

- the inhomogeneous plasma of the Jovian magnetosphere and, in general, are not in charge equilibrium with the local plasma environment. At any time, the charge on the grain is determined by currents due to the local plasma environment and the charge acquired in other environments encountered earlier. This charge memory makes the grain charging history irreversible and therefore nonconservative, so that energy and angular momentum may be gained from or lost to the magnetosphere.
11. M. Baguhl *et al.*, *Space Sci. Rev.* **72**, 471 (1995).
 12. Here, grain size refers to the radius of spherical particles. We used smaller interstellar dust particles than interplanetary dust particles, because the higher velocity (energy) of the interstellar grains relative to Jupiter requires a stronger electromagnetic interaction with the magnetosphere for capture to occur. Numerical integrations of test particles confirmed that larger interstellar dust particles cannot lose enough energy to become captured, whereas smaller interstellar dust particles have enough kinetic energy to penetrate deep into the jovian magnetosphere, where energy and angular momentum exchange can lead to capture.
 13. A $10^{-4} R_J$ target area was used for interstellar dust, because their higher velocities relative to Jupiter mean that only particles on near-miss trajectories can be captured. Numerical integrations of test particles with larger impact parameters showed that they are not captured. A smaller area was used to improve capture statistics.
 14. J. A. Burns, M. R. Showalter, G. E. Morfill, in *Planetary Rings*, R. Greenberg and A. Brahic, Eds. (Univ. of Arizona Press, Tucson, AZ, 1984), pp. 200–272.
 15. E. J. Öpik, *Proc. R. Ir. Acad.* **54**, 165 (1951).
 16. J. E. Colwell, *Icarus* **106**, 536 (1993).
 17. N. Divine, *J. Geophys. Res.* **98**, 17029 (1993).
 18. The cometary orbits are for long-period Oort-cloud comets. Velocities of cometary dust relative to Jupiter in this model range from 8 to 34 km/s (planetary dust approach velocities are between 6 and 10 km/s). Heliocentric orbital elements are appropriate for dust released from an Oort-cloud comet: $a = 20,000$ to $40,000$ AU, $e = 1$ to $1 - (5.2 \text{ AU}/a)$, and $\cos i$ randomly distributed between -1 and 1 (4).
 19. N. Asada, *J. Geophys. Res.* **90**, 12445 (1985).
 20. H. A. Zook and S. Su, *Lunar Planet. Sci.* **XIII**, 893 (1982). The four retrograde satellites are Ananke ($a = 21.2 \times 10^{11}$ cm, $e = 0.169$, $i = 147^\circ$, $r = 15$ km), Carme ($a = 22.6 \times 10^{11}$ cm, $e = 0.207$, $i = 163^\circ$, $r = 22$ km), Pasiphae ($a = 23.5 \times 10^{11}$ cm, $e = 0.378$, $i = 148^\circ$, $r = 35$ km), and Sinope ($a = 23.7 \times 10^{11}$ cm, $e = 0.275$, $i = 153^\circ$, $r = 20$ km).
 21. H. A. Zook, S. Su, D. D. Humes, *Lunar Planet. Sci.* **XVI**, 946 (1985).
 22. E. Grün, H. A. Zook, H. Fechtig, R. H. Giese, *Icarus* **62**, 244 (1985).
 23. D. E. Gault, *Moon* **6**, 32 (1973).
 24. D. Stöffler, D. E. Gault, J. Wedekind, G. Polkowski, *J. Geophys. Res.* **80**, 4062 (1975).
 25. K. R. Housen, R. M. Schmidt, K. A. Holsapple, *ibid.* **88**, 2485 (1983).
 26. Particles launched from the retrograde jovian satellites are already bound to Jupiter, and their orbital evolution is much slower than that of exogenic grains that pass through the entire magnetosphere in one orbit. Computation time constraints limited the number of such particles at each size to 2500. Enough integrations were performed to arrive at an upper limit of capture efficiency of such grains into small retrograde orbits like those detected by the Galileo DDS.
 27. M. Horányi, *Geophys. Res. Lett.* **21**, 1039 (1994).
 28. _____ and J. A. Burns, *J. Geophys. Res.* **96**, 19283 (1991).
 29. Partially supported by NASA. G. Linkert and H. Krüger assisted in Galileo data reduction and A. Heck assisted in figure preparation.

24 October 1997; accepted 18 February 1998

Time Scales and Heterogeneous Structure in Geodynamic Earth Models

Hans-Peter Bunge,* Mark A. Richards,
Carolina Lithgow-Bertelloni, John R. Baumgardner,
Stephen P. Grand, Barbara A. Romanowicz

Computer models of mantle convection constrained by the history of Cenozoic and Mesozoic plate motions explain some deep-mantle structural heterogeneity imaged by seismic tomography, especially those related to subduction. They also reveal a 150-million-year time scale for generating thermal heterogeneity in the mantle, comparable to the record of plate motion reconstructions, so that the problem of unknown initial conditions can be overcome. The pattern of lowermost mantle structure at the core-mantle boundary is controlled by subduction history, although seismic tomography reveals intense large-scale hot (low-velocity) upwelling features not explicitly predicted by the models.

