Earth and Planetary Science Letters 421 (2015) 98–106

Contents lists available at ScienceDirect

Earth and Planetary Science Letters

www.elsevier.com/locate/epsl

Differential contraction of subducted lithosphere layers generates deep earthquakes

Lijun Liu *, Jin S. Zhang

Department of Geology, University of Illinois at Urbana–Champaign, United States

ARTICLE INFO

Article history:
Received 13 February 2015
Received in revised form 30 March 2015
Accepted 31 March 2015
Available online xxxx
Editor: P. Shearer

Keywords:
differential volume reduction
phase transformation
focal mechanism
slab internal stress
down-dip compression
outboard earthquakes

ABSTRACT

The origin of forces for generating deep earthquakes remains elusive. We propose a new mechanical model that involves constant release of shear strain accumulated between the crust and mantle lithosphere caused by differential volume changes during phase transformations as a slab sinks. This generates increasing down-dip compression inside the slab with depth, consistent with the global distribution of deep earthquake focal mechanisms. Using experimentally calibrated slab rheology, we show that the estimated distribution of slab internal stress agrees well with the depth–frequency relationship of deep earthquakes in both cold and warm slabs, with the peak in seismicity at ~600 km depth corresponding to an induced stress maximum. This mechanism of residual stress within a cold slab also provides a solution to deep earthquakes occurring beneath former subduction zones such as western Tonga and Spain. We further suggest that this model may reconcile existing geodynamic paradoxes concerning the absolute strength of slabs, discrepancies on mantle rock strength between laboratory and geophysical inversions, as well as the competition between earthquake-generating stresses and slab-pull forces.

1. Introduction

Earthquakes are generally considered a consequence of tectonic deformation at shallow depth (<50 km). However, the origin of forces generating earthquakes at greater depths, especially below 300 km (hereafter referred to as deep earthquakes), is less clear. Traditionally, deep earthquakes are considered to be distinct from those at shallower depths, based on an exponential decay with depth of the number frequency for shallow and intermediate earthquakes (e.g., Frohlich, 2006; Houston, 2007), with the rest representing deep earthquakes (Fig. 1A). Several unique characters of deep earthquakes include their maximum occurrence rate at ~600 km (Frohlich, 2006), the dominant down-dip compressional stress (Fig. 1B) inferred from global focal mechanism studies (Isacks and Molnar, 1971; Engdahl et al., 1998), and their apparent difference in distribution of foci depth in cold (western Pacific) versus warm (South America) slabs (Hayes et al., 2012).

Proposed mechanical models for stresses that generate deep earthquakes largely fall into two categories: (1) buoyancy force associated with thermal (Isacks and Molnar, 1971; Alisic et al., 2010; Myhill, 2013) and/or compositional anomalies (Bina, 1997), and (2) volume reductions from phase transformations (Devaux et al., 2000; Guest and Schubert, 2003). The magnitude of stresses resulting from buoyancy mechanisms is usually no more than 100 MPa (Alisic et al., 2010), reflecting the limited buoyancy due to either a low-density metastable olivine wedge (e.g., Bina, 1997) or a lower-than-ambient slab temperature (e.g., Alisic et al., 2010). Another reason buoyancy-induced stresses have limited amplitudes reflects their counter-action of the downward directed slab pull force, whose magnitude, when averaged over depth, is on the level of 200–400 MPa (e.g., Conrad and Lithgow-Bertelloni, 2002). In stark contrast, stresses caused by volumetric strain during phase transformations are usually at a level of 1 GPa (Devaux et al., 2000; Guest and Schubert, 2003), an order of magnitude higher than those from buoyancy-originated stresses.

Observed static stress drops of deep earthquakes sometimes reach >100 MPa (Kanamori et al., 1998); however, these likely represent a lower-limit of the actual background slab stress. Deep earthquakes usually have lower seismic radiation efficiencies (as small as <0.1) compared to shallower events (Houston, 2007), suggesting that the radiated energy represents only a minor portion of the total energy released (Kanamori and Heaton, 2000). Furthermore, the high ambient pressure at EQ depths would likely result in large cohesion and frictional stress along fault planes (Brace and Kohlstedt, 1980).

* Corresponding author.
E-mail address: ljliu@illinois.edu (L.J. Liu).

http://dx.doi.org/10.1016/j.epsl.2015.03.053
0012-821X/

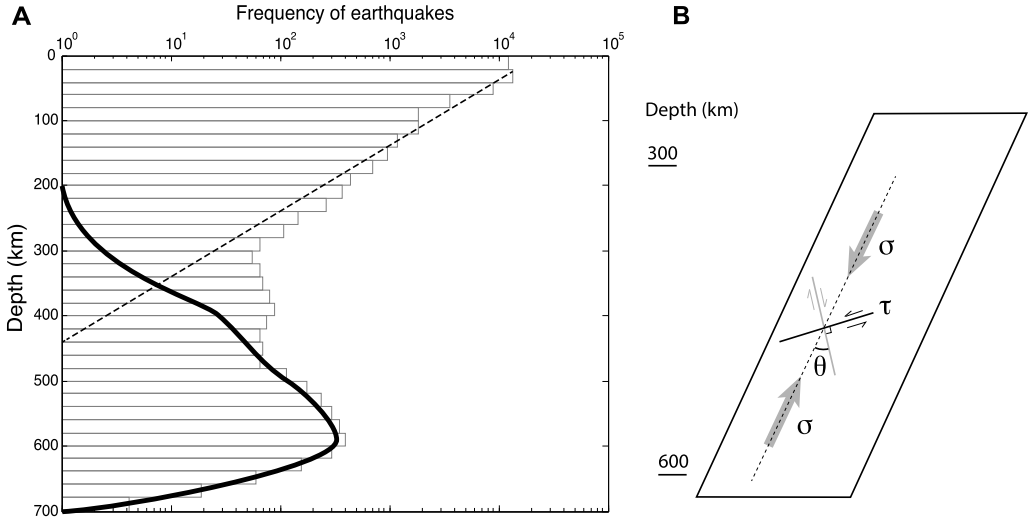

Fig. 1. A) Depth distribution of global seismicity. Earthquakes with magnitudes greater than 5 in the EHB catalog (Engdahl et al., 1998) for the period from 1964 to 2004 are plotted assuming a 20-km thick depth interval. Deep earthquake distribution (bold black line) is estimated by removing the shallow ones (black dashed line) that follow an empirical exponential decay with depth (Frohlich, 2006). B) Current understanding of typical deep earthquake fault geometry, with two possible fault plans orthogonal to each other. σ: maximum principle stress; τ: shear stress along fault plane.

The viscous nature of mantle deformation would probably maintain a significant level of background stress during an earthquake, even with melt present along the fault surface (Kanamori et al., 1998). Fig. 1B also shows that the relation between down-dip compression (σ) and the net earthquake-promoting stress along a fault plan (τ) is as below

$$\tau = \sigma(\sin\theta - \mu\cos\theta) \quad (1)$$

Here μ is viscous friction coefficient. If we assume $\theta = 45°$, Eq. (1) suggests that with a moderate value (e.g., 0.5) of μ, only ~30% of σ can convert to τ, the nominal value of static stress drop during an earthquake. In cases when θ is smaller (e.g., with the assistance of pre-existing weak zones, Warren et al., 2007), or μ is larger, or significant initial cohesion exists, the earthquake-generation efficiency of σ can be further reduced. Overall, a significantly larger background slab stress than observed static stress drops seems necessary.

