *Bunge et al.* [2003] case. As Ra decreases and thermal diffusion increases (curves 2–4) the performance of the algorithm becomes poor: more iterations are needed to achieve the prescribed ε. All curves illustrate that the first 4 to 7 iterations contribute mainly to the reduction of $\delta\varphi_n$. The convergence drops after a relatively small number of iterations. The curves approach the horizontal line with an increase in the number of iterations, because β_k tends to zero with a large number of iterations (see equation (5)). The increase of $\|\nabla J\|$ at $k = 2$ is associated with uncertainty of this gradient at $k = 1$.

[38] Implementation of minimization algorithms requires the evaluation of both the objective functional and its gradient. Each evaluation of the objective functional requires an integration of the model equation (1) with the

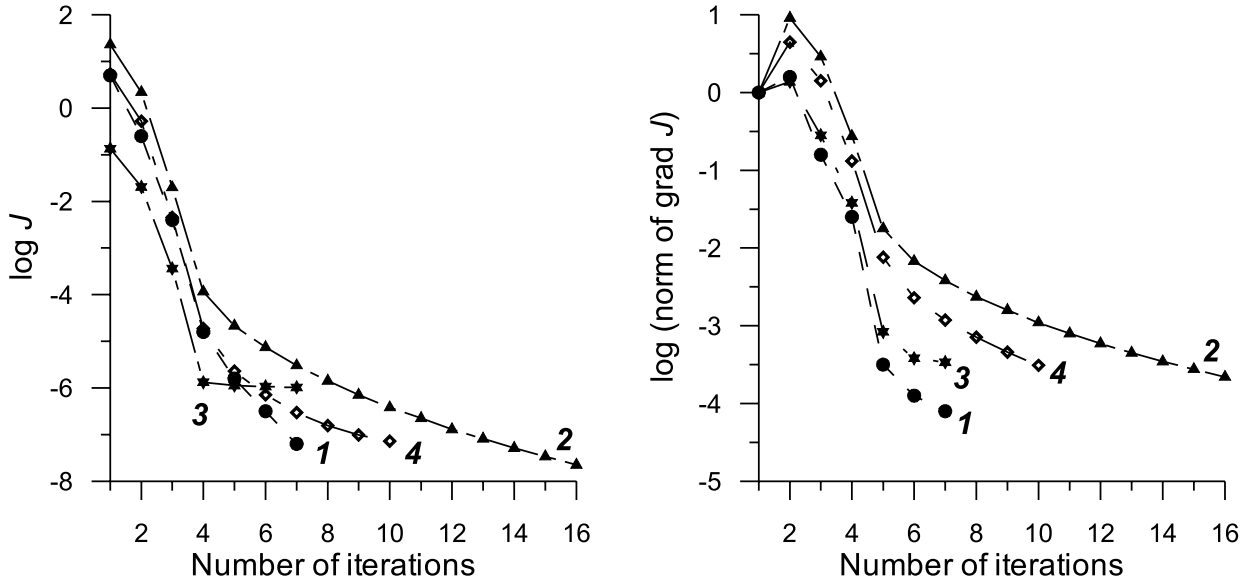

Figure 4. Relative reductions of (left) the objective functional J and (right) the norm of the gradient of J as functions of the number of iterations. Curves indicate 1, $r = 20$, $Ra = 9.5 \times 10^5$; 2, $r = 20$, $Ra = 9.5 \times 10^2$; 3, $r = 200$, $Ra = 9.5 \times 10^3$; and 4, $r = 200$, $Ra = 9.5 \times 10^2$.

appropriate boundary and initial conditions, whereas the gradient is obtained through the backward integration of the adjoint equations (6). The performance analysis shows that the CPU time required to evaluate the gradient J is about the CPU time required to evaluate the objective functional itself, and this is because the direct and adjoint heat problems are described by the same equations.

[39] Despite its simplicity, the minimization algorithm used in this study provides for a rapid convergence and good quality of optimization at high Rayleigh numbers (low thermal diffusion). The convergence rate and the quality of optimization become worse with the decreasing Rayleigh number. The use of the limited-memory quasi-Newton algorithm L-BFGS [*Liu and Nocedal*, 1989] might provide for a better convergence rate and quality of optimization [*Zou et al.*, 1993]. Meanwhile, we note that although an improvement of the convergence rate by using another minimization algorithm (e.g., L-BFGS) will reduce the computational expense associated with the solving of the problem under question, this reduction would be not significant, because the large portion (about 70%) of the computer time is spent to solve the three-dimensional (3-D) Stokes equations.

7. Discussion

7.1. Mantle Plume Heads Yes, Tails No

[40] A plume is hot, narrow mantle upwelling that is invoked to explain hot spot volcanism. In a temperature-dependent viscosity fluid such as the mantle, a plume is characterized by a mushroom-shaped head and a thin tail. Upon impinging under a moving lithosphere, such a mantle upwelling should therefore produce a large amount of melt and successive massive eruption, followed by smaller but long-lived hot spot activity fed from the plume tail [*Morgan*, 1972; *Richards et al.*, 1989; *Sleep*, 1990]. Meanwhile, slowly rising plumes (a buoyancy flux of less than 10^3 kg s^{-1}) coming from the core-mantle boundary should have cooled so much that they would not melt beneath old lithosphere [*Albers and Christensen*, 1996].

[41] A mantle plume is a well-established geological structure in computer modeling and laboratory experiments. Numerical experiments on dynamics of mantle plumes [*Trompert and Hansen*, 1998; *Zhong*, 2005] showed that the number of plumes increases and the rising plumes become thinner with an increase in Rayleigh number. Disconnected thermal plume structures appear in thermal convection at Ra greater than 10^7 [*Hansen et al.*, 1990; *Malevsky et al.*, 1992]. At high Ra (in the hard turbulence regime) thermal plumes are torn off the boundary layer by the large-scale circulation or by nonlinear interactions between plumes [*Malevsky and Yuen*, 1993]. Plume tails can also be disconnected when the plumes are tilted by plate scale flow [e.g., *Olson and Singer*, 1985; *Steinberger and O'Connell*, 1998]. Here we discuss an alternative mechanism for the disconnected mantle plume heads and tails.

[42] Mantle plumes are generated at the top of the thermal boundary layer (TBL), which is produced by conductive heating of the material at the core-mantle boundary (or upper and lower mantle boundary). When the TBL becomes unstable, any perturbation of the TBL top leads to upwelling. Injection of hot material from the source TBL layer into the colder mantle generates strong plumes that are fed for a while from the layer. Colder material overlying the source layer (e.g., portions of lithospheric slabs subducted to the core-mantle boundary) replaces hot material at the locations where the source material is fed into mantle plumes. Some time is required to recover the volume of source material depleted due to plume feeding [*Howard*, 1966]. Because the volume of upwelling material is comparable to the volume of the TBL feeding the mantle plumes, hot material could eventually be exhausted, and mantle plumes would be starved thereafter.