Geodynamic Earth models were pioneered by Hager and O'Connell (1), who calculated mantle flow by imposing present-day plate motions as a surface boundary condition. With the advent of global seismic tomography (2), these models were extended to predict the geoid and dynamic topog-

raphy (3). However, these Earth models are "static," because they solve for instantaneous mantle flow in response to boundary conditions, internal loads, or both.

Time-dependent Earth models are required to understand how the evolution of mantle flow affects Earth processes that occur on geologic time scales. For example, continental shelf and platform stratigraphy are controlled by vertical motions of the continental lithosphere in response to mantle convection (4). True polar wandering is caused by changes in the inertia tensor as a result of mantle convection (5), and the alternation between periods of rapid and slow magnetic field reversals is probably related to mantle-controlled changes at the core-mantle boundary (CMB).

The development of time-dependent Earth models has been delayed for several reasons: (i) Sufficient computer power to

resolve the narrow thermal boundary layers in global mantle convection models has not been available; (ii) it is not obvious how the internal mantle density structure can be related to plate motion observations at the surface; and (iii) it is not known how time-dependent Earth models can be initialized at some starting point in the past, because the mantle density structure is known only for the present day (6).

Some of these difficulties have been overcome. (i) Advances in computer power allow three-dimensional (3D) spherical convection to be simulated at a resolution on the order of 50 to 100 km (7, 8). At the same time, large-scale mantle velocity heterogeneity structure has been mapped in greater detail (9, 10), and seismic tomography has imaged subducted slabs (11–13). (ii) The connection of internal mantle density structure to the history of subduction (14, 15) has allowed estimation of the internal buoyancy forces that drive plates (16). These developments allow convection models to be combined with plate motion reconstructions and such models to be tested with seismic data.

Figure 1B shows an Earth model obtained with the TERRA convection code (17, 18). More than 10 million finite elements provide an element resolution of about 50 km throughout the mantle, which allowed us to model convection at a Rayleigh number of 10^8 (19). The history of plate motion is imposed as a time-dependent velocity boundary condition (20) starting in the mid-Mesozoic at 119 to 100 million years ago (Ma). We chose this starting time because well-constrained reconstructions exist only as far back in time as the 119 to 100 Ma period.

In computing the Earth model (Table 1) we assumed that (i) the mantle is of uniform

H.-P. Bunge, Institut de Physique du Globe de Paris, Laboratoire de Sismologie, 4 place Jussieu, 75252 Paris Cedex 05, France.

M. A. Richards, Department of Geology and Geophysics, University of California, Berkeley, CA 94720, USA.

C. Lithgow-Bertelloni, Department of Geological Sciences, University of Michigan, Ann Arbor, MI 48109, USA.

J. R. Baumgardner, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87544, USA.

S. P. Grand, Department of Geological Sciences, University of Texas, Austin, TX 78713, USA.

B. A. Romanowicz, Berkeley Seismological Laboratory, and Department of Geology and Geophysics, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: bunge@ipgg.jussieu.fr

chemical composition and (ii) is heated primarily from within by radioactivity, with a modest 20% component of bottom heating from an isothermal core. (iii) Viscosity increases by a factor of 40 in the lower mantle and in the lithosphere (21) relative to the asthenosphere. (iv) Mantle minerals undergo two phase transitions at a depth of 410 and 670 km. (v) The CMB supports no shear stresses. (vi) The motion of the upper boundary layer can be deduced from plate-motion reconstructions. (vii) The initial condition for thermal heterogeneity in the mantle is approximated by a quasi steady-state of convection derived by imposing mid-Mesozoic plate motions for all time before 119 Ma.

The first six assumptions roughly consti-

tute a “standard model” for whole mantle convection (22). An internally heated mantle is consistent with cosmochemical and geochemical constraints on abundances of radioactive elements in the mantle (23), but a larger component of core heating is possible. Postglacial rebound and geoid modeling suggest an increase in viscosity of one to two orders of magnitude from the upper to the lower mantle (24, 25).

Our seventh assumption is the most problematic, so we tested how mantle convection responds to a realistic change in plate movement. We first imposed 119 to 100 Ma plate motions for 2 billion years to produce an initial condition (Fig. 1A). We then imposed present-day plate motions for 500 million years (My). The global cross-correlation

of the evolving mantle temperature field with the initial condition starts from exactly one [perfect correlation (Fig. 2)] and evolves toward a final stage, at which nearly all initial-condition information is lost. Most of the adjustment occurs during the first 150 My, and the correlation falls to about 0.3. It reaches a value of about 0.2 in the remaining 350 My, which reflects the correlation of the Mesozoic and the present-day plate configuration. The CMB correlation also declines rapidly to about 0.4 during the initial 150 My, accelerated by our inclusion of about 20% core heating (18, 26).