Phase transformation induced slab stresses that may reach ~1 GPa do provide a larger background stress than observed static stress drops. However, the observed frequency–depth relationship (Fig. 1A) remains unexplained in previous phase transformation based models. In a model that only considers equilibrium phase changes (Guest and Schubert, 2003), the predicted high stresses localize at interfaces between major phase boundaries such as 410 km and 520 km. Models involving a metastable olivine wedge (e.g., Devaux et al., 2000) suggest a continuous and flat spectrum of stress distribution along the slab interior. Therefore, these models do not provide a consistent explanation for the peak occurrence of earthquakes at ~600 km (Fig. 1A). Furthermore, the discontinuous Benioff zone with deep earthquakes clustering around 600 km depth in warm slabs such as those beneath South America also needs explanation (Wiens, 2001; Houston, 2007).

In this paper, we propose that the slab stress state is controlled more by its entire subduction history than only the present-day snapshot. We present a new model for generating large internal slab stresses, which takes into account all major phase transitions within both the crust and mantle lithosphere occurring during subduction. This model can potentially help to explain most characteristics of deep earthquakes discussed above, as well as to possibly reconcile several geodynamic paradoxes on the strength of slabs and mantle rocks, as well as the competition between earthquake-generating stress and subduction-facilitating slab-pull force.

2. Phase-change induced volume contractions within a subducting slab

In this section, we consider the volume contractions due to major phase transformations inside a slab that consists of a crust and mantle lithosphere. In particular, we consider the accumulative mechanical effects of these volume contractions along the subducting slab since initial subduction. For simplicity, we first consider an idealized case, in which we treat the crust and mantle lithosphere as evolving independently while being subject to the various volume contractions.

Following the subduction trajectory of a slab, the basaltic crust (MORB) experiences several phase transformations (Aoki and Takahashi, 2004; Xu et al., 2008), as shown in Fig. 2A. First, a diffuse basalt–eclogite transformation occurs between 30 and 100 km depth, followed by another more diffuse transformation between 250 and 450 km where majorite garnet forms (remaining stable down to ~700 km). In addition, the sharp coesite–stishovite transformation occurs at 300 km (Aoki and Takahashi, 2004). Net volume changes due to these transformations are −15%, −5% and −4%, respectively.

The mantle lithosphere (harzburgite) also experiences a sequence of phase changes during subduction (Xu et al., 2008; Irifune and Ringwood, 1987). The major boundaries (Fig. 2A) include orthopyroxene dissolving into majorite at 200–410 km with a −3% volume change, the olivine–wadsleyite transition around 410 km inducing −5% volume change, the wadsleyite–ringwoodite transition at 520 km causing −1.5% volume change, and finally, the formation of perovskite and ferropericlase below 660 km causes another −9% change. In Fig. 2A, phase transformation related unit-volume (v) reductions within these two lithosphere layers (crust + mantle) are shown as sharp changes superimposed on the overall volume decreases due to adiabatic compression.

The difference in contraction rates within the two lithospheric layers can be quantitatively illustrated by calculating their respective total subducted volume and length accompanying subduction. Fig. 2B displays the integrated volume (V) and length (L) of each lithosphere layer as a function of depth, normalized by their volume and length at surface conditions (without adiabatic compres-

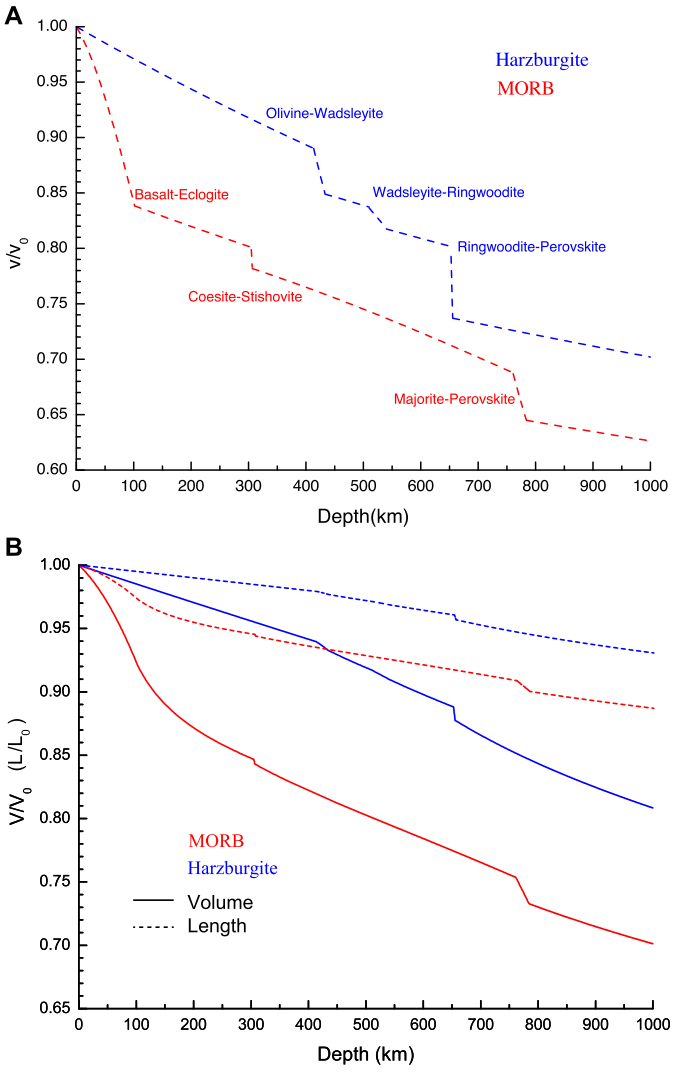

Fig. 2. A) Unit volume (v) reduction due to adiabatic compression with depth within the subducting oceanic crust (MORB, red) and mantle lithosphere (harzburgite, blue), superimposed by volume changes from various phase transformations as annotated. B) Accumulative volume (V) and length (L) reduction (relative their original state denoted by subscript 0) with depth in a vertical slab. Solid lines represent volume and dashed lines represent length. Red indicates oceanic crust (MORB) and blue mantle lithosphere (harzburgite). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sion and phase transformations) since the time subduction started. We assume an isotropic strain partitioning within all three directions when converting volume to length (this is not necessarily true when the slab is mechanically weak, as discussed later). Fig. 2B shows that the total volume reductions of the crust and harzburgite mantle from initiation of subduction to 660 km depth are ~25% and ~13%, respectively. Correspondingly, the total length shrinks by ~8% and ~4%, respectively. The net effect of the differential contractions is to gradually separate the crust and mantle that were adjacent when subduction started. In other words, subduction leads to an increasing differential displacement (therefore, strain) between the crust and mantle with depth (Fig. 3A). To translate the length ratios (Fig. 2B) to differential displacement in (Fig. 3A), one needs to first multiply these ratios with their respective original lengths to get the deformed (actual) length, and then take the difference between the two (crust and mantle) deformed lengths. Fig. 3A also considers a dipping slab at an angle of 60°, where the overall greater slab length within the upper mantle causes a larger differential displacement between the crust and mantle.

