The effects of a composite non-Newtonian and Newtonian rheology on mantle convection

Arie P. van den Berg,[1] Peter E. van Keken[1] and David A. Yuen[2]

[1]*Department of Theoretical Geophysics, University of Utrecht, PO Box 80.021, Budapestlaan 4, 3508 TA Utrecht, The Netherlands*
[2]*Department of Geology and Geophysics, Minnesota Supercomputer Institute, and Army High Performance Computing Research Center, University of Minnesota, Minneapolis, MN 55415, USA*

SUMMARY

A finite element method based on a primitive variables formulation is used to model both steady-state and time-dependent mantle convection with a composite Newtonian and non-Newtonian (power-law) rheology. The rheological model employs the transition stress as a means of partitioning the relative importance of the two rheologies. Results show that there is no direct correlation between viscosity and temperature anomalies. Fluctuations of the velocity fields are much greater and faster than for Newtonian flows. Fluctuations with amplitudes several times the background velocity are quite common. Intermittency effects with quiescent periods punctuated by chaotic bursts are observed. From scaling arguments temporal fluctuations of the volume-averaged viscosity are comparable in magnitude to the variations in the surface heat flow for the non-Newtonian flows, but are smaller than the variations in the velocity field. At larger transition stress the Newtonian behaviour becomes dominant and the temporal variations of the viscosity diminish. Both steady-state and time-dependent results show that for a given transition stress the non-Newtonian behaviour prevails to a greater extent with increasing Rayleigh number. Implications of this non-Newtonian tendency for Archaean tectonics are discussed.

Key words: mantle convection, rheology, viscosity.

1 INTRODUCTION

Mantle convection has been studied for Newtonian rheological (McKenzie, Roberts & Weiss 1974; Hansen & Ebel 1988; Hansen, Yuen & Kroening 1990) and non-Newtonian models (Parmentier, Turcotte & Torrance 1976; Cserepes 1982; Christensen 1984; Christensen & Yuen 1989; King & Hager 1990; Weinstein 1991; Malevsky & Yuen 1992). With the increasing availability of supercomputers, investigation of computationally intensive time-dependent non-Newtonian convection has now become feasible. Improved graphics facilities have increased the capabilities to capture interesting phenomena inherent in time-dependent non-Newtonian convection. Recent experimental work (Karato & Li 1992) has shown that both linear and non-linear creep mechanisms may be significant in mantle flow processes. In this work we investigate the time-dependent mantle dynamics of a model from combining both Newtonian and non-Newtonian rheologies, as this aspect has not been studied before. We study the

influences from varying the relative contribution of the two creep mechanisms on the convective style of the mantle at various levels of convective vigor.

In Section 2 we describe the rheological model and the governing equations to be solved in the numerical modelling. In Section 3 we present the modelling results for steady-state and time-dependent models separately. In Section 4 we give the concluding remarks. Appendices A and B provide both the more detailed technical aspects and benchmark results.

2 MODEL PRESENTATION AND MATHEMATICAL FORMULATION

2.1 Rheological models

Most previous work on mantle convection with non-Newtonian fluids (Parmentier 1978; Cserepes 1982; Christensen 1984; Christensen & Yuen 1989) dealt with the power-law or Ostwald de Waele rheological models (Bird, Stewart & Lightfoot 1960; Schowalter 1978; Bird,

Armstrong & Hassager 1987, Ranalli 1987). This steady-state creep model is defined by the constitutive equation in dimensional form:

$$e_{ij} = B^{-1}\tau^{n-1}\tau_{ij}. \tag{1}$$

The strain-rate tensor e_{ij} in (1) is defined in terms of the velocity field u_i as

$$e_{ij} = \partial_i u_j + \partial_j u_i. \tag{2}$$

The scalar stress τ is defined as the second invariant of the deviatoric stress tensor τ_{ij}

$$\tau = [1/2\tau_{ij}\tau_{ij}]^{1/2}. \tag{3}$$

The power-law index n is greater than one in shear-thinning, pseudo-plastic fluids. A typical value of $n = 3$ is used in geophysical models (e.g. Parmentier, Turcotte & Torrance 1976). A form of (1) more convenient for computational purposes is obtained by inversion of (1);

$$\tau_{ij} = B^{1/n}e^{(1-n)/n}e_{ij} \tag{4}$$

where the scalar strain-rate e is defined to be the second invariant of the strain-rate tensor

$$e = [1/2e_{ij}e_{ij}]^{1/2}. \tag{5}$$

A generalized Newtonian viscosity model can be given as

$$\tau_{ij} = \eta(p, T, e)e_{ij}. \tag{6}$$

In the following we will neglect the temperature and pressure dependence of the viscosity η because we wish to focus only on the interaction between the non-linear and linear aspects of mantle rheology. Otherwise the parameter space would be greatly expanded. From the definitions given above we have for the effective viscosity in the power-law model

$$\eta = \tau_{ij}/e_{ij} = B^{1/n}e^{(1-n)/n}. \tag{7}$$

For $n > 1$—relevant in geophysical applications—this definition corresponds to a 'pseudo-plastic' material (Schowalter 1978) with a shear-thinning behaviour, i.e. the viscosity decreases with increasing strain rate. In the case of pure power-law rheology the viscosity is unbounded in the limiting case of vanishing strain rate. Thus stagnation points with very large viscosity values may occur in a power-law fluid. Experimental results show that the viscosity will be bounded because different flow processes will take over at low strain rates. Karato & Li (1992) found from experiments on perovskite analogues that diffusion creep with linear stress–strain-rate relation (for constant grain size) takes over from dislocation climb/glide mechanisms, which correspond to the power-law behaviour. Several viscosity models are used in the literature on fluid dynamics which include a Newtonian viscosity—independent of the strain rate at low stress levels and power-law viscosity at higher stress values. A composite stress-dependent rheology with Newtonian and power-law end members has been used in the context of plate tectonic models (Turcotte & Oxburgh 1972; Meisner & Vetter 1976). We will follow (Parmentier *et al.* 1976) and use the Ellis model (Bird, Stewart & Lightfoot 1960; Bird, Armstrong & Hassager 1987), introduced by Parmentier *et al.* (1976) in the context of mantle convection as a 'transition-stress' model. Another more recent model

with similar asymptotic behaviour at low strain rate is the Careau–Yasuda model (Bird *et al.* 1987).

The transition stress model can be interpreted in a straightforward way in terms of two simultaneous creep deformation mechanisms which dominate at low- and high-stress respectively. The two simultaneous deformation processes can be assumed to be cumulative.

$$e_{ij} = e_{ij}^* + e_{ij}^{(p)} = \left[\frac{1}{\eta^*} + \frac{1}{\eta^{(p)}}\right]\tau_{ij} \tag{8}$$

where the asterisk denotes the diffusion creep mechanism and the superscript (p) refers to the dislocation creep mechanisms. We note that (8) includes also the temperature and pressure dependence of the individual flow mechanisms. The activation energy for dislocation creep is much higher than that for the diffusion mechanism, at least for olivine (e.g. Ranalli 1987, 1991). This difference in the activation energies can have important geodynamical implications regarding the dominating style of creep in mantle convection. From (7) and (8) the generalized effective viscosity is

$$\eta(\tau) = \left[\frac{1}{\eta^*} + B^{-1}\tau^{n-1}\right]^{-1}. \tag{9}$$

The scalar stress invariant τ in (9) can be expressed in terms of the strain rate as

$$\eta(e) = \left[\frac{1}{\eta^*} + B^{-1/n}e^{(n-1)/n}\right]^{-1}. \tag{10}$$

We note here that (10) is an approximation to (9) as discussed by Parmentier *et al.* (1976). This equation is valid for both the Newtonian and non-Newtonian end-members but is only of an approximate nature near the transition stress (see Fig. 1). This expression is equivalent with the
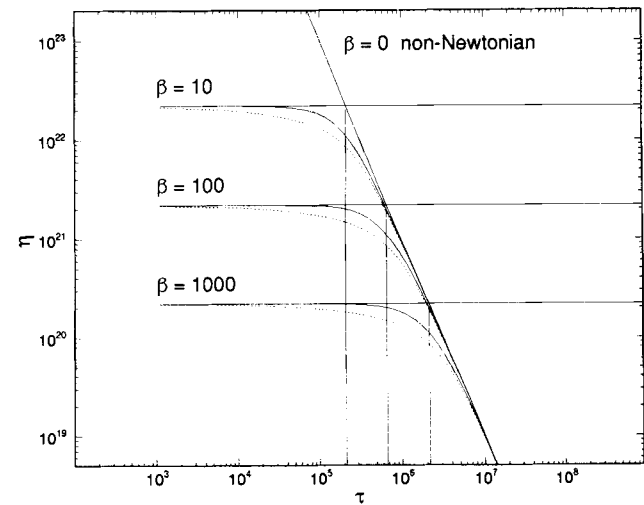

Figure 1. Dimensional viscosity versus scalar stress τ (mks units) for parameter values given in Table 1. $\beta = 0$ (inclined straight line): power-law case. Horizontal lines for $\beta = 10$, 100 and 1000 denote a constant diffusion-creep viscosity component. Solid curved lines display the effective stress-dependent visocosity computed from (9), approaching the Newtonian and power-law limits for low and high-stress values respectively. Dashed lines show the effective viscosity calculated from (10) using strain-rate values based on (9). Vertical lines denote the τ_T values given in Table 1.

Ellis model as described in Bird *et al.* (1960). The effective viscosity is the geometric mean of a Newtonian viscosity and a strain-rate dependent, non-Newtonian, power-law viscosity.

This composite rheology has also been employed in postglacial rebound studies (Gasperini, Yuen & Sabadini 1992). For non-dimensionalizing the equations for numerical modelling we will use the scaling applied in (Christensen & Yuen 1989) for a power-law fluid,

$$\mathbf{x} = x_0\bar{\mathbf{x}} = h\bar{\mathbf{x}}, \qquad t = t_0\bar{t} = h^2\kappa^{-1}\bar{t}, \qquad e = e_0\bar{e} = \kappa h^{-2}\bar{e} \qquad (11)$$

where h is the depth of the rectangular domain, κ the thermal diffusivity. We will consider fluids with depth-dependent rheological parameters, in particular, horizontally layered media. Henceforth, the two rheologies are assumed to be not thermally activated [$\eta(T,p)$ not assumed]. As in van den Berg *et al.* (1991), we will use $B = B(z)$ and define the pre-factor scale value as the value in the bottom (Nth) layer of the model (lower mantle),

$$B(z) = B_0\bar{B}(z) = B_N\bar{B}(z). \qquad (12)$$

The viscosity scale factor is defined as

$$\eta_0 = B_0^{1/n}e_0^{(1-n)/n} = B_0^{1/n}t_0^{(n-1)/n} = B_0^{1/n}(\kappa^{-1}h^2)^{(n-1)/n}. \qquad (13)$$

In the following we will use the simplified notation for the dimensional pre-factor

$$A(z) = B^{1/n}(z), \qquad A_0 = B_0^{1/n}. \qquad (14)$$

The transition value of the scalar stress, τ_T is defined as the stress level where we have an equi-partitioning between the Newtonian and non-Newtonian components of the strain rate. Its value follows from (9) as

$$\eta^{*-1} = B^{-1}\tau_T^{n-1}, \qquad \tau_T = [B/\eta^*]^{1/(n-1)}. \qquad (15)$$

In the following we will make use of the non-dimensional Newtonian viscosity. We define its inverse as a new parameter,

$$\beta = \frac{1}{\bar{\eta}^*} = \frac{\eta_0}{\eta^*}. \qquad (16)$$

The relation between β and τ_T follows from the definition of η_0 (13) and the transition stress (15)

$$\beta = \eta_0/\eta^* = \eta_0 B^{-1}\tau_T^{n-1}. \qquad (17)$$

There are no parameter values available from direct measurements of power-law creep in mantle silicates under lower mantle conditions. Assuming a power-law index $n = 3$, a strain rate $e = 10^{-15}\,\mathrm{s}^{-1}$ and a power-law viscosity value $\eta^{(p)} = 10^{21}\,\mathrm{Pa\,s}$ is consistent with a pre-factor value $B = 10^{33}\,\mathrm{Pa^3\,s}$. Table 1 presents values of η^* and τ_T for a number of values of the parameter β controlling the Newtonian component of the viscosity.

In Fig. 1 the stress dependence of the effective viscosity is plotted for the parameter combinations given in Table 1. Diffusional creep in silicates is strongly dependent on the grain size. The corresponding viscosity increases with the grain size d (Karato & Li 1992). Parmentier *et al.* (1976) report transition-stress values for olivine with a grain size $d \approx 1\,\mathrm{mm}$, $0.03\,\mathrm{MPa} < \tau_T < 0.1\,\mathrm{MPa}$. Recent experimental work by Katato & Li (1992) on perovskite analogue shows the transition stress to be much higher. For grain size 1 to

Table 1. Transition stress.

Model parameters (MKS units): $n = 3$, $B = 10^{33}$, $\kappa = 10^{-6}$, $h = 1.8 \times 10^6 \rightarrow \eta_0 = 0.219 \times 10^{24}$.