The modeled response time of 150 My is comparable to the vertical transit time in

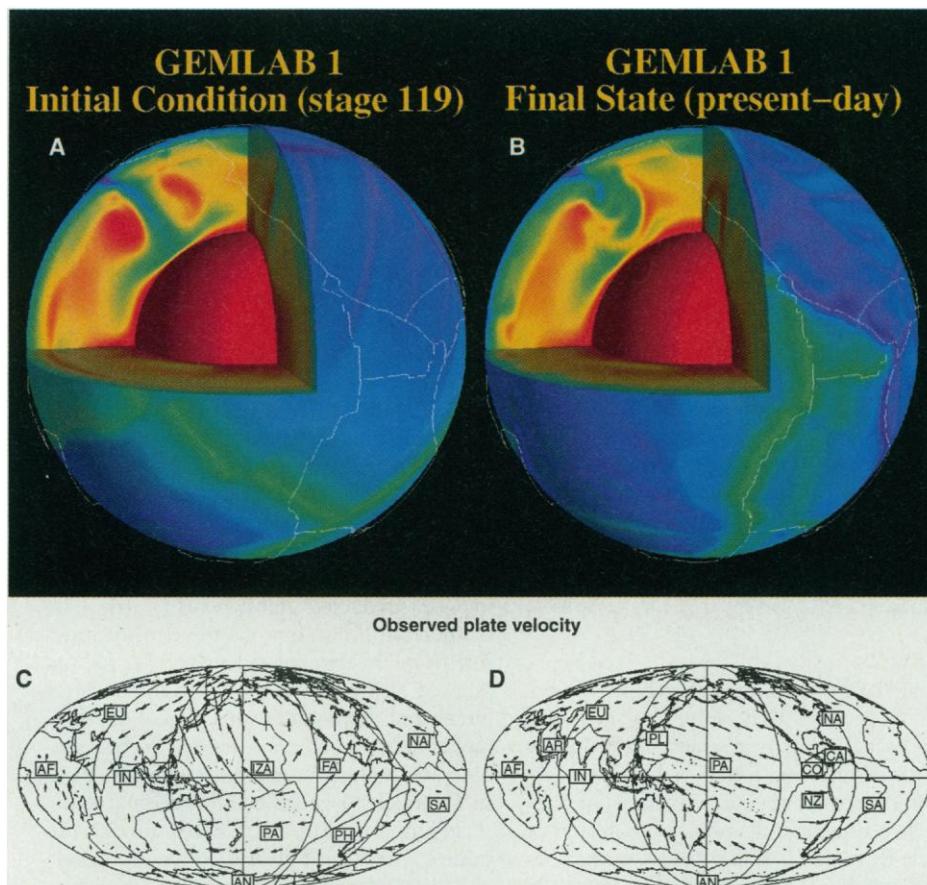


Fig. 1. (A) Cut-away of the 3D temperature field for a starting model seen from the Pacific hemisphere. (GEMLAB: Geodynamic Earth Model of Los Alamos and Berkeley). The model was obtained by imposing mid-Mesozoic plate motions until quasi steady-state was reached (see text). Blue is cold and red is hot. The upper 50 km of the mantle are removed in order to show the convective planform. Present-day plate boundaries are drawn for geographic reference. Narrow hot zones near the surface reflect passive mantle upwelling at the Izanagi, Farallon, and Phoenix spreading centers. The prominent cold downwelling in the cross-sectional view results from subduction of the Izanagi and Farallon plates in the northwestern Pacific. (B) Same as (A) but after the 119-Ma through present-day plate motion record has been imposed. The major difference lies in the more complicated downwelling structure of (B), which reflects the history of subduction beneath the northwestern Pacific. (C) Map of plate boundaries for the 119 to 100 Ma stage. Arrows indicate the direction of plate motion. Their length is proportional to the plate speed. The ancient Izanagi (IZA), Farallon (FA), and Phoenix (PH) plates occupy most of the Pacific basin. (D) Same as (C), but for the present day. The Izanagi, Farallon, and Phoenix plates have largely disappeared.

Table 1. Model parameters.