[43] We evaluate the volume V_p of source material that moves into a single plume from the core-mantle boundary

Table 1. Model Parameters and Values

Symbol	Parameter	Value
h	depth of domain, km	2800
g	acceleration due to gravity, m s^{-2}	9.8
P	pressure, Pa	
r	viscosity ratio	20, 200
$t \in [0, \vartheta]$	time, years	
T	temperature, K	
T_{ref}	reference temperature, K	3270
T_{surf}	surface temperature, K	270
T_0	initial dimensionless temperature	
$T*$	initial dimensionless temperature guess	
$\Delta T = T_{\text{ref}} - T_{\text{surf}}$	temperature drop, K	3000
δT	dimensionless temperature residual	
$\mathbf{u} = (u_1, u_2, u_3)$	velocity, cm yr^{-1}	
Pe	Peclet number	
Ra	Rayleigh number	
α	thermal expansivity, K^{-1}	3×10^{-5}
χ	dimensionless temperature at time $t = \vartheta$	
κ	thermal diffusivity, m^2 s^{-1}	10^{-6}
η	viscosity, Pa s	
η_{ref}	reference viscosity, Pa s	8×10^{22}
ρ_{ref}	reference density, kg m^{-3}	4000

over the time interval t_p (required by the plume to reach the base of the lithosphere), and compare that with the volume V_{TBL} of the TBL material conductively generated over the same time interval t_p. For plume height h_p = 2600 km and tail radius r_p ranging from 100 to 200 km, the volume $V_p = \pi r_p^2 h_p$ is estimated to be 0.8 to 3.3 $\times$ 10^8 km^3.

[44] The velocity of plume upwelling

$$w_p = \left(\frac{\alpha g Q}{4\pi c \eta_p} \right)^{1/2}$$

can be estimated analytically from a solution to the boundary layer equations for the steady state flow above a source of heat in a fluid whose viscosity is a temperature-dependent [*Olson et al.*, 1993]. The velocity w_p and the time t_p ($=h_p/w_p$) depend on the plume viscosity η_p as a function of depth and the heat flux $Q = Bc/\alpha$, where B is the buoyancy flux of the plume and c is specific heat. For the typical mantle values given in section 4 (Table 1), B = 3000 to 6000 kg s^{-1}, and $\eta_p = 10^{20}$ to 10^{21} Pa s, the volume $V_{\text{TBL}} = \frac{4}{3}\pi[(r_c + \delta r)^3 - r_c^3]$ (r_c is the radius of the Earth's core, and

$$\delta r = \left(\pi \kappa t_p\right)^{1/2} = \left[\frac{4\pi^3 \kappa^2 h_p^2 \eta}{gB} \right]^{1/4}$$

is the TBL thickness) would range from about 6.3 $\times$ 10^8 to 1.4 $\times$ 10^9 km^3 for the time range of 11 to 48 Myr and the δr range of about 35 to 70 km.

[45] The seismic tomography study [*Montelli et al.*, 2004] has revealed 32 present mantle plumes with radii ranging from 100 to 400 km. Even if only half of the seismically imaged plumes are assumed to have deep mantle roots, we can conclude that the material of the TBL is insufficient to simultaneously feed them. This suggests that only a few mantle plumes can be fed from the TBL at any time and that other plumes are in a phase of thermal diffusive decay. While the discrimination of low-velocity anomalies (seen in seismic tomography models) in active and less active plumes is a challenging problem, laboratory and numerical experiments can provide us with the information.

[46] Recent laboratory experiments on convective instabilities in a layer of fluid with temperature-dependent viscosity and heated from below have shown the generation and evolution of thermal plumes and the transient features of the plumes [*Davaille and Vatteville*, 2005; *Silveira et al.*, 2006]. The temperature difference applied at the lower boundary was chosen such that the Rayleigh number is comparable to that of the Earth's mantle. Initially, a TBL forms at the hot boundary, its thickness increasing by diffusion. When the local Rayleigh number based on the TBL thickness reaches a critical value, the TBL becomes unstable and breaks up to produce plumes [*Howard*, 1966]. A plume reaches the top boundary and spreads laterally. Once the hot TBL has been emptied, the plume tail begins to disappear from the bottom up, leaving only the cooling and shrinking sublithospheric overhangs. The cycle of plume development repeats once the critical thickness of the TBL is reached. The analogue experiments have shown that the mean velocity of the fluid decreases with the maturity of the plumes and hence thermal diffusion becomes a major agent in the heat transfer.

[47] Our numerical results on the diffusive decay of mantle plumes with depleted source regions are in a good agreement with the results of the laboratory experiments. They may have important implications for the interpretation of seismic tomographic images of mantle plumes. Finite frequency seismic tomography images [*Montelli et al.*, 2004] show that a number of plumes extend to midmantle depths but are not visible below these depths. From seismological point of view, the absence of the plume tails could be explained as a combination of several factors [*Romanowicz and Gung*, 2002]: elastic velocities are sensitive to composition as well as temperature; the effect of temperature on velocities decreases with increasing pressure [*Karato*, 1993]; and wavefront healing effects make it difficult to accurately image low-velocity bodies [*Nolet and Dahlen*, 2000]. The "disappearance" of the plume tails can hence be explained as effects of poor tomographic resolution at deeper levels. Apart from this, our results demonstrate the plausibility of finding a great diversity in the morphology of seismically imaged mantle plumes, including plume heads without tails and plumes with tails that are detached from their sources.

[48] The mathematical model of mantle plume dynamics described by a set of equations (1)–(3) is simple, and many complications are omitted. A viscosity increase from the upper to the lower mantle is not included in the model, although it is suggested by studies of the geoid [*Ricard et al.*, 1993], postglacial rebound [*Mitrovica*, 1996], and joint inversion of convection and glacial isostatic adjustment data [*Mitrovica and Forte*, 2004]. The adiabatic heating/cooling term in the heat equation can provide more realistic distribution of temperature in the mantle, especially near the thermal boundary layer. Our model does not include phase transformations [e.g., *Liu et al.*, 1991; *Honda et al.*, 1993a, 1993b; *Harder and Christensen*, 1996], although the phase changes can influence the evolution of mantle plumes retarding/accelerating their ascent. The coefficient of thermal expansion [e.g., *Chopelas and Boehler*, 1989; *Hansen et al.*, 1991, 1993] and the coefficient of thermal conduc-

tivity [e.g., *Hofmeister*, 1999] are not constant in the mantle and vary with depth and temperature. Moreover, if *Badro et al.* [2004] findings of a significant increase in the radiative thermal conductivity at high pressure are relevant to the lower mantle, plume tails should diffuse away even faster than it is predicted by our models.