β	η^* Pa s	τ_T MPa
0	∞	0
10	0.2×10^{23}	0.214
100	0.2×10^{22}	0.676
1000	0.2×10^{21}	2.14

10 mm they report values of τ_T of approximately 10 MPa and 2 MPa respectively. As a consequence the influence of diffusional creep in lower mantle flow can be expected to be significant.

Applying the scaling defined in (13), (14) to (7), the dimensionless power-law component of the viscosity is obtained

$$\bar{\eta}^{(p)} = \frac{\eta^{(p)}}{\eta_0} = \bar{A}(z)\bar{e}^{(n-1)/n}. \qquad (18)$$

For the non-dimensional transition stress viscosity we find from applying the same scaling to (10), that

$$\bar{\eta}(e) = \frac{\eta}{\eta_0} = \left[\beta + \frac{1}{\bar{A}\bar{e}^{(1-n)/n}}\right]^{-1}. \qquad (19)$$

In the following we will drop the overbar for the non-dimensional quantities.

The formal Rayleigh number derived from the physical parameters used is (Christensen & Yuen 1989; van den Berg, Yuen & Van Keken 1991)

$$Ra = \frac{\alpha\rho_0\Delta Tgh^3}{\eta_0\kappa} = \frac{\alpha\rho_0\Delta Tgh}{A_0t_0^{-1/n}} = \frac{\alpha\rho_0\Delta Tgh^{1+2/n}}{A_0\kappa^{1/n}} \qquad (20)$$

where α is the thermal expansivity, ρ_0 is the density, g is the gravitational acceleration and ΔT is the difference between the uniform top and bottom temperature. The viscosity pre-factor scale A_0, defined as the lower mantle value, is contained in the viscosity scale factor η_0 in the definition of the Rayleigh number. Besides the formal Rayleigh number Ra, we also use an effective Rayleigh number Ra_c defined by replacing the viscosity scale parameter in (20) by the volume averaged viscosity $\langle\eta\rangle$, $Ra_c = \langle\eta\rangle^{-1}Ra$ (Malevsky & Yuen 1992).

We will consider layered media with depth-dependent values of the parameters A and β. In the general case of finite contribution of diffusion creep an extra parameter, the dimensionless, constant linear viscosity $\bar{\eta}^*$, appears in the non-dimensional equations through the depth distribution of the non-dimensional parameter β. Similarly for models with a depth-dependent pre-factor, the depth distribution $A(z)$ defines an extra set of control parameters. The set of control parameters for Rayleigh–Benard convection with infinite Prandtl number for the transition-stress rheology model and layered models of A and β is $\{Ra, n, A(z), \beta(z), \lambda\}$, where λ is the aspect ratio.

2.2 Governing equations and numerical method

We consider Rayleigh–Benard convection without internal heating for infinite Prandtl number in an incompressible

(Broussinesq) fluid. The 2-D domain is a rectangular box of height h and aspect ratio $\lambda = 2.5$. We apply the transition-stress rheological model defined in the previous section. The governing non-dimensional time-dependent equations in primitive variable ($\mathbf{u}, p, T$) form are

$$\partial_j \tau_{ij} - \partial_i p = RaT\delta_{i3} \tag{21}$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T - \nabla^2 T = 0 \tag{22}$$

$$\tau_{ij} = \eta(n, \beta, e)e_{ij} \tag{23}$$

where Ra is the Rayleigh number defined in (18), p the dynamical pressure, $\eta(n, \beta, e)$ the transition stress viscosity, n the power-law index, e the strain rate parameter. t is time, T is the temperature perturbation with respect to the background reference temperature.

The coordinates are given by x and z with the gravity vector aligned with the z axis. Constant temperature boundary conditions are applied at the top and bottom boundaries. Along the lateral boundaries the zero heat-flux symmetry condition is applied. We impose free-slip, non-permeable mechanical conditions on all sides of the domain.

We solve (21) and (22) using a penalty-function finite element method (Cuvelier, Segal & van Steenhoven 1986; Hughes 1987), with a penalty-function parameter value L of 10^6. Within the framework of the penalty function approach the pressure p is set to $L\nabla \cdot \mathbf{u}$, with L the penalty function parameter.

The computer codes used were developed using the finite element tool-kit package SEPRAN (Segal & Praagman 1984). We use nested grids of triangular elements for the velocity and temperature fields. A description of the mesh and element definition used is provided in Appendix A. We have used a Bubnov–Galerkin formulation without using upwinding techniques to solve the energy equation. Instead we used the appearance of oscillatory temperature solutions as an indicator for insufficient grid points.

We have applied the finite element method to discretize both eqs (21) and (22), using quadratic triangular elements for the Stokes eq. (21) and linear triangular elements of smaller size (see above) for the temperature eq. (22). This yields the coupled system of equations (Cuvelier *et al.* 1986; Hughes 1987)

$$\mathbf{S(U)U} = \mathbf{F(T)} \tag{24}$$

$$\frac{d}{dt}\mathbf{T} = -\mathbf{M}^{-1}\mathbf{A}[\mathbf{U(T)}]\mathbf{T} \tag{25}$$

where $\mathbf{U}$ and $\mathbf{T}$ are the vectors of nodal point values of the velocity and temperature fields. Since we use the same nodal points for both fields, $\mathbf{U}$ has twice as many vector elements as $\mathbf{T}$ in the 2-D case. The vector $\mathbf{F(T)}$ is derived from the right-hand side of (21). Through this dependence of the temperature-solution vector $\mathbf{T}$, the Stokes equation is coupled to the energy equation. $\mathbf{S(U)}$ is the stiffness matrix which depends on the viscosity η (Cuvelier *et al.* 1986). In the non-Newtonian model this makes the matrix coefficients depend on the velocity-solution vector, turning (24) into a non-linear equation. $\mathbf{M}$ is a diagonal-lumped mass matrix, $\mathbf{A}$ is the sum of two matrices, derived from the advection and Laplacian terms in (22). (25) is a non-linear equation

through the implicit temperature dependence of the advective velocity. The coupled set of equations therefore contains two different non-linearities: first, the material non-linearity introduced through the transition-stress viscosity and second, the non-linearity introduced by the temperature dependence of the advective velocity.

The time-dependent system (24) and (25) is integrated from the initial condition using a predictor-corrector scheme which combines an implicit Euler predictor step and a Crank–Nicolson corrector step (Hansen & Ebel 1988). The initial condition is defined as a solution of the steady-state problem for the same Rayleigh number, with a perturbation δT added to the temperature field only,

$$\delta T(x, z) = \epsilon \sin(\pi z) \cos(\pi x/\lambda) \tag{26}$$

where $\lambda = 2.5$ is the aspect ratio of the domain and $\epsilon = 0.01$. A detailed description of the algorithm used is presented in Appendix A. We have benchmarked this non-Newtonian code for both the steady-state and time-dependent versions. The results for the time-dependent code can be found in Appendix B.

3 NUMERICAL RESULTS

3.1 Steady-state models

To obtain physical insight in the behaviour of simple convection models for a range of transition stress viscosity parameter values, we will first show results for a number of steady-state convection models, as they can lend physical insight into the problem at a very low computational cost. Stationary solutions were computed by solving eqs (21) and (22) with the time derivative dropped. The equations are solved iteratively, using successive substitution as in (van den Berg *et al.* 1991)—see also Appendix A. Starting solution vectors ($\mathbf{U}, \mathbf{T}$) for the interation were obtained either as the result of a previous case in a running computation by varying a single parameter, or from a Newtonian model computation with comparable Nusselt number and rms (root mean squared) velocity (Vrms), as expected for the non-Newtonian case (Christensen & Yuen 1989).

Results were computed on a mesh of N columns and M rows of rectangular primary cells (see Appendix A). Mesh refinement has been applied at the horizontal boundaries to resolve the boundary layers.

The effect of varying the relative contribution of the linear diffusion creep mechanism is illustrated in Fig. 2, for Rayleigh number 1000 and 2000 and power-law index $n = 3$. Results were computed on a mesh of 48×20 cells (3977 nodal points). Each symbol represents the result of a stationary convection model computation, characterized by the Rayleigh number Ra and the transition stress parameter β. The left-hand frame shows the volume average of the effective viscosity versus the transition stress parameter β. For small values of β the fluid is predominantly non-Newtonian. In the power-law limit ($\beta = 0$) we find values 1.41×10^{-2} ($Ra = 2000$), and 2.04×10^{-2} ($Ra = 1000$). $\eta^* = 1/\beta$ (upper straight line) is the upper limit of the effective viscosity corresponding to a stagnation point ($e = 0$). $\eta_T = 1/(2\beta)$ (lower straight line) corresponds to an equi-partitioning between linear and non-linear creep in the
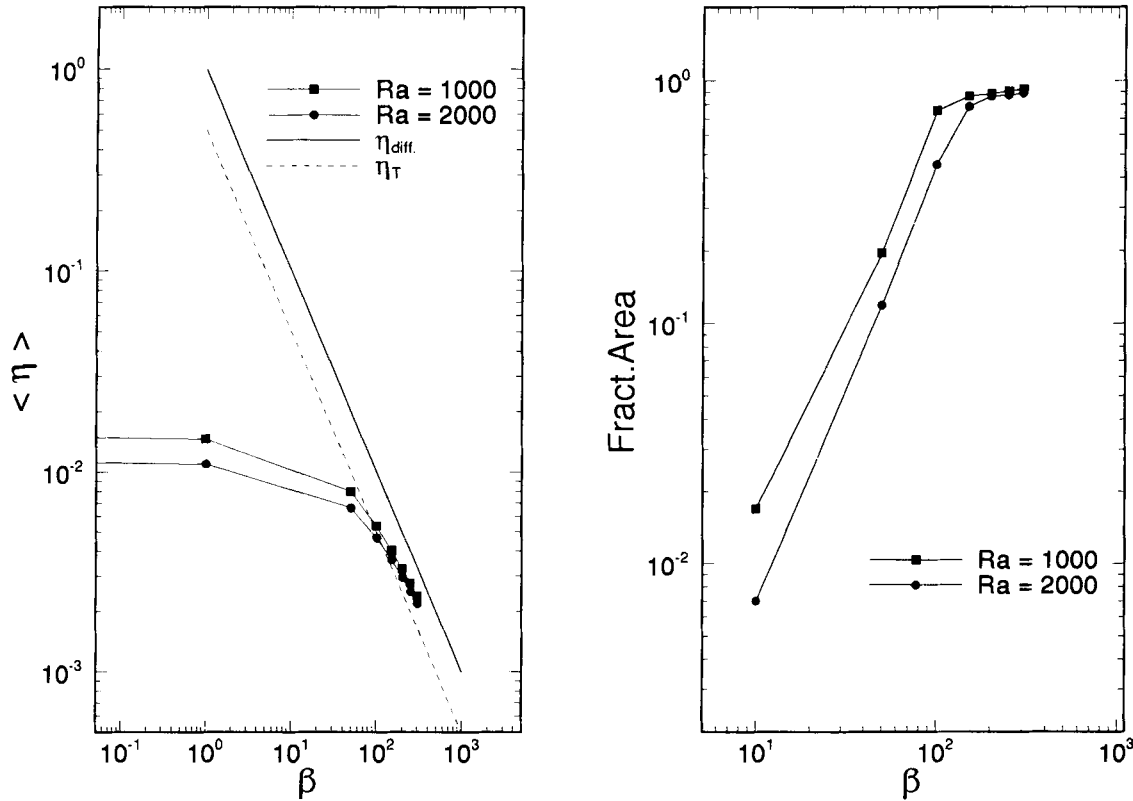
Figure 2. Results of two series of steady-state models for Rayleigh number $Ra = 1000$ and $Ra = 2000$ for varying values of β. Left: dimensionless volume-averaged viscosities. The straight lines denote the diffusive viscosity $\eta^* = 1/\beta$ (upper) and the viscosity value at the transition-stress value $\eta_T = \eta(\tau_T)$, $\eta_T = \eta^*/2$. Right: fractional area of the domain where the fluid is effectively Newtonian: $\eta > \eta_T$.

effective strain rate. For high values of β the linear diffusional creep dominates and the fluid becomes increasingly more Newtonian, as the curves approach the limiting value $\eta \rightarrow \eta^*$.

For the same models, the right-hand frame represents the fraction of the area where the fluid is predomiantly Newtonian, i.e. the effective viscosity is larger than the transition value, $\eta(\mathbf{x}) > \eta_T$. The transition to Newtonian average viscosity for increasing values of β is more gradual for the higher Rayleigh number. The β value of the 50 per cent fractional area in the right-hand frame coincides approximately with the transition points $\eta = \eta_T$ in the left-hand frame. This results suggests that at high Ra non-Newtonian behaviour is retained to a greater degree.