Outer shell radius	6370 km
Inner shell radius	3480 km
T , surface	300 K
T , CMB	2800 K
ρ , surface	3500 kg m^{-3}
ρ , CMB	5568 kg m^{-3}
α , surface	$4.011 \times 10^{-5} \text{ K}^{-1}$
α , CMB	$1.256 \times 10^{-5} \text{ K}^{-1}$
η (upper mantle)	$8.0 \times 10^{21} \text{ Pa s}$
η (lithosphere and lower mantle)	$40 \times \eta$ upper mantle
Thermal conductivity k	$6.0 \text{ W m}^{-1} \text{ K}^{-1}$
Internal heating rate Q_{int}	$6.0 \times 10^{-12} \text{ W kg}^{-1}$
Heat capacity	$1.134 \times 10^3 \text{ J Kg}^{-1} \text{ K}^{-1}$
γ , 410 km	2 MPa K^{-1}
γ , 670 km	-4 MPa K^{-1}

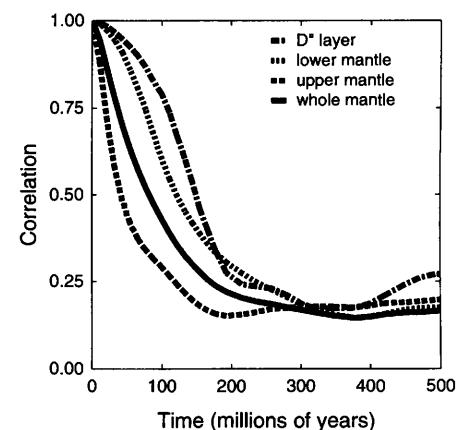


Fig. 2. Global cross-correlation of the evolving mantle convection model with the initial condition as a function of time. Correlation is shown separately for the upper, the lower, and the whole mantle, and for the CMB. A correlation of 1 means perfect correlation. Zero means no correlation. The model evolves from perfect correlation at the beginning toward a final state corresponding to the present-day velocity boundary conditions. Most of the adjustment occurs within the first 150 My. The final correlation value of 0.2 reflects the correlation between the Mesozoic and the present-day plate tectonic regime.

our convection model, that is, the time it takes for thermal disturbances to be advected from the surface plates to the CMB. Thus, our simulation demonstrates that most of the initial-condition “information” from past plate-motion regimes is lost after about 150 Ma, suggesting that the plate motion record is probably just adequate for modeling the genesis of present-day mantle heterogeneity (27).

We compared the Earth model with the seismic shear body-wave velocity structure of Grand (12), which is similar to the compressional body-wave study of van der Hilst *et al.* (13), and with the long-period horizontal shear-wave (SH) study (SAW 12) of Li and Romanowicz (10). Convection at 1100 and 2000 km (Figs. 3A and 4A) is dominated by narrow, linear cold downwellings under North America and the western Pacific caused by subduction of the Farallon, Izanagi, and Pacific plates. Temperature variations away from downwellings are small, as expected for mostly internally heated convection. The Farallon slab is displaced eastward rela-

tive to the simple model of Ricard *et al.* (15), because of return flow from sub-Pacific into sub-Atlantic mantle beneath America.

Smaller scale downwellings away from plate boundaries are the result of boundary-layer instabilities beneath slow-moving plates such as Africa. However, their role is minor as a result of the relatively stiff lithosphere and the radial mantle viscosity structure (28). There is a lack of active hot upwellings, owing to the small amount of core heating. Near the CMB at 2800 km (Fig. 5A), heterogeneity is dominated by large-scale structure as the cold downwellings spread laterally at the CMB.

The body-wave tomography shows narrow, sheetlike downwellings corresponding to old subducted slabs beneath the Americas (Farallon plate), the Tethys, and the western Pacific. There is a prominent low-velocity anomaly under Africa. Other mantle regions show minor heterogeneity. The CMB is characterized by broadscale fast-velocity anomalies in the circum-Pacific and a very large amplitude low-velocity anomaly be-

neath Africa. The long-period SH study, which gives a particularly good fit to the nonhydrostatic geoid, agrees in the overall location of these anomalies.

Some characteristics of GEMLAB1 are similar to those of Grand’s S-wave study, but there are important differences. For example, the cold subduction-related CMB temperature pattern in GEMLAB1 correlates well with the seismic models, but there are no prominent hot regions with temperatures substantially above the mean. This result is expected inside a hot thermal boundary layer and suggests that the very low wave-speed anomalies may not represent purely thermal effects. They may, however, result from chemical heterogeneity, which may also help to explain why some of the low wave-speed anomalies are confined to the deeper mantle.

