Rayleigh-Taylor Instabilities from Hydrous Melting Propel "Cold Plumes"
at Subduction Zones

T.V. Gerya*
Institute of Experimental Mineralogy Russian Academy of Sciences, 142432 Chernogolovka, Moscow, Russia
Institut für Geologie, Mineralogie und Geophysik, Ruhr-Universität Bochum, 44780 Bochum, Germany
D.A. Yuen
University of Minnesota Supercomputing Institute and Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota, 55455-0219, USA

Submitted to Earth and Planetary Science Letters,
December, 2002.

* Corresponding author. Tel.: +49-234-3223518;
fax: +49-234-3214433.
E-mail address: taras.gerya@ruhr-uni-bochum.de (T. Gerya)
Abstract. It is commonly thought hot diapiric flows prevail in the mantle wedge above the subducting slab. However, hydration and partial melting along the slab can create a situation in which a Rayleigh-Taylor instability can develop at the top of a cold subducting slab. We have modelled numerically with a high-resolution 2-D regional model this paradoxically interesting geological phenomenon in which rising diapiric structures, colder than the asthenosphere by 300 to 400 degrees, are driven upward by compositional buoyancy. These "cold plumes" with a compositional, hydrous origin, launched from a depth of greater than 100 km, are lubricated by viscous heating, have an upward velocity in excess of 10 cm/yr, penetrate the relatively hot asthenosphere in the mantle wedge within a couple of million years and thus can cool the surroundings. These cold plumes are fueled by partial melting of the hydrated mantle and subducted oceanic crust due to fluid release from dehydration reactions within the slab, including the decomposition of serpentine. There may be a spatial correlation between seismicity and the particular depth of cold plume initiation, which may be related to the initial age of the subducting lithosphere.

Keywords: mantle plumes; subduction zones; magmatic activity; seismicity; 2-D numerical modeling
1. Introduction

Today it is well accepted that upwellings in the mantle can be driven by either thermal or chemical buoyancy and they can occur in various morphologies, such as cylindrical structures, diapiric shaped objects (e.g. Olson and Nam, 1986, Bercovici and Mahoney, 1994) or even lineated features (Ni and Helmberger, 2003). Mantle plumes are commonly held to come from either the core-mantle boundary (Morgan, 1972) with a thermal origin or chemical origin (Hansen and Yuen, 1988) or from the transition zone (e.g. Cserepes and Yuen, 2000). Yet there are enigmatic volcanic features near the vicinity of the subducting slab which require their sources to be associated with fluid dynamic instabilities due to melting developed on the top layer of the subducting slab and the unconventional temperature field with tongue-like features produced by a subducting slab (McKenzie, 1969, Marsh, 1979a,b, Davies and Stevenson, 1992, Zhao et al., 1995, Gorbatov et al., 1999, Hall and Kincaid, 2001, Tamura et al., 2002, Tsumura et al., 2000, Iwamori and Zhao, 2000; Zhao, 2001). Any sort of melting would introduce a source of chemical buoyancy because of the differentiation process. From kinematic thermal models (e.g. Oxburgh and Turcotte, 1971), it is well known from geometrical considerations that the temperature gradient normal to the surface of the descending lithosphere is inverted because of the presence of secondary upward mantle circulation (e.g. Hsui and Töksöz, 1979) flowing toward the wedge. This type of thermal-compositional anomalies leads favorably to thermal-chemical Rayleigh-Taylor instabilities (Hall and Kincaid, 2001; Gerya and Yuen, 2002a). They would also lead to localized melting and attendant magmatic and seismic activities caused by dehydration of the serpentine crust (Tamura et al., 2001, Seno et al., 2001; Zhao et al., 2001; Rüpke et al., 2002). In Figure 1 we depict the outcome from this unusual thermal-chemical dynamical situation as a 3-D conceptual model, where the core of the slab is colder than the top part of slab surface consisted of oceanic crustal material. Three-dimensional diapiric structures are suggested to form from thermal-chemical instabilities developed at a depth of around 100 km to 200 km, as a result of the dehydration of serpentines and the subsequent hydration reactions involving peridotites (e.g. Ulmer and Tromsdorff, 1995; Plank and Langmuir, 1998).

In this paper we will investigate from 2-D numerical modelling this geodynamically plausible scenario of cold but compositionally lighter plumes emerging from the upper surface of the slab due to a supercritically unstable situation in thermal-chemical convection (e.g. Hansen and Yuen, 1990). This unusual situation is caused by the colder subducting lithosphere consisting of
an intrinsically lighter hydrated peridotite lying underneath a hotter but compositionally heavier dry overlying mantle.

In section 2 we will go over the initial and boundary conditions of the 2-D model. In section 3 we will describe the overall hydration and dehydration processes involved in this regional model consisting of the slab and the mantle wedge. In section 4 we describe the implementation of a melting model. The complex rheological model associated with the slab and the mantle wedge is described in section 5. We present in section 6 the fundamental equations of 2-D thermal-chemical convection and also the numerical method. The results of our numerical simulation illustrating the dynamical feasibility of cold plumes formation are shown in section 7. Finally, in section 8 we summarize our findings and present our conclusions.

2. Initial and Boundary Conditions of 2-D Model

From our 3-D conceptual model we have designed a regional 2-D model based on finite-differences (Gerya et al., 2002b) that takes into account the process of hydration of the mantle wedge by the fluid released from a kinematically prescribed subducting plate. Figure 2 shows the initial conditions (panel a), the imposed kinematic boundary conditions for both the thermal and velocity fields (panel b) and the hydration model used (panel c). We assume that dehydration of the subducting slab liberates an upward migration of aqueous fluid, resulting in the hydration of the mantle wedge near the slab (see panel c). The hydration leads to a sharp decrease in density and viscosity of mantle rocks, creating very favorable conditions for the development of compositionally driven Rayleigh-Taylor instability along the propagating hydration front (Figures 1 and 2). This compositional difference arises from the water contents.

The initial conditions and other essential features, such as the lithologic stratification and isotherms are displayed in Figure 2a. The subduction zone is prescribed by a weak layer (e.g., [Kincaid and Sacks, 1997; Schmeling et al., 1999; van Hunen et al., 2000) with a uniform thickness of 8 km and with a changing inclination caused by the slab bending. In view of the possible role played by water in starting subduction [Regenauer-Lieb et al., 2001], this weak layer is prescribed as a brittle fault within mantle rocks, characterized by selected wet olivine rheological parameters, and a high pore fluid pressure is assumed. During subduction, this predefined fault zone is substituted by weak subducted crustal rocks and hydrated mantle, implying a decoupling along the plate interface (e.g., [Zhong and Gurnis, 1995]).
The initial temperature field in the subducting plate, with a uniform descending rate of 2 cm/yr, (Figure 2a) is defined by an oceanic geotherm \( T_0(z) \) due to a cooling half-space model with a specified age. Initial temperature distribution in the overriding plate corresponds to the equilibrium thermal profile with 0°C at the surface and 1350°C at 32 km depth. During the course of the numerical simulation this geotherm \( T_1(z) \) on the right in Fig. 1b is maintained and not affected by the mantle circulation in the wedge region.

Figure 2b shows the boundary conditions used in our model. These kinematic boundary conditions correspond to the corner flow model (e.g. Stevenson and Turner, 1977, Tovish et al., 1978); modified to account for asthenospheric mantle flow at temperatures exceeding 1000°C in the mantle wedge (e.g., [Davies and Stevenson, 1992; Kincaid and Sacks, 1997; Gerya et al., 2002]). The following conditions were defined for the lower-right corner of the model (see panel b): (i) a given incoming horizontal velocity from the right boundary and (ii) a compensating outward velocity inclined parallel to the subducting plate at the lower boundary. To insure a smooth continuity of the temperature field across the lower boundary of the truncated regional model in the absence of a strong local heat-source, such as viscous heating, we have used a boundary condition involving the vanishing of the vertical Lagrangian heat flux, \( q_z \), with depth corresponding to the boundary condition \( \partial q_z / \partial z = 0 \) [Turcotte and Schubert, 1973].

The top surface is calculated dynamically at each timestep like a free surface (e.g. Poliakov and Podladchikov, 1992; Gurnis et al., 1996). We have used an approximate scheme to account for the changes in the topography, we used a layer with a lower viscosity (10^{18} \text{ Pa s}) whose initial thickness is 8 km on the top of the oceanic crust. The density of this layer is taken to be 1 kg/m^3 (air) at \( z < 4 \) km and 1000 kg/m^3 (sea water) at \( z > 4 \) km. An interface between this layer and the top of the oceanic crust is considered to be the erosion/sedimentation surface, which evolves according to the following transport equation

\[
\frac{\partial z_{es}}{\partial t} = v_z - v_x \frac{\partial z_{es}}{\partial x} - v_s + v_e,
\]

where \( z_{es} \) is the vertical position of the erosion/sedimentation surface as a function of the horizontal distance, \( x \); \( v_z \) and \( v_x \) are the vertical and horizontal components of the material velocity vector at the erosion/sedimentation surface; \( v_s \) and \( v_e \) are, respectively sedimentation and erosion rates corresponding to the relation

\[ v_s = 0 \text{ mm/a}, \ v_e = 1 \text{ mm/a} \text{ when } z < 4 \text{ km}, \]
\[ v_s = 0.3 \text{ mm/a}, \quad v_c = 0 \text{ mm/a when } z>4 \text{ km}. \]

At each time step equation 1 is solved numerically to calculate the corresponding vertical displacement of the erosion/sedimentation surface. There is a slight dynamical feedback from the topographical variations to the underlying mantle circulation due to changes in the dynamic horizontal pressure gradients (e.g. Poliakov et al., 1993; Ellis et al., 2001).

3. Dehydration and Hydration Models

We have used a simple model (Gerya et al., 2002) for describing the dehydration of the subducting slab and the hydration of the mantle wedge (Fig. 2c). To account for the hydration process in a viscously deforming medium, we describe the vertical displacement of the hydration front with respect to the upper interface of the subducting slab (in Eulerian coordinates) by the following transport equation [Gerya et al., 2002]

\[
\frac{\partial z_h}{\partial t} = v_z - v_x \frac{\partial z_h}{\partial x} - v_h, \quad (2)
\]

where \( z_h \) is the vertical location of the hydration front as a function of the horizontal distance, \( x \), measured from the trench; \( v_h \) is the hydration rate; \( v_z \) and \( v_x \) are the vertical and horizontal components of the material velocity vector at the hydration front. At each time step, equation 2 is solved for obtaining the corresponding vertical displacement of the hydration front with respect to the upper surface of subducting plate.