3. Slab internal stress originating from differential contractions

First we emphasize that adiabatic volume reduction within a homogeneous slab would not result in internal stress accumulation (Fig. 3B). If, in the above calculation of differential contraction, the two slab layers (crust and mantle) remain fully decoupled, there would not be any internal stress either. In reality, however, the cohesion (or strength) of mantle rocks within a cold slab would prevent this differential displacement to occur, which results in an internal stress distributed across the slab (Fig. 3B). Intuitively, one can imagine that this stress is trying to extend the crust but compress the mantle lithosphere along their interface, with the maximum principal stress parallel to the slab down-dip direction (Fig. 4). In reality, both the magnitude and spatial distribution of stress will depend on several factors including rheology of the slab, temperature profile, and the history of slab internal deformation.

3.1. Internal strain accumulation during subduction

Similar to earlier studies that estimated slab stress distribution (e.g., Guest and Schubert, 2003), we start with calculating the internal strain (ultimately strain rate) and temperature profile of the slab. The first-order pattern of slab internal strain can be illustrated with an end-member model by assuming a purely elastic slab, where an analytical solution exists. In this case, the internally stressed slab is similar to a bimetallic strip where different thermal expansion coefficients of the two metal layers result in an internal contraction upon a temperature change. The analytical solution to this classic engineering problem (Zhang and Chen, 2008) locates the zero-strain (therefore zero-stress) axis at 2/3 of the substrate (i.e., harzburgite mantle) thickness, where the substrate above this axis is compressed and the part below extended (strain distribution shown as dashed lines in Fig. 4A; the strain profiles also account for the fact that stresses vanish on both upper and lower slab surfaces). In effect, the apparent shear deformation between the crust and mantle layer generates dominant down-dip compression within the upper 2/3 of the mantle lithosphere, consistent with the focal mechanism and likely location of deep earthquakes (e.g., Houston, 2007). The difference in a subducting slab from a bimetallic strip is that internal contraction (therefore strain) grows with depth in the former (Figs. 2–4) but remains constant along strip in the latter.

In the purely elastic regime, the amplitude of slab internal strain will grow with depth due to increasing differential displacement (Figs. 2, 3). According to the analytical solution (Zhang and Chen, 2008), the maximum shear strain occurs close to the top surface of the slab and its amplitude can be approximated by the amount of differential displacement (Fig. 3A) divided by 2/3 of the slab thickness (e.g., $2/3 \times 90$ km $= 60$ km; Fig. 4A). This gives a strain of 8.3% at 200 km depth (5 km/60 km) and ~40% at 600 km depth (25 km/60 km). Assuming a shear modulus of 100 GPa (Cammarano et al., 2003), these strains convert to a stress of 8.3 and 40 GPa at 200 and 600 km, respectively. Apparently, such high stresses cannot be elastically sustained in mantle rocks, because mantle rocks yield at shear stresses that are at least one order of magnitude smaller (Brace and Kohlstedt, 1980; Kawazoe et al., 2009; Fig. 5). Therefore, the accumulative strain between the crust and mantle should be plastically dissipated during the subduction process (Fig. 4B). However, the above elastic analysis provides an estimate on the spatial distribution and total amount of potential energy (i.e., strain) accumulated during subduction. This strain profile is also useful for estimating strain rates, as discussed later.

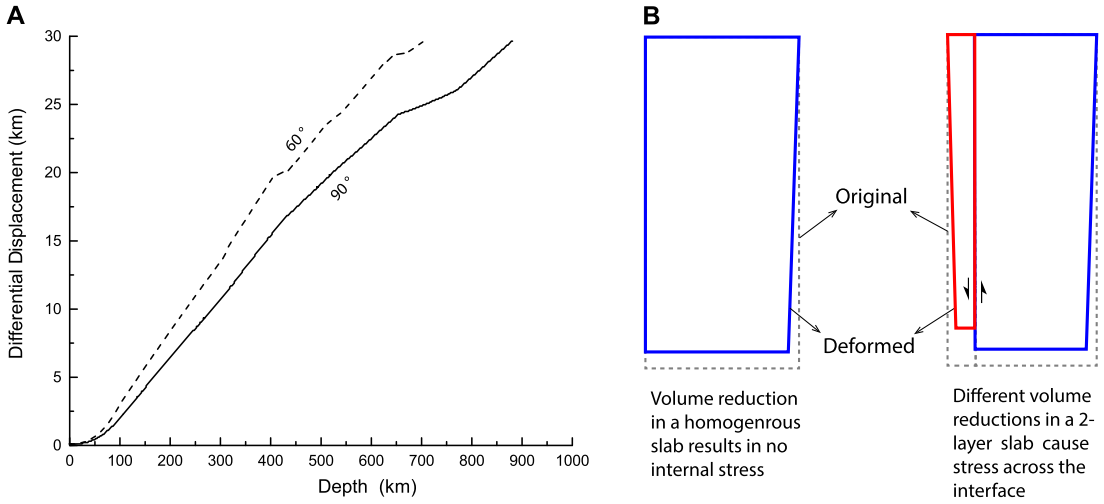

Fig. 3. A) Accumulative differential displacements between the crust and mantle lithosphere over depth for two slabs subducting at different dip angles. The one with a smaller dip (e.g., 60°) accumulates more differential displacement over depth than the one with a larger dip (e.g., 90°). B) Sketch of slab volume reduction due to adiabatic compression during subduction. This process leads to no internal stress accumulation in a homogeneous one-layer slab but will generate depth-accumulating internal stress within a two-layer slab.

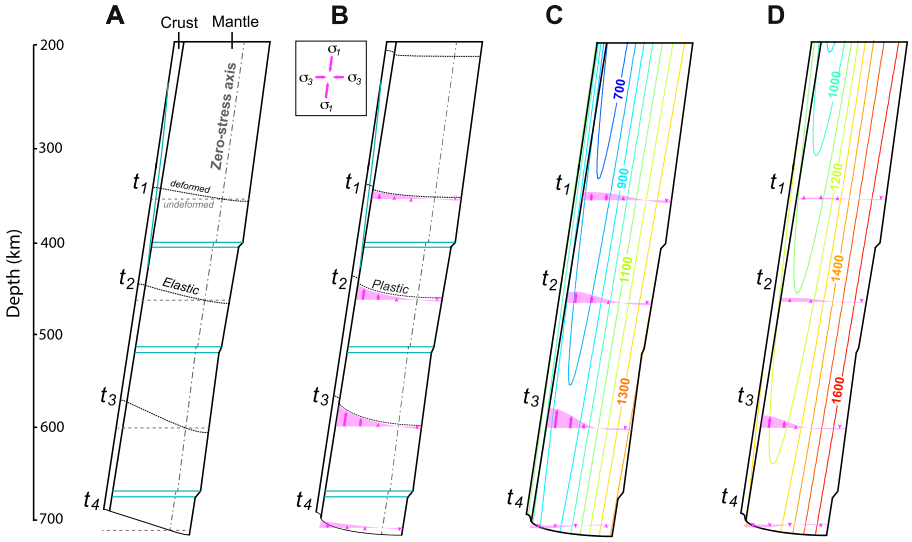

Fig. 4. A) Strain distribution within a purely elastic slab, where internal deformation is shown with two dashed lines. The strain distribution follows an analytical solution (see main text) for the slab interior and also assumes both upper and lower slab surfaces are stress free. Phase transformation interfaces are shown as blue lines. The gray dot-dashed line represents the zero-stress axis. B) Same as A, except with a different strain profile due to plastic deformation. The magenta area with arrows illustrates the down-dip compressional stress within the harzburgite mantle. C) Temperature distribution inside a 100-Ma slab subducting at 10 cm/yr. Annotations of the geotherms are for Kelvin (K). The colder slab core may concentrate more stress compared to B) where a uniform temperature is assumed. D) Temperature and stress distribution inside a 40-Ma slab subducting at 5 cm/yr. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Laboratory experiments suggest that at upper mantle depths rocks deform via diffusion and dislocation creep (i.e., the power-law rheology) when temperature is close to the ambient mantle condition (1600 K in uppermost mantle to 1800 K at the base of the transition zone) (Hirth and Kohlstedt, 2003; Kawazoe et al., 2009). At lower temperatures (e.g., >300 K lower than ambient), mantle rocks start to deform plastically following the Peierls regime (Raterron et al., 2004; Kawazoe et al., 2009, 2010). Although the transitional temperature between the power-law and Peierls regimes is not well defined, the mechanical behaviors of these flow laws are in fact similar within this temperature range (e.g., Kawazoe et al., 2010).

3.2. Slab thermal states during subduction

Since the temperature of mantle rocks largely determines its rheological behavior (e.g., Hirth and Kohlstedt, 2003; Kawazoe et al., 2009), we need to quantitatively calculate the temperature profile within a subducting slab, while considering different slab ages and subduction speeds. In practice, we solve the equation of conservation of energy inside a compressible mantle (Supplementary text). During subduction, the slab gradually cools down the surrounding mantle due to thermal diffusion, at the cost of raising slab internal temperature (Fig. S1).

We first consider a 100 Ma old slab subducting at a speed of 10 cm/yr, an analog of western Pacific subduction. The initial (before subduction) slab temperature profile follows the half-space cooling model. The evolution of slab temperature over depth is due to thermal diffusion and adiabatic heating. In this case, both the upper (top of the crust) and bottom sides of the slab are diffusion boundaries instead of a boundary with a fixed mantle temperature. Thus calculated slab temperature profile is strongly asymmetric with the crustal side significantly cooler than the bottom side of the slab, as is the case even at 700 km depth (Fig. S1A–B).

A similar conclusion is found for a younger (40 Ma) slab with a slower (5 cm/yr) subduction rate, although with an overall higher temperature distribution (Fig. 4C, Fig. S1C). These results are consistent with previous calculations of slab temperature profile (e.g., Peacock, 2002).

For the old and fast subducting slab (Fig. 4B), temperature along the upper side of the slab remains as cool as 1100 K down to 700 km depth. This suggests that the above mentioned crust–mantle deformation is well within the Peierls regime (Kawazoe et al., 2009). In young and slow subducting slabs (Fig. 4C), the upper side of the slab can reach ~1300 K at depths below 400 km. Additional consideration of the increase of thermal conductivity with depth (Xu et al., 2004; Tosi et al., 2010) will increase slab temperature by ~50 K (Fig. S1D). In all these calculations, we neglect the effect of viscous heating and latent heat due to phase transformations, which may further elevate slab temperature at transition zone depths by another 50 K (Guest and Schubert, 2003). In reality, the total length of subducting slabs should be slightly longer, due to curving at both the top and bottom of the upper mantle; the increase of total slab length (therefore diffusion) will also slightly elevate slab temperature, especially for young slabs. All together, we suggest that average temperature along the upper portion (including the crust) of the slab at the base of upper mantle may range at 1100–1200 K and 1400–1500 K for the old and young slabs, respectively.

3.3. Stress distribution within subducting slabs

As discussed in Section 3.1, the slab cannot elastically carry all the internal deformation during subduction. A significant portion of this strain (or potential energy) has to be released via plastic flow as subduction continues. According to the experimentally calibrated Peierls' flow regime (Kawazoe et al., 2009, 2010), one needs to know strain rates and temperature at a given ambient pressure (depth) in order to estimate the stress magnitudes (Fig. 5). Furthermore, the formulation of the Peierls rheology naturally evolves to the power-law rheology at higher temperatures, so we don't need to explicitly distinguish the difference between the two rheological regimes. Since Section 3.2 discussed slab temperature profiles, here we will focus on the estimate of internal strain rates. To do this, we have to understand the rheological difference of different minerals, and the deformation history of the slab during subduction.

At shallow depths, both the CPX-rich crust and the presence of water inside the upper half slab thickness above ~200 km depth (Muramotoa et al., 2011) will result in an overall weak slab (Zhang and Green, 2007). Therefore, the increasing differential contraction between the crust and mantle due to the basalt–eclogite transformation would likely be locally dissipated, preserving little internal strain (and stress) within the harzburgite slab (Fig. 4B). As the slab subducts deeper, both dehydration reactions and the CPX-garnet transformation will strengthen the slab, especially the crustal layer. Eventually (at 300–400 km), the increasing amount of garnet in the crust will make it stiffer than the mantle lithosphere (Zhang and Green, 2007). As a consequence, the differential displacement that grows with depth would result in more strain accumulated inside the mantle lithosphere, building up internal stress (Fig. 4B). To calculate the lithosphere stress, we should in theory separately consider olivine and wadsleyite across the 410 km phase transformation, where the latter displays a slightly stronger rheology (by a factor of two) than the former (Kawazoe et al., 2010). But since the wadsleyite rheology is also influenced (likely weakened) by grain sizes and hydrogen content whose effects are less certain (Karato, 2007), we take the olivine rheology as a proxy for the harzburgite mantle lithosphere in the entire upper mantle (Fig. 5).

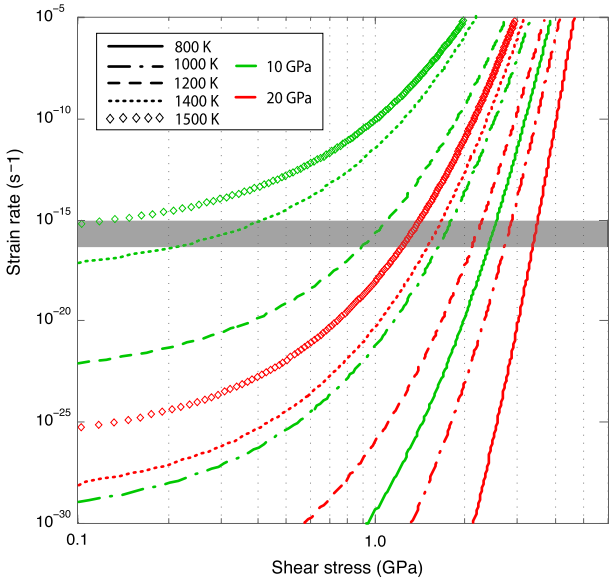

Fig. 5. Experimentally calibrated Peierls rheology, showing the stress and strain rate relationship at two different pressures and five different temperatures, relevant for a subducting slab within the upper mantle (see Fig. 4). The gray shaded region marks the plausible strain rate values present in the upper half thickness of a subducting slab. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

To convert the slab strain distribution with depth into strain rates is a challenging task, since it is hard to accurately track the temporal evolution of slab internal deformation due to uncertainties in rheology. However, we can estimate the magnitude of strain rates from the total amount of strain (i.e., differential displacement, Figs. 3A, 4A) and the stress–strain rate relationship (Fig. 5). On the one hand, an upper bound of strain rate value could be estimated by instantaneously releasing all the strain accumulated during early subduction (more details later). On the other hand, we realize that some residual strain must remain during subduction. This is because strain release is a self-regulating process: Initially subduction accumulates internal strain, which leads to increasing stress; this stress then increases strain rate, releasing more strain and reducing the magnitude of stress; this reduced stress then slows down strain rate and speeds up strain accumulation. Eventually, there is a balance between the remaining strain and the ongoing strain rate. In effect, estimating this remaining strain provides a lower bound on the value of instantaneous strain rate.