In the midmantle, the strong endothermic phase change delays the sinking of material through the transition zone, and GEMLAB1 provides a relatively poor match to Grand’s image of subducted slabs

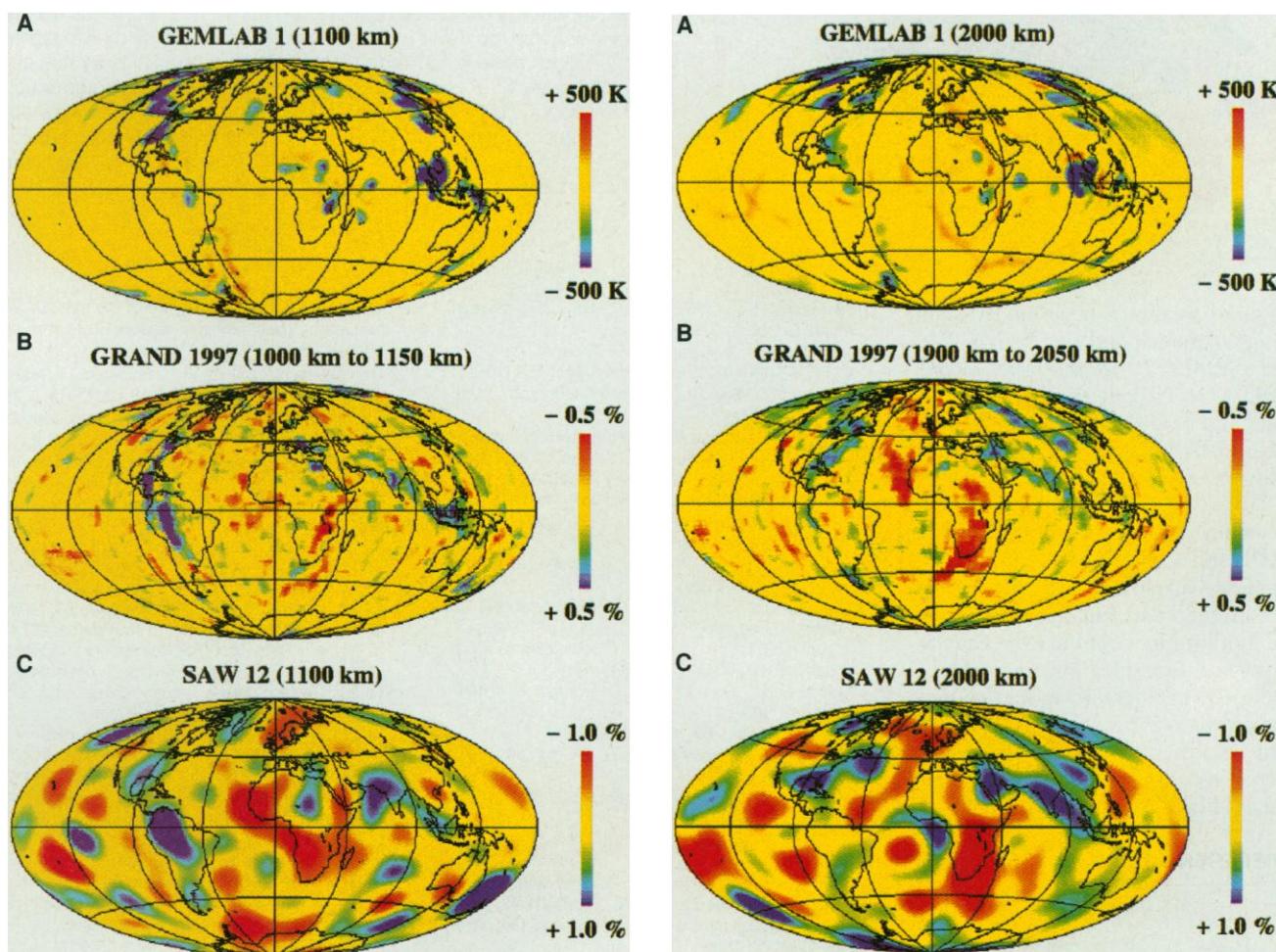


Fig. 3 (left). Maps of de-meaned lateral heterogeneity at a depth of 1100 km. (A) The standard Earth model GEMLAB1 described in the text. The seismic shear wave models are from Grand (B) and Li and Romanowicz (C). For the

convection model, blue is cold and red is hot. For the seismic models, blue is fast and red is slow. **Fig. 4 (right).** Same as Fig. 3 but at a depth of 2000 km.