Figure 2(a) illustrates the fact that at a fixed value of the formal Rayleigh number Ra, the average viscosity is reduced by an increase of the Newtonian creep component in the strain rate. This suggests the use of an effective Rayleigh number Ra_c based on the volume average of the effective viscosity as defined in Section 2.1 to characterize the vigour of convection (Parmentier *et al.* 1976; Malevsky & Yuen 1992). The effective Rayleigh number depends on the actual flow field through the effective viscosity and has to be calculated *a posteriori* for a given model. In particular Ra_c will be time dependent for time-dependent flows. Fig. 3(a) shows the relation between the effective and formal Rayleigh number for several values of the rheological control parameter β and for aspect ratio $\lambda = 2.5$. The filled symbols with error bars for $\beta = 0$, 10, 100 correspond to

time-averaged values and rms fluctuation about the mean for the time-dependent models discussed in Section 2.2. Fig. 3(b) shows the Nusselt number versus effective Rayleigh number for the same models of Fig. 3(a). Disconnected symbols correspond to time-dependent models treated in Section 2.2. For higher values of β the curves merge, illustrating the approach to an effective Newtonian flow. For large values of β the average viscosity of the flow approaches $1/\beta$ and we have for the effective Rayleigh number $Ra_c \approx \beta Ra$ illustrated by the parallel curves in Fig. 3(a).

The models discussed up to now have been limited to constant values of the power-law index $n = 3$ and transition stress parameter β. However, recent experimental work (Karato & Li 1992) has established evidence for a vertical layering in the rheological parameter $n(z)$. Van den Berg *et al.* (1991) have investigated in particular the generation of pronounced plate-like behaviour in numerical modelling results by the introduction of layered models with a non-Newtonian power-law upper mantle on top of a Newtonian lower mantle. Within the framework of the composite rheology discussed here this type of layering can be represented by the introduction of a step-function distribution $\beta = (\beta_1, \beta_2)$ with the top-layer value β_1 set to zero and β_2 a finite value representing a finite transition stress in accordance with the rheology found by Karato & Li (1992) for a perovskite analogoue.

Figure 4 displays the results of a layered model with a rheological interface at depth $z = 0.23$. The parameters for
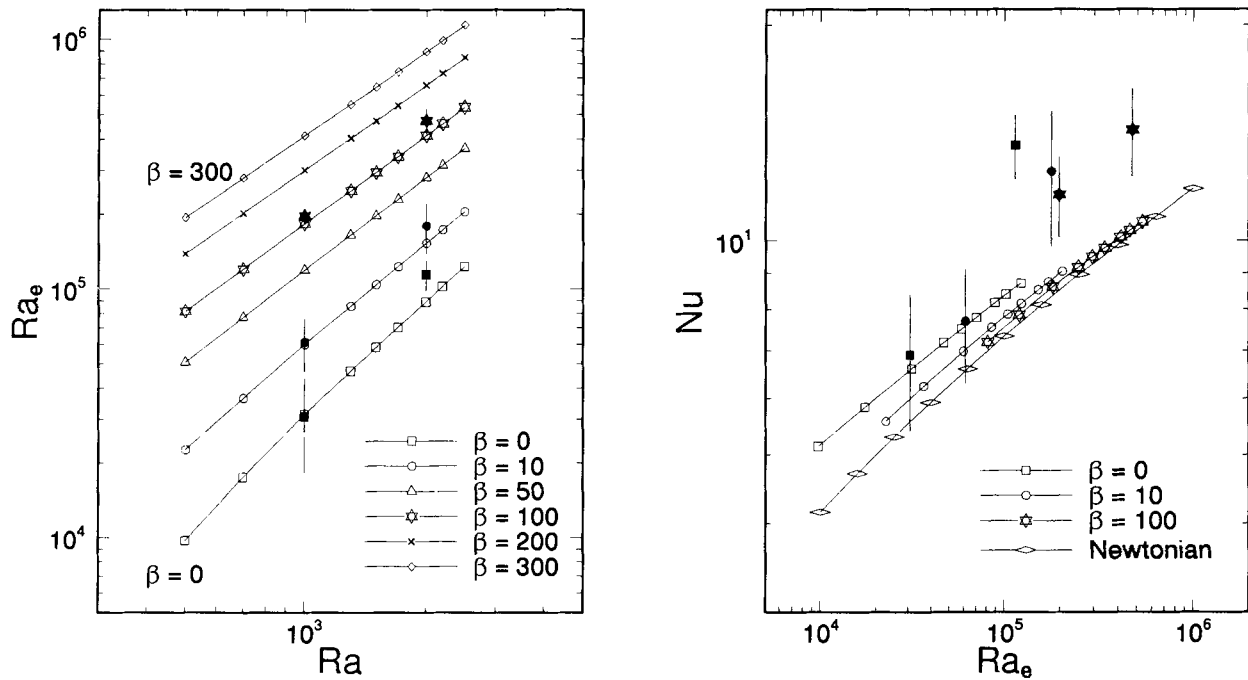
Figure 3. Left (a): effective Rayleigh number based on the volume-average viscosity versus formal Rayleigh number (20) for six values of β. Connected (open) symbols corresponds to steady-state models. Disconnected (filled) symbols with error bars represent time-averaged values of time-dependent models. The error bars denote the rms fluctuation. [Same symbol convention in (b)]. Right (b): Nusselt number versus effective Rayleigh number for several values of β. The lower curve represents the limiting purely Newtonian case.

the two layes are $n = (3, 3)$, $\beta = (0, 200)$ and $A = (1/30, 1)$. Grid refinement near the interface has been applied in order to accommodate the large contrast in the rheological parameters at the interface. Refinement is also necessary at the lateral boundaries to capture the jet-like structure of the solution at the lateral boundaries in the upper mantle. The upper layer in this model is highly mobile and exhibits pronounced plate-like behaviour illustrated most clearly by the concentration of streamlines near top boundary.

The interaction between the two layers is characterized by two narrow jets at both ends of the interface between the layers displayed in the temperature field in the middle panel. A tongue of hot fluid is drawn out from the left-hand jet by the high horizontal velocity in the upper layer. The top panel displays the resulting effective viscosity field. The lower layer is effectively Newtonian, as shown by the largely uniform viscosity values. Low viscosity values are found in the upper layer in regions with high strain rate near the jets at both ends of the layer, and in a narrow shear zone between the mobile upper layer and the high-viscosity lower layer. A further enhancement of plate-like behaviour can be obtained by the introduction of a thin layer with a higher power-law index to mimic a strong lithospheric plate on top of an upper mantle with smaller power-law index and a purely Newtonian lower mantle. Values $n = 5$ and $n = 3$ were used for lithosphere and upper mantle respectively in van der Berg *et al.* (1991). Other types of physical parameterization can also lead to plate-like behaviour at the top boundary layer (Kopitzke 1979; Schmeling & Jacoby 1981; Jacoby & Schmeling 1982; King & Hager 1990; King, Gable & Weinstein 1992; Weinstein 1991; Weinstein & Olson 1992). Using the shallow-water approximation.

Weinstein (1991) and Weinstein & Olson (1992) found plate-like behaviour with a thin power-law layer on top of a fully Newtonian mantle by requiring n to be at least 6.

3.2 Time-dependent simulations

We have investigated the influences of varying the amount of diffusional creep in the creep law on time-dependent mantle convection. For Rayleigh numbers $Ra = 1000$ and $Ra = 2000$ we investigated models for values of β: 0, 10 and 100. We consider the same aspect ratio 2.5 domain as before. The mesh for all cases with $Ra = 1000$ is characterized by: $N = 48$, $M = 20$ and number of nodal points $NP = 3977$ (see Appendix A). For the more vigorous convection cases with $Ra = 2000$ we used $N = 70$, $M = 30$, ($NP = 8601$). The six models presented here are characterized by different values of control parameters Ra and β. In order to put the results into context with Newtonian studies we have computed the time-dependent effective Rayleigh number $Ra_c(t)$ for each of the models. The effective Rayleigh number is calculated as a derived quantity in a post-processing step from the time-dependent viscosity field derived from the velocity results. The velocity field is obtained from the integration of the system of eqs (24) and (25) (see Appendix A.2). The time-averaged effective Rayleigh number $\mu(Ra_c)$ and the rms fluctuation about the mean $\sigma(Ra_c)$ for the different models are listed in Table 2.

The effective Rayleigh numbers in Table 2 are also plotted as disconnected filled symbols in Fig. 3(a) where the error bars denote the rms fluctuation. The time-averaged effective Rayleigh numbers are, in general, close to the
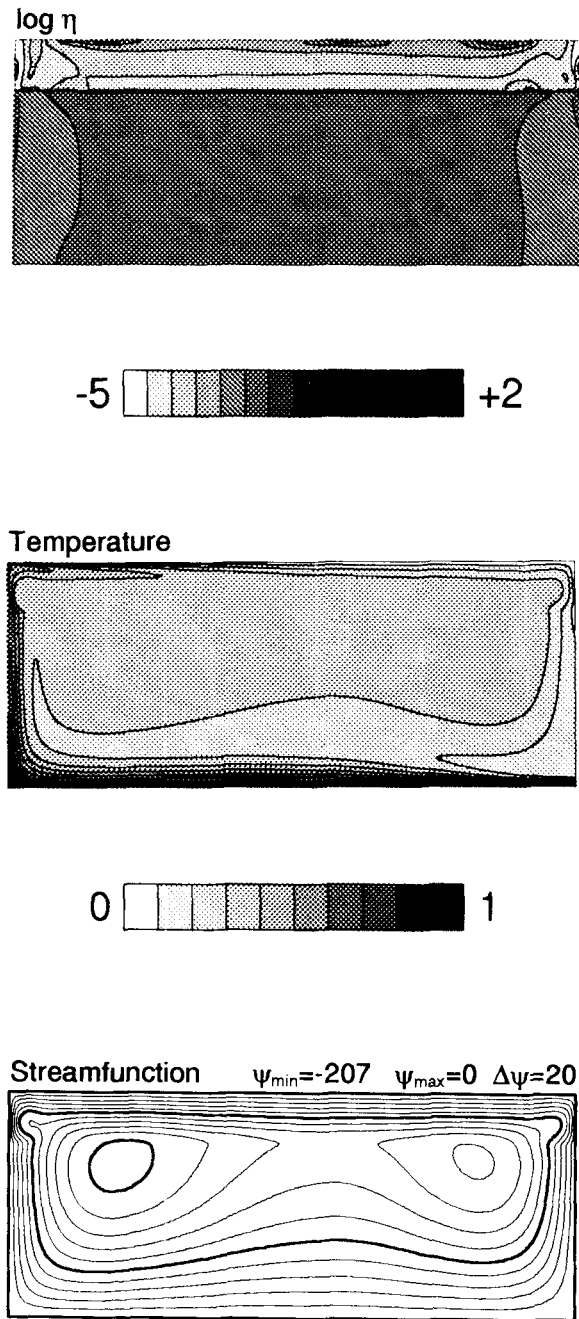
log η

-5 ▮▮▮▮▮▮▮▮▮▮ +2

Temperature

0 ▮▮▮▮▮▮▮▮▮▮ 1

Streamfunction ψ_{min}=-207 ψ_{max}=0 $\Delta\psi$=20

Figure 4. Results for a layered rheology of a non-Newtonian uppermantle on top of an effectively Newtonian lower mantle. The Rayleigh number is 500. The depth-dependent rheological parameters are defined as $\beta = (0, 200)$, and $A = (1/30, 1)$. The Nusselt number and rms velocity are $Nu = 11.84$, $Vrms = 519.4$. The ratio of the layer averaged voscosities $\eta_{lower}/\eta_{upper}$ is 11.4.

Table 2. Nusselt number and effective Rayleigh number.

β	$\mu(Ra_e)$	$\sigma(Ra_e)$	$\mu(Nu)$	$\sigma(Nu)$	$\mu(Ra_e)$	$\sigma(Ra_e)$	$\mu(Nu)$	$\sigma(Nu)$
	Ra = 1000				*Ra* = 2000			
0	3.08×10^4	1.31×10^4	6.9	1.5	1.14×10^5	1.61×10^4	13.6	1.4
10	6.13×10^4	1.48×10^4	7.7	1.4	1.79×10^5	4.18×10^4	12.5	2.7
100	1.97×10^5	1.25×10^4	11.6	1.5	4.74×10^5	5.46×10^4	14.3	2.

values for the corresponding steady-state model characterized by the same control parameters. Clearly the relative rms fluctuation decreases for increasing values of β as the viscosity approaches a uniform Newtonian limit. Time-averaged Nusselt numbers and rms fluctuation from Table 2 are displayed in Fig. 3(b). The average values are higher than the corresponding steady-state cases. This can be explained by the fact that the steady-state data represent single-cell, aspect ratio $\lambda = 2.5$ flow fields, whereas in the time dependent cases the single-cell pattern is disturbed and often close to a two-cell flow field with a correspondingly higher Nusselt number (Hansen, Yuen & Malevsky 1992).

In Fig. 5 snapshots are given for the temperature, stream function and viscosity fields for three instants of time for $Ra = 1000$ and $\beta = 0$. Time increases in the upward direction. The left-hand (temperature) column shows the evolution of a rising hot plume, colliding into the left-hand boundary. It also shows the development of a cold downwelling descending from the top and cutting through the remains of the hot plume (top panel). In the middle column the stream function for the same time instants is displayed. The bottom panel shows the large velocity amplitude generated near the left-hand boundary when the hot plume impinges on the vertical boundary. In the next snapshot the velocity amplitude has already been reduced considerably. This phenomenon shows that strong plume–plume collisions at the left-hand boundary dominate the dynamics of the solution. The rms velocity shows a sharp, large amplitude peak during the collision (see Fig. 10). The increased upward flow of hot material at the time of collision also causes a spike in the Nusselt number time series. The spikes in the Nusselt number time series are, however, less sharp than their rms velocity counterparts. This is due to the fact that hot patches of fluid are injected into the top layer in short bursts from the plume collision, resulting in a sharp increase in the local heat flow and the Nusselt number (van Keken, Yuen & van den Berg 1992). After the burst, the influx of hot material into the top boundary layer diminishes. The hot remnant of the injection is entrained in the large-scale flow while the peak temperature decays with time and its contribution to the surface heat flow is integrated into the Nusselt number for a period of time that is of the order of one overturn period of the large-scale flow. The right-hand column of Fig. 5 shows the viscosity. For this pure power-law case the viscosity variations are strong and concentrated near the stagnation area ($e = 0$) of the flow. Viscosity minima coincide with the area of high shear rate in the corners of the domain and along the boundary layers of the rapidly evolving plumes.

Figure 6 displays the results for a model with an increased contribution of the diffusional creep strain rate for the same Rayleigh number, $Ra = 1000$, $\beta = 100$. The vigour of convection has increased compared with the previous case, due to the decrease in average viscosity. The rate at which boundary-layer instabilites are generated has increased and the hot plumes rise quicker with their heads effectively detaching before reaching the surface. This change in the dynamics also results in a smaller number of plume–plume collisions and the attendant reduced level of fluctuation of the time series of the rms velocity (see Fig. 12). The right-hand column of logarithmic viscosity snapshots illustrates a general reduction in the viscosity. The

Ra = 1000 β = 0

0 [gradient bar] 1

Temperature
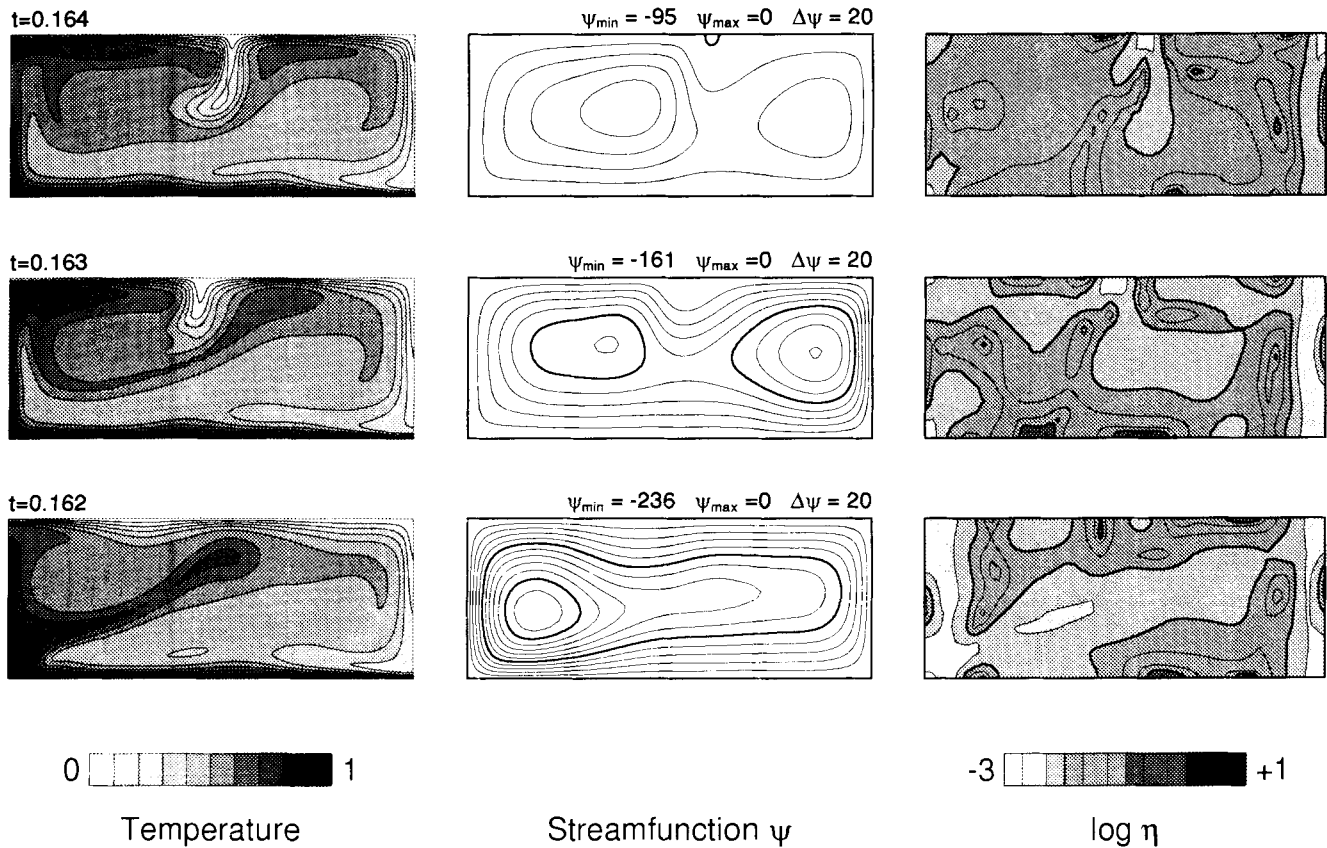
Streamfunction ψ

-3 [gradient bar] +1

log η

Figure 5. Snapshots of time-dependent fields for $Ra = 1000$, $\beta = 0$ for 3 instants of time. Left column: temperature, displayed between 0.0 and 1.0. Gray-scale interval 0.1. Middle column: stream function, contour interval $\Delta\psi = 20$, heavy contours at -200 and -100. Right column: log viscosity. Grey-scale contour boundaries are at 2.6, -2.3, -2.0, -1.6, -1.3, -1.0, ..., 0.3, 0.6. Contours are drawn at the same boundaries. We note 0.001 in t corresponds to 100 Ma for the parameters given in Table 1.

logarithmic viscosity values in the dark pattern area lie within the range $(-2.6, -2.0)$ where the fluid is predominantly Newtonian. From the viscosity time series plot (Fig. 12, top panel) we see that the volume average of the viscosity is near the transition value $\eta_T = 1/200$.

In Figs 7 to 9 results are shown for higher Rayleigh number $Ra = 2000$ for the same values of β. Fig. 7 illustrates the pure power-law case $Ra = 2000$, $\beta = 0$. The temperature field and stream-function plot reveal a two-cell counter-rotating flow pattern, with a cold downwelling in the middle. The logarithmic power-law viscosity reaches values up to 0.0 in the stagnant area near the centre of the middle panel. The time dependence of the maximum viscosity value is illustrated in Fig. 13. Fig. 8 shows results for $Ra = 2000$, $\beta = 10$. The snapshots show solution fields that have not yet lost the symmetry from the initial condition. The temperature plots illustrate the evolution of a colliding, symmetric couple of a hot plume and cold downwelling. The stream function and viscosity plot (middle panels) clearly illustrate a narrow high-shear zone were the logarithmic viscosity drops below -3.0. Maximum viscosity values are

reached in small areas with predominantly Newtonian viscosity. The fractional area of these regions is only a few per cent—a characteristic for this low value of β. The time series of the volume-averaged viscosity illustrates the same phenomenon (Fig. 14, top panel). In Fig. 9 we show results for $Ra = 2000$, $\beta = 100$. The maximum viscosity is reduced below -2.0. The dark pattern areas of the viscosity plots represent viscosities above the transition value $\eta_T = 1/200$. The fractional area of predominantly Newtonian fluid is smaller—for the time instants shown—than in the case of $Ra = 1000$, $\beta = 100$. The viscosity shows a more granular structure than in the previous cases. A similar trend towards small-scale granular viscosity distribution was found by Malevsky & Yuen (1992) for higher Rayleigh number for power-law fluids.

From the snapshots of viscosity and temperature fields shown in Figs 5 to 9 we conclude a clear absence of direct correlation between variations in temperature and viscosity. Viscosity maxima are found in hot rising plume heads and viscosity minima exist in cold downwellings. Although we did not use a temperature-dependent rheology in our

Ra = 1000   β = 100

t=0.163
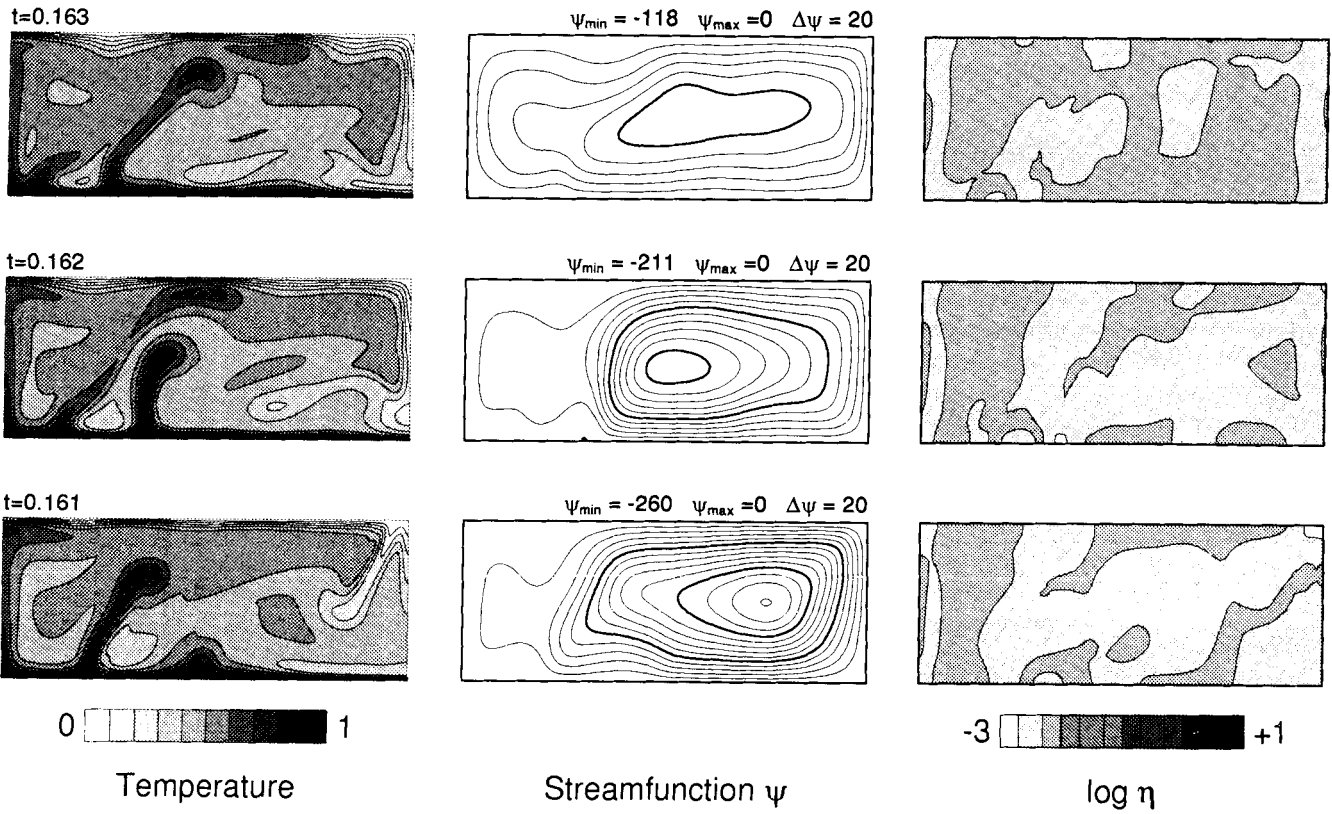
$\psi_{min} = -118$ $\psi_{max} = 0$ $\Delta\psi = 20$

t=0.162

$\psi_{min} = -211$ $\psi_{max} = 0$ $\Delta\psi = 20$

t=0.161

$\psi_{min} = -260$ $\psi_{max} = 0$ $\Delta\psi = 20$

0 [   ] 1
Temperature

Streamfunction ψ

-3 [   ] +1
log η

Figure 6. Same as in Fig. 5, for $Ra = 1000$, $\beta = 100$.