In practice, an upper bound of strain rate can be approximated by first measuring the rate of differential displacement accumulation with depth (Fig. 6A), which is then normalized by 2/3 of the slab thickness (Fig. 4A). If we further assume the total slab thickness remains unchanged (~90 km) with depth, the strain rate should follow the gradient of the differential displacement curve (Fig. 6A). For a subduction rate of 10 cm/yr, the differential displacement increases by about 5 km per million year (~100 km slab length). And this suggests a maximum (close the upper side of the slab) internal strain rate of $\sim 4 \times 10^{-15}$ s^{-1} above 400 km depth, and slightly decreases (by 20%) below this depth. Because this strain rate represents the total deformation within the crust and mantle, the effective strain rate occurring within the mantle lithosphere would be further lowered by a small amount, but should still be no less than 10^{-15} s^{-1}, given that the crust is likely stronger than the mantle below 400 km depth. Therefore, a plausible upper bound of strain rates across the upper half thickness of the slab (Fig. 4B) is close to 10^{-15} s^{-1}. However, the maximum strain rate would decrease sharply toward the bottom of the upper mantle due to the significant slowdown of differential displacement growth below 660 km (Fig. 6A, case *1*). Because the

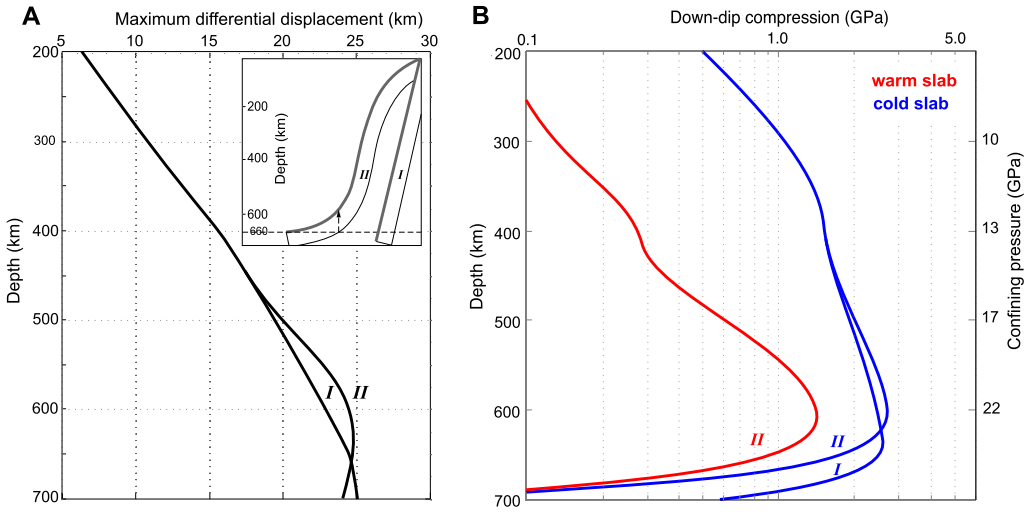

Fig. 6. A) Depth distributions of maximum differential displacements between the crust and mantle lithosphere. Two subduction scenarios with different slab geometry (I and II) are considered. The upward arrow marks the turning point of strain accumulation to strain ablation. B) Slab stress envelopes calculated from Fig. 5 and Fig. 6A. Both a cold slab (blue, Fig. 4B) and a warm slab (red, Fig. 4C) are considered. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

slab tends to lie flat on top of the lower mantle due to increased resistance from below (Fig. 6A, case *II*), the growth rate of the differential displacement would drop sharply as the greater volume of harzburgite sinks below 660 km causing significant volume reductions, canceling the earlier accumulated strain. This turning point in strain accumulation is around 600 km, given the geometry of the slab. As a result, the down-dip compression below 600 km would decrease rapidly and may eventually switch to down-dip extension (Figs. 4, 6B).

The lower bound of strain rate values can be subsequently estimated using the slab strain rates obtained above, which require a pre-existing stress level of ~1 GPa for most temperature and pressure conditions (Fig. 5). In order to elastically maintain this stress, a 1% strain (again assuming a shear modulus of 100 GPa) at any time is needed. Fig. 4 suggests that this strain requires 0.6 km (normalized by 60 km slab thickness) of differential displacement sustained for 1 million years. This is only a minor portion (~10%) of the total accumulated differential displacement during the same time period of subduction. Holding this strain from dissipation only affects the estimated strain rate by a small amount (10%), well within the uncertainties of absolute strain rates. Importantly, this uncertainty will not notably affect the stress magnitude, because of its weak dependence on strain rates (Fig. 5). Therefore, we could take the value of 10^{-15} s^{-1} to represent the typical internal strain rate during subduction. However, the sudden volume reduction of the harzburgite mantle at 410 km (Fig. 6A) would temporally slow down stress accumulation with depth (Fig. 6B).

From these estimated strain rate values and slab temperature profiles (Fig. 4), we can further estimate the amount of internal stress release during subduction. Fig. 5 illustrates the stress–strain rate relation for the Peierls' rheology (Kawazoe et al., 2009). The gray shaded region (Fig. 5) locates the corresponding stress magnitude for estimated slab internal strain rates and temperatures at two given depths (pressures). It is clear that stress increases with pressure: from 10 to 20 GPa at a constant strain rate, stress will increase by a factor of 2 to 10, depending on temperature. At 300 km depth (~10 GPa, green curves), for a given temperature, the estimated strain rates (10^{-15}–10^{-16} s^{-1}) correspond to stresses whose magnitudes vary by a factor of two or less. However, at the same depth, for a given strain rate, stresses from the plausible upper slab temperature values (1100–1500 K) vary by more than one order of magnitude. In contrast, at ~530 km depth (20 GPa, red curves), the effects of both strain rates and temperature are restrained to a much smaller extent and the estimated stress is always above 1 GPa.

Thus estimated stress magnitude along the down-dip direction in both a cold and a warm slab is shown in Fig. 6B. For both cases, the maximum stress occurs around 600 km depth, due to the dominant effect of pressure-dependence. In a cold slab (Fig. 4B), the maximum stress of ~2.5 GPa decreases slowly toward shallower depth, and is still ~1 GPa at 300 km. In contrast, in a warm slab (Fig. 4C), the maximum stress of >1 GPa at 600 km decreases more rapidly to <200 MPa at 300 km depth. Significantly, the difference in these stress envelopes could potentially explain the continuous depth-distribution of deep earthquake in western Pacific vs. the sporadic distribution in South America with earthquakes clustering at ~600 km. Laterally, the strong effect of temperature, especially at shallower depth (Fig. 5), should concentrate stresses toward the cooler slab interior (Fig. 4C, D) relative to a slab with a uniform temperature (Fig. 4B). Fig. 5 also suggests that, at depths >500 km, large residual stress (~1 GPa) could be retained over geological times as long as the slab remains cool enough (<1400 K), because strain rates necessary to maintain this stress can be as small as 10^{-20} s^{-1}, allowing internal strain to survive for millions of years.

4. Discussion

4.1. Model validation

The physical validity of our model may be further verified by comparing with independent calculations. First, slab viscosity provides one test: taking a nominal value of 10^{24} Pa s (Karato et al., 2001; Billen, 2008; Alisic et al., 2010), the strain rate induced by differential contraction (~10^{-15} s^{-1}) suggests a stress of ~1 GPa, similar to the stress level calculated in Fig. 6. Second, models considering either equilibrium phase changes (Guest and Schubert, 2003) or a metastable olivine wedge (Devaux et al., 2000) both arrived at a peak stress of 1–2 GPa, again consistent with our estimated stress magnitude. We therefore suggest that our estimation of slab stress magnitude is not off scale. The fact that this model can potentially explain more observations than earlier ones suggests this is a plausible mechanical model for generating deep earthquakes.

One more way to further test our proposed model is to closely examine focal mechanisms at a typical deep earthquake region

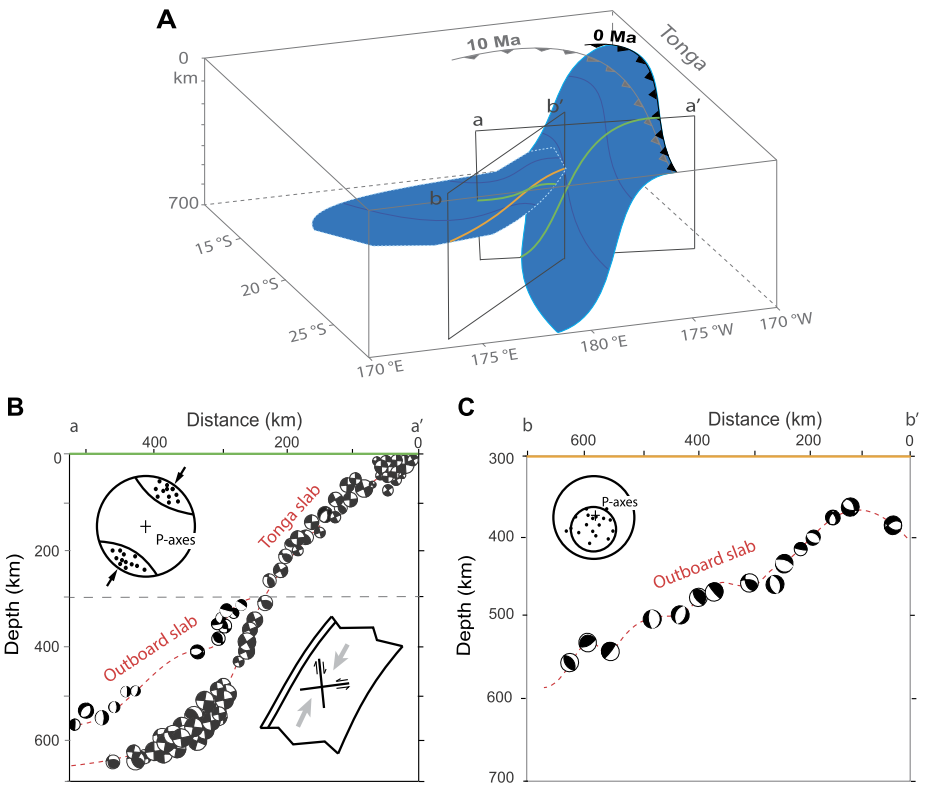

Fig. 7. A) Geometry of two separate slabs to the west of the Tonga trench (based on Chen and Brudzinski, 2001), where the blue surfaces represent slabs, dark blue lines within slabs mark material subducted at the same time, and black boxes with yellow (or orange) lines locate slab cross sections shown in B & C. The outboard slab likely results from the E–W oriented subduction zone at 10 Ma at northern end of Tonga. B) Earthquake focal mechanisms within the two slab cross sections along profile a–a′, assuming a northern hemisphere projection (Chen and Brudzinski, 2001). The circular inset in the upper left shows the P-axis orientation for earthquakes deeper than 300 km inside the active subducting Tonga slab. The graph in the lower right illustrates the stress and faulting schemes. The red dashed lines delineate slab geometry. C) The left shows the earthquake focal mechanisms along b–b′ within the outboard slab, also assuming a northern hemisphere projection (Brudzinski and Chen, 2005). The P-axis of earthquakes from the outboard slab in both B and C are largely north south oriented, consistent with the tectonically inferred down-dip direction in A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Fig. 7). We use a detailed compilation of focal mechanisms within the Tonga subduction zone (Chen and Brudzinski, 2001) where seismic studies delineate multiple slabs (Fig. 7A). Earthquakes along the actively subducting Tonga slab define a continuous Benioff zone from the base of the lithosphere to the upper-lower mantle boundary (Fig. 7B). In this slab, both the down-dip compressional stress regime (from 200 to 650 km) and depth–frequency distribution of deep earthquakes are consistent with our mechanical model (Figs. 4, 6B). Similar focal mechanisms are also found along the Japan slab and are confirmed by detailed fault mapping using waveform modeling (Chen et al., 2014; Zhan et al., 2014).

The continuity of down-dip compression above 300 km depth within the Tonga slab (Fig. 7B) suggests that phase change induced internal contraction may also influence intermediate-depth earthquakes. In this case, the down-dip compressional stress competes with the slab-bending induced down-dip extensional stress (e.g., Alisic et al., 2010) at shallow depth. Along cold subducting slabs, the phase-change induced slab stress is still strong above 300 km (Fig. 6B), and may continue to generate down-dip compressional earthquakes (Fig. 7B). In contrast, the phase-change induced stress is weak at shallow depths within warm slabs (Fig. 6B), resulting in down-dip extension dominant earthquakes (e.g., Brudzinski and Chen, 2005).

4.2. Implications for outboard deep earthquakes

At Tonga, some deep earthquakes occurred outside the actively subducting slab (Fig. 7C). Formation of these earthquakes remains enigmatic and has been attributed to reflect the outboard slab being petrologically anomalous (Chen and Brudzinski, 2001). Plate reconstruction (Fig. 7A; Müller et al., 2008) suggests that this outboard slab is likely derived from a former subduction zone striking east west at the north end of Tonga since 10 Ma, and that the gradual conversion of this trench to a transform boundary has subsequently abandoned the former slab into the present-day transition zone. While focal mechanisms along the interior of this abandoned slab (Fig. 7B, C) significantly differ from those within the Tonga slab, they all demonstrate a north-south oriented P-axis (compression). This is in remarkable consistency with the tectonically inferred down-dip direction of this fossil slab (Fig. 7A).

Furthermore, these outboard earthquakes suggest that the internal strain accumulated during earlier subduction should have been preserved over geological times, confirming our earlier analysis (e.g., Fig. 5). In fact, the same mechanism could be used to explain enigmatic deep earthquakes occurring at other places on Earth. One such example is the solitary but repeating deep earthquakes (at 650 km) beneath Spain, the Mediterranean region (Buforn et al., 1991, 2011), where residual internal strain could have survived millions of years within the possibly oldest (>140 Ma) slabs on Earth that were subducted during the Cenozoic closure of the Tethys Ocean (Müller et al., 2008).

4.3. Reconciling geodynamic paradoxes

Finally, by attributing deep earthquakes to slab internal stress from phase transformations, we can potentially reconcile several outstanding paradoxes in geodynamics: (1) the competition of

earthquake-generating stress and the subduction-facilitating slab pull force, (2) ongoing debates on the strength of subducting slabs, and (3) discrepancy on mantle rheological strength inferred from laboratory experiments and geophysical inversions.

Models that generate large down-dip compressional stresses due to slab buoyancy (Bina, 1997; Alisic et al., 2010) also consume a significant portion of the slab pull force (Conrad and Lithgow-Bertelloni, 2002), and could lead to slowdown of subduction (Tetzlaff and Schmeling, 2000). This suggests a potential anti-correlation between the rate of deep seismicity and subduction velocity. However, most deep earthquakes occur along fast subducting slabs, such as Tonga. In our model, the phase-change induced slab stresses are internal forces, which, by definition, do not affect the net slab pull force. Our model also implies a positive relation between the frequency of deep seismicity and the speed of subduction, as generally observed.

Another set of debates are on the mechanical strength of lithosphere and subducting slabs, where large amount of discrepancies exists between theoretical predictions based on laboratory measurements of rock rheology and geophysical inversions of lithosphere/slab viscosity. On the one hand, subduction models using the power-law rheology produce very strong slabs (Billen, 2008; Jadamec and Billen, 2010; Stader et al., 2010; Rolf et al., 2012), but studies constrained by seismically imaged slab geometry seem to favor slabs that are much weaker (Stegman et al., 2006; Ribe et al., 2007; Liu and Stegman, 2011; Capitanio et al., 2011). On the other hand, laboratory experiments (e.g., Brace and Kohlstedt, 1980; Meade and Jeanloz, 1990; Karato et al., 2001) usually suggest rock yield strengths or slab viscosities that are orders of magnitude higher than inferences from geophysical inversions (Billen et al., 2003; Capitanio et al., 2011; Zhong and Watts, 2013).

If we assume deep earthquakes release slab internal stress through deformation along narrow weak zones or faults (e.g., Warren et al., 2007) inside the slab, these processes could efficiently weaken the slab at the macroscopic level. As a result, the subducting slab would be locally (slab segments bounded by faults) strong, but overall (across upper mantle depths) weak, allowing the slab to bend or buckle in response to external forces. This is also consistent with the geodynamic inference that oceanic plates largely lose their macroscopic-scale strength after entering the subduction zone (Billen et al., 2003).

5. Conclusion

We propose a new mechanical model (Fig. 8) that generates relatively large internal stresses within subducting slabs. Central to our model is the continuous accumulation and release of potential energy between the crust and mantle lithosphere due to differential volume reductions from phase transformations accompanying subduction. The model predicts consistent down-dip compression within the upper half thickness of the slab below 300 km depth, in agreement with global studies on deep earthquake focal mechanisms (Isacks and Molnar, 1971; Chen et al., 2004). A maximum internal stress of 1–2 GPa occurs at 600 km depth (Fig. 6B) in both cold and warm slabs, explaining the peak occurrence rate of deep earthquakes at the same depth (Fig. 1A). The model provides an explanation for the distinct depth-dependence of earthquake distribution observed in cold (e.g., western Pacific) vs. warm (e.g., South America) slabs; the model suggests that the former have a consistently high stress envelope below 200 km, while the latter reflects a solitary high stress peak at the base of the upper mantle. The capability of cold slabs to retain large residual stress at the base of the transition zone also provides a solution to enigmatic deep earthquakes occurring beneath former subduction zones such as western Tonga and Spain.

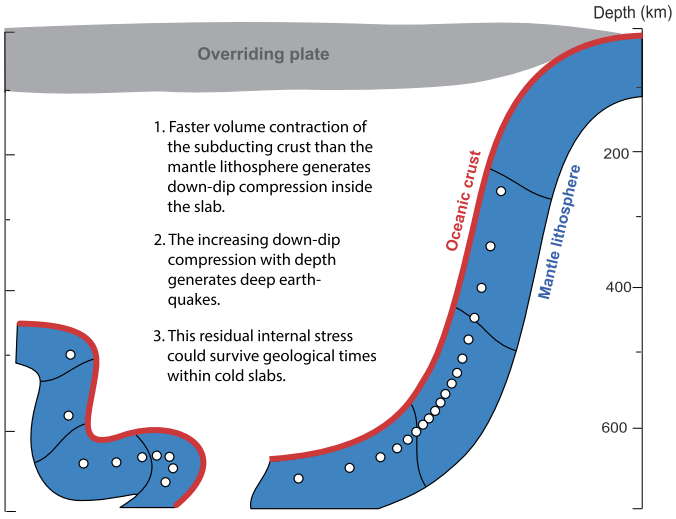

Fig. 8. Schematic representation of the proposed mechanical model for generating slab internal stress and deep earthquakes. Black lines track slab internal deformation during subduction. White dots indicate earthquakes.

Acknowledgements

We thank Xiaodong Song, Jay Bass, and Craig Lundstrom for helpful discussions during development of the model. L.L. was supported by NSF grant EAR-1345135. J.S.Z. was supported by NSF grant EAR-07-38871.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.epsl.2015.03.053.

References

Alisic, L., Gurnis, M., Stadler, G., Burstedde, C., Wilcox, L.C., Ghattas, O., 2010. Slab stress and strain rate as constraints on global mantle flow. Geophys. Res. Lett. 37, L22308.
Aoki, I., Takahashi, E., 2004. Density of MORB eclogite in the upper mantle. Phys. Earth Planet. Inter. 143–144, 129–143.
Billen, M.I., Gurnis, M., Simons, M., 2003. Multiscale dynamics of the Tonga–Kermadec subduction zone. Geophys. J. Int. 153, 359–388.
Billen, M.I., 2008. Modeling the dynamics of subducting slabs. Annu. Rev. Earth Planet. Sci. 36, 325–356.
Bina, C., 1997. Patterns of deep seismicity reflect buoyancy stresses due to phase transitions. Geophys. Res. Lett. 24, 3301–3304.
Brace, W.F., Kohlstedt, D.L., 1980. Limits on lithospheric stress imposed by laboratory experiments. J. Geophys. Res. 85, 6248–6252.
Brudzinski, M.R., Chen, W.-P., 2005. Earthquakes and strain in subhorizontal slabs. J. Geophys. Res. 110, B08303.
Buforn, E., Udias, A., Madariaga, R., 1991. Intermediate and deep earthquakes in Spain. Pure Appl. Geophys. 136, 375–392.
Buforn, E., Pro, C., Cesca, S., Udias, A., del Fresno, C., 2011. The 2010 Granada, Spain, deep earthquake. Bull. Seismol. Soc. Am. 101, 2418–2430.
Cammarano, F., Goes, S., Vacher, P., Giardini, D., 2003. Inferring upper-mantle temperatures from seismic velocities. Phys. Earth Planet. Inter. 138, 197–222.
Capitanio, F.A., Faccenna, C., Zlotnik, S., Stegman, D.R., 2011. Subduction dynamics and the origin of Andean orogeny and the Bolivian orocline. Nature 480, 83–86.
Chen, P.-F., Bina, C.R., Okal, E.A., 2004. A global survey of stress orientations in subducting slabs as revealed by intermediate-depth earthquakes. Geophys. J. Int. 159, 721–731.
Chen, W.-P., Brudzinski, M.R., 2001. Evidence for a large-scale remnant of subducted lithosphere beneath Fiji. Science 292, 2475–2479.
Chen, Y., Wen, L., Chen, J., 2014. A cascading failure during the 24 May 2013 great Okhotsk deep earthquake. J. Geophys. Res. 119, 3035–3049.
Conrad, C., Lithgow-Bertelloni, C., 2002. How mantle slabs drive plate tectonics. Science 298, 207–209.
Devaux, J.P., Fleitout, L., Schubert, G., Anderson, C., 2000. Stresses in a subducting slab in the presence of a metastable olivine wedge. J. Geophys. Res. 105, 13,365–13,373.