[49] Mantle plumes exist within the large-scale convective flow, which may disrupt the plumes before they diffuse thermally [e.g., *Richards and Griffiths*, 1988]. *Steinberger* [2000] performed numerical experiments to clarify an interplay between a large-scale mantle flow and mantle plume and hot spot dynamics and showed that during the rise plume tails can be tilted toward large-scale mantle upwellings. Meanwhile, we believe that the possible deformation of plume tails should not significantly alter our results on thermal diffusion of the plumes.

[50] Several reasons constrain us to consider in the present study the simplified mathematical model as the first (principal) step to sophisticated models. The use of the variational data assimilation techniques for the problems of mantle convection began only recently [*Bunge et al.*, 2003; *Ismail-Zadeh et al.*, 2003a]. This technique requires derivation of adjoint equations (to estimate initial temperature conditions in the mantle) each time when the set of the equations is changed. The cost to be paid is in software development since an adjoint model has to be developed. Moreover, since we analyze effects of thermal diffusion on the fate of mantle plumes, we avoid many complications and considered only the most essential component of mantle plume dynamics, namely, temperature-dependent viscosity. While inclusion of these complications and other model refinements are worthwhile, our experiments do show that thermal diffusion plays an important role in the fate of mantle plumes and it provides an explanation for the ''variety'' of mantle plumes observed in seismic tomographic images.

7.2. Assimilation of Present Temperature Derived From Seismic Tomography

[51] The variational assimilation of synthetic data (mantle plumes generated by computer simulations) showed a possibility to restore strong features of the plumes after their thermal diffusion. In this section we illustrate how real (no synthetic) present crust/mantle temperature can be assimilated into the geological past. For this aim we use recent teleseismic body wave tomography data, which image the lithosphere and asthenosphere for the southeastern Carpathians [*Martin et al.*, 2005]. We should note that the region is not associated with a mantle plume activity and chosen because of high-resolution seismic tomography data made available to the authors.

[52] The seismic tomographic model of the region consists of eight layers of different thickness (from 15 to 50 km), which are each subdivided laterally into 42×42 km^2 blocks [*Martin et al.*, 2005]. To restrict numerical errors in our data assimilation we smooth the velocity anomaly data using spline interpolations between the blocks and the layers. To convert the *P wave* seismic velocity anomalies beneath the region into temperature we model initially synthetic *P wave* seismic velocities considering the effects of anharmonicity (composition), anelasticity and partial melting on the seismic velocities [*Ismail-Zadeh et al.*, 2005]. The anharmonic (frequency-independent and nonattenuating) part of the synthetic velocities is calculated on the basis of published data on laboratory measurements of density and elastic parameters of the main rock-forming minerals [*Bass*, 1995] at various thermodynamic conditions for the composition of the crust and mantle (57.9% Ol, 16.3% CPx, 13.5% Opx, and 12.3% Gt [*Green and Falloon*, 1998]) and the slab (69% Ol, 10% CPx, 19% Opx, and 2% Gt [*Agee*, 1993]). Once the synthetic velocities are calculated for a first-guess temperature, an iteration process is used to find the ''true'' temperature, minimizing the difference between the synthetic and ''observed'' (in seismic tomography experiments) velocities. The temperature in the shallow levels of the region is constrained from measured surface heat flux corrected for paleoclimate changes and for the effects of sedimentation [*Demetrescu et al.*, 2001]. Figure 5a illustrates several depth slices of the present temperature model derived from the seismic tomography data.

[53] We assimilate the present temperature data into the geological past to restore the prominent thermal features of the Earth's structures in the region. We use the following parameters in this case study: $h = 670$ km, the aspect ratio (ratio between horizontal and vertical lengths of the model) is 1.5, $r = 1000$, $\Delta T = 1700$ K, $\rho_{\mathrm{ref}} = 3400$ kg m^{-3}, $\eta_{\mathrm{ref}} = 10^{21}$ Pa s, $Ra = 5.2 \times 10^5$. Other parameters are the same (see Table 1). The equations and boundary conditions are defined in sections 2 and 3. To reduce the numerical noise in the data assimilation, we regularize the solution by using the quasi-reversibility method by *Lattes and Lions* [1969]. Figure 5b shows the temperature restored to 22 Myr ago.

[54] Early Miocene subduction beneath the Carpathian arc and subsequent gentle continental collision transported cold and dense lithospheric material into the hotter mantle [*Sperner and The CRC 461 Team*, 2005]. The cold (blue) region seen at the 20 km slice of the restored temperature (Figure 5b) can be interpreted as a crustal portion of a lithospheric slab. The structure is almost invisible at the relevant slice of the present temperature, because the slowly descending cold slab has been warmed up (and hence has faded away) due to thermal diffusion since an active slab subduction in the region has ended about 10 Myr ago [*Csontos et al.*, 1992]. Thermal conduction in the shallow Earth (where viscosity is high) plays a significant part in heat transfer compared to thermal convection. The deeper we look into the region (see the slices at depths of 60 km and 130 km in Figure 5b), the larger are effects of thermal advection compared to diffusion: the cold (dense) lithosphere has moved upward to the place where it has been in the Miocene times. At 280 km depth a shape of the colder slab is clearly visible at the slice of the present temperature (Figure 5a) and practically invisible at the slice of the restored temperature (Figure 5b), because the slab did not reach the depth 22 Myr ago.

[55] Thus the assimilation of the present temperature derived from seismic tomography data shows that prominent thermal mantle structures can be restored from their present diffused stage.