Ra = 2000   β = 0

t=0.105

$\psi_{min} = -53$ $\psi_{max} = 135$ $\Delta\psi = 20$

t=0.095

$\psi_{min} = -143$ $\psi_{max} = 28$ $\Delta\psi = 20$

t=0.085

$\psi_{min} = -124$ $\psi_{max} = 140$ $\Delta\psi = 20$

0 [   ] 1
Temperature

Streamfunction ψ

-3 [   ] +1
log η
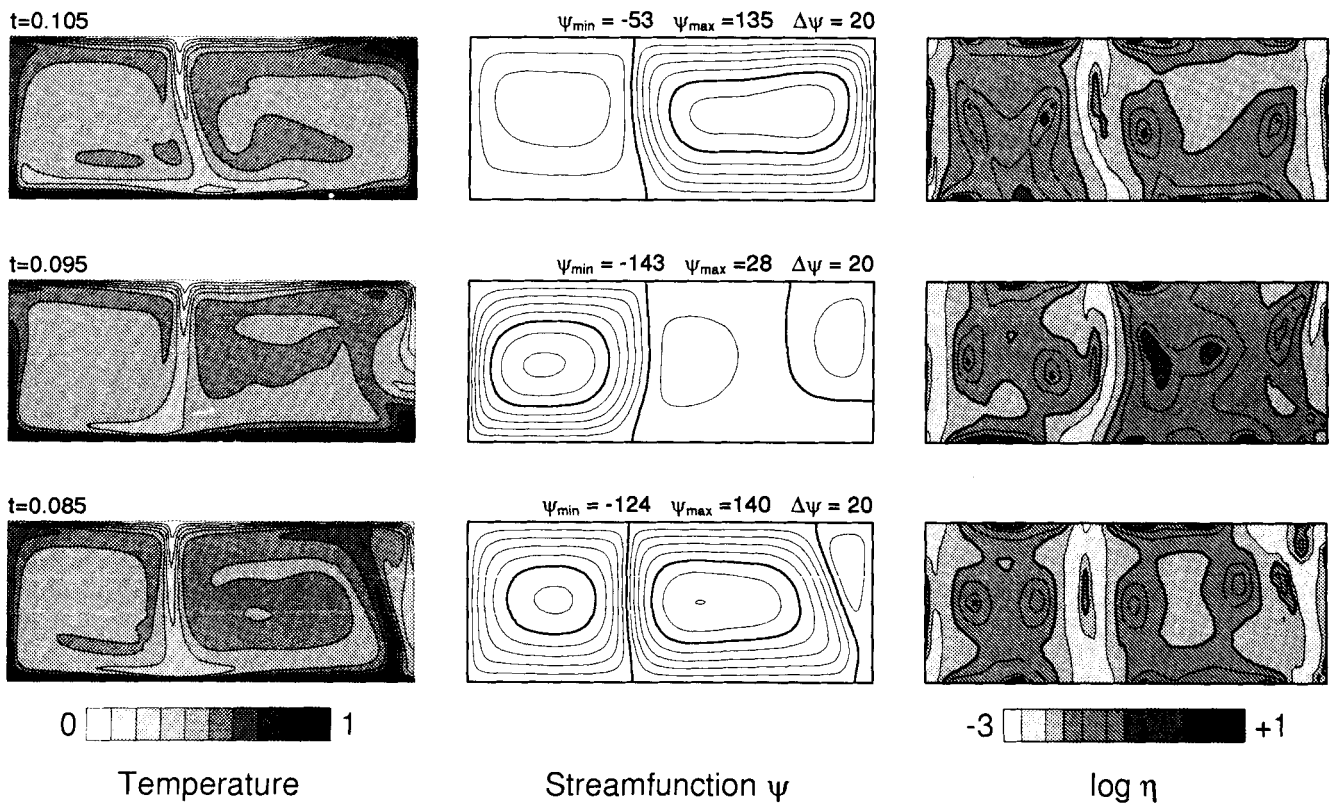
Figure 7. Same as in Fig. 5, for $Ra = 200$, $\beta = 0$.

Ra = 2000   β = 10

t=0.0235

ψ_min = 0   ψ_max = -117   Δψ = 20

t=0.0225

ψ_min = -188   ψ_max = 82   Δψ = 20

t=0.0215

ψ_min = -84   ψ_max = 12   Δψ = 20

0 ▭ 1
Temperature

Streamfunction ψ

-3 ▭ +1
log η
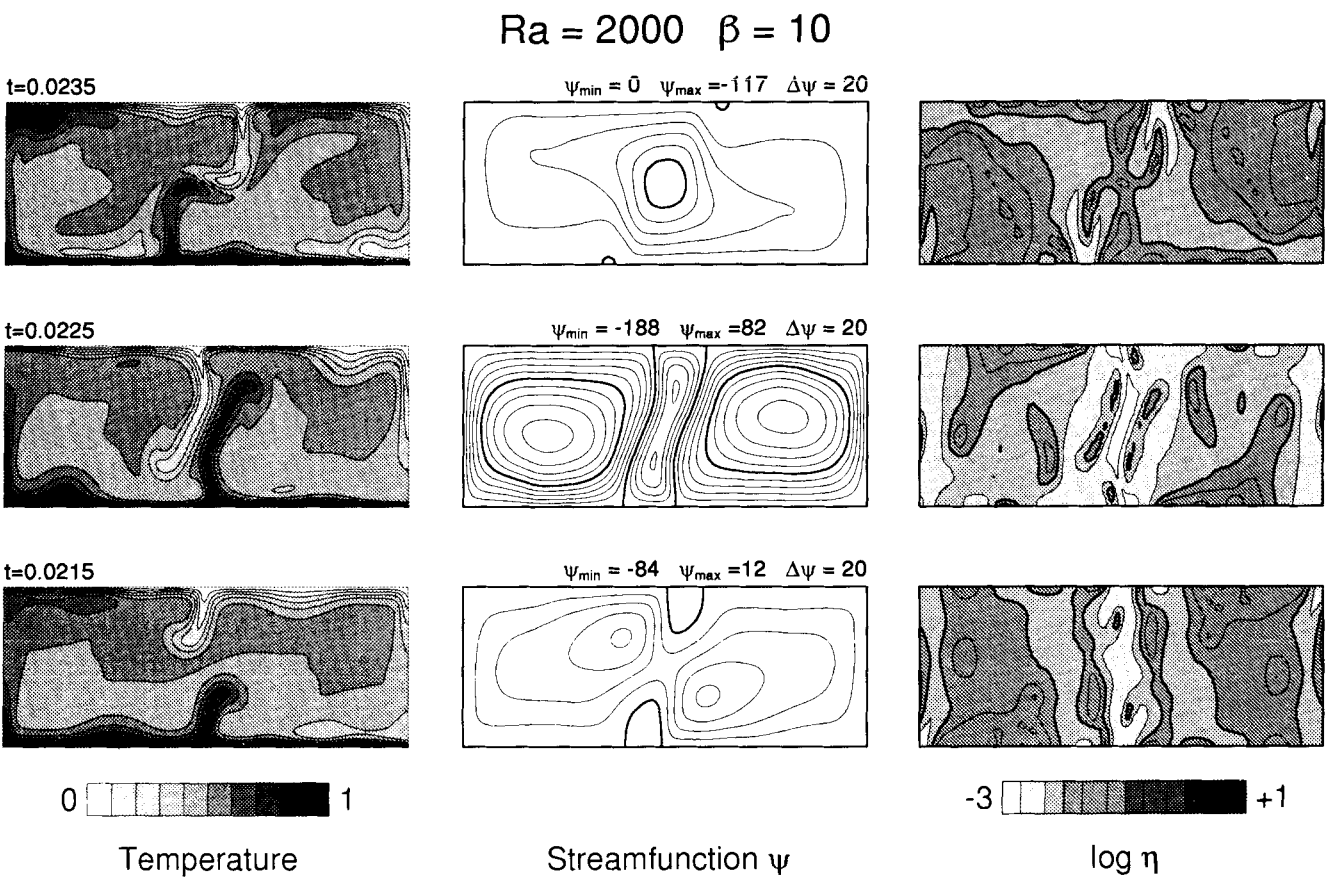
Figure 8. Same as in previous figure, for *Ra* = 2000, *β* = 10.

Ra = 2000   β = 100

t=0.021

ψ_min = -301   ψ_max = 0   Δψ = 20

t=0.0205

ψ_min = -245   ψ_max = 0   Δψ = 20

t=0.02

ψ_min = -224   ψ_max = 11   Δψ = 20

0 ▭ 1
Temperature

Streamfunction ψ

-3 ▭ +1
log η
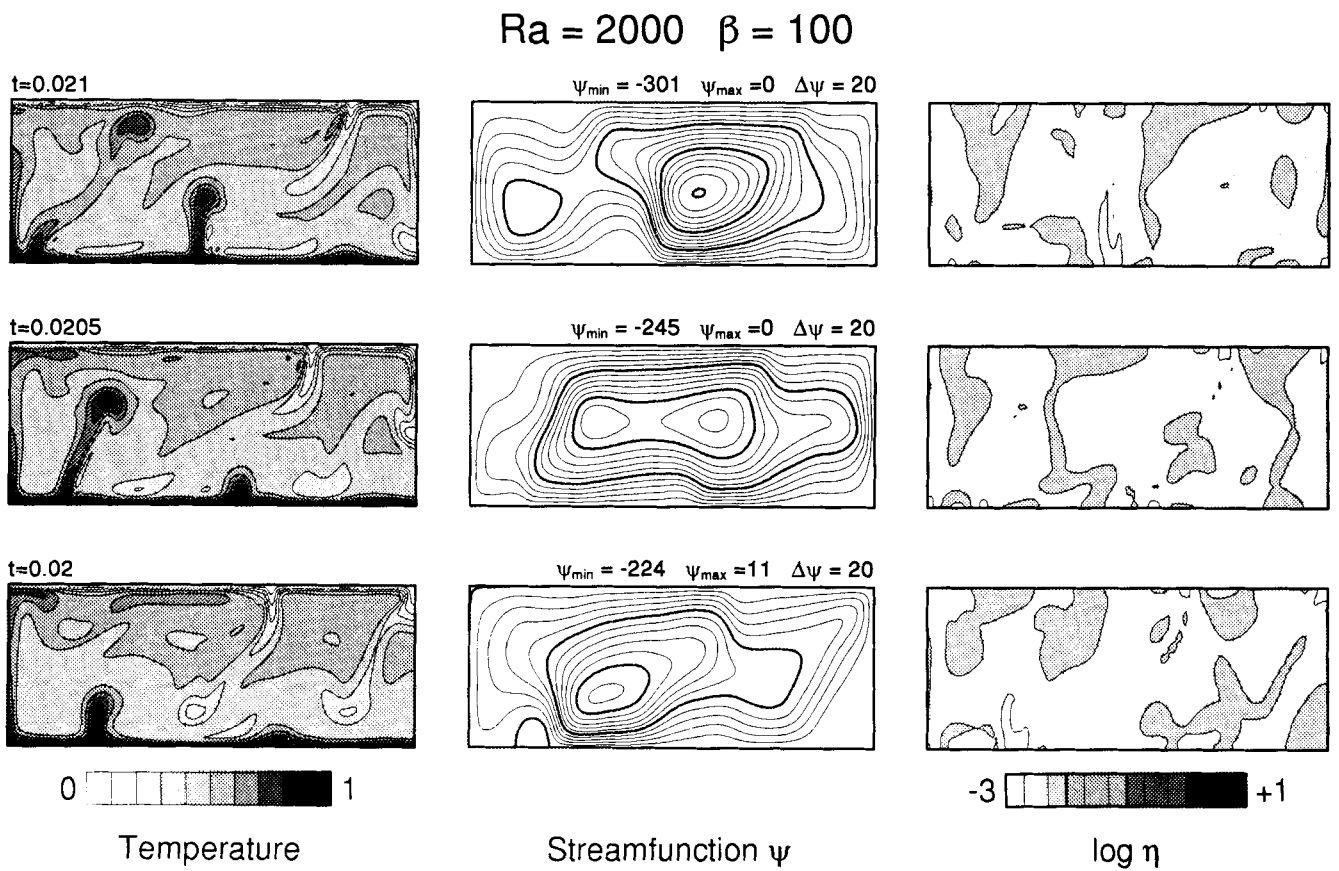
Figure 9. Same as in previous figure, for *Ra* = 2000, *β* = 100.

Ra = 1000 β = 0

Ra = 1000 β = 10

Figure 11. Time series as in Fig. 10, for $Ra = 1000$, $\beta = 10$.

Figure 10. Time series of integral quantities for $Ra = 1000$, $\beta = 0$. Bottom: Nusselt number. Middle: rms velocity. Top: maximum viscosity (top trace), volume-averaged viscosity (middle trace), minimum viscosity (bottom trace).

models, the results clearly indicate that direct correlation caused by such temperature dependence will be strongly reduced by the strain-rate dependence. This effect will contaminate any interpretation of seismic anomalies deduced from tomography in terms of viscosity fluctuations.

In Figs 10 to 15 we show the time series of integrated quantities for the time-dependent models corresponding to Figs 5 to 9. Here we used different lengths for the displayed time windows, resulting in different time scales on the plot. Fig. 10 represents results for $Ra = 1000$, $\beta = 0$ (purely power law). The bottom panel shows the Nusselt number versus integration time. The middle panel displays the rms velocity Vrms. Both time series show the same characteristics, a build-up in amplitude between $t = 0.0$ and $t = 0.8$ which is followed by a series of co-eval spikes in amplitude at irregular intervals. The spikes in the rms velocity (Vrms) time series imply that large amplitude, rapid fluctuations in the surface viscosity occur. This behaviour would facilitate frequent and fast plate re-organizations in non-Newtonian models of mantle dynamics. The spikes in the Nusselt number are less sharp than their counterparts in the Vrms for reasons mentioned in the discussion of Fig. 5. The top panel displays time series of maximum and minimum value of the logarithmic viscosity (top and bottom trace respectively) and of its volume average (middle trace). The
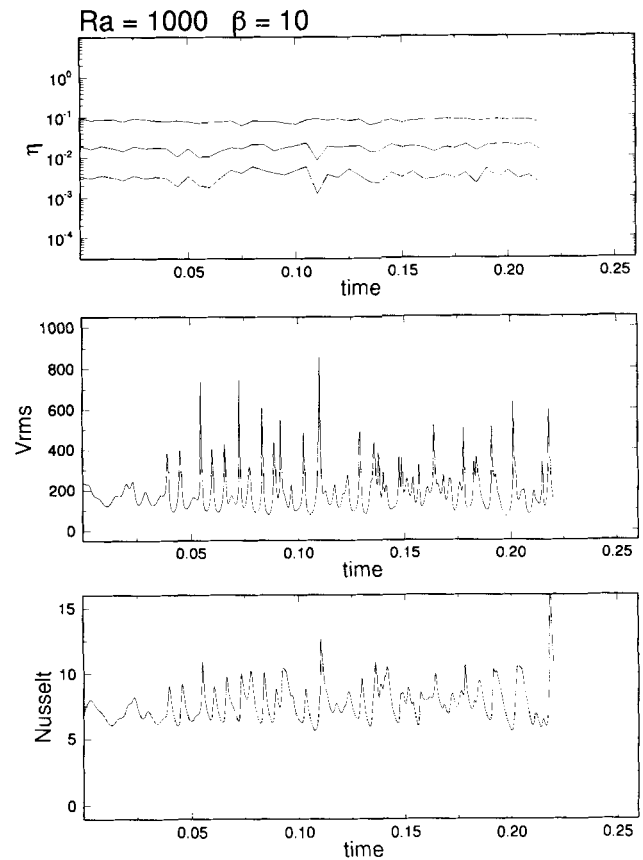
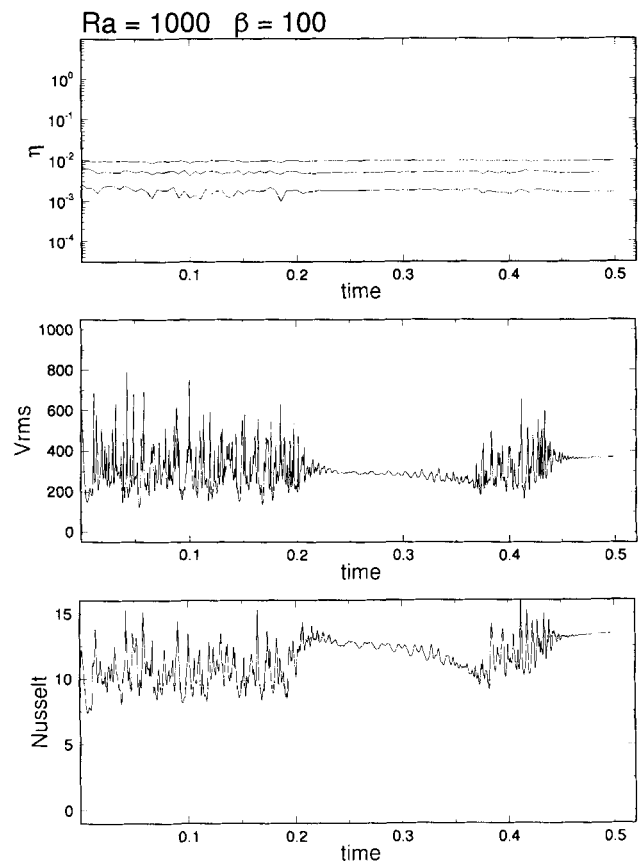## Ra = 1000 β = 100

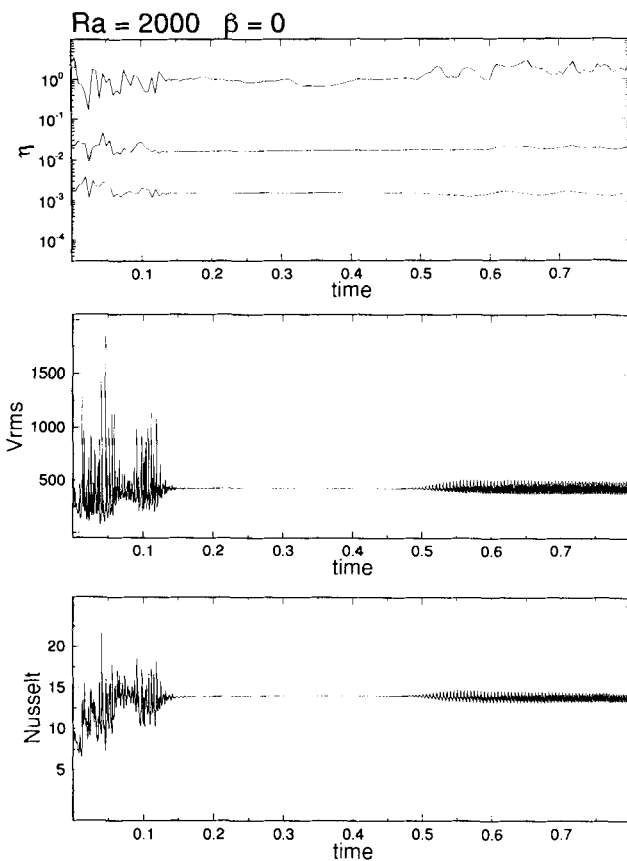**Figure 12.** Time series as in Fig. 10, for $Ra = 1000$, $\beta = 100$.

Figure 13. Time series as in Fig. 10, for $Ra = 2000$, $\beta = 0$.
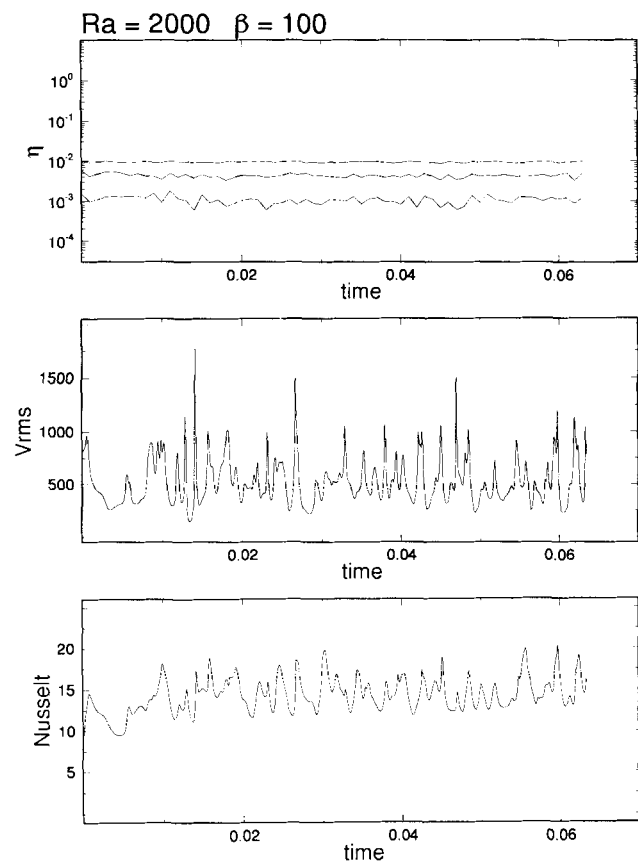

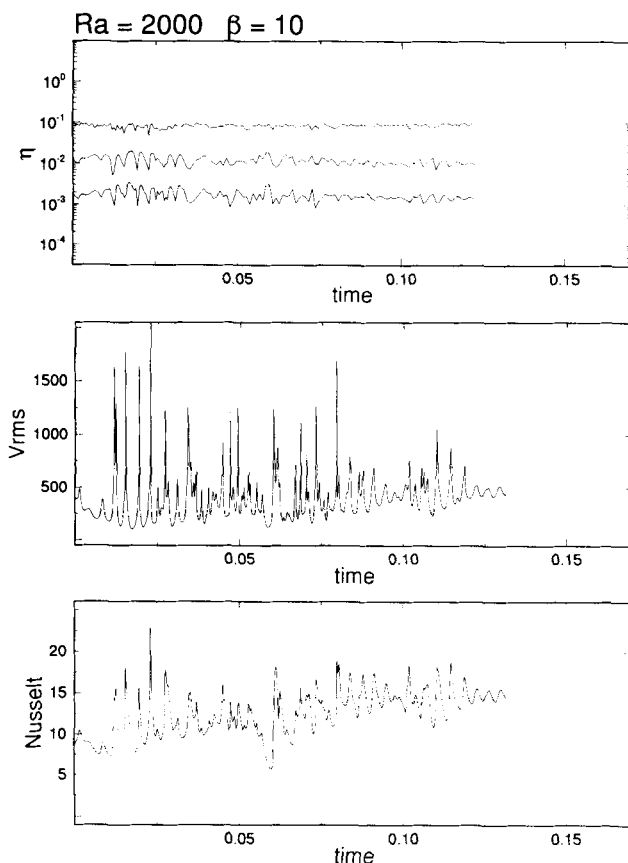**Figure 15.** Time series in Fig. 10, for $Ra = 2000$, $\beta = 100$.

Figure 14. Time series as in Fig. 10, for $Ra = 2000$, $\beta = 10$.

time series of Nusselt number and Vrms were sampled at every integration time step. The viscosity time series have been calculated as a post-processing step from the field solutions, which were output from the program with a constant time interval. As a result, the viscosities are sampled at a lower rate. In this power-law case the viscosity becomes large in stagnant parts, the maxima display fluctuations of almost two orders of magnitude. Fluctuations of the volume average and minimum values are small than for the maxima. Spikes in the Vrms line-up with the minimum values in the viscosity traces, illustrating the shear-thinning behaviour of the fluid with power-law index $n = 3$. Fig. 11 shows the same quantities as in the previous figure for $Ra = 1000$, $\beta = 10$. The introduction of a finite diffusional creep contribution has truncated the viscosity to values below $\eta^* = 1/10$ increasing the mobility of the fluid. As a consequence, the average value of the Nusselt number and Vrms are slightly increased. The viscosity time series clearly show the truncation effect in the maximum values to $1/10$. The level of fluctuation in the maxima has decreased drastically compared with the previous case. The volume average of the viscosity is clearly below 0.05, corresponding to the transition viscosity η_T. The fluid behaves predominantly in a non-Newtonian manner.

Figure 12 shows a similar result for parameter values $Ra = 1000$, $\beta = 100$. The amplitude level of the Nusselt number has increased from approximately 8 to 11. Similarly Vrms increases from approximately 200 to 300. The flow field is chaotic in the first interval up to about $t = 0.2$, characterized by a large-scale single cell consistent

convection pattern with intermittent boundary layer instabilities causing spikes in the time series (see Fig. 6). The quiescent interval corresponds to a nearly stationary two-cell flow pattern, gradually decaying into a single-cell pattern again shortly before $t = 0.4$. The viscosity traces show a further decrease in the maximum values, with $\eta^* = 1/100$. The minimum has been reduced to a lesser extent, resulting in a drop of the spatial variation of the viscosity to within one order of magnitude. The volume average is close to the value of the transition viscosity $\eta_T = 1/200$. Fig. 13 shows the same time series as before for $Ra = 2000$, $\beta = 0$. The flow field corresponds to single-cell convection with intermittent boundary layer instabilities up to approximately $t = 0.15$, (see Fig. 7). After this time a quiescent period with a two-cell pattern prevails up to $t = 0.55$. Such a behaviour is typical of intermittency in turbulent flows (e.g. Grossmann & Lohse 1992). The occurrence of intermittency in which the timescale of intermittency is longer than the overturn timescale is a new type of time-dependent behaviour in mantle convection. Later on the flow becomes periodic; the downwelling limb dividing the two counter-rotating cells oscillates and a regular pattern of Nusselt number and Vrms develops. The Nusselt number and Vrms have increased to approximately 13 and 400 respectively. Fig. 14 shows similar results for $Ra = 2000$, $\beta = 10$. The time window is shorter than the previous case. The overall characteristics are similar to the first chaotic time interval of the previous case (Fig. 13). This is also illustrated by the snapshots of the solution at $t = 0.215$ (Fig. 8). Fig. 15 displays the same time series in a short time window for the case $Ra = 2000$, $\beta = 100$. The level of Nusselt number and Vrms have increased further to about 15 and 500 respectively. A comparison with Fig. 12 shows that the increase of Ra has delayed the transition to Newtonian behaviour.

From boundary-layer scaling arguments applied to the dependence of the Nusselt number on the Rayleigh number, the following scaling relation for the fluctuations in the average viscosity $\langle \eta \rangle$ can be derived:

$$d \log (Nu) \sim pd \log (\langle \eta \rangle), \tag{27}$$

where p is the exponent in the scaling relation for the Nusselt number. A rough estimate of the exponent is determined from our stationary convection results to be $p = 0.25$. The results show that the level of fluctuation in the volume-averaged viscosity is of the same magnitude as the fluctuations in the Nusselt number for the pure power-law case. This is a result which is consistent with the above scaling relation. However, we found that the scaling relationship between Vrms and $\langle \eta \rangle$ does not hold. This can be explained by the larger dominance of the large amplitude spikes in the Vrms time series, due to the presence of plume–plume collisions. For increasing values of β, the temporal variations of the viscosity decrease, as the fluid becomes more Newtonian like.

4 CONCLUSIONS

In this work we have presented a numerical method for modelling thermal convection in both stationary and time-dependent models with a hybrid rheology including Newtonian and power-law non-Newtonian as end-member

cases. The computer codes used were developed using the general-purpose finite element tool-kit package SEPRAN, (Segal & Praagman 1984) which makes adaptations for various rheologies relatively easy through implemenation of new finite elements. The steady-state results show a gradual transition from, non-Newtonian (power-law) to linear Newtonian behaviour as a function of increasing β. The transition is effectively delayed for higher Rayleigh number values.

Our results display strong spatial variations—especially for short wavelengths—in the non-Newtonian viscosity. We found that there is no direct correlation between the viscosity and temperature fields. Strong plates of highly viscous material coexist in both hot and cold boundary layers. Weak-high shear zones can coexist in hot upwelling and cold downwelling limbs of the convection cells. The presence of non-Newtonian rheology will contaminate purely temperature-dependent viscosity estimates (Colin & Fleitout 1992) calculated from a direct relationship between thermal anomalies and long (Tanimoto 1990; Zhang & Tanimoto 1992) and short (Zhang & Tanimoto 1992; van der Hilst *et al.* 1991) wavelength seismic anomalies.

These time-dependent results generally show a chaotic flow pattern of large-scale single-cell convection with intermittent boundary-layer instabilites, resulting in strong temporal fluctuations of global quantities such as the Nusselt number and Vrms. This result is in agreement with the work of Christensen & Yuen (1989) and of Malevsky & Yuen (1992). The fluctuations in the Nusselt number are smaller than those in Vrms. Variations in the local surface heat flow, however, can exceed the surface average considerably (van Keken *et al.* 1992) indicating intermittency in the flow behaviour. The intermittency time scale found in the model is longer than the overturn timescale. This represents a new type of time-dependent behaviour in mantle convection. The high-amplitude spikes in the time series of Vrms illustrate the tendency towards greater spatial temporal-velocity fluctuations for a non-Newtonian upper mantle, allowing for rapid changes in plate velocities and facilitating plate re-organizations. This is an important feature required of dynamical models describing the earth's mantle. With respect to the correlation of the viscosity and temperature fields, similar remarks can be made for our time-dependent solutions as for the steady-state results. The heterogeneity of the viscosity field increases with the complexity of the flow field. Low-viscosity values are not directly correlated to high-temperature regions: viscosity minima are found in cold downwellings.

For pure power-law rheology the temporal fluctuations in the volume-averaged viscosity are comparable in magnitude to the fluctuations in the Nusselt number and smaller than the fluctuations in rms velocity. The effect of increasing β results in smaller fluctuations in the volume-averaged viscosities, as the fluid becomes more Newtonian.

The transition from power-law rheology to Newtonian rheology is delayed by an increasing Rayleigh number. Here we conjecture that in the early Archaean, when the Rayleigh number and the temperature were higher than at present, mantle rheology might have been predominantly non-Newtonian because of the compensating effect due to a larger activation energy in dislocation creep (Ranalli 1991). The dominance of non-Newtonian flow will ensure a more

chaotic time-dependent convection regime than otherwise expected using just a Newtonian temperature-dependent rheology. This again may explain the rapid cooling of the mantle in the Archaean, resulting in different styles of differentiation in the upper mantle at the end of the Archaean and the beginning of plate tectonics (Vlaar 1985, 1986; Vlaar & van den Berg 1991).

ACKNOWLEDGMENTS

REFERENCES

Bird, R. B., Stewart, W. E. & Lightfoot, E. N., 1960. *Transport Phenomena*, Wiley, New York.

Bird, R. B., Armstrong, R. C. & Hassager, A. O., 1987. *Dynamics of Polymeric Liquids*, Wiley, New York.

Christensen, U. R., 1984. Convection with pressure and temperature dependent non-Newtonian rheology, *Geophys. J. R. astr. Soc.*, **77**, 343–384.

Christensen, U. R. & Yuen, D. A., 1989, Time-dependent convection with non-Newtonian viscosity, *J. geophys. Res.*, **94**, 814–820.

Colin, P. & Fleitout, L., 1992. Lateral viscosity contrasts and the geoid and topography anomalies associated with mantle mass heterogeneities, *Ann. Geophy.*, (Supplement 1) **10**, C54.

Crochet, M. J., Davies, A. R. & Walters, A. K., 1984. *Numerical Simulation of Non-Newtonian Flow*, Elsevier, Amsterdam, The Netherlands.

Cserepes, L., 1982. Numerical studies of non-Newtonian mantle convection, *Phys. Earth planet. Inter.*, **30**, 49–61.

Cuvelier, C., Segal, A. & van Steenhoven, A. A., 1986. *Finite Element Methods and Navier–Stokes Equations*, Reidel, Dordrecht, The Netherlands.

Damsteegt, J., Segal A. & van der Zanden, J., 1986. On the convergence of numerical computations of the flow of a power-law and a Carreau liquid, *Technical Report WTHD*, **184**, Laboroaty for Aero- and Hydrodynamics Technical University, Delft.

Gasperini, P., Yuen, D. A. & Sabadini, R., 1992. Postglacial rebound with upper-mantle non-Newtonian and lower mantle Newtonian rheology, *Geophys. Res. Lett.*, **19**, 1711–1714.

Grossmann, S. & Lohse, D., 1992. Intermittency in the Navier–Stokes dynamics, *Z. Phys. B. Condens. Matt.*, **89**, 11–19.

Hansen, U. & Ebel, A., 1988. Time-dependent thermal convection—a possible explanation for multiscale flow in the earth's mantle, *Geophys. J. Int.*, **94**, 181–192.

Hansen, U., Yuen, D. A. & Kroening, S. E., 1990. Transition to hard turbulence in thermal convection at infinite Prandtl number, *Phys. Fluids*, **A2(12)**, 2157–2163.

Hansen, U., Yuen, D. A. & Malevsky, A. V., 1992. A comparison of steady state and strongly chaotic thermal convection at high Rayleigh number, *Phys. Rev. A*, **46**, 4742–4754.

Hughes, T. J. R., 1987. The Finite Element Method—Linear Static and Dynamic Finite Element Analysis, Prentice Hall, Englewood Cliffs, New Jersey.

Jacoby, W. R. & Schmeling, H., 1982. On the effects of the lithosphere on mantle convection and evolution, *Phys. Earth planet. Inter.*, **29**, 305–319.

Karato, S. & Li, P., 1992. Diffusion creep in perovskite and the linear rheology of the earth's lower mantle, *Science*, **255**, 1238–1240.

King, S. D. & Hager, B. H., 1990. The relationship between plate velocity and trench viscosity in Newtonian and power-law subduction calculations, *Geophys. Res. Lett.*, **17**, 2409–2412.

King, S. D., Gable, C. W. & Weinstein, S. A., 1992. Models of convection-driven tectonic plates: a comparison of methods and results, *Geophys. J. Int.*, **109**, 481–487.

Kopitzke, U., 1979. Finite element convection models; comparison of shallow and deep mantle convection and temperatures in the mantle, *J. Geophys.*, **46**, 97–121.

Malevsky, A. V. & Yuen, D. A., 1992. Strongly chaotic non-Newtonian mantle convection, *Geophys. Atrophys. Fluid Dyn.*, **65**, 149–171.

McKenzie, D. P., Roberts, J. M. & Weiss, N. O., 1974. Convection in the Earth's mantle: towards a numerical simulation, *J. Fluid Mech.*, **62**, 465–538.

Meisner, R. O. & Vetter, U. R., 1976. Isostatic and dynamic processes and their relation to viscosity, *Tectonophysics*, **35**, 137–148.

Parmentier, E. M., 1978. A study of thermal convection in non-Newtonian fluids. *J. Fluid. Mech.*, **84**, 1–11.

Parmentier, E. M., Turcotte, D. L. & Torrance, K. E., 1976. Studies of finite amplitude non-Newtonian thermal convection with application to convection in the earth's mantle, *J. geophys. Res.*, **81**, 1839–1846.

Ranalli, G., 1987. *Rheology of the Earth*, Allen and Unwin, Boston.

Ranalli, G., 1991. The microphysical approach to mantle rheology, in *Glacial Isostasy, Sea-Level and Mantle Rheology*, pp. 343–378, eds Sabadini, R., Lambeck, E. K. & Boschi, E. V., Kluwer Academic Publishers, Dordrecht.

Schmeling, H. & Jacoby, W. R., 1981. On modelling the lithosphere in mantle convection with non-linear rheology, *J. Geophys.*, **50**, 89–100.

Schowalter, W. R., 1978. *Mechanics of Non-Newtonian Fluids*, Pergamon Inc., Oxford, England.

Segal, A. & Praagman, N., 1984. *SEPRAN-user manual*, Sepra, Leidschendam, The Netherlands.

Tanimoto, T., 1990. Long wavelength S-wave velocity structure throughout the mantle, *Geophys. J.*, **100**, 327–336.

Turcotte, D. L. & Oxburgh, E. L., 1972. Mantle convection and the new global tectonics, *Ann. Rev. Fluid Mech.*, **4**, 33–68.

van den Berg, A. P., Yuen, D. A. & van Keken, P. E., 1991. Effects of depth variations in creep laws on the formation of plates in mantle dynamics, *Geophys. Res. Lett.*, **18**, 2197–2200.

van der Hilst, R., Engdahl, R., Spakman, W. & Nolet, G., 1991. Tomographic imagining of subducted lithosphere below northwest Pacific island arcs, *Nature*, **353**, 37–43.

van Keken, P. E., 1993. Numerical models of thermo-chemically driven fluid flow with non-linear rheology applied to the Earth's lithosphere and mantle, *PhD thesis*, Utrecht University, Utrecht.

van Keken, P. E., Yuen, D. A. & van den Berg, A. P., 1992. Pulsating diapiric flows; consequences of vertical variations of mantle creep laws, *Earth planet. Sci. Lett.*, **112**, 179–194.

Vlaar, N. J., 1985. Precambrian geodynamical constraints, in *The Deep Proterozoic Crust in the North Atlantic Provinces*, pp. 3–20, eds Tobi, A.C. & Touret, J. L. R., D. Reidel Publishing Company, Dordrecht.

Vlaar, N. J., 1986. Archaean global dynamics, *Geologie en Mijnbouw*, **65**, 91–101.

Vlaar, N. J. & van den Berg, A. P., 1991. Continental evolution and archaeo-sea-levels, in *Glacial isostasy, sea-level and mantle rheology*, pp. 637–662, eds Sabadini, R., Lambeck, K. & Boschi, E. V., Kluwer Academic Publishers, Dordrecht.

Weinstein, S. A., 1991. Laboratory and numerical studies of dynamical processes related to thermal convection in the

earth's mantle, *PhD thesis*, The Johns Hopkins University, Baltimore, MD.

Weinstein, S. A. & Olson, P. L., 1992. Thermal convection with non-Newtonian plates, *Geophys J. Int.*, **111**, 515–530.

Zhang, Y. S. & Tanimoto, T., 1992. Ridges, hotspots and their interaction as observed in seismic velocity maps, *Nature*, **355**, 45–49.

APPENDIX A

A.1 Mesh definition

For solving the coupled set of equations—Stokes momentum—and energy eqs (21) and (22)—we have used separate, interleaved finite element meshes defined on a single set of nodal points. The 2-D rectangular domain is divided into N columns and M rows of rectangular cells. A gradual mesh refinement towards the horizontal boundaries is applied. Each rectangular cell is divided into two quadratic triangular six-node elements for the velocity. Each velocity triangle is further divided by its midpoints into four triangular three-node elements for the temperature. The subdivision of a single rectangular mesh cell into two velocity and eight temperature elements is displayed in Fig. A1. The two velocity elements are spanned by the primary nodal points P_1, P_2, P_4 and P_2, P_3, P_4 respectively. The sets of nodal points of the two interleaved meshes are identical. The number of nodal points is $NP = (2N + 1) \times (2M + 1)$. The degrees of freedom of the problem are the nodal-point values of velocity and temperature. The number of degrees of freedom is $2 \times NP$ for the Stokes equation and NP for the energy equation. This mesh definition has the advantage that the degrees of freedom are related to a single set of nodal points, thus facilitating the transport of nodal-point field values from one equation of the coupled set to the other. On the other hand, the choice of linear elements of reduced element size for the energy equation provides a greater stability of the solution. This last feature is important, since we have used a Bubnov–Galerkin formulation without using upwinding techniques. Instead we used the appearance of oscillatory tempeature solutions as an indicator for insufficient grid density. In developing the codes we have used the general purpose finite element took-kit SEPRAN (Segal & Praagman 1984). This finite element tool-kit offers a great flexibility in developing computer codes for a wide range of problems by simply adding new element routines to a library of available standard elements. Furthermore it relieves the researcher from much of the task of software engineering connected with the development of a special purpose code from scratch.