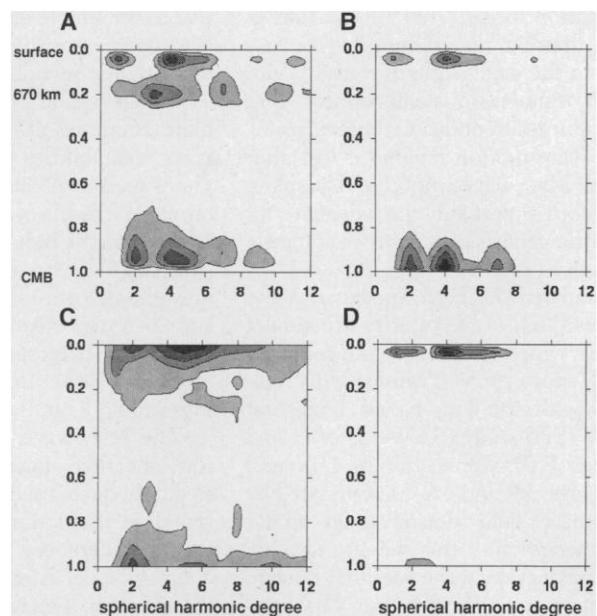
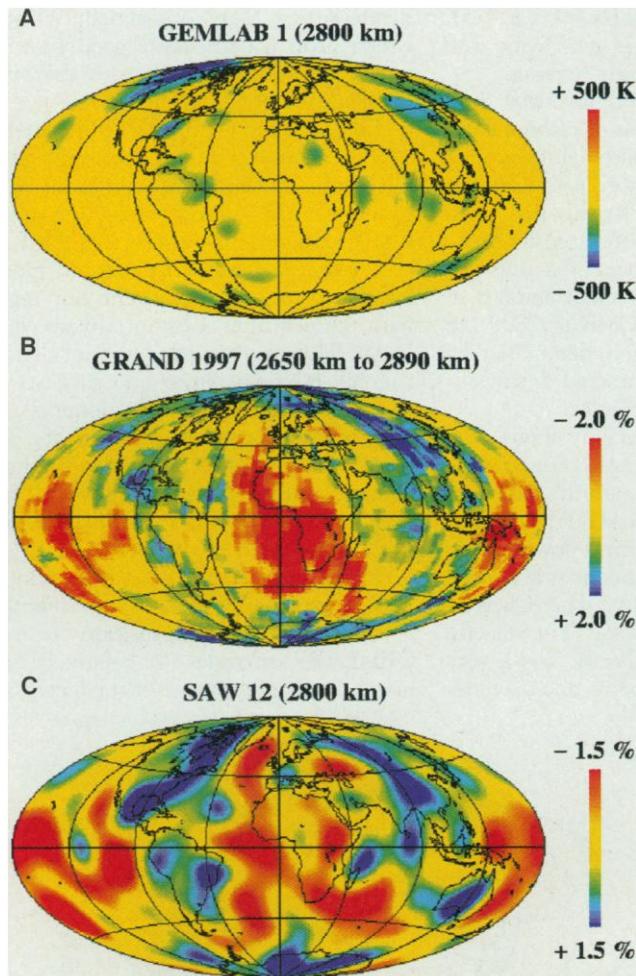


Fig. 5 (left). Same as Fig. 3 but at the CMB. **Fig. 6 (right).** Spectral heterogeneity maps (SHMs) for GEMLAB1 (A), and the seismic model of Grand (D) and Li and Romanowicz (C). GEMLAB2 (B) is similar to GEMLAB1, but without the strong phase transitions. Root-mean-square spectral amplitude is contoured as a function of nondimensional mantle depth (surface at the top, CMB at the bottom) and spherical harmonic degree (0 to 12). Each panel is normalized to the maximum amplitude for that panel, and there are five contour intervals. The SHMs show a low-degree structure in the upper and lower boundary layer, which gives way to a whiter spectrum in the midmantle. In GEMLAB1, an additional pronounced heterogeneity is seen in the transition zone because of the strong endothermic phase transition at a depth of 670 km.

beneath some weaker subduction systems under South America and the Tethys. We compared GEMLAB1 to an Earth model GEMLAB2 (29) without the strong phase transitions, and found that GEMLAB2 provides a better fit to Grand's study, as evidenced by the spectral heterogeneity plots (Fig. 6).

The midmantle differences between GEMLAB1 and GEMLAB2 illustrate the effect of model parameter changes (phase transitions in this case), but uncertainties in the plate motion input data may be equally important. For example, our reconstructions indicate considerably more subduction beneath the northwestern Pacific than is evident in either seismic model, which suggest that reconstructions in GEMLAB models should be refined.

REFERENCES AND NOTES

1. B. H. Hager and R. J. O'Connell, *J. Geophys. Res.* **84**, 1031 (1979).
2. A. M. Dziewonski, *ibid.* **89**, 5929 (1984); J. H. Woodhouse and A. M. Dziewonski, *ibid.*, p. 5953; B. H. Hager and R. W. Clayton, in *Mantle Convection: Plate Tectonics and Global Dynamics*, W. R. Peltier, Ed. (Gordon Breach, New York, 1989), vol. 4, chap. 9.
3. M. A. Richards and B. H. Hager, *J. Geophys. Res.*