Following the arguments of Schmidt and Poli [1998], who favored a continuous rather than a single phase dehydration model for the subducting lithosphere, we assume that the rate of fluid release along the slab (e.g., [Kerrick and Connolly, 2001]) and, consequently, the hydration rate can be roughly approximated by a linear function of the horizontal distance, \( x \), measured from the trench (Figure 2c)

\[
\frac{v_h}{v_s} = A[1-Bx/x_{\text{lim}}] \quad \text{when } x< x_{\text{lim}}
\]

\[
\frac{v_h}{v_s} = 0 \quad \text{when } x> x_{\text{lim}}
\]

where \( v_s \) is the subduction rate; \( x_{\text{lim}} \) is the limiting distance from the trench (Fig. 2c), corresponding to the horizontal distance to the right of which fluid release from the subducting plate is negligible (see [Gerya et al., 2002]). The parameter \( A \) in equation 3 may be interpreted as a non-dimensional intensity in the hydration process of the mantle wedge. The non-dimensional parameter \( B \) may vary from -1 to 1, characterizing either the increase (\( B<0 \)) or the decrease
(B>0) in hydration rate with depth. As shown by Gerya et al. [2002], typical values of the parameter A range between 0.01 and 0.30.

4. Partial Melting Model

To model the changes in the physical properties of rocks from partial melting, we have adopted the algorithm developed by Bittner and Schmeling [1995] for modelling magmatic underplating. According to this model, melting of hydrated peridotite and subducted rocks (oceanic crust and sediments) occurs in the \( PT \)-region between the wet solidus and dry liquidus of all petrological components. As a first approximation, the degree of melting is taken to increase linearly with the temperature according to the relations

\[
M = \begin{cases} 
0 & \text{at } T < T_{\text{solidus}}, \\
(T - T_{\text{solidus}})(T_{\text{liquidus}} - T_{\text{solidus}}) & \text{at } T_{\text{solidus}} < T < T_{\text{liquidus}}, \\
1 & \text{at } T > T_{\text{liquidus}},
\end{cases}
\]

where \( M \) is the volumetric fraction of melt with temperature; \( T_{\text{solidus}} \) and \( T_{\text{liquidus}} \) are, respectively, wet solidus and dry liquidus temperature at given pressure and rock composition (Johannes, 1985; Hess, 1989; Schmidt and Poli, 1998; Poli and Schmidt, 2002). Because the solidus curve varies nonlinearly with pressure, the melting interval varies nonlinearly with pressure.

An effective density, \( \rho_{\text{eff}} \), of partially molten rocks is calculated from the formula

\[
\rho_{\text{eff}} = \rho_{\text{solid}} - M(\rho_{\text{solid}} - \rho_{\text{molten}}),
\]

where \( \rho_{\text{solid}} \) and \( \rho_{\text{molten}} \) are, respectively, densities of solid and molten rock varying with pressure and temperature according to the relation

\[
\rho_{P,T} = \rho_0 \left[ 1 - \alpha(T - T_0) \right] \times \left[ 1 + \beta(P - P_0) \right],
\]

where \( \rho_0 \) is the density at the standard pressure (0.1 MPa) \( P_0 \) and temperature \( T_0 \) (298 K); \( \alpha \) and \( \beta \) are, respectively, the thermal expansion and compressibility coefficients, which are taken to be constant.

The effects of latent heating are accounted for by an increased effective heat capacity of partially molten rocks (\( C_{\text{p,eff}} \)) calculated according to the equation

\[
C_{\text{p,eff}} = C_p + H_L/(T_{\text{liquidus}} - T_{\text{solidus}}),
\]

where \( C_p \) is heat capacity of the rock assemblage; \( H_L \) is latent heat of melting for the lithological unit, J/kg.

5. Rheological Model
We are dealing with a composite rheological set-up in our model, because of many constituents and lithological phases. At temperatures lower than about 700 K, we have followed the approach developed by Schott and Schmeling (1998), wherein the creep rheology for the solid rocks ($M<0.1$) is combined with a quasi-brittle rheology to yield an effective rheology. For this purpose the Mohr-Coulomb law [Brace and Kohlstedt, 1980; Ranalli, 1995] is simplified to the yield stress, $\sigma_{\text{yield}}$ criterion and implemented by a “Mohr-Coulomb-viscosity”, $\eta_{\text{MC}}$ as follows

$$\eta_{\text{MC}} = \sigma_{\text{yield}}/(4\varepsilon_{\text{II}})\frac{1}{2}, \quad (8)$$

$$\sigma_{\text{yield}} = (M_1P_{\text{lith}} + M_2)(1-\lambda).$$

where $\varepsilon_{\text{II}}=\frac{1}{2}e_{ij}e_{ij}$ is the second invariant of the strain rate tensor, with dimension $s^{-2}$; $\lambda = P_{\text{fluid}}/P_{\text{lith}}$ is the pore fluid pressure coefficient, i.e. the ratio between pore fluid pressure, $P_{\text{fluid}}$, and lithostatic pressure, $P_{\text{lith}}$; $M_1$ and $M_2$ MPa are empirical constants ($M_1=0.85, M_2=0$ when $\sigma_{\text{yield}} < 200$ MPa and $M_1=0.6, M_2=60$ MPa when $\sigma_{\text{yield}} > 200$ MPa [Brace and Kohlstedt, 1980]).

The total effective viscosity, $\eta$, is then defined by the following criterion

$$\eta = \eta_{\text{creep}} \quad \text{when} \quad 2(\varepsilon_{\text{II}})^{\frac{1}{2}} \eta_{\text{creep}} < \sigma_{\text{yield}},$$

$$\eta = \eta_{\text{MC}} \quad \text{when} \quad 2(\varepsilon_{\text{II}})^{\frac{1}{2}} \eta_{\text{creep}} > \sigma_{\text{yield}}, \quad (9a) \quad (9b)$$

where $\eta_{\text{creep}}$ is the creep viscosity, Pa s. The creep viscosity depending on the stress and temperature, is defined in term of deformation invariants by [Ranalli, 1995]

$$\eta_{\text{creep}} = (\varepsilon_{\text{II}})^{1-n} F (A_D)^{-1/n} \exp(E/nRT), \quad (10)$$

where $F$ is a dimensionless coefficient depending on the type of experiments on which the flow law is based (e.g., $F=2(1-n)/3(1+n)^{2n}$ for triaxial compression and $F=2(1-2n)/n$ for simple shear).

High pressure metamorphic rocks exhumed from subduction zones reveal little evidence of dislocation deformation of their principal minerals at depths from 40 to 150 km [Stöckhert and Renner, 1998; Renner et al., 2001; Mauler et al., 2001; Stöckhert, 2002]. Therefore, a constant Newtonian creep viscosity $\eta_{\text{creep}}$ (see Gerya et al. [2002]) is assigned to the subducted sediment, the upper hydrated portion of the basaltic oceanic crust, and to the serpentinized mantle, while a power-law rheology is used for the lower gabbroic portion of the oceanic crust, and for the unserpentinized mantle material. As the appropriate flow laws allowing for more realistic predictions are not yet available, the Newtonian rheology is implemented by using a constant effective viscosity $\eta_{\text{creep}}$ with absolute values of $10^{18}$, $10^{19}$, or $10^{20}$ Pa·s, used in the different numerical experiments.
The rheology of the lower gabbroic layer of the oceanic crust is taken to be represented by a flow law for dislocation creep of plagioclase with composition An\textsubscript{75}, with $E = 238$ kJ·mol\textsuperscript{-1}, $n = 3.2$, and $\log A_D = -3.5$ ($A_D$ given in MPa\textsuperscript{n}·s\textsuperscript{-1}), as quoted in Ranalli [1995]. For this material a high pore fluid pressure is assumed ($\lambda = 0.90$ in equation 8).

The rheology of the unhydrated mantle is represented by a flow law for dislocation creep of dry olivine, with $E = 532$ kJ·mol\textsuperscript{-1}, $n = 3.5$, and $\log A_D = 4.4$ ($A_D$ given in MPa\textsuperscript{n}·s\textsuperscript{-1}) [Ranalli, 1995]. The brittle strength is assumed to be high, because of the absence of a free pore fluid ($\lambda = 0$ in equation 8).

The rheology of the partially hydrated mantle beyond the antigorite stability field is represented by a flow law for dislocation creep of wet olivine, with $E = 470$ kJ·mol\textsuperscript{-1}, $n = 4$, and $\log A_D = 3.3$ ($A_D$ given in MPa\textsuperscript{n}·s\textsuperscript{-1}) [Ranalli, 1995], assuming a high pore fluid pressure ($\lambda = 0.90$ in equation 8).

Lastly, for the melts we have adopted the scheme proposed by Bittner & Schmeling (1995), the effective viscosity, $\eta$, of molten rocks ($M > 0.1$) was calculated by using the following formula [Pinkerton and Stivenson, 1992]

$$\eta = \eta_0 \exp\{2.5 + [(1-M)/M]^{0.48}(1-M)\},$$

(11)

where $\eta_0 = 10^{13}$ Pa s is taken for the basic oceanic crust and hydrated peridotite and $\eta_0 = 5 \times 10^{14}$ Pa s is adopted from subducted felsic sedimentary rocks [Bittner and Schmeling, 1995].

6. Mathematical Modelling and Numerical Implementation

We have considered 2-D creeping flow wherein both thermal and chemical buoyant forces are included along with mechanical heatings from adiabatic compressional/expansion work and viscous dissipation in the temperature equation. We have elected to employ the primitive variables of velocity, $\mathbf{v}$, and pressure $P$ in the momentum equation. The conservation of mass is approximated by the incompressible continuity equation

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} = 0$$

(12)

which should be valid for the shallow upper-mantle regime considered here. The two-dimensional Stokes equations with both thermal and chemical buoyancies take the form:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} = \frac{\partial P}{\partial x}$$

(13)

$$\frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \sigma_{xz}}{\partial x} = \frac{\partial P}{\partial z} - g \rho(T,C,M)$$

(14)

The density in the vertical component of the momentum equation depends explicitly on the temperature $T$, the composition $C$, and the degree of melting $M$. We do not explicitly consider
the sub-grid process of percolation of melts (e.g. Scott and Stevenson, 1986) and aqueous fluids, which would require a different set of governing equations (e.g. Connolly and Podladchikov, 1998, Vasilyev et al, 1998). We note that in this kinematically-driven problem a Rayleigh number in the momentum equation cannot be properly defined. The buoyancy ratio \( R_\rho \) is defined to be \( \Delta \rho/\rho \alpha \Delta T \), where \( \Delta \rho \) is the compositional change from the reference density \( \rho \), \( \Delta T \) is the temperature difference and \( \alpha \), the thermal expansivity. The buoyancy ratio is an important control parameter in thermal-chemical convection (Hansen and Yuen, 1990).

We also employ realistic rheological constitutive relationships between the stress and strain-rate, whose coefficient \( \eta \) represents the effective viscosity \( \eta(P,T,C,\varepsilon_{II},M) \), which depends on the composition, temperature, pressure, strain-rate and degree of melting (see section 5)

\[
\begin{align*}
\sigma_{xx} &= 2\eta \varepsilon_{xx} \\
\sigma_{xz} &= 2\eta \varepsilon_{xz} \\
\sigma_{zz} &= 2\eta \varepsilon_{zz} \\
\varepsilon_{xx} &= \partial v_x / \partial x \\
\varepsilon_{xz} &= \frac{1}{2}(\partial v_x / \partial z + \partial v_z / \partial x) \\
\varepsilon_{zz} &= \partial v_z / \partial z \\
\end{align*}
\]

As part of our new computational strategy (Gerya and Yuen, 2002b), instead of using the Eulerian frame of reference, we have adopted the computationally easier Lagrangian frame of reference in which the temperature equation with a temperature-dependent thermal conductivity \( k(T) \) takes the form

\[
\rho C_p (D T/D t) = \partial q_x / \partial x + \partial q_z / \partial z + H_r + H_a + H_s 
\]

(15)

\[
\begin{align*}
q_x &= k(T) \times (\partial T / \partial x) \\
q_z &= k(T) \times (\partial T / \partial z) \\
H_r &= \text{constant} \\
H_a &= T \alpha [v_x (\partial P / \partial x) + v_z (\partial P / \partial z)] \approx T \alpha \rho v_z g \\
H_s &= \sigma_{xx} \varepsilon_{xx} + \sigma_{zz} \varepsilon_{zz} + 2\sigma_{xz} \varepsilon_{xz} \\
k(T) &= M + N/(T+77)
\end{align*}
\]

where \( D/D t \) represents the substantive time derivative.

Other notations are: \( x \) and \( z \) denote, respectively, the horizontal and vertical coordinates, in m, \( v_x \) and \( v_z \) are components of velocity vector \( \mathbf{v} \) in m/s\(^{-1}\); \( t \) time in s; \( \sigma_{xx}, \sigma_{xz}, \sigma_{zz} \) are components of the viscous deviatoric stress tensor in Pa; \( \varepsilon_{xx}, \varepsilon_{xz}, \varepsilon_{zz} \) are components of the strain rate tensor in s\(^{-1}\).
\( P \) the pressure in Pa; \( T \) the temperature in K; \( q_x \) and \( q_z \) are horizontal and vertical heat fluxes in \( \text{W} \cdot \text{m}^{-2} \); \( \eta \) the effective viscosity (see section 5) in \( \text{Pa} \cdot \text{s} \); \( \rho(T,C,M) \) the density in \( \text{kg} \cdot \text{m}^{-3} \), dependent on the temperature, composition and degree of melting (see section 4) \( g=9.81 \, \text{m} \cdot \text{s}^{-2} \) is the gravitational vector; \( k(T) \) is the temperature-dependent thermal conductivity \( \text{W} \cdot \text{m}^{-1} \cdot \text{K}^{-1} \), depending on the chemical composition, and temperature with empirical coefficients \( M \) and \( N \) taken for different rocks according to \( \text{Clauser & Huenges, 1995} \); \( C_p \) is the isobaric heat capacity in \( \text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1} \); \( H_r, H_a, \) and \( H_s \) denote, respectively, radioactive, adiabatic and shear heat production in \( \text{W} \cdot \text{m}^{-3} \). The dissipation number of this regional model is around 0.04.

We have employed a recently developed 2-D code (I2VIS) based on arbitrary order of accuracy finite-differences with a marker technique allows for the accurate conservative solution of the governing equations on rectangular fully staggered Eulerian grid for multiphase viscoplastic flow \( \text{Gerya and Yuen, 2002b} \). This control volume based Eulerian/Lagrangian formulation of momentum and temperature equations provides conservation of both stress fields and heat flows. Advective transport solution for variable material properties and temperature is based on a non-diffusive characteristics-based marker technique. This numerical technique with the direct solver of the momentum and temperature equations \( \text{Gerya and Yuen, 2002b} \) can handle the variable P-T dependent thermal conductivity and density and P-T- and strain-rate dependent viscosity. This code can also cope with rapid advection, sharp discontinuity of both thermal and material properties, high viscosity contrast varying by at least 8 orders in magnitude, adiabatic and shear heating and thermomechanical influence from partial melting. A detailed description for the numerical method and algorithmic tests is provided in \( \text{Gerya and Yuen (2002b)} \).

7. Numerical and Visualization Results

All models presented below have been run on a finite-difference grid of 200x100 regularly spaced Eulerian points and with 500,000 markers for portraying the fine details of the temperature, material and viscosity fields. We have carried out over fifty runs to map out the proper parameter space for the occurrence of cold plumes. Lateral viscosity contrasts, up to \( 10^7 \), are maintained in the models shown here. We call attention to the color bar consisting of 9 colors, which represent the different constituents, consisting of sediments, partially molten sediments, oceanic crust (including basaltic and gabbroic portions, Fig. 2a), partially molten oceanic crust, dry mantle, serpentinized mantle, hydrated serpentine-free mantle and partially
molten hydrated mantle. In Fig. 3 we show the temperature field portrayed as isotherms and the distinct lithologic structures of the subducted oceanic crust and melt in the left-hand column of Fig. 3 and the effective viscosity+velocity fields in the right-hand column, with the sequence of events going from top to bottom. The elapsed time shown is around 30 Ma. Henceforth, we will refer to Figs. 3 and Fig. 4 as the reference model.