Engdahl, E.R., van der Hilst, R., Buland, R., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. Bull. Seismol. Soc. Am. 88, 722–743.
Frohlich, C., 2006. Deep Earthquakes. Cambridge University Press, Cambridge, UK.
Guest, A., Schubert, G., 2003. Stress field in the subducting lithosphere and comparison with deep earthquakes in Tonga. J. Geophys. Res. 108 (B6), 2288.
Hayes, G.P., Wald, D.J., Johnson, R.L., 2012. Slab1.0: a three-dimensional model of global subduction zone geometries. J. Geophys. Res. 117, B01302.
Hirth, G., Kohlstedt, D., 2003. Rheology of the upper mantle and the mantle wedge: a view from the experiments. In: Geophys. Monograph, vol. 138. AGU, pp. 83–105.
Houston, H., 2007. Deep earthquakes. In: Schubert, G. (Ed.), Treatise on Geophysics, vol. 4. Elsevier, pp. 321–350.
Irifune, T., Ringwood, A.E., 1987. Phase transformations in a harzburgite composition to 26 GPa: implications for dynamical behaviour of the subducting slab. Earth Planet. Sci. Lett. 86, 365–376.
Isacks, B., Molnar, P., 1971. Distribution of stress in the descending lithosphere from a global survey of focal-mechanism solutions of mantle earthquakes. Rev. Geophys. Space Phys. 9, 103–174.
Jadamec, M., Billen, M.I., 2010. Reconciling rapid 3-D mantle flow and surface plate motions near the eastern Alaska slab edge. Nature 465, 338–341.
Kanamori, H., Anderson, D.L., Heaton, T.H., 1998. Frictional melting during the rupture of the 1994 Bolivian earthquake. Science 279 (5352), 839–842.
Kanamori, H., Heaton, T., 2000. Microscopic and macroscopic physics of earthquakes. In: GeoComplexity and the Physics of Earthquakes. In: Geophys. Monograph. AGU, pp. 147–163.
Karato, S., Riedel, M., Yuen, D.A., 2001. Rheological structure and deformation of subducted slabs in the mantle transition zone: implications for mantle circulation and deep earthquakes. Phys. Earth Planet. Inter. 127, 83–108.
Karato, S., 2007. Microscopic models for the effects of hydrogen on physical and chemical properties of Earth materials. In: Yuen, D.A., Maruyama, S., Karato, S.-I., Windley, B.F. (Eds.), Superplumes: Beyond Plate Tectonics. Springer, Netherlands, pp. 321–356.
Kawazoe, T., Karato, S., Otsuka, K., Jing, Z., Mookherjee, M., 2009. Shear deformation of dry polycrystalline olivine under deep upper mantle conditions using a rotational Drickamer apparatus (RDA). Phys. Earth Planet. Inter. 174, 128–137.
Kawazoe, T., Karato, S., Ando, J., Jing, Z., Otsuka, K., Hustoft, J.W., 2010. Shear deformation of polycrystalline wadsleyite up to 2100 K at 14–17 GPa using a rotational Drickamer apparatus (RDA). J. Geophys. Res. 115, B08208.
Liu, L., Stegman, D., 2011. Segmentation of the Farallon slab. Earth Planet. Sci. Lett. 311, 1–10.
Meade, C., Jeanloz, R., 1990. The strength of mantle silicates at high pressures and room temperature: implications for the viscosity of the mantle. Nature 348, 533–535.
Müller, R.D., Sdrolias, M., Gaina, C., Roest, W.R., 2008. Age, spreading rates and spreading asymmetry of the world's ocean crust. Geochem. Geophys. Geosyst. 9, Q04006.
Muramotoa, M., Michibayashia, K., Andob, J., Kagic, H., 2011. Rheological contrast between garnet and clinopyroxene in the mantle wedge: an example from Higashi-akaishi peridotite mass, SW Japan. Phys. Earth Planet. Inter. 184, 14–33.
Myhill, R., 2013. Slab buckling and its effect on the distributions and focal mechanisms of deep-focus earthquakes. Geophys. J. Int., 1–17. http://dx.doi.org/10.1093/gji/ggs054.
Peacock, S.M., 2002. Thermal structure and metamorphic evolution of subducting slabs. In: Eiler, J., Abers, G.A. (Eds.), Geophys. Monograph, vol. 138. AGU, Washington, DC.
Raterron, P., Wu, Y., Weidner, D.J., Chen, J., 2004. Low-temperature olivine rheology at high pressure. Phys. Earth Planet. Inter. 145, 149–159.
Ribe, N.M., Stutzmann, E., Ren, Y., van der Hilst, R.D., 2007. Buckling instabilities of subducted lithosphere beneath the transition zone. Earth Planet. Sci. Lett. 254, 173–179.
Rolf, T., Coltice, N., Tackley, P.J., 2012. Linking continental drift, plate tectonics and the thermal state of the Earth's mantle. Earth Planet. Sci. Lett. 351–352, 134–146.
Stadler, G., Gurnis, M., Burstedde, C., Wilcox, L.C., Alisic, L., Ghattas, O., 2010. The dynamics of plate tectonics and mantle flow: from local to global scales. Science 329, 1033–1038.
Stegman, D.R., Freeman, J., Schellart, W.P., Moresi, L.-N., May, D.A., 2006. Influence of trench width on subduction hinge retreat rates in 3-D models of slab. Geochem. Geophys. Geosyst. 7, Q03012.
Tetzlaff, M., Schmeling, H., 2000. The influence of olivine metastability on deep subduction of oceanic lithosphere. Phys. Earth Planet. Inter. 120, 29–38.
Tosi, N., Yuen, D.A., Cadek, O., 2010. Dynamical consequences in the lower mantle with the post-perovskite phase change and strongly depth-dependent thermodynamic and transport properties. Earth Planet. Sci. Lett. 298, 229–243.
Warren, L.M., Hughes, A.N., Silver, P.G., 2007. Earthquake mechanics and deformation in the Tonga–Kermadec subduction zone from fault plane orientations of intermediate- and deep-focus earthquakes. J. Geophys. Res. 112, B05314.
Wiens, D.A., 2001. Seismological constraints on the mechanism of deep earthquakes: temperature dependence of deep earthquake source properties. Phys. Earth Planet. Inter. 127, 145–163.
Xu, W., Lithgow-Bertelloni, C., Stixrude, L., Ritsema, J., 2008. The effect of bulk composition and temperature on mantle seismic structure. Earth Planet. Sci. Lett. 275, 70–79.
Xu, Y., Shankland, T.J., Lihardt, S., Rubie, D.C., Langenhorst, F., Klasinski, K., 2004. Thermal diffusivity and conductivity of olivine, wadsleyite and ringwoodite to 20 GPa and 1373 K. Phys. Earth Planet. Inter. 143–144, 321–336.
Zhan, Z., Kanamori, H., Tsai, V., Helmberger, D., Wei, S., 2014. Rupture complexity of the 1994 Bolivia and 2013 Sea of Okhotsk deep earthquakes. Earth Planet. Sci. Lett. 385, 89–96.
Zhang, J., Green, H., 2007. Experimental investigation of eclogite rheology and its fabrics at high temperature and pressure. J. Metamorph. Geol. 25, 97–115.
Zhang, N., Chen, J., 2008. An alternative two-variable model for bending problems of multilayered beams. J. Appl. Mech. 75, 044503.
Zhong, S., Watts, A.B., 2013. Lithospheric deformation induced by loading of the Hawaiian Islands and its implications for mantle rheology. J. Geophys. Res. 118, 6025–6048.