- 89**, 5987 (1984); Y. Ricard, L. Fleitout, C. Froidevaux, *Annal. Geophys.* **2**, 267 (1984).
4. J. X. Mitrovica, C. Beaumont, G. T. Jarvis, *Tectonics* **8**, 1079 (1989); M. Gurnis, *Nature* **344**, 754 (1990).
5. Y. Ricard, G. Spada, R. Sabadini, *Geophys. J. Int.* **113**, 284 (1993).
6. Because heat transport involves diffusion, mantle convection models cannot be run backward in time from a relatively well-known present-day state, even if a perfect forward model could be constructed.
7. P. J. Tackley, D. J. Stevenson, G. A. Glatzmaier, G. Schubert, *Nature* **361**, 699 (1993).
8. H.-P. Bunge, M. A. Richards, J. R. Baumgardner, *ibid.* **379**, 436 (1996).
9. W.-J. Su, R. L. Woodward, A. M. Dziewonski, *J. Geophys. Res.* **99**, 6945 (1994); G. Masters, S. Johnson, G. Laske, H. Bolton, *Philos. Trans. R. Soc. London Ser. A* **354**, 1385 (1996).
10. X. D. Li and B. Romanowicz, *J. Geophys. Res.* **101**, 22245 (1996).
11. S. P. Grand, *ibid.* **99**, 11591 (1994).
12. _____, R. D. van der Hilst, S. Widiyantoro, *GSA Today* **7**, 1 (1997).
13. R. D. van der Hilst, S. Widiyantoro, E. R. Engdahl, *Nature* **386**, 578 (1997).
14. M. A. Richards and D. C. Engebretson, *ibid.* **355**, 437 (1992).
15. Y. Ricard, M. A. Richards, C. Lithgow-Bertelloni, Y. Le Stunff, *J. Geophys. Res.* **98**, 21895 (1993).
16. C. Lithgow-Bertelloni and M. A. Richards, *Geophys. Res. Lett.* **22**, 1317 (1995).
17. H.-P. Bunge and J. R. Baumgardner, *Comput. Phys.* **9**, 207 (1995).
18. H.-P. Bunge, M. A. Richards, J. R. Baumgardner, *J. Geophys. Res.* **102**, 11991 (1997).
19. The Rayleigh number based on internal heating ap-

proaches the dynamic regime of the mantle, which probably convects with an effective Rayleigh number on the order of 10^9 to 10^9 .

20. The plate tectonic information consists of 11 plate-motion stages extending back to 119 Ma. During a stage, plate motions are relatively constant. Stage boundaries correspond to episodes of large or sudden changes in plate motions, that is, plate rearrangements. Plate stage information consists of the complete plate boundaries for all the major plates and the Euler rotation vectors of each plate. The first six plate stages (0 to 10, 10 to 25, 25 to 43, 43 to 48, 48 to 56, and 56 to 64 Ma) span the Cenozoic and are based on the global plate boundaries and rotation poles of Gordon and Jurdy (30); the remaining five Mesozoic plate stages (64 to 74, 74 to 84, 84 to 94, 94 to 100, and 100 to 119) are based on the global compilation of Lithgow-Bertelloni and Richards (31). All plate boundaries and rotation poles are in the hotspot reference frame. The inherent uncertainties associated with using plate tectonic information in geodynamic models are described in (31). Many of the Mesozoic reconstructions, particularly for plates that have disappeared (Izanagi, Phoenix, Kula) or are in the process of disappearing (Farallon), are only approximate, especially the positions of subduction zones and to a lesser extent the rotation poles.
21. In the Earth the lithospheric viscosity is many orders of magnitude larger than the upper mantle viscosity, but the numerical capability to resolve such large viscosity contrast is not available. Simulations (28) demonstrate that a 20-fold viscosity increase in the lithosphere is sufficient to suppress boundary-layer instabilities.
22. G. F. Davies, and M. A. Richards, *J. Geol.* **100**, 151 (1992).
23. G. J. Wasserburg, G. J. F. MacDonald, F. Hoyle,

W. A. Fowler, *Science* **143**, 465 (1964).
 24. J. X. Mitrovica, *J. Geophys. Res.* **101**, 555 (1996).
 25. B. H. Hager and M. A. Richards, *Philos. Trans. R. Soc. London Ser. A* **328**, 309 (1989).
 26. This CMB response time is accelerated by a factor of 2, relative to a model without core heating.
 27. The correspondence of the mantle response time and the time period for reliable reconstructions is not coincidental. Reconstructions are largely dependent on the magnetic isochrons of sea-floor spreading, which is limited to the characteristic maximum age for oceanic lithosphere. Simple boundary-layer convection theory (22) predicts that this characteristic boundary-layer age should be similar to the vertical transit time for convection. Based on the prediction that advection velocities scale roughly with the natural logarithm of the viscosity contrast (18), the mantle response time should be accurate within a factor of 2.
 28. H.-P. Bunge and M. A. Richards, *Geophys. Res. Lett.* **23**, 2987 (1996).
 29. H.-P. Bunge *et al.*, data not shown.
 30. R. G. Gordon and D. M. Jurdy, *J. Geophys. Res.* **91**,

12389 (1986).
 31. C. Lithgow-Bertelloni and M. A. Richards, *Rev. Geophys.* **36**, 72 (1998).
 32. We thank G. Davies and R. van der Hilst for constructive reviews, J. Painter for supporting the 3D graphics, and the Los Alamos Branch of the Institute of Geophysics and Planetary Physics for continuing support. Computing resources were provided by the Advanced Computing Laboratory of Los Alamos National Laboratory.