The initial stages, younger than 25 Ma, are characterized by the progressive hydration of the mantle wedge resulted in the propagation of hydration front (light bright-blue + light gray-blue) from the subducting slab. The hydrated mantle is clearly subdivided into two parts: an upper, serpentinized subduction channel (light gray-blue) developing within an upper colder lithospheric portion of the wedge and a lower hydrated serpentine-free peridotite zone (light bright-blue) developing within a lower hotter asthenospheric portion outside of $P-T$ stability field of antigorite. These two zones are separated by the zone of narrowing and are also associated with strong lateral thermal gradient. This narrowing is caused by the forced flow of the hot asthenosphere toward the subducting plate, which is controlled by the conservation of mass. The narrowing also produces a wedge-shaped closure of the serpentinized subduction channel, producing a circulation involving many petrologic constituents within the channel (see Gerya et al., 2002).

Afterwards, from 25 to 34 Myr, there emerges a Rayleigh-Taylor instability driven by compositional variations, which are caused by hydration along the upper surface of the serpentine-free (light bright-blue) hydrated peridotite zone. This instability is caused by the strong (100-200 kg/m$^3$) density contrast (e.g., [Hall and Kincaid, 2001]) between the hydrated lower density peridotite and the higher density of the overriding dry asthenosphere. The density contrast is partially compensated, but not completely cancelled, by the significant (300-400°C) temperature difference between the hydrated and dry mantle. The buoyancy ratio $R_\rho$ for these models lies between 3.0 and 4.5 and is very supercritical and chemically dominated. The diapirs generated by hydrated peridotites rise upward from the subducting plate and then go up to the hotter mantle wedge. Viscous heating of the diapirs facilitates the partial melting and further maintains their buoyancy. Melting also results in a strong lowering of the diapir viscosity, thus producing rapidly upwelling finger-shaped “cold plume”, characterized by an "inverted" thermal structure, where the core is considerably colder than the the rims.
The upward plume velocity can be very significant (around tens of centimeters per year) and exceeds the subduction speed by a long shot. In several numerical experiments most intensive development of the “cold plumes” may trap significant amount of subducted oceanic crust and sedimentary rocks from the top of the subducting slab, thus producing a complex mixture of partially molten rocks with both subducted and mantle origins lodged within the wedge. The depth for the appearance of the diapirs is defined by a critical time window available for the growth of the Reyleigh-Taylor instabilities, which are displaced downward along the hydration front by an intense flow. In case of smaller $R_p$, these instabilities do not have enough time to produce plumes at the depth of 200 km.

Concerning the viscosity field, a pronounced low viscosity channel is developed along the upper surface of the slab (right-hand column of Fig. 3) and the core of the rising cold plume has a low viscosity because of strong strain-rate dependence of the viscosity and the partial-melt rheology. We note that the viscosity fields in the core of slab in the left-hand side of the panels in the right column of Fig. 3 cannot be considered to be realistic because of the kinematic nature of the slab velocity boundary conditions (see the regular patterns of the arrows), which give rise to the sharp edges in the viscosity field. Inside the slab the temperature field evolves according to advection-diffusion heat equation with adiabatic heating.

In Fig. 4 we display the temperature field (left-hand column) visualized in a continuous manner (see the color bar) and the distribution of viscous heating (right-hand column), which has been normalized by heating due to radioactive elements present in peridotite (Turcotte and Schubert, 1982). The paradoxical nature of the cold rising plume is illustrated clearer by the upward protrusion of the blue finger in the temperature field. There is a great amount of viscous heating produced by this thermal-chemical plume, due to a decrease in the gravitational potential energy from the rising of lighter material, which in turn is converted efficiently to mechanical heating by the large $R_p$ ratio (Hansen and Yuen, 1995, 1996; Myasnikov et al., 1999). Again, because of the kinematic nature of the velocity boundary condition of the slab, we have not considered the shear heating associated with slab bending (e.g. Conrad and Hager, 1999) in the temperature equation. Therefore, one should only consider the shear heating distribution at the right-hand side of the panels in Fig. 4, where the fluid motions are not prescribed but solved as part of the creep momentum equations (eqns. 13 and 14).
There are many other parameters to vary in this highly nonlinear problem. Data-assimilation treatment of this problem by adjoint equations (e.g. Marchuk, 1995) can be handled by using the cost-function formalism developed for finite Prandtl number plumes at high Rayleigh number by Hier-Majumder et al. (2002) but this venture would be beyond the scope of this paper. In Fig. 5 we show four examples illustrating the sensitivities of the solution to the variations in parameter. We note that in Fig. 5 the buoyancy ratio $R_\rho$ has been increased by 20%. First, we test the sensitivity of the solution to a doubling of the hydration rate in Fig. 5a. The instability triggered by greater fluxing of hydrous fluids develops much earlier at time of 26.3 Myr than the reference model (see Fig. 3) and results in a thick serpentinized layer (grey-blue). A cold plume propagates powerfully from the descending slab right under the lithosphere, leading to the possibility of underplating. Next, we study the influence of having a younger lithosphere in Fig. 5b (upper-right panel), where a younger lithosphere with an age of 60 Myr has been used as the initial condition. Many more instabilities than the reference model are developed along with a thick serpentinized layer (grey-blue). The tips of the two upwellings have different melting components, showing the diversity of the magmatic outcomes in this model. We show in Fig. 5c that the instability can also develop without shear heating but this would require a greater amount of chemical buoyancy than in the reference model. Because of the lack of shear heating, the instabilities can only develop deeper down, leading to more melting of the peridotite (red color) inside the plume, due to the crossing of the wet peridotite solidus curve by the temperature field near the slab interface. Finally, in Fig. 5d (lower right panel) we look at the effects of an older age slab of 100 Ma. After 21.6 Myr the thermal-chemical instability reaches an apex characterized by a thick head filled with partially molten sediments and basalts embedded in hydrated peridotite. Afterwards this upwelling loses its potency and is bent by the downward circulation. There is a great richness in the petrologic possibilities of the melt developed in these thermal-chemical plumes.

We focus our attention on the influence of shear heating on the development of these compositionally driven instabilities. We go back to the same buoyancy ratio $R_\rho$ of 4.0, used in Fig. 3 and 4. In Fig. 6 we illustrate the effects of varying amounts of shear heating by displaying the isotherms superimposed on the petrological structure fields. The amount of shear heating employed is greatest in Fig. 6a, where shear heating along the length of the entire slab is accounted for and includes the brittle deformation field at low temperatures lower than 600 K as
well, intermediate in Fig. 6b, where only shear heating at temperatures greater than 600 K is included in the temperature equation and no shear heating at all in Fig. 6c. A comparison of these three situations show clearly the importance of shear heating in promoting these cold plumes, since in the case without shear heating the instability is hardly excited. in contrast to the strong instability found in the case where shear heating is everywhere (Fig. 6a). More work is needed in studying the role of frictional heating in the brittle regime. This comparison show that at $R_\rho$ of 4.0 even small addition of frictional heat at shallow depths within a subduction zone is crucial for the generation of the cold plumes within a critical time window of around 10 to 20 Myr. Another possible source of heat at these depths may be due to the latent heat of the serpentinization of mantle wedge that is not considered in our model.

8. Conclusions and Perspectives

In this work we have demonstrated the dynamical feasibility for upwellings, around 300 to 400 degrees colder than the ambient mantle, to emerge from the top surface of the descending lithosphere and to pass through the mantle wedge. These counter-intuitive cold hydrous plumes are powered primarily by compositional buoyancy, which is maintained by the presence of partially molten material due to melting triggered by fluxing of hydrous fluids released from many realistic (Schmidt and Poli, 1998) dehydration reactions, such as serpentine decomposition, at depths greater than around 100 km. These ascending plumes with a diverse origins of melting components, are also lubricated by a sheath of reduced viscosity layer due to the heat produced by the large amount of viscous dissipation, which is caused by the efficient conversion of gravitational potential energy into mechanical dissipation (Myasnikov et al., 1999). The amount of viscous dissipation owing to the upward plume movement increases nonlinearly with the buoyancy ratio $R_\rho$ between compositional and thermal buoyancy (Hansen and Yuen, 1995, Myasnikov et al., 1999). In spite of the small dissipation number and the relatively short vertical trajectory, at most 200 km, these chemical plumes produce much more viscous heating than their thermal counterparts from the lower mantle.

The upward passage of these cold plumes would exert important consequences on the thermomechanical evolution of subducting zones. First, they should penetrate rapidly through the hot asthenosphere in the mantle wedge and, second, this would cause local cooling. Instead of being interpreted as hot anomalies (Tamura et al., 2002) these low seismic velocity anomalies
(Zhao et al., 2002) near Japan may be interpreted as having some water content (Jung and Karato, 2001), which would induce anisotropy on the seismic waves and contaminate the common seismic interpretation of hot thermal anomalies to be associated with slow seismic waves. The dynamical consequences of this chilling of the mantle wedge by the cold plumes would lead to a more rapid descent of the slab, which would result in a more rapid subduction of this portion of the wedge. This would induce a local replenishment of hot asthenospheric material, which is then followed by another short period of cold plume activity after tens of millions of years.

The regional mantle circulation and the thermal structure developed in our numerical models are quite complex and differ significantly from those obtained by only considering thermal effects (e.g., [Peacock 1990, ; Davies and Stevenson, 1992; van Hunen et al., 2000]). Our findings truly illustrate the multiscale nature of mantle convection by displaying another hierarchy of secondary plumes emerging from the slab, a primary boundary layer in mantle convection. Our results also appear to be more realistic when they are compared with the seismic data revealing the complexity of seismological signatures in the vicinity of subduction zones, particularly near Japanese trench area and at various depths (e.g., [Tamura et al., 2001, 2002; Seno et al., 2001; Omori et al., 2002, Zhao et al., 2002]). Therefore, we believe that our numerical approach and additional comparison with geophysical and petrological constraints may lead to a deeper understanding of the cause and effect relationships in the seismic and magmatic activities at convergent plate boundaries and with the age of the subducting lithosphere.

Besides the subducting slab, we may speculate further that similar “cold plumes” can also be produced at much greater depths, such as the transition zone, a site for the development of secondary plumes (e.g. Cserepes and Yuen, 2000). Another set of cold plumes can develop from the hydrated portions of deeply subducted or even detached slabs, lying atop the 660 km discontinuity. In summary, we argue strongly that a significant part of the magmatic and seismic activities in the Earth’s interior may be related to development of cold plumes with compositional, hydrous origins. Therefore, we would need to investigate further this new intrinsically nonlinear phenomenon by a multi-facted approach involving the combined efforts of seismology, rheology, petrology and numerical modelling, since there is a great sensitivity of the geological and geophysical outcomes to the various input parameters.
Acknowledgements. This work was supported by RFBR grant # 00-05-64939, by an Alexander von Humboldt Foundation Research Fellowship to TVG, the geophysics program of the National Science Foundation and by the Deutsche Forschungsgemeinschaft within the scope of Sonderforschungsbereich 526 “Rheology of the Earth - from the upper crust into the subduction zone” at Ruhr-University. The authors appreciate discussions with J. Van Decar, concerning the recognition of the “cold plume” phenomena, Tetsu Seno, Shigenori Maruyama and S. Omori, concerning Japanese plate tectonics and other matters. We thank Lilli Yang for drawing Figure 1.

References


B.D Marsh, Island-arc development: some observations, experiments, and speculations, J. Geol. 87 (1979a) 687-713.

B.D Marsh, Island arc volcanism, American Scientist 67 (1979b) 161-172.


T.V. Gerya, Institut für Geologie, Mineralogie und Geophysik, Ruhr-Universität Bochum, 44780 Bochum, Germany (e-mail taras.gerya@ruhr-uni-bochum.de)

D.A. Yuen, Minnesota Supercomputer Institute and Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota, 55455-0219 USA (e-mail davey@msi.umn.edu)

(Received ; revised ; accepted )
Figure 1. Conceptual 3-D model of the cold plume generation process. The perceived depth of the model extends down to the olivine to spinel transition. The cold plume is generated from a source layer of hydrated (non-serpentinized) peridotite located on the top of the subducting slab and produced by hydration and cooling of material of the mantle wedge. Aqueous fluid, responsible for the hydration process, is produced by dehydration reactions within the slab, including the decomposition of serpentine.
Figure 2. Design of (a) initial, (b) boundary conditions and (c) hydration model of the mantle wedge used in our 2-D numerical experiments. The computational domain is regional in character, and kinematic boundary conditions must be imposed. The grid configuration is 200x100 regularly spaced points and 500,000 markers are employed. The subduction speed is assumed to be 2 cm/yr with an initial angle of 45 degrees. See the text for other details.
Figure 3. Development of the geometry of the mantle wedge during early (0-15 Myr) stages of the ongoing subduction. The inset separate sketches represent enlarged 320x200 km areas of the
original 400 km x 200 km models. Left: Evolution of the temperature field (isotherms labelled in °C) and distribution of rock types (color code): 1 = weak layer on the top of the model (see Fig. 2a), 2 and 3 = solid partially molten sedimentary rocks, respectively; 4 and 5 = solid and partially molten oceanic crust (both basaltic and gabbroic, see Fig. 2a), respectively; 6 = unhydrated mantle, 7 = serpentinized mantle, 8 = hydrated unserpentinized (i.e. beyond stability field of antigorite) mantle; 9 = partially molten hydrated mantle. Right: Evolution of viscosity (color code) and velocity field (arrows).

Model parameters: box size – 400km x 200 km; grid resolution 200x100 points; geometry representation – 500,000 markers; hydration rate parameters; \( A = 0.05 \), \( B = 0 \), \( x_{\text{lim}} = 200 \) km; subducting lithosphere age 80 Myr; subduction rate 2cm/year; temperature-dependent thermal conductivity \([\text{Clauser & Huenges}, 1995]\), \( k, \text{W m}^{-1}\text{K}^{-1}\) – sedimentary rocks = \( 0.64 + 807/(T+77) \), oceanic crust = \( 1.18 + 474/(T+77) \), mantle rocks = \( 0.73 + 1293/(T+77) \); isobaric heat capacity, J/K – for all rock types = 1000; standard density, \( \rho_0, \text{kg/m}^3\) – sedimentary rocks = 3000(solid)/2400(molten), oceanic crust = 3200(solid)/2900(molten); mantle = 3300(dry)/3000(serpentinized)/3200(hydrated)/3000(molten); thermal expansion coefficient, \( \alpha, \text{K}^{-1}\), for all rock types = \( 3 \times 10^{-5} \); compressibility coefficient, \( \beta, \text{MPa}^{-1}\) for all rock types = \( 1 \times 10^{-5} \); newtonian viscosity of serpentinised mantle, sedimentary rocks and basaltic crust = \( 10^{19} \text{ Pa s} \); viscosity variations = \( 10^{18}-10^{25} \text{ Pa s} \).
Figure 4. Development of the temperature field (left column) and shear heating distribution (right column) of the model shown in Figure 3. $R_p$ lies between 3.0 and 4.5. Shear heating has been normalized by heating due to radioactive elements present in peridotite (Turcotte and Schubert, 1982). Temperature field is continuous in character and varies according to the color bar.
Figure 5. Sensitivity of the solutions to some input parameters. The buoyancy ratio $R_\rho$ used for this figure is 5.0 which is due to the lower standard density of solid sedimentary rocks (2700 kg/m$^3$) and oceanic crust (3100 kg/m$^3$). (a) doubling the amount of hydration rate. (b) younger lithospheric age of 60 Myr. (c) no shear heating at all. (d) older lithospheric age of 100 Myr. Other parameters of the models and color code are the same as in Figure 3.
Figure 6 Influence of viscous dissipation on the development of cold plumes. $R_p$ is 4.0 for all models. (a) viscous heating is included at all temperatures. (b) viscous heating is accounted for only for temperatures above 600 K. (c) no shear heating is considered. Other parameters of the models and color code are the same as in Figure 3.