8. Conclusion

[56] In this paper, models of mantle plume weakening due to thermal diffusion have been analyzed. Injection of hot material from the thermal source layer into the colder

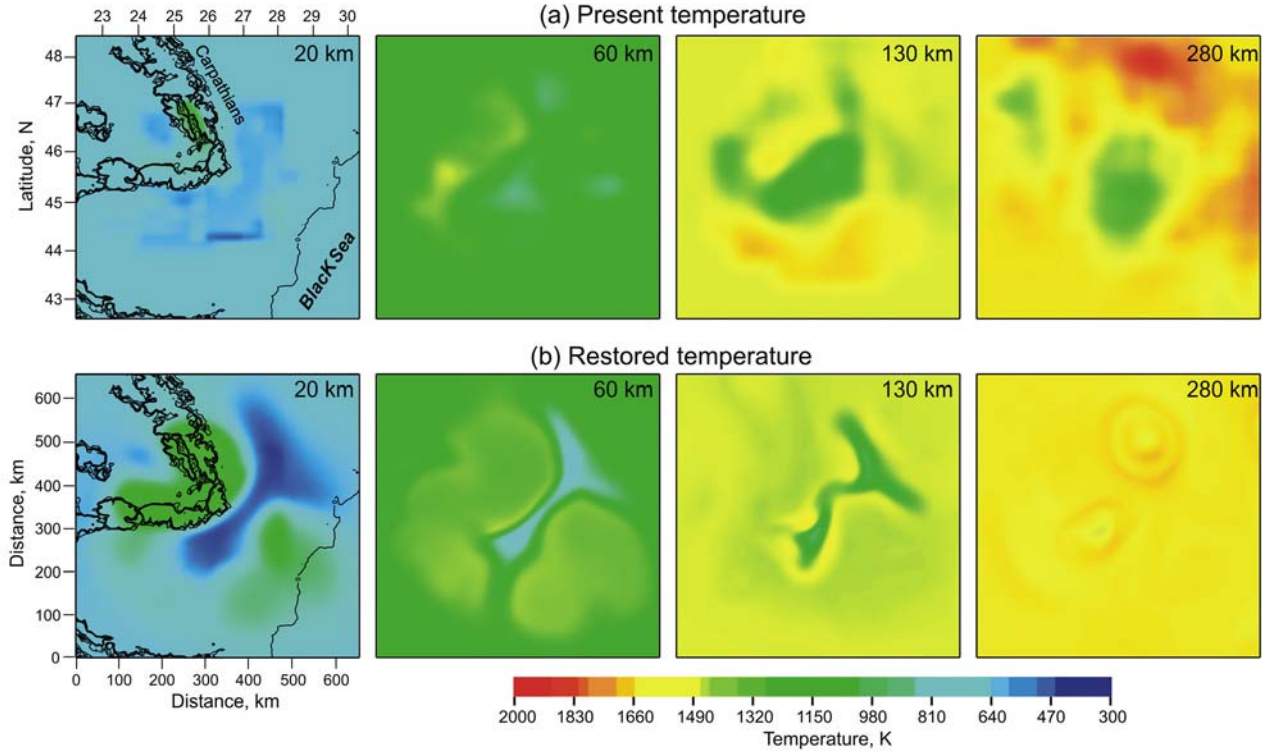

Figure 5. Present and restored (to 22 Myr ago) temperature beneath the southeastern Carpathians at depths of 20, 60, 130, and 280 km. (a) Temperatures derived from *P* wave velocity anomalies. (b) Temperature restored by data assimilation. Isolines present the surface topography.

mantle generates strong plumes that are fed from the source layer for a while. However, the feeding from the source layer can weaken with time and then thermal diffusion takes over and controls the subsequent evolution of the mantle plumes. The plumes begin to diffuse away and the plume tails are the first structures to disappear. The tails of different plumes vanish at different times depending on the geometry of the tails. The morphological diversity of the plumes predicted by the numerical experiments is similar to the plume diversity observed in seismic tomographic images [*Montelli et al.*, 2004; *Zhao*, 2004].

[57] We have also studied how the restoration process (data assimilation algorithm) works in recovering strong features of mantle plumes after they have weakened by thermal diffusion and in the presence of a large depth gradient of mantle viscosity. The restoration process becomes poor as both diffusion and viscosity gradient increase. For a given range of Rayleigh number and two values of the viscosity gradient, the convergence rate of the objective functional shows a large variation, which implies that the performance is very sensitive to the magnitude of both diffusion and viscosity gradient.

[58] The present temperature obtained from high-resolution teleseismic tomography data for the southeastern Carpathians has been assimilated into the geological past. Results of this case study suggest that the data assimilation can be used to restore initial mantle temperatures and can allow revealing prominent thermal structures in the mantle from their present diffused stage. A part of the geophysical community may maintain skepticism about the assimilation of present mantle-related data to the geological past. This skepticism may partly have its roots in our poor knowledge of the Earth's present structure and its physical properties, which cannot allow for rigorous numerical paleoreconstructions of the mantle evolution. An increase in the accuracy of seismic tomography inversions and geodetic measurements, improvements in the knowledge of gravity and geothermal fields, and more complete experimental data on the physical and chemical properties of mantle rocks will facilitate mantle reconstructions.

Appendix A: Challenges in Variational Data Assimilation for Thermoconvective Flow in the Mantle

[59] Although the variational data assimilation technique described above can theoretically be applied to many problems in mantle and lithosphere dynamics, a practical implementation of the technique for modeling of real geodynamic processes backward in time (to restore the temperature and flow pattern in the past) is not a simple task. Smoothness of the initial data (present temperature) and of the target temperature (restored temperature in the past) is an important factor in backward modeling. Moreover, a choice of the initial temperature guess φ_0 in iteration scheme (4) is not trivial.

A1. On the Smoothness of the Initial Temperature

[60] The solution $T(\vartheta, \cdot\,; \varphi)$ of the heat problem (1) is a sufficiently smooth function and belongs to space L_2 (Ω). The present temperature χ_δ derived from the seismic tomography is a representation of the exact temperature χ of the Earth and so it must also belong to this space and

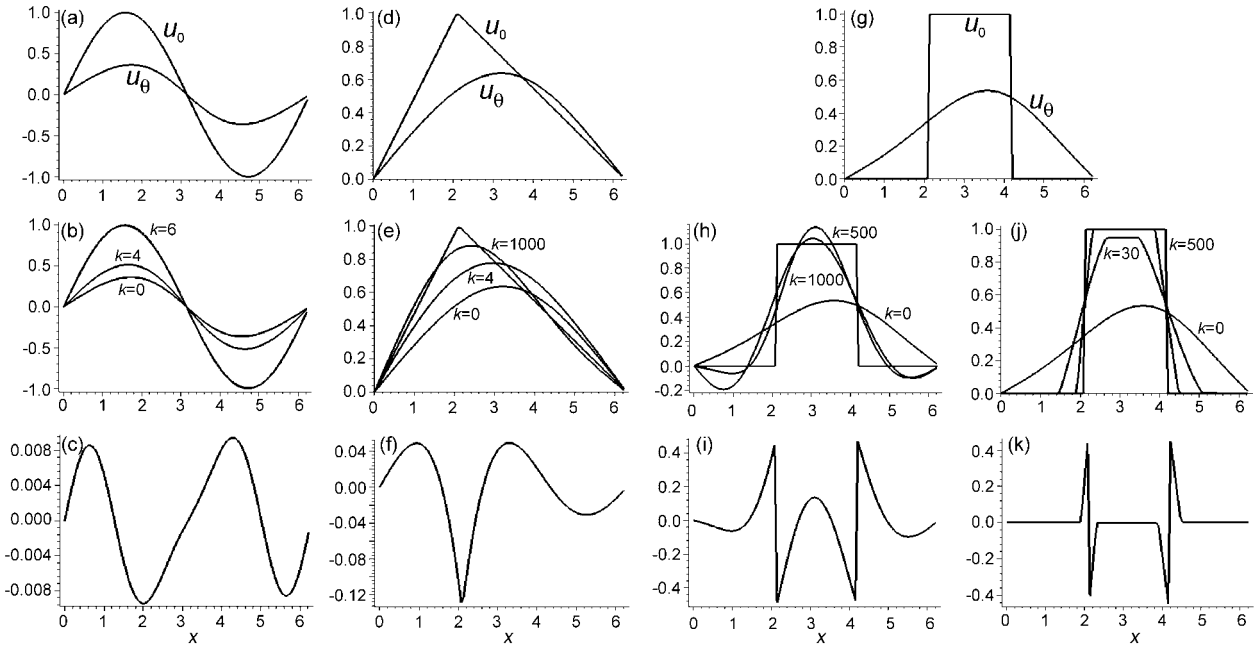

Figure A1. Recovering function u_0 from the smooth guess function u_θ. (a–c) The sufficiently smooth u_0; (d–f) continuous piecewise smooth function u_0; and (g–k) discontinuous function u_0. Plots of u_0 and u_θ are presented in Figures A1a, A1d, and A1g; successive approximations to u_0 at Figures A1b, Figures A1e, Figures A1h, and Figures A1j; and the residual functions in Figures A1c, Figures A1f, Figures A1i, and Figures A1k.

hence be rather smooth; otherwise, the objective functional J cannot be defined. Therefore before any assimilation of the present temperature data can be attempted, the data must be smoothed. The smoothing of the present temperature improves the convergence of the iterations. However, there are still some numerical issues associated with the solution of the improperly posed problem (we remind the reader that the inverse problem of thermal convection is improperly posed [e.g., *Tikhonov and Arsenin*, 1977]).

[61] If the initial temperature guess φ_0 is a smooth function, all successive temperature iterations φ_k in scheme (4) should be smooth functions too, because the gradient of the objective functional ∇J is a smooth function since it is the solution to the adjoint problem (6). The temperature iterations φ_k are disturbed by small computational errors, which are inherent in any numerical experiment (see Appendix B). These perturbations grow with time unless the iteration scheme (7) or a similar one [*Samarskii and Vabischevich*, 1995] is used as discussed in section 5. Another possibility is to use the quasi-reversibility method [*Lattes and Lions*, 1969] to regularize a temperature field or high-order adjoint techniques [*Alekseev and Navon*, 2001].

[62] A choice of the initial temperature guess φ_0 (smooth versus discontinuous functions) influences the convergence of the iterations. There are however no general "recipes" for the choice of the initial temperature guess, and this depends mainly on the experience of computer modelers in solving such numerical problems.

A2. On the Smoothness of the Target Temperature

[63] If mantle temperature in the geological past was not a smooth function of space variables, recovery of this temperature using the technique described in this paper is not effective because the iterations converge very slowly to the target temperature. Here we explain the problem of recovering the initial mantle temperature at the time of plume onset on the basis of three one-dimensional model tasks: restoration of a smooth, piecewise smooth and discontinuous target function. We note that the temperature in the Earth's mantle is not a discontinuous function but its shape can be close to a step function.

[64] We consider that the dynamics of a physical system is described by the Burgers equation $u_t + uu_x = u_{xx}$, $0 \le t \le 1$, $0 \le x \le 2\pi$ with the boundary conditions $u(t, 0) = 0$, $u(t, 2\pi) = 0$, $0 \le t \le 1$ and the condition $u_\theta = u(1, x; u_0)$, $0 \le x \le 2\pi$ at $t = 1$, where the variable u can denote temperature. The problem is to recover the function $u_0 = u_0(x)$, $0 \le x \le 2\pi$ at $t = 0$ (the state in the past) from the function $u_\theta = u_\theta(x)$, $0 \le x \le 2\pi$ at $t = 1$ (its present state). The finite difference approximations and the variational method are applied to the Burgers equation with the appropriate boundary and initial conditions.

A2.1. Task 1

[65] Consider the sufficiently smooth function $u_0 = \sin(x)$, $0 \le x \le 2\pi$. The functions u_0 and u_θ are shown in Figure A1a. Figures A1b and A1c illustrate the iterations φ_k using the iterative scheme similar to (4) for $k = 0$, 4, 6 and the residual $r_6(x) = u_0(x) - \varphi_6(x)$, $0 \le x \le 2\pi$, respectively. We see that iterations converge rather rapid for the sufficiently smooth target function.

A2.2. Task 2

[66] Now consider the continuous piecewise smooth function $u_0 = 3x/(2\pi)$, $0 \le x \le 2\pi/3$ and $u_0 = 3/2 - 3x/(2\pi)$, $2\pi/3 \le x \le 2\pi$. Figure A1 presents the functions u_0 and

u_θ (Figure A1d), the successive approximations φ_k for $k = 0$, 4, 1000 (Figure A1e), and the residual $r_{1000}(x) = u_0 x) - \varphi_{1000}(x)$, $0 \le x \le 2\pi$ (Figure A1f), respectively. This example shows that a large number of iterations is required to reach the target function.

A2.3. Task 3

[67] Consider the discontinuous function u_0, which takes 1 at $2\pi/3 \le x \le 4\pi/3$ and 0 in other points of the closed interval $0 \le x \le 2\pi$. Figure A1 presents the functions u_0 and u_θ (Figure A1g), the successive approximations φ_k for $k = 0$, 500, 1000 (Figure A1h), and the residual $r_{1000}(x) = u_0(x) - \varphi_{1000}(x)$, $0 \le x \le 2\pi$ (Figure A1e), respectively. We see that convergence to the target temperature is very poor.

[68] To improve the convergence to the target function, a modification of the variational method based on a priori information about a desired solution was suggested by *Korotkii and Tsepelev* [2003]. Figure A1j shows the successive approximations $\widetilde{\varphi}_k$ for $k = 0$, 30, 500, and Figure A1k shows the residual $\widetilde{r}_{500}(x) = u_0(x) - \widetilde{\varphi}_{500}(x)$, $0 \le x \le 2\pi$, respectively. The approximations $\widetilde{\varphi}_k$ based on the method of gradient projection [*Vasiliev*, 2002] converge to the target solution better than approximations generated by equation (4).

Appendix B: Errors in Forward and Backward Modeling

[69] A numerical model has three kinds of variables: state variables, input variables, and parameters. State variables describe the physical properties of the medium (velocity, pressure, temperature) and depend on time and space. Input variables have to be provided to the model (initial or boundary conditions), most of the time these variables are not directly measured but they can be estimated through data assimilation. Most models contain also a set of parameters (e.g., viscosity, thermal diffusivity), which have to be tuned to adjust the model to the observations. All the variables can be polluted by errors.

[70] There are three kinds of systematic errors in numerical modeling of geodynamical problems: model, discretization, and iteration errors. Model errors are associated with the idealization of Earth dynamics by a set of conservation equations governing the dynamics. The model errors are defined as the difference between the actual Earth dynamics and the exact solution of the mathematical model. Discretization errors are defined as the difference between the exact solution of the conservation equations and the exact solution of the algebraic system of equations obtained by discretizing these equations. Also, iteration errors are defined as the difference between the iterative and exact solutions of the algebraic system of equations. It is important to be aware of the existence of these errors, and even more to try to distinguish one from another.

[71] Apart from the errors associated with the numerical modeling, another two components of errors are essential when mantle temperature data are assimilated into the past: (1) data misfit associated with the uncertainties in the present temperature distribution in the Earth's mantle and (2) errors associated with the uncertainties in initial and boundary conditions. Since there are no direct measurements of mantle temperatures, the temperatures can be estimated indirectly either from seismic wave (and their anomalies), geochemical analysis or through the extrapolation of surface heat flow observations. Many models of mantle temperature are based on the conversion of seismic tomography data into temperature. Meanwhile, a seismic tomography image of the Earth's mantle is a model indeed and incorporates its own model errors. Another source of uncertainty comes from the choice of mantle compositions in the modeling of mantle temperature from the seismic velocities. Therefore, if the present mantle temperature models are biased, information on temperature can be improperly propagated to the geological past.

[72] The temperature at the lower boundary of the model domain we used in forward and backward numerical modeling is, of course, an approximation to the real temperature, which is unknown and may change over time at this boundary. Hence errors associated with the knowledge of the temperature (or heat flux) evolution at the core-mantle boundary are another essential component of errors, which can be propagated into the past during the data assimilation.

[73] In numerical modeling sensitivity analysis assists in understanding the stability of the model solution to small perturbations in input variables or parameters. For instance, if we consider mantle temperature in the past as a solution to the backward model, what will be its variation if there is some perturbation on the inputs of the model (e.g., present temperature data)? The gradient of the objective functional with respect to input parameters in variational data assimilation gives the first-order sensitivity coefficients. The second-order adjoint sensitivity analysis presents some challenge associated with cumbersome computations of the product of the Hessian matrix of the objective functional with some vector [*Le Dimet et al.*, 2002], and hence it is omitted in our study. *Hier-Majumder et al.* [2006] performed the first-order sensitivity analysis for two-dimensional problems of thermoconvective flow in the mantle. See *Cacuci* [2003] and *Cacuci et al.* [2005] for more detail on sensitivity and uncertainty analysis.

[74] **Acknowledgments.** We thank A. Davaille, M. Ghil, S. Honda, N. Ribe, H. Wilhelm, and T. Yanovskaya for discussions on thermal plume diffusion, data assimilation, and on geothermal and seismic tomography modeling. We are very grateful to H.-P. Bunge, I. M. Navon, anonymous reviewers and the Associate Editor, who provided careful reviews that significantly improved an initial version of the manuscript. The research was supported by the German Research Council (DFG-Wi-687/18-1), Russian Foundation of Basic Research (RFBR-05-01-00098), and the Russian Academy of Sciences (OH3-260603-879) grants.

References

Agee, C. B. (1993), Petrology of the mantle transition zone, *Annu Rev. Earth Planet. Sci.*, *21*, 19–41.

Albers, M., and U. R. Christensen (1996), The excess temperature of plumes rising from the core-mantle boundary, *Geophys. Res. Lett.*, *23*, 3567–3570.

Alekseev, A. K., and I. M. Navon (2001), The analysis of an ill-posed problem using multiscale resolution and second order adjoint techniques, *Comput.Methods Appl. Mech. Eng.*, *190*, 1937–1953.

Badro, J., J.-P. Rueff, G. Vanko, G. Monaco, G. Figuet, and F. Guyot (2004), Electronic transitions in perovskite: Possible nonconvecting layers in the lower mantle, *Science*, *305*, 383–386.

Bass, J. D. (1995), Elasticity of minerals, glasses, and melts, in *Mineral Physics and Crystallography, A Handbook of Physical Constants*, *AGU Ref. Shelf*, vol. 2, edited by T. J. Ahrens, pp. 45–63, AGU, Washington, D. C.

Bennett, A. F. (1992), *Inverse Methods in Physical Oceanography*, 346 pp., Cambridge Univ. Press, New York.

Boussinesq, J. (1903), *Theorie Analytique de la Chaleur*, vol. 2, 172 pp., Elsevier, New York.

Bunge, H.-P., M. A. Richards, C. Lithgow-Bertelloni, J. R. Baumgardner, S. P. Grand, and B. Romanowicz (1998), Time scales and heterogeneous structure in geodynamic Earth models, *Science*, *280*, 91–95.

Bunge, H.-P., M. A. Richards, and J. R. Baumgardner (2002), Mantle circulation models with sequential data-assimilation: Inferring present-day mantle structure from plate motion histories, *Philos. Trans. R. Soc. London, Ser. A*, *360*, 2545–2567.

Bunge, H.-P., C. R. Hagelberg, and B. J. Travis (2003), Mantle circulation models with variational data assimilation: Inferring past mantle flow and structure from plate motion histories and seismic tomography, *Geophys. J. Int.*, *152*, 280–301.

Busse, F. H., et al. (1993), 3D convection at infinite Prandtl number in Cartesian geometry–a benchmark comparison, *Geophys. Astrophys. Fluid Dyn.*, *75*, 39–59.

Cacuci, D. G. (2003), *Sensitivity and Uncertainty Analysis*, vol. I, *Theory*, 285 pp., CRC, Boca Raton, Fla.

Cacuci, D. G., M. Ionescu-Bujor, and I. M. Navon (2005), *Sensitivity and Uncertainty Analysis. Volume II: Applications to Large-Scale Systems*, 368 pp., CRC, Boca Raton, Fla.

Chandrasekhar, S. (1961), *Hydrodynamic and Hydromagnetic Stability*, 654 pp., Oxford Univ. Press, New York.

Chopelas, A., and R. Boehler (1989), Thermal expansion measurements at very high pressure, systematics and a case for a chemically homogeneous mantle, *Geophys. Res. Lett.*, *16*, 1347–1350.

Condie, K. C. (2001), *Mantle Plumes and Their Record in Earth History*, 306 p., Cambridge Univ. Press, New York.

Csontos, L., A. Nagymarosy, F. Horvath, and M. Kovac (1992), Tertiary evolution of the intra-Carpathian area: A model, *Tectonophysics*, *208*, 221–241.

Davaille, A. (1999), Simultaneous generation of hotspots and superswells by convection in a heterogeneous planetary mantle, *Nature*, *402*, 756–760.

Davaille, A., and J. Vatteville (2005), On the transient nature of mantle plumes, *Geophys. Res. Lett.*, *32*, L14309, doi:10.1029/2005GL023029.

Demetrescu, C., S. B. Nielsen, M. Ene, D. Z. Serban, G. Polonic, M. Andreescu, A. Pop, and N. Balling (2001), Lithosphere thermal structure and evolution of the Transylvanian Depression: Insight from new geothermal measurements and modelling results, *Phys. Earth Planet. Inter.*, *126*, 249–267.

Fletcher, R., and C. M. Reeves (1964), Function minimization by conjugate gradients, *Comput. J.*, *7*, 149–154.

Forte, A. M., and J. X. Mitrovica (2001), Deep-mantle high-viscosity flow and thermochemical structure inferred from seismic and geodynamic data, *Nature*, *410*, 1049–1056.

Ghil, M., and P. Malanotte-Rizzoli (1991), Data assimilation in meteorology and oceanography, *Adv. Geophys.*, *33*, 141–266.

Green, D. H., and T. J. Falloon (1998), Pyrolite: A Ringwood concept and its current expression, in *The Earth's Mantle*, edited by I. Jackson, pp. 311–378, Cambridge Univ. Press, New York.

Hansen, U., D. A. Yuen, and S. E. Kroening (1990), Transition to hard turbulence in thermal convection at infinite Prandtl number, *Phys. Fluids A*, *2*(12), 2157–2163.

Hansen, U., D. A. Yuen, and S. E. Kroening (1991), Effects of depth-dependent thermal expansivity on mantle circulations and lateral thermal anomalies, *Geophys. Res. Lett.*, *18*, 1261–1264.

Hansen, U., D. A. Yuen, S. E. Kroening, and T. B. Larsen (1993), Dynamical consequences of depth-dependent thermal expansivity and viscosity on mantle circulations and thermal structure, *Phys. Earth Planet. Inter.*, *77*, 205–223.

Harder, H., and U. R. Christensen (1996), A one-plume model of Martian mantle convection, *Nature*, *380*, 507–509.

Hier-Majumder, C. A., B. J. Travis, E. Belanger, G. Richard, A. P. Vincent, and D. A. Yuen (2006), Efficient sensitivity analysis for flow and transport in the Earth's crust and mantle, *Geophys. J. Int.*, in press.

Hofmeister, A. M. (1999), Mantle values of thermal conductivity and the geotherm from photon lifetimes, *Science*, *283*, 1699–1706.

Honda, S., S. Balachandar, D. A. Yuen, and D. Reuteler (1993a), Three-dimensional mantle dynamics with an endothermic phase transition, *Geophys. Res. Lett.*, *20*, 221–224.

Honda, S., D. A. Yuen, S. Balachandar, and D. Reuteler (1993b), Three-dimensional instabilities of mantle convection with multiple phase transitions, *Science*, *259*, 1308–1311.

Howard, L. N. (1966), Convection at high Rayleigh number, in *Applied Mechanics, Proc. of the 11th Intl Congress of Applied Mechanics, Munich, Germany 1964*, edited by H. Goertler and P. Sorger, pp. 1109–1115, Springer, New York.

Ismail-Zadeh, A. T., A. I. Korotkii, B. M. Naimark, and I. A. Tsepelev (2001), Numerical simulation of three-dimensional viscous flows with gravitational and thermal effects, *Comput. Math. Math. Phys.*, *41*, 1399–1415.

Ismail-Zadeh, A. T., A. I. Korotkii, B. M. Naimark, and I. A. Tsepelev (2003a), Three-dimensional numerical simulation of the inverse problem of thermal convection, *Comput. Math. Math. Phys.*, *43*, 587–599.

Ismail-Zadeh, A. T., A. I. Korotkii, and I. A. Tsepelev (2003b), Numerical approach to solving problems of slow viscous flow backwards in time, in *Computational Fluid and Solid Mechanics*, edited by K. J. Bathe, pp. 938–941, Elsevier, New York.

Ismail-Zadeh, A., G. Schubert, I. Tsepelev, and A. Korotkii (2004), Inverse problem of thermal convection: Numerical approach and application to mantle plume restoration, *Phys. Earth Planet. Inter.*, *145*, 99–114.

Ismail-Zadeh, A., B. Mueller, and G. Schubert (2005), Three-dimensional modeling of present-day tectonic stress beneath the earthquake-prone southeastern Carpathians based on integrated analysis of seismic, heat flow, and gravity observations, *Phys. Earth Planet. Inter.*, *149*, 81–98.

Jellinek, M. A., and M. Manga (2002), The influence of a chemical boundary layer on the fixity, spacing and lifetime of mantle plumes, *Nature*, *418*, 760–763.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, 341 pp., Cambridge Univ. Press, New York.

Karato, S.-I. (1993), Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, *20*, 1623–1626.

Korotkii, A. I., and I. A. Tsepelev (2003), Solution of a retrospective inverse problem for one nonlinear evolutionary model, in *Proceedings of the Steklov Institute of Mathematics, Suppl. 2*, pp. 80–94, Nauka, Moscow.

Lattes, R., and J. L. Lions (1969), *The Method of Quasi-Reversibility: Applications to Partial Differential Equations*, 388 pp., Elsevier, New York.

Le Dimet, F.-X., I. M. Navon, and D. N. Daescu (2002), Second-order information in data assimilation, *Mon. Weather Rev.*, *130*, 629–648.

Leitch, A. M., V. Steinbach, and D. A. Yuen (1996), Centerline temperature of mantle plumes in various geometries: Incompressible flow, *J. Geophys. Res.*, *101*, 21,829–21,846.

Liu, D. C., and J. Nocedal (1989), On the limited memory BFGS method for large scale optimization, *Math. Program.*, *45*, 503–528.

Liu, M., D. A. Yuen, W. Zhao, and S. Honda (1991), Development of diapiric structures in the upper mantle due to phase transitions, *Science*, *252*, 1836–1839.

Malevsky, A. V., and D. A. Yuen (1993), Plume structures in the hard-turbulent regime of three-dimensional infinite Prandtl number convection, *Geophys. Res. Lett.*, *20*, 383–386.

Malevsky, A. V., D. A. Yuen, and L. M. Weyer (1992), Viscosity and thermal fields associated with strongly chaotic non-Newtonian thermal convection, *Geophys. Res. Lett.*, *19*, 127–130.

Martin, M., J. R. R. Ritter, and the CALIXTO Working Group (2005), High-resolution teleseismic body-wave tomography beneath SE Romania–I. Implications for three-dimensional versus one-dimensional crustal correction strategies with a new crustal velocity model, *Geophys. J. Int.*, *162*, 448–460.

McDonald, A. (1984), Accuracy of multi-upstream semi-Lagrangian advective schemes, *Mon. Weather Rev.*, *112*, 1267–1279.

McLaughlin, D. (2002), An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and forecasting, *Adv. Water Resour.*, *25*, 1275–1286.

Mitrovica, J. X. (1996), Haskell (1935) revisited, *J. Geophys. Res.*, *101*, 555–569.

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

Montelli, R., G. Nolet, F. A. Dahlen, G. Masters, E. R. Engdahl, and S.-H. Hung (2004), Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, *303*, 338–343.

Moore, W. B., G. Schubert, and P. Tackley (1998), Three-dimensional simulations of plume–lithosphere interaction at the Hawaiian Swell, *Science*, *279*, 1008–1011.

Morgan, W. J. (1972), Plate motions and deep convection, *Geol. Soc. Am. Mem.*, *132*, 7–22.

Navon, I. M., X. Zou, J. Derber, and J. Sela (1992), Variational data assimilation with an adiabatic version of the NMC spectral model, *Mon. Weather Rev.*, *120*, 1433–1446.

Nolet, G., and F. A. Dahlen (2000), Wave front healing and the evolution of seismic delay times, *J. Geophys. Res.*, *105*, 19,043–19,054.

Olson, P., and H. Singer (1985), Creeping plumes, *J. Fluid Mech*, *158*, 511–531.

Olson, P., G. Schubert, and C. Anderson (1993), Structure of axisymmetric mantle plumes, *J. Geophys. Res.*, *98*, 6829–6844.

Ribe, N. M., and U. Christensen (1994), Three-dimensional modeling of plume–lithosphere interaction, *J. Geophys. Res.*, *99*, 669–682.

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

Richards, M. A., and R. W. Griffiths (1988), Deflection of plumes by mantle shear flow: Experimental results and a simple theory, *Geophys. J. R. Astron. Soc.*, *94*, 367–376.

Richards, M. A., R. A. Duncan, and V. Courtillot (1989), Flood basalts and hot spot tracks: Plume heads and tails, *Science*, *246*, 103–107.

Ritsema, J., H. J. van Heijst, and J. H. Woodhouse (1999), Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science*, *286*, 1925–1928.

Romanowicz, B., and Y. Gung (2002), Superplumes from the core-mantle boundary to the lithosphere: Implications for heat flux, *Science*, *296*, 513–516.

Samarskii, A. A., and P. N. Vabishchevich (1995), *Computational Heat Transfer*, vol. 1, *Mathematical Modelling*, 370 pp., John Wiley, Hoboken, N. J.

Samarskii, A. A., and P. N. Vabishchevich (2004), *Numerical Methods for Solving Inverse Problems of Mathematical Physics*, 478 pp., URSS, Moscow.

Samarskii, A. A., P. N. Vabishchevich, and V. I. Vasil'ev (1997), Iterative solution of a retrospective inverse problem of heat conduction, *Math. Model.*, *9*(5), 119–127.

Schubert, G., D. L. Turcotte, and P. Olson (2001), *Mantle Convection in the Earth and Planets*, 940 p., Cambridge Univ. Press, New York.

Silveira, G., E. Stutzmann, A. Davaille, J.-P. Montagner, L. Mendes-Victor, and A. Sebai (2006), Azores hotspot signature in the upper mantle, *J. Volcanol. Geotherm. Res.*, in press.

Sleep, N. H. (1990), Hotspots and mantle plumes: Some phenomenology, *J. Geophys. Res.*, *95*, 6715–6736.

Sperner, B., and The CRC 461 Team (2005), Monitoring of slab detachment in the Carpathians, in *Perspectives in Modern Seismology*, *Lect. Notes Earth Sci.*, vol. 105, edited by F. Wenzel, pp. 187–202, Springer, New York.

Steinberger, B. (2000), Plumes in a convecting mantle: Models and observations for individual hotspots, *J. Geophys. Res.*, *105*, 11,127–11,152.

Steinberger, B., and R. J. O'Connell (1998), Advection of plumes in mantle flow: Implications for hotspot motion, mantle viscosity and plume distribution, *Geophys. J. Int.*, *132*, 412–434.

Talagrand, O. (1997), Assimilation of observations: An introduction, *J. Meteorol. Soc. Jpn.*, *75*, 191–209.

Tikhonov, A. N., and V. Y. Arsenin (1977), *Solution of Ill-Posed Problems*, 258 pp., Winston, Washington, D. C.

Trompert, R. A., and U. Hansen (1998), On the Rayleigh number dependence of convection with a strongly temperature-dependent viscosity, *Phys. Fluids*, *10*, 351–360.

Vasiliev, F. P. (2002), *Methods of Optimization* (in Russian), 824 pp., Factorial Press, Moscow.

Wunsch, C. (1996), *The Ocean Circulation Inverse Problem*, Cambridge Univ. Press, 442 pp., New York.

Zhao, D. (2004), Global tomographic images of mantle plumes and subducting slabs: Insight into deep Earth dynamics, *Phys. Earth Planet. Inter.*, *146*, 3–34.

Zhong, S. (2005), Dynamics of thermal plumes in three-dimensional isoviscous thermal convection, *Geophys. J. Int.*, *162*, 289–300.

Zou, X., I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. Le Dimet (1993), Numerical experience with limited-memory quasi-Newton and truncated Newton methods, *SIAM J. Optimization*, *3*(3), 582–608.

A. Ismail-Zadeh, Geophysikalishes Institut, Universität Karlsruhe, Hertzstr. 16, Karlsruhe D-76187, Germany. (alik.ismail-zadeh@gpi.uka.de)

A. Korotkii and I. Tsepelev, Institute of Mathematics and Mechanics, Ural Branch, Russian Academy of Sciences, S. Kovalevskoy ul. 16, Yekaterinburg 620219, Russia.

G. Schubert, Department of Earth and Space Sciences, University of California, 3806 Geology Building, 595 Charles Young Drive East, Los Angeles, CA 90095-1567, USA.