A.2 Algorithms used to solve the convection equations

For the solution of the steady-state models we used the following Picard-type (successive substitution) iterative scheme to solve the system of non-linear equations.
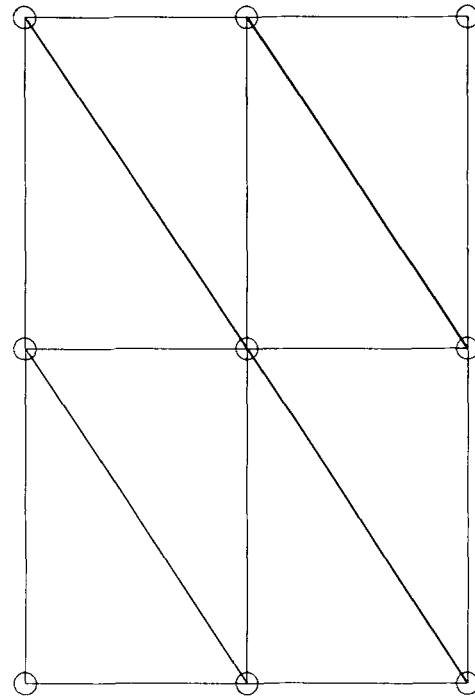
From given starting vectors for both the temperature and the velocity U_0 and T_0:

(a) solve the Stokes equation for the velocity field U_{k+1}

$$S(U_k)U_{k+1} = F(T_k).$$

Figure A1. Rectangular primary cell of the finite element mesh, spanned by the primary nodal points $P_1, \ldots, P_4$. Each rectangular cell is divided in two six-node triangles for the velocity and eight three-node triangles for the temperature.

(b) Solve the steady-state energy equation for the temperature field T_{k+1}

$$A(U_k)T_{k+1} = G.$$

(c) Repeat (a) and (b) until both fields have converged sufficiently i.e. until,

$$\max[\varepsilon_u, \varepsilon_T] < \varepsilon$$

where

$$\varepsilon_u = \frac{\|U_{k+1} - U_k\|}{\|U_k\|}, \qquad \varepsilon_T = \frac{\|T_{k+1} - T_k\|}{\|T_k\|}$$

where the vector norm used is the maximum norm, i.e. the maximum of the absolute value of the vector elements. No subiterations are performed for the solution of the non-linear Stokes equation. The convergence criterion was typically set to $\varepsilon = 0.001$.

To solve the time-dependent problem we used a predictor-corrector algorithm expressed in the following scheme: for each time step do steps (a) to (d)

(a) predictor—energy equation (implicit Euler)

$$M\Delta t^{-1}[T^{*(n+1)} - T^{(n)}] + A[U^{(n)}]T^{*(n+1)} = 0.$$

(b) Predictor—Stokes equation

$$SU^{*(n+1)} = F[T^{*(n+1)}].$$

(c) Corrector—energy equation (Crank–Nicolson)

$$M\Delta t^{-1}[T^{(n+1)} - T^{(n)}] + \tfrac{1}{2}A[U^{*(n+1)}]T^{(n+1)}$$
$$+ \tfrac{1}{2}A[U^{(n)}]T^{(n)} = 0.$$

(d) Corrector—Stokes equation

$$SU^{(n+1)} = F[T^{(n+1)}].$$

The non-linear Stokes equation is solved in each step by Picard iteration.

$$S(U)U = F$$

iterate until convergence:

$$\|U_{k+1} - U_k\| < \varepsilon_U \|U_k\|$$

$$S(U_k)U_{k+1} = F.$$

The superscripts (n) and $(n + 1)$ denote time indices. The asterisk denotes the outcome of the predictor step, Δt denotes the time-step value. The value of the convergence criterion was typically $\varepsilon_U = 0.01$, resulting in two iterations on the Stokes momentum equation in most cases, only when the flow field changes rapidly in time more iterations—up to four were needed. We used an adaptive time step, half the value of the time step based on the Courant–Friedrichs–Levy creterion Δt_{cfl}. We define,

$$\Delta t_{cfl} = \min \left[\frac{h_x}{u_x}, \frac{h_y}{u_y} \right],$$

h_x, h_y are the dimensions of the rectangular mesh cells and u_x, u_y the magnitudes of the velocity components. This value of the time step was verified to be a conservative choice. Test runs showed that the results showed little difference with results obtained with a larger fraction (up to 1.0) of the Courant time step.

The linear equations in the above scheme were solved using a sparse matrix solver based on a direct method of LU decomposition. Some benchmark results of the algorithm are given in Appendix B.

A.2 Linearization of the viscous stress term in the Stokes momentum equation

Linearization of the expression for the viscous stress in the Stokes equation can be based on a Taylor expansion of the stress in terms of the strain rate (Damsteegt, Segal & van der Zanden 1986). For a given non-linear function of y, $f(y)$ we have

$$f(y) = f[y^{(n)}] + [y - y^{(n)}]\, \partial f / \partial y \,\big|_{y=y^{(n)}} + O(y - y^{(n)})^2.$$

The viscous stress term $\tau_{ij} = \eta e_{ij}$ can be interpreted as a non-linear (tensor) function of e_{ij}. Applying the expansion formula above we have

$$\tau_{ij} = \eta e_{ij} = \eta^{(n)} e_{ij}^{(n)} + [e_{ij} - e_{ij}^{(n)}]\left[\eta^{(n)} + \frac{\partial \eta}{\partial e_{ij}}\Big|_{e_{ij}^{(n)}} e_{ij}^{(n)} \right]$$
$$+ O(e_{ij} - e_{ij}^{(n)})^2. \tag{A.1}$$

The derivative of the vicosity gives

$$\frac{\partial \eta}{\partial e_{ij}} = \frac{d\eta}{de}\frac{\partial}{\partial e_{ij}}[1/2 e_{kl}e_{kl}]^{1/2} = \frac{1}{2}\frac{d\eta}{de}e^{-1}e_{ij}.$$

In the Taylor expansion (1) this gives

$$\eta e_{ij} = \eta^{(n)} e_{ij} + [e_{ij} - e_{ij}^{(n)}]\frac{d\eta}{de}\big|_{e^{(n)}} e^{(n)} + O(\cdots)^2.$$

The first two terms represent an (approximate) linear relationship in e_{ij}. Further specialization follows through a specification of $\eta(e_{ij})$. For a power-law fluid we get

$$\eta(e_{ij}) = A e^{(1-n)/n}$$

and

$$e\frac{d\eta}{de} = (1/n - 1)A e^{(1-n)/n} = (1/n - 1)\eta.$$

Substitution in the linearized stress term results in

$$\tau_{ij} \approx \eta^{(n)}e_{ij} + [e_{ij} - e_{ij}^{(n)}](1/n - 1)\eta^{(n)}$$
$$= \eta^{(n)}\{(1 - 1/n)e_{ij}^{(n)} + (1/n)e_{ij}\}. \tag{A.2}$$

For a Newtonian fluid with power-law index $n = 1$ this expression gives the correct result. Note that the first term in the linearized expression contains the result of the (previous) nth iteration for the strain-rate tensor. This known term will result in a contribution to the right-hand side of the finite element equations to be solved. In practice it appears that the Newton–Raphson-type linearization scheme (A.2) does not converge properly for powerlaw index $n > 1$ (Crochet, Davies & Walters 1984). Instead of (A.2), we have used the zeroth-order term of the Taylor expansion,

$$\eta e_{ij} = \eta^{(n)}e_{ij} + O[e_{ij} - e_{ij}^{(n)}]. \tag{A.3}$$

This iteration formula suggests a 'successive substitution' method also known as Picard iteration.

In calculating the stiffness matrix for the Stokes equation we use a seven-point Gauss integration scheme (Cuvelier *et al.* 1986). Interpolation for the Gauss-point values is applied to the scalar strain rate e and the viscosity is expressed analytically in terms of e. Newton–Cotes integration schemes for the stiffness matrix (Cuvelier *et al.* 1986) involve quadratic interpolation of the viscosity itself. Care has to be taken to avoid negative viscosity values in areas of large gradients that may cause instability of the finite elements equation, (Christensen 1984; Malevsky & Yuen 1992). We have compared the Gauss integration scheme with a Newton–Cotes scheme that employed linear-viscosity interpolation (avoiding negative viscosity values). The seven-point Gauss scheme proved to be stable and display faster convergence for increasing grid density—consistent with the higher-order accuracy of the Gauss scheme.

APPENDIX B

B.1 Time-dependent benchmark results

In order to validate the code for time-dependent models, we have performed a bench-mark calculation for a time-dependent convection problem and compared the results with those reported in (Malevsky & Yuen 1992). In the bench mark the time-dependent equations for Rayleigh–Benard convection—with a pure power-law ($\beta = 0$) fluid with power-law index $n = 3$, are solved for a box of aspect ration $\lambda = 1$. The Rayleigh number as defined in (22) was

Table B1. Benchmark results.

Case	Mesh	Δt	ε_U	$t = 0.0$		$t = 0.002$	
				ψ_{max}	$Vrms$	ψ_{max}	$Vrms$
1	10x10	$0.5 \times \Delta t_{cfl}$	10^{-2}	5.120	11.70	17.00	38.88
2	10x10	$0.1 \times \Delta t_{cfl}$	10^{-2}	5.120	11.70	19.85	45.40
3	10x10	5×10^{-5}	10^{-2}	5.120	11.70	20.27	46.36
4	10x10	5×10^{-5}	2×10^{-3}	5.180	11.83	20.94	47.90
5	20x20	$0.5 \times \Delta t_{cfl}$	10^{-2}	5.175	11.82	20.01	45.77
6	20x20	$0.1 \times \Delta t_{cfl}$	10^{-2}	5.175	11.82	20.93	47.88
7	20x20	5×10^{-5}	10^{-2}	5.175	11.82	21.05	48.15
8	20x20	5×10^{-5}	2×10^{-3}	5.235	11.96	21.72	49.69
9	30x30	$0.5 \times \Delta t_{cfl}$	10^{-2}	5.183	11.843	20.55	47.01
10	30x30	$0.1 \times \Delta t_{cfl}$	10^{-2}	5.183	11.843	21.02	48.10
11	30x30	5×10^{-5}	10^{-2}	5.183	11.843	21.15	48.40
12	30x30	5×10^{-5}	2×10^{-3}	5.243	11.98	21.83	49.96
13	32x32	5×10^{-5}	10^{-4}			21.594	49.592
14	32x32	5×10^{-5}	10^{-5}			22.141	50.664
15	30x30	$0.1 \times \Delta t_{cfl}$	10^{-2}			21.470	49.160

$2^{-1/3} \times 500 = 396.85$. This value of the Rayleigh number was chosen in order to match the bench-mark computation reported by Malevsky & Yuen (1992) who used a slightly different definition of the Rayleigh number. The same boundary conditions are applied as in the rest of this paper, i.e. free slip and impermeable mechanical conditions, fixed temperature 0 and 1 at the top and bottom boundaries and insulating ($\nabla T \cdot \mathbf{n} = 0$) conditions at the vertical boundaries. The initial temperature is defined as a perturbed conductive field,

$$T(x, z) = z + 0.1 \times \sin(\pi z) \cos(\pi x)$$

where z measure the depth below the top boundary. Instead of defining a zero-velocity starting vector, we specify the viscosity to be uniform $\eta = 1.0$, for the first Picard iteration in the solution of the Stokes equation at time $t = 0$. In this way we avoid the singularity in the power-law viscosity in regions with negligible strain rates. In subsequent Picard iterations the viscosity is computed from the velocity outcome from the previous iteration. In Table B1 we present Vrms and the maximum value of the stream function ψ_{max} at time $t = 0$, and $t = 0.002$ which is close to the first maximum of the time series $\psi_{max}(t)$.

The result of case (13), (14) and (15) of Table B1 are taken from (Malevsky & Yuen 1992). Results (13) through (15) were obtained using independent stream-function codes by Malevsky & Yuen (1992), Christensen & Yuen (1989) and van Keken (1993). The integration time step Δt was either taken as a fraction of the Courant time step (Δt_{cfl}) or as a fixed value. Different convergence criteria for the Picard iteration for the Stokes equation (ε_U) have been applied—see Appendix A2. The mesh-size in the second column of Table B1 refers to the number of rectangular ($N \times M$) mesh cells discussed in Appendix A1. Taking case (12) as a reference, we see that convergence occurs when using a denser mesh or when either the integration time step or the Picard convergence criterion is reduced. The results of the three different methods agree to within a few per cent.