27 October 1997; accepted 18 February 1998

Strong Regularities in World Wide Web Surfing

Bernardo A. Huberman, Peter L. T. Piroli, James E. Pitkow, Rajan M. Lukose

One of the most common modes of accessing information in the World Wide Web is surfing from one document to another along hyperlinks. Several large empirical studies have revealed common patterns of surfing behavior. A model that assumes that users make a sequence of decisions to proceed to another page, continuing as long as the value of the current page exceeds some threshold, yields the probability distribution for the number of pages that a user visits within a given Web site. This model was verified by comparing its predictions with detailed measurements of surfing patterns. The model also explains the observed Zipf-like distributions in page hits observed at Web sites.

The exponential growth of the World Wide Web is making it the standard information system for an increasing segment of the world's population. The Web allows inexpensive and fast access to unique and novel services, including electronic commerce, information resources, and entertainment, provided by individuals and institutions scattered throughout the world (1). Despite the advantages of this new medium, the Internet still fails to serve the needs of the user community in a number of ways. Surveys of Web users find that slow access and inability to find relevant information are the two most frequently reported problems (2). The slow access is at least in part a result of congestion (3), whereas the difficulty in finding useful information is related to the balkanization of the Web structure (4). Because it is difficult to solve this fragmentation problem by designing an effective and efficient classification scheme, an alternative approach is to seek regularities in user patterns that can then be used to develop technologies for increasing the density of relevant data for users.

A common way of finding information on the Web is through query-based search engines, which enable quick access to information that is often not the most relevant. This lack of relevance is partly attributable to the impossibility of cataloging an exponentially growing amount of information in ways that anticipate users' needs.

But because the Web is structured as a hypermedia system, in which documents are linked to one another by authors, it also supports an alternative and effective mode of use in which users surf from one document to another along hypermedia links that appear relevant to their interests.

Here, we describe several strong regularities of Web user surfing patterns discovered through extensive empirical studies of different user communities. These regularities can be described by a law of surfing, derived below, that determines the probability distribution of the depth—that is, the number of pages a user visits within a Web site. In conjunction with a spreading activation algorithm, the law can be used to simulate the surfing patterns of users on a given Web topology. This leads to accurate predictions of page hits. Moreover, it explains the observed Zipf-like distributions of page hits to Web sites (5).

We start by deriving the probability $P(L)$ of the number of links L that a user follows in a Web site. This can be done by considering that there is value in each page a user visits, and that clicking on the next page assumes that it will be valuable as well. Because the value of the next page is not certain, we can assume that it is stochastically related to the previous one. In other words, the value of the current page V_L is the value of the previous one V_{L-1} plus or minus a random term. Thus, the page values can be written as

$$V_L = V_{L-1} + \epsilon_L \quad (1)$$

where the values ϵ_L are independent and identically distributed Gaussian random variables. A particular sequence of page valuations is a realization of a random process and thus is different for each user. Within this formulation, an individual will continue to surf until the expected cost of continuing is perceived to be larger than the discounted expected value of the information to be found in the future. This can be thought of as a real option in financial economics, for which it is well known that there is a threshold value for exercising the option to continue (6, 7). Even if the value of the current page is negative, it may be worthwhile to proceed, because a collection of high-value pages may still be found. If the value is sufficiently negative, however, then it is no longer worth the risk to continue. That is, when V_L falls below some threshold value, it is optimal to stop.

The number of links a user follows before the page value first reaches the stopping threshold is a random variable L . For the random walk of Eq. 1, the probability distribution of first passage times to a threshold is given asymptotically by the two-parameter inverse Gaussian distribution

$$P(L) = \sqrt{\frac{\lambda}{2\pi L^3}} \exp\left[\frac{-\lambda(L - \mu)^2}{2\mu^2 L}\right] \quad (2)$$

(8), with mean $E(L) = \mu$ and variance $\text{Var}(L) = \mu^3/\lambda$, where λ is a scale parameter. This distribution has two characteristics worth stressing in the context of user surfing patterns. First, it has a very long tail, which extends much farther than that of a normal distribution with comparable mean and variance. This implies a finite probability for events that would be unlikely if described by a normal distribution. Consequently, large deviations from the average number of user clicks computed at a site will be observed. Second, because of the asymmetry of the distribution function, the typical behavior of users will not be the same as their average behavior. Thus, because the mode is lower than the mean, care must be exercised with available data on the average number of clicks, as this average overestimates the typical depth being surfed.

To test the validity of Eq. 2, we analyzed data collected from a representative sample

Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA.