Mantle flow in subduction systems: The subslab flow field and implications for mantle dynamics

Maureen D. Long\textsuperscript{1,2} and Paul G. Silver\textsuperscript{1,3}

Received 6 November 2008; revised 8 June 2009; accepted 7 July 2009; published 28 October 2009.

[1] The character of the mantle flow field in subduction zones remains poorly understood, despite its importance for our understanding of subduction dynamics. In particular, little attention has been paid to mantle flow beneath subducting slabs. In order to identify processes that make first-order contributions to the global pattern of subslab mantle flow, we have compiled shear wave splitting measurements from subduction zones worldwide from previously published studies and estimated average splitting parameters for the subslab region. Globally, the subslab region is overwhelmingly dominated by trench-parallel fast splitting directions. We tested for relationships between splitting delay time, a measure of the strength of anisotropy, and parameters that are indicators of tectonic processes, such as trench migration velocity, convergence velocity and obliquity, age of subducting lithosphere, and slab dip, curvature, seismicity, and thickness. We used several different plate motion models to describe plate and trench motions and evaluated the differences among the models. We find that only one parameter, namely, the trench migration velocity in a Pacific hot spot reference frame, appears to correlate well with the strength of the subslab splitting signal. This finding supports a model in which the mantle beneath subducting slabs is dominated by three-dimensional flow induced by trench migration. We explore several implications of our model for various aspects of mantle dynamics, including the choice of a suitable reference frame(s) for mantle flow, the existence of a thin decoupling zone between slabs and the subslab mantle, and consequences for mass transfer between the upper and lower mantle.


1. Introduction

[2] Subduction zones, where cold, negatively buoyant lithosphere is recycled into the Earth’s mantle, represent both key components of the plate tectonic system and prime sites of volcanic and seismological hazard [e.g., Stern, 2002]. Subduction zones represent perhaps the most tectonically complicated regions on the Earth, and although much progress has been made in understanding their structure and dynamics, several crucial aspects of the subduction process remain enigmatic. In particular, the dynamic interaction between sinking slabs and the surrounding mantle flow field remains poorly understood. The simplest model for this flow field invokes two-dimensional flow, with corner flow in the mantle wedge above the slab and entrained flow beneath it; however, this model is increasingly contradicted by seismological (and other) observations. The character of the subduction zone flow field and the nature of the forces that drive it have implications for our understanding of the formation and migration of melt, volcanic production and hazard, the tectonics of the fore-arc and back-arc region, and the style of mantle convection and mixing. In particular, the characterization of mantle flow beneath downgoing slabs, and the associated implications for how slabs interact with surrounding mantle material, has received little attention. In this study, we focus on anisotropy and mantle flow in the subslab region, for several reasons. First, despite its importance for a complete understanding of the mantle flow field associated with subduction, the subslab region has been paid comparatively little attention in studies of mantle flow in subduction systems, in contrast to the mantle wedge. Second, as discussed later in this paper, the relationship between seismic anisotropy and mantle flow is much more straightforward for the subslab region than for the wedge, where unusual olivine fabrics [Jung and Karato, 2001] or effects from melt [Holtzman et al., 2003] are possible and the interpretation of seismic anisotropy can be more ambiguous. Third, as we will show, observations of anisotropy in the subslab region are much simpler and exhibit much less regional variation than do comparable observations for the wedge region.

[3] The most direct constraints on the geometry of mantle flow are obtained through observations of seismic anisotropy. In particular, the observation that seismic shear waves...
are birefringent, and undergo “splitting” as they propagate through anisotropic regions of the Earth, has provided a tool for studying the local geometry of mantle flow beneath seismic stations (see overviews by Silver [1996], Savage [1999], and Park and Levin [2002]). Shear wave splitting studies, which measure the fast direction $\phi$ and the delay time $\delta t$ for each shear arrival, have increased in popularity and now represent a standard tool for probing mantle deformation. Subduction zones have been a popular target for shear wave splitting practitioners, and since early studies by Ando et al. [1983], Fukao [1984], and Bowman and Ando [1987], several dozen splitting studies in subduction zone regions have been performed. With the wealth of such data now available, it is both important and timely to undertake comprehensive studies of shear wave splitting behavior in subduction zones globally, with the aim of determining the dominant properties of the mantle flow field associated with the subduction process.

Although shear wave splitting in subduction zone regions has been extensively characterized, consensus on how it relates to mantle processes has not been forthcoming, for several reasons. First, the inherent structural complexity of subduction zones (which comprise the overriding plate, the superslab mantle wedge, the slab itself, and the subslab mantle) not only makes the interpretation of splitting measurements difficult, but it makes the measurements themselves challenging; complex anisotropic structure beneath a station leads to complicated splitting patterns which exhibit significant variations [e.g., Long and van der Hilst, 2005] and good data coverage is needed to fully characterize such complexity. Second, the wide variety of splitting behaviors observed in subduction zones has posed challenges for interpretation: large variations in both $\phi$ and $\delta t$ have been observed, significant lateral heterogeneity is often present, and the preponderance of fast directions that are parallel to the trench are contrary to the expectation of simple two-dimensional (2-D) entrained flow.

A promising approach to this problem lies in a systematic study of splitting constraints in subduction zone regions and a search for relationships between splitting parameters and parameters that describe subduction. We recently proposed a model [Long and Silver, 2008] that successfully explains many aspects of global subduction zone splitting patterns. This model, in which the subduction zone flow field is controlled by 3-D flow induced by the migration of the trench with respect to the underlying mantle, explains the preponderance of trench-parallel fast directions beneath the slab as well as the observed correlation between trench migration velocity and delay times for both the subslab and wedge regions that we observed in a global compilation of splitting measurements.

Here, we expand our previous work and present a more complete investigation of the relationships between splitting parameters and parameters that describe the subduction process, with a focus on the subslab mantle flow field. The work presented here extends our previous investigation in four ways. First, we have augmented our global database of subduction zone splitting measurements with characterizations of source-side anisotropy from teleseismic measurements of intermediate-depth earthquakes occurring in subduction zones [e.g., Wrokey et al., 2005; Rokosky et al., 2006; Müller et al., 2008]. These data provide a more direct constraint on anisotropy and flow beneath downgoing slabs, as the measurements are uncontaminated by anisotropy in the wedge. Second, we present comparisons between subslab splitting parameters and a wide variety of parameters that describe subduction, such as trench migration velocity, convergence velocity and obliquity, age and spreading history of subducting lithosphere, slab dip, curvature, seismicity, thickness, and morphology, and arc length. These comparisons allow us to test alternative hypotheses concerning factors that control anisotropy beneath the mantle wedge. Third, we revisit the comparisons between subslab split times and trench migration velocities shown by Long and Silver [2008] using a variety of global reference frames and plate motion models [Schellart et al., 2008; Lallemand et al., 2008]; in principle, this should allow for the identification of an ideal “rest frame” for the convecting mantle, at least in the Pacific, around which most subduction zones in our compilation are located. Finally, we address in detail some of the implications of the Long and Silver [2008] model for mantle dynamics, including the nature of a likely decoupling zone beneath subducting slabs and constraints on mass exchange between the upper and lower mantle.

2. A Global Data Set of Subslab Shear Wave Splitting Parameters in Subduction Zones

2.1. Methods for Isolating the Subwedge Contributions

Isolating the contribution to splitting observed at the surface from anisotropy beneath subducting slabs is not completely straightforward; here we describe the approach taken by Long and Silver [2008] and also applied in this study. Teleseismic shear phases such as SKS, SKKS, and direct S from deep events are commonly used in splitting studies to characterize upper mantle anisotropy beneath the station. In the case of a subduction zone, however, such phases sample several components of the subduction system, including the subslab mantle, the slab itself, the mantle wedge, and the overriding plate (Figure 1). In contrast, direct S waves from earthquakes originating in the subducting slab recorded on the overriding plate will only be affected by anisotropy in the mantle wedge and in the
overriding plate (and perhaps a small portion of the slab itself). If splitting from both local S phases and teleseismic S or SK(S) phases is measured, therefore, and if a careful correction for splitting in the wedge and overriding plate is applied to teleseismic measurements is applied, then an estimate may be obtained for the anisotropy beneath the mantle wedge. This is the approach we have taken to obtain estimates of the subwedge splitting signal; where possible, we have focused on stations located close to the trench, where teleseismic phases sample a large volume of subslab mantle material and little, if any, of the mantle wedge. This approach can isolate the contribution from the subwedge mantle, but an ambiguity remains as to the relative contributions to this signal from the downgoing slab itself versus the subslab mantle. As we discuss in section 2, several lines of evidence argue for a primary contribution to the subwedge splitting signal from active mantle flow beneath the slab; slab anisotropy does not appear to dominate subwedge splitting on a global scale. For now, however, we restrict our analysis to determining the subwedge splitting signal without distinguishing between anisotropy in the slab versus the subslab mantle.

[8] The correction of teleseismic splitting measurements to remove the effects of wedge anisotropy is only straightforward for the simplified case where the layers of anisotropy (in the wedge, the subslab mantle, and perhaps in the slab itself) are oriented either parallel or perpendicular to each other. For this geometric configuration, the delay times can simply be added (for layers with the same orientation) or subtracted (for layers with a perpendicular orientation). In practice, however, this correction is usually more difficult, for several reasons. First, splitting patterns in the mantle wedge are often very complicated [e.g., Pozgay et al., 2007; Wiens et al., 2008; Long and Silver, 2008] and the inferred anisotropy is often complex and changes on short length scales; in such cases, care is needed to choose appropriate local S raypaths to correct the corresponding teleseismic phases. Second, for the case where the measured fast directions for teleseismic and local are not either parallel or perpendicular to each other, correcting the teleseismic splitting for the contribution from wedge anisotropy is more complicated than simply subtracting (or adding) the wedge split time from the teleseismic split time. Ideally, in this case, the back azimuthal variations in teleseismic splitting should be carefully characterized and forward modeling of multiple-layer anisotropy [Silver and Savage, 1994] should be carried out to correctly identify the splitting parameters associated with the lower (subslab) anisotropic layer. Fortunately, however, in many cases a simple addition or subtraction of split times can be carried out, as measured fast directions for teleseismic phases at stations tend to be (mostly) trench-parallel or (in a few cases) trench-perpendicular, and many wedges display a characteristic transition from trench-parallel fast directions to trench-perpendicular fast directions as well.

[9] A third challenge to correctly accounting for the wedge contribution to teleseismic phases is a possible dependence of measured splitting parameters on the frequency content of the waves under study. Many studies have shown that shear wave splitting of local S phases in subduction zones can be frequency-dependent [Fouch and Fischer, 1996; Marson-Pidgeon and Savage, 1997; Long and van der Hilst, 2006; Wirth and Long, 2008] and since the frequency bands used in local S splitting studies often vary from study to study and are generally higher than those used in studies of teleseismic splitting, a direct comparison among studies is difficult. Despite these challenges, however, our method of accounting for contributions to teleseismic splitting from the mantle wedge and overriding plate and attributing the rest of the splitting signal to anisotropy in the subslab mantle and the slab itself has proven successful at providing a rough estimate of subslab splitting, although the errors on individual estimates may be large [Long and Silver, 2008].

[10] Estimates of source-side anisotropy from splitting measurements on teleseismic phases from earthquakes that originate in subducting slabs (Figure 1) provide another type of data that is useful for characterizing subwedge anisotropy [Russo and Silver, 1994]. In this approach, splitting due to anisotropy beneath the source, splitting due to upper mantle anisotropy near the receiver, and (in some cases) splitting due to anisotropy in the lowermost mantle are characterized separately; detailed descriptions of the methods used to isolate each contribution are given by Wookey et al. [2005], Rokosky et al. [2006], and Müller et al. [2008]. This type of data was not included in our original compilation; its inclusion here has allowed us to add new subduction zones (the Marianas and Scotia) to our compilation of subwedge anisotropy. In many ways, this method provides a more direct constraint on subwedge anisotropy, since the mantle wedge and overriding plate of the subduction system are not sampled by this type of measurement (although, of course, anisotropy on the receiver side and in the lowermost mantle must be properly accounted for). The characterization of source-side splitting from slab earthquakes, along with the use of unconventional raypath configurations such as the use of sS phases to study anisotropy in the wedge near the source [e.g., Anglin and Fouch, 2005], has the potential to add further to our global compilation in the future.

[11] A key consideration when assigning average delay times to the subwedge portion of the subduction system is accounting for potential differences in slab and subslab path lengths among different regions. These path lengths are controlled by the positions of the seismic stations under study relative to the trench and by the dip and thickness of the subducting slab. We have preferentially chosen stations located close to the trench in order to minimize these effects, but some differences in slab and subslab path lengths exist. Because the depth of the bottom of the anisotropic layer beneath subducting slabs is poorly known and may vary in different regions, it is difficult to explicitly correct subwedge delay times for the effect of path length. We do, however, estimate the range of subwedge path lengths associated with each range of delay time estimates, as discussed below.

2.2. Description of Regional Studies and Results

[12] Our compilation is based on over two dozen previously published studies and consists of 22 individual estimates for subwedge shear wave splitting in 15 different subduction zones, as described below. For subduction systems with long arc lengths where data are available for stations located at different locations along the arc,
such as South America or the Aleutians, we have made separate estimates of subwedge splitting (and other tectonic parameters, described in section 4) for individual segments of the arc. A map that summarizes our estimates for subwedge splitting for subduction zones worldwide is shown in Figure 2. Error bars are estimated by using the range of split times reported by authors of each study; the error bars are, necessarily, fairly approximate, but they are designed to encompass most of the range of subwedge delay times allowed by the data and roughly represent a 95% confidence region [see also Long and Silver, 2008]. The average subwedge delay time estimates, error bars, range of path lengths, and data sources are summarized in Table 1.

[13] Shear wave splitting in the northern part of the Tonga-Kermadec subduction zone has been studied extensively, although most studies have focused on measuring local S splitting [Fischer and Wiens, 1996; Fischer et al., 1998; Smith et al., 2001]. SKS splitting was measured at a few stations in northern Tonga by Fischer et al. [1998], but the stations were located fairly far away from the trench and the measurements likely sample the wedge much more than the subwedge mantle [see also Fischer et al., 2000]. For our compilation of subwedge anisotropy, therefore, we relied on measurements for permanent station RAO documented by Long and Silver [2008]. A comparison of SKS and local S splitting at this station reveals a trench-parallel fast direction and a delay time of 1.7 ± 0.7 s that can be attributed to anisotropy beneath the slab. Station RAO is located approximately 100 km above the subducting Tonga slab, so the subwedge upper mantle path lengths are roughly 300 km. Further to the south, shear wave splitting in the Hikurangi subduction zone off the east coast of the North Island of New Zealand has been extensively documented [e.g., Gledhill and Gubbins, 1996; Brisbourne et al., 1999; Matcham et al., 2000; Audoine et al., 2004; Marson-Pidgeon and Savage, 2004; Morley et al., 2006]. This region was not included in the original Long and Silver [2008] compilation because of its proximity to the strike-slip Alpine Fault and the associated complicated tectonic setting, but here we have chosen to use results from the northeastern part of the North Island, which is located closest to the Hikurangi trench and furthest from the Alpine Fault. Comparison of SKS splitting with local S measurements at low frequencies [Audoine et al., 2004; see also Greve et al., 2008] yields an estimate for subwedge splitting beneath Hikurangi of \( t = 1.5 \pm 0.5 \) s, with a trench-parallel fast direction. The northeasternmost stations used by Audoine et al. [2004] are located \( \sim 20–40 \) km above the Hikurangi slab, so subwedge upper mantle path lengths are roughly \( \sim 360 \) km.

[14] For the Sumatra/Indonesia subduction zone, we relied on measurements reported by Hammond et al. [2006], as well as measurements from permanent station PSI reported by Long and Silver [2008]. Hammond et al. [2006] reported a nearly isotropic wedge, with a small contribution to local S splitting from anisotropy in the overriding lithosphere, and significant SKS splitting with a fast direction oblique to the trench that they attribute to anisotropy within the slab. Our own local and teleseismic measurements at station PSI suggest that between 0.8 s and 1.4 s of subwedge splitting contributes to the teleseismic splitting signal, with little or no splitting due to wedge anisotropy. The subwedge fast splitting direction is close to trench-parallel. Station PSI is located \( \sim 200 \) km above the subducting Indonesia slab, so subwedge path lengths in the upper mantle beneath this station are \( \sim 200 \) km.

[15] The tightest constraints on subwedge anisotropy beneath the Mariana subduction zone come from source-side splitting for intermediate depth earthquakes, studied by Wookey et al. [2005]. This work yielded split times that range from 0.5 to 1.0 s for the Marianas, with consistently trench-parallel fast directions, but this represents measurements from only two earthquakes. These events originated at
Table 1. Summary of the Constraints on Subwedge Splitting

<table>
<thead>
<tr>
<th>Subduction Zone</th>
<th>Type of Data</th>
<th>Subwedge $\delta t$ (s)</th>
<th>Subwedge Path Lengths in Upper Mantle (km)</th>
<th>Number of Stations (Events)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonga-Kermadec</td>
<td>SKS, local S</td>
<td>~1.0–2.4 Trench-parallel</td>
<td>~300</td>
<td>1</td>
<td>Long and Silver [2008]</td>
</tr>
<tr>
<td>Hikurangi</td>
<td>SKS, local S</td>
<td>~1.0–2.0 Trench-parallel</td>
<td>~360</td>
<td>2</td>
<td>Adsivne et al. [2004]</td>
</tr>
<tr>
<td>Sumatra/Indonesia</td>
<td>SKS, local S</td>
<td>~0.8–1.4 Trench-parallel, oblique</td>
<td>~200$^a$</td>
<td>~20</td>
<td>Long and Silver [2008] and Hammond et al. [2006]</td>
</tr>
<tr>
<td>Izu-Bonin</td>
<td>SKS</td>
<td>~0.5–1.0 Trench-parallel, some oblique; splitting patterns complex</td>
<td>~300</td>
<td>4</td>
<td>Wirth and Long [2008]</td>
</tr>
<tr>
<td>Marianas</td>
<td>Source-side S</td>
<td>~0.5–1.0 Trench-parallel</td>
<td>~200</td>
<td>(2)</td>
<td>Wookey et al. [2005]</td>
</tr>
<tr>
<td>Ryukyu</td>
<td>SKS, local S</td>
<td>~0.0–0.2 Trench-parallel, if any</td>
<td>~280–350</td>
<td>8</td>
<td>Long and van der Hilst [2005, 2006], Nakajima and Hasegawa [2004], and Wirth and Long [2008]</td>
</tr>
<tr>
<td>Japan (N. Honshu, Hokkaido)</td>
<td>SKS, local S</td>
<td>~0.2–0.8 Oblique, some trench-parallel; splitting patterns complex</td>
<td>~200–320</td>
<td>~20</td>
<td>Long and van der Hilst [2005, 2006], Nakajima and Hasegawa [2004], and Wirth and Long [2008]</td>
</tr>
<tr>
<td>Kamchatka</td>
<td>SKS, local S</td>
<td>~0.4–0.8 Trench-parallel</td>
<td>~200–340</td>
<td>~5</td>
<td>Pustin et al. [2001] and Levin et al. [2004]</td>
</tr>
<tr>
<td>Aleutians</td>
<td>SKS, local S</td>
<td>~0.0–0.3 Trench-parallel, if any</td>
<td>~300</td>
<td>3</td>
<td>Long and Silver [2008]</td>
</tr>
<tr>
<td>Cascadia</td>
<td>SKS</td>
<td>~1.0–1.5 Trench-perpendicular</td>
<td>~300–350</td>
<td>~15</td>
<td>Currie et al. [2004]</td>
</tr>
<tr>
<td>Caribbean</td>
<td>SKS, limited local S</td>
<td>~1.0–2.0 Trench-parallel</td>
<td>~300</td>
<td>~7</td>
<td>Pitier- Feliciangeli and Kendall [2008]</td>
</tr>
<tr>
<td>Calabria</td>
<td>SKS</td>
<td>~1.0–2.0 Trench-parallel</td>
<td>~200–350</td>
<td>~30</td>
<td>Baccheschi et al. [2007]</td>
</tr>
<tr>
<td>Middle America</td>
<td>SKS, local S</td>
<td>~1.0–1.7 Trench-parallel</td>
<td>~250–360</td>
<td>~20</td>
<td>Abt et al. [2005, 2009]</td>
</tr>
<tr>
<td>South America</td>
<td>SKS, local S, source-side S</td>
<td>~0.5–1.5 Variable, but mostly trench-parallel</td>
<td>~0–200 (source-side), ~100–300 (SKS)</td>
<td>~20 (~10)</td>
<td>Rokosky et al. [2006] (source-side), Polet et al. [2000] (SKS, local S), Russo and Silver [1994] (all types)</td>
</tr>
<tr>
<td>Scotia</td>
<td>Source-side S</td>
<td>~1.0–1.6 Trench-parallel</td>
<td>~125–390</td>
<td>~20 (~10)</td>
<td>Müller [2001]</td>
</tr>
</tbody>
</table>

$^a$From the published literature, as described in section 2.2. We have listed the most relevant citations, but additional relevant studies are discussed in the text. For each subduction zone, the range of subwedge delay times and fast directions allowed by the data are listed, along with the approximate number of stations (or events for source-side splitting) used to obtain the estimate. We estimate the subwedge upper mantle path length by assuming that the depth of the bottom of the anisotropic layer is at ~400 km and taking into consideration the depth to the slab beneath the stations used (for SKS/local S studies) and/or the depth of the earthquakes (for source-side studies).

$^b$For PSI; large range for other stations.
depths of \( \sim 200 \) km in the slab, so the rays sample \( \sim 200 \) km of subslab upper mantle. The splitting of local S phases has been studied in the Marianas by Fouch and Fischer [1998], Volli et al. [2006], and Pozgay et al. [2007], but constraints on SKS splitting are very sparse [Fouch and Fischer, 1998], largely due to the unfavorable distribution of seismicity in the epicentral distance range used to study SKS. For the Izu-Bonin region, there are published constraints on mantle wedge anisotropy from Anglin and Fouch [2005], but few constraints on SKS. A preliminary investigation of SKS and local S splitting at F-net stations located in the northern part of the Izu-Bonin arc by Wirth and Long [2008] yielded generally trench-parallel SKS fast directions, but some SKS splitting patterns exhibit back azimuthal variations and the interpretation of the subwedge splitting signal requires more detailed forward modeling. The few local S measurements are from deep events [Wirth and Long, 2008] and are not well constrained; the local raypath distribution is not ideal for isolating the subslab contribution to anisotropy. We therefore assign \( \sim 0.5–1.0 \) s of trench-parallel subslab splitting to Izu-Bonin in our complication, but we emphasize that this is based on preliminary results.

[16] Shear wave splitting at stations in the Ryukyu arc was studied by Long and van der Hilst [2005, 2006], who carried out a comparison between teleseismic and local shear wave splitting in the same frequency bands and concluded that most of the teleseismic splitting signal can be attributed to the mantle wedge, with little or no splitting due to subwedge anisotropy required to explain the data. Further modeling work [Long et al., 2007; Kneller et al., 2008] has confirmed this interpretation; in particular, Kneller et al. [2008] found that a model with significant anisotropy in the mantle wedge due to B-type olivine fabric with a small contribution to trench-parallel splitting (most likely less than 0.2 s) from the slab or subslab mantle explains the data well. Therefore, in our compilation we attribute between 0 and 0.2 s of trench-parallel splitting to the subwedge mantle beneath Ryukyu. Ryukyu stations are located mainly on the arc islands and the subwedge raypaths lengths range from approximately 280–340 km.

[17] Both teleseismic and local splitting beneath Japan has been extensively studied [e.g., Fouch and Fischer, 1996; Sandvol and Ni, 1997; Fischer et al., 1998; Nakajima and Hasegawa, 2004; Long and van der Hilst, 2005; Nakajima et al., 2006; Wirth and Long, 2008] and the observed splitting patterns in Japan are perhaps the most complicated of any subduction zone in the world. This is particularly true at stations located near the Kanto region near the triple junction of the Pacific, Philippine Sea, and Eurasian plates, where the slab morphology (and presumably the resulting flow patterns) is complicated [e.g., Long and van der Hilst, 2005; Wirth and Long, 2008]. With the caveat that the causes for the complex splitting patterns in Japan are not yet well understood, we have focused on stations in northern Honshu and in Hokkaido to obtain a general estimate for subwedge splitting beneath Japan. Teleseismic splitting in this region is complex, although stations closest to the trench in both Honshu and Hokkaido exhibit clear trench-parallel fast directions and split times of up to \( \sim 0.8–1.0 \) s. Fast directions measured at other stations in the region are generally oblique to the trench. Local splitting for northern Honshu and Hokkaido was studied by Nakajima and Hasegawa [2004] and Nakajima et al. [2006], although these studies examined energy in much higher frequency bands than corresponding studies of teleseismic splitting, making direct comparisons difficult. These workers documented small split times (\( \sim 0.1 \) s) and a striking transition in fast direction from trench-parallel close to the trench to trench-perpendicular in the back arc, which they attribute to B-type olivine fabric in the corner of the mantle wedge. Wirth and Long [2008] recently measured local splitting in northern Honshu at frequencies comparable to those used to measure teleseismic splitting [Long and van der Hilst, 2005], and found split times of \( \sim 0.3 \) s. Given these observations, our best estimate for northern Japan is that \( 0.5 \pm 0.3 \) s of splitting can be attributed to the subwedge region, with fast directions that range from trench-parallel to oblique to the trench. Raypath lengths used in this estimate range from \( \sim 200 \) to 320 km.

[18] Further to the north, in Kamchatka, both teleseismic and local splitting have been studied using data from the Side Edge of the Kamchatka Slab experiment [Peyton et al., 2001; Levin et al., 2004]. Peyton et al. [2001] found significant ( \( 0.1–1 \) s) trench-parallel SKS splitting at stations in southern Kamchatka with much smaller splitting ( \( 0.1–0.3 \) s) from local S phases. Mantle wedge anisotropy from local S splitting was studied in greater detail by Levin et al. [2004]; they found slightly higher delay times ( \( 0.2–0.6 \) s) and complex spatial patterns, with a suggestion that fast directions exhibit a change from trench-perpendicular close to the trench to trench-parallel farther in the back arc (the opposite trend was observed in Japan by Nakajima and Hasegawa [2004]). Taken together, these studies suggest that between 0.4 and 0.8 s of trench-parallel splitting can be attributed to the subwedge mantle beneath Kamchatka. The stations used to obtain the SKS measurements of Peyton et al. [2001] were generally located between 60 and 200 km above the subducting slab, which roughly corresponds to subwedge upper mantle path lengths between \( \sim 200 \) and 340 km. Our constraints on subwedge anisotropy for the Aleutian subduction zone come from measurements made by Long and Silver [2008] for three permanent stations in the Aleutian island arc (ATKA, NIKO, and SMY). A comparison of local and teleseismic splitting at these stations suggests that there is little, if any, splitting from the subwedge region. SKS splitting measurements reveal \( \sim 0.5–1.0 \) s of (nearly) trench-parallel splitting, but a comparison with local S splitting suggests that the anisotropy being sampled by SKS waves at these stations is in the mantle wedge. Therefore, we estimate a contribution to splitting from the subwedge region of \( 0–0.3 \) s at Aleutian stations. ATKA, NIKO, and SMY are all located approximately 100 km above the subducting slab, which corresponds to \( \sim 300 \) km of subwedge upper mantle path length.

[19] The Cascadia subduction zone in the northwestern United States is characterized by a dearth of seismicity beneath \( \sim 80 \) km, so unlike most subduction zones worldwide, direct comparisons between local and teleseismic shear wave splitting are difficult. Shear wave splitting in the Pacific Northwest is becoming increasingly well characterized, thanks to new data from the USArray component of Earthscope and other temporary deployments in the region, but only a few studies have examined splitting for stations close to the trench that sample large volumes of
subwedge mantle. We have relied mainly on SKS measurements made by Currie et al. [2004]; they found that stations in the Cascadia fore-arc regions exhibit 1.0–1.5 s of splitting, with a fast direction that is roughly perpendicular to the strike of the trench. There are few constraints on wedge splitting in this region; one early study by Cassidy and Bostock [1996] found local split times of \(-0.3\) for the deepest earthquakes beneath Vancouver Island, but even the deepest events in their data set may not sample much mantle wedge material. However, Currie et al. [2004] argue that most of the splitting they measure can be attributed to anisotropy beneath the subducting slab, based on the large split times observed at stations close to the trench for raypaths that mainly sample the subwedge region, and we follow their interpretation. The fore-arc stations used by Currie et al. [2004] to estimate SKS splitting were located \(\sim 50–100\) km above the subducting slab, corresponding to subwedge upper mantle path lengths between \(\sim 300\) and \(350\) km. Their observation that subwedge anisotropy beneath Cascadia exhibits a trench-perpendicular fast direction is highly unusual in the global data set.

[20] Constraints for the Caribbean subduction zone came from the work of Piñero-Felicangeli and Kendall [2008], who measured both local and teleseismic splitting, although local measurements are somewhat sparse. Piñero-Felicangeli and Kendall [2008] argue for \(\sim 1.0\) s of trench-parallel splitting due to subwedge anisotropy in the Caribbean; we have assigned a large error bar to the estimate (due to the weak constraints on wedge anisotropy) and we use a range of \(0.5\) to \(1.5\) s in our compilation. We have focused on the stations located in the Caribbean arc itself, which are located approximately \(100\) km above the subducting slab. We take a similar approach to assigning an estimate for subwedge splitting beneath the Calabrian arc in Italy; in this region, SKS splitting has been studied by several workers [Schmidt et al., 2004; Civello and Margheriti, 2004; Baccheschi et al., 2007] but there are few constraints on wedge anisotropy, so it is not entirely clear how much of the SKS signal can be attributed to mantle flow beneath the wedge. Baccheschi et al. [2007] argue that most of the SKS splitting in their data set is due to trench-parallel anisotropy beneath the slab, as the large split times (up to \(\sim 3\) s) are too large to be attributed solely to anisotropy in the lithosphere or the mantle wedge We assign a subwedge splitting time of \(1.5 \pm 0.5\) s with a trench-parallel fast direction for the Calabrian arc in our compilation. The stations are located generally between \(50\) and \(200\) km above the subducting Calabria slab, which corresponds roughly to subwedge upper mantle path lengths of \(\sim 200–350\) km.

[21] Splitting constraints from the Middle America subduction zone come from studies using data from the TUCAN experiment [Abt et al., 2005, 2009; Hoernle et al., 2008]. Mantle wedge anisotropy in this region has been well characterized, as the three-dimensional distribution of seismic anisotropy was probed using a tomographic imaging method for local S splitting measurements [Hoernle et al., 2008; Abt et al., 2009]. SKS splitting has been investigated at a subset of the TUCAN stations; consistently trench-parallel SKS fast directions with large split times (up to \(\sim 2.0–2.5\) s) were documented by Abt et al. [2005]. A comparison of local S split times in this region, which generally range up to \(\sim 0.3–0.5\) s, with SKS measurements indicates that there is significant trench-parallel splitting beneath the subducting Middle America slab; we have assigned a range of \(1.0–1.7\) s for \(\delta t\) in our compilation. We focused on results from TUCAN stations located less than \(150\) km above the subducting slab, which corresponds to subwedge upper mantle path lengths between \(\sim 250\) and \(360\) km.

[22] Further to the south, in the South America subduction zone, shear wave splitting has been investigated by several different groups using both permanent stations and temporary deployments [Russo and Silver, 1994; Bock et al., 1998; Polet et al., 2000; Anderson et al., 2004]. Using data from stations located along the South American margin and using a variety of raypath combinations, Russo and Silver [1994] inferred significant splitting beneath the subducting Nazca slab, with mostly trench-parallel fast directions. Later studies [Bock et al., 1998; Polet et al., 2000] found a mix of trench-parallel, trench-perpendicular, and oblique SKS fast directions for South America; a comparison of the large split times observed for SKS phases with the generally small split times observed for local S phases indicates that there is significant subwedge splitting beneath South America. In particular, Polet et al. [2000] identified a localized region of trench-normal fast directions but concluded that elsewhere in South America, trench-parallel fast directions dominated the SKS signal. Using data from a temporary deployment in Chile and Argentina, Anderson et al. [2004] concluded that there is a large contribution \(\sim 1\) s or more) to the observed trench-parallel SKS splitting from within or below the slab, with only a small \(\sim 0.15\) s on average) contribution from the wedge and overlying plate (M. Anderson, personal communication, 2009). Subwedge anisotropy beneath Chile has also been studied using observations of source-side splitting for earthquakes originating in the South American slab by Rokosky et al. [2006]; they identified consistently trench-parallel fast directions with subwedge split times between \(0.5\) and \(1.6\) s. All of these studies paint a fairly consistent picture of subwedge splitting beneath South America, which appears to be dominated by trench-parallel fast directions with localized regions of trench-perpendicular or oblique directions. Subwedge split times are on the order of \(\sim 0.5–1.5\) s, depending on locality. Subwedge path lengths vary greatly depending on the type of data; for source-side measurements, event depths vary from \(\sim 220\) km to transition zone depths. We preferentially selected stations located closer to the trench for SKS and local S measurements, but also included measurements for a few stations located farther away; for this type of data, the subwedge path lengths vary from \(100\) to \(300\) km.

[23] The South Sandwich subduction zone lies to the south of South America; shear wave splitting has been measured at a few stations located in the arc itself [Müller, 2001] but the station coverage is extremely sparse and the lack of high-quality SKS and local S measurements for stations located in the arc itself makes interpretation difficult. More direct constraints on subwedge splitting are provided by a recent study of source-side anisotropy using earthquakes in the downgoing slab and receivers located in Antarctica [Müller et al., 2008]. This work demonstrated that significant trench-parallel splitting occurs beneath the wedge, with split times generally between \(0.4\) and \(1.6\) s, depending on earthquake depth. Measurements from shal-
lower earthquakes are most relevant for our work, since they have the longest paths lengths through the subwedge mantle, and the split times from such events in the Müller et al. [2008] study were generally greater than ∼1 s. Therefore, we assigned the South Sandwich subduction zone a range of subwedge split times from 1.0 to 1.6 s, with trench-parallel fast directions. The hypocentral depths used in the Müller et al. [2008] study ranged from ∼10 km to ∼275 km, which corresponds to a range of subwedge path lengths in the upper 400 km of the Earth of ∼125–390 km.

[24] Our compilation of subwedge splitting parameters makes use of two distinctly different types of data; we exploit both SKS splitting estimates which have been corrected for the effect of wedge anisotropy (inferred from local S splitting) and estimates of source-side splitting from slab earthquakes measured at distant stations. The successful comparison of results using these two different approaches in the same region provides a good compatibility check and argues for the robustness of our SKS correction procedure. South America represents the best opportunity to check the two different types of measurements against each other; the source-side splitting estimates from Russo and Silver [1994] and Rokosky et al. [2006] are very consistent with our estimates obtained from the SKS and local S splitting measurements of Russo and Silver [1994], Polet et al. [2000], and Anderson et al. [2004]. Specifically, Rokosky et al. [2006] document subslab splitting of ∼1 s with fast directions that fall within ∼25° of the local trench strike at latitudes between 26°–28°S; just to the south, the northernmost stations of the CHARME experiment exhibit trench-parallel SKS splitting with delay times between ∼0.5 s and ∼1.5 s [Anderson et al., 2004], while local S phases exhibit much smaller splitting (∼0.15 s (M. Anderson, personal communication, 2009)). Rokosky et al. [2006] find similar subslab splitting at latitudes between 20°–24°S; again, this is generally consistent with SKS and local S splitting [Russo and Silver, 1994; Polet et al., 2000], particularly for stations located close to the trench. As estimates of source-side anisotropy from slab earthquakes in a variety of regions become available, we expect to perform additional comparisons between different types of data, but the existing estimates suggest that our SKS correction procedure correctly isolates the subwedge splitting contribution.

2.3. Subwedge Splitting Signal: Distinguishing Between Slab and Subslab Contributions

[25] As discussed above, the subwedge splitting signal could be due to contributions from anisotropy in the slab itself, anisotropy in the subslab mantle, or both. While a contribution to splitting from the downgoing slab cannot be ruled out in any given subduction zone, anisotropy in the slab does not appear to represent a significant contribution to the global splitting signal, for several reasons. First, the overwhelming preponderance of trench-parallel fast directions in the global data set argues strongly for a mechanism that is related to trench/slab geometry (such as subslab mantle flow) rather than a mechanism that is related to fossil anisotropy in the downgoing lithospheric slab. Second, as we will discuss in section 4, there is no global correlation between subwedge split times and the age (and thus thickness) of the downgoing slab; this suggests that the slab itself does not contribute significantly to the observed global subwedge splitting signal. This lack of a correlation argues against models that invoke slab anisotropy to explain global SKS splitting patterns, as discussed further in section 8. Third, in regions where estimates of subwedge splitting are available from multiple data sets (both from a comparison of local and teleseismic splitting and from source-side splitting from slab earthquakes), these estimates are very consistent, despite the fact that different raypath configurations sample the slab in different ways. Therefore, we interpret the subwedge splitting signal as mainly representing anisotropy and mantle flow beneath the subducting slab.

2.4. Interpreting Subslab Fast Directions: Mineral Physics Considerations

[26] Until very recently, the available mineral physics data have uniformly suggested that the subslab mantle is dominated by A-, C-, or E-type olivine and that the conditions needed for B-type olivine fabric (high stress, low temperature, and significant hydrogen content) are not present beneath the slab [Jung and Karato, 2001; Jung et al., 2006; Karato et al., 2008]. This would imply that fast directions beneath the slab should be interpreted as being roughly parallel to the local mantle flow direction. Recent experiments, however, have provided some evidence for a pressure-induced transition to B-type olivine fabric [Jung et al., 2009]; taken at face value, these experiments suggest that the mantle flow direction would be orthogonal to the fast direction for depths greater than ∼80–90 km. This prediction, however, is difficult to reconcile with existing petrological and seismological evidence, which strongly suggests that only limited regions of the upper mantle are dominated by B-type olivine fabric [see Long, 2009, and references therein]. For example, modeling studies that rely on the A-type (or similar) fabric paradigm have been very successful in explaining anisotropy observed beneath ocean basins [e.g., Behn et al., 2004; Conrad et al., 2007; Becker, 2008]; in particular, studies which have compared azimuthal anisotropy from surface wave observations to the predictions from geodynamical models have shown that A-type (or similar) fabric likely predominates at asthenospheric depths, contradicting the prediction of Jung et al. [2009] that B-type fabric should be present below ∼90 km. The relationship between strain and anisotropy that holds in the asthenosphere beneath ocean basins very likely holds beneath subducting slabs as well. Because Jung et al. [2009] covered only a limited set experimental conditions and because deviatoric stresses in their experiments were very high, the depth at which any B-type olivine fabric transition might occur in the mantle remains poorly constrained. Therefore, we interpret the observed subslab fast polarization directions as being parallel to the direction of mantle flow; for the vast majority of subduction zones globally, this corresponds to trench-parallel subslab flow.

3. A Synoptic Model for Subduction Zone Anisotropy and Mantle Flow

[27] Our working model for the interaction of subducting slabs with mantle flow is described in detail by Long and Silver [2008]; here we provide a brief summary. Two
striking observations, strengthened by the additional observations made in this study, provided the starting point for this previously proposed model. First, the global compilation of subslab splitting parameters presented here reveals that the subslab fast directions are overwhelmingly trench-parallel (Figures 2 and 3). A few systems, such as Japan and South America, exhibit both trench-parallel directions and directions that are oblique or perpendicular to the trench, but even these regions tend to be dominated by trench-parallel anisotropy. The one notable exception in the published literature is Cascadia (this exception is discussed in more detail in sections 6 and 8). The trench-oblique points correspond to Japan and to a segment of the Indonesia-Sumatra subduction zone.

Our model is summarized in Figures 4 and 5, which shows a schematic diagram of subslab mantle flow, as well as a plot of subslab split time versus trench migration velocity similar to that of Long and Silver [2008]. We have added new data to this plot, taking advantage of measurements from the Scotia [Müller et al., 2008], Hikurangi [Audoin et al., 2004], Izu-Bonin [Wirth and Long, 2008], the Marianas [Wookey et al., 2005], and South America [Rokosky et al., 2006] subduction zones that were not included in our original compilation. For the plot shown in Figure 5, we have computed a correlation coefficient of $R = 0.72$ between $|V_t|$ and $dt$; this is significant at greater than a 99% confidence level (that is, the range of $R$ values allowed by the data at a 99% confidence level does not include zero).

Our model requires trench-parallel flow beneath the slab without producing significant slab-entrained flow and thus entails three significant properties. First, the suppression of slab-entrained flow requires a mechanism for decoupling downgoing slabs from the mantle material beneath it. Possible mechanisms for and implications of this decoupling, along with the exceptions represented by regions such as Cascadia, are explored in section 6. Second, our model requires a barrier to flow at depth, likely at either the top or the base of the mantle transition zone, which is permeable to subducting slabs but does not permit the flow

![Figure 3](image-url)  
**Figure 3.** Histogram of angles between the dominant subslab fast direction and local trench strike for the global subslab splitting data set. The two trench-perpendicular data points correspond to Cascadia and to the small segment of the South America subduction zone at a latitude of $\sim 20^\circ$S. (Preliminary SKS splitting results from Mexico [Stubailo and Davis, 2007; Léon Soto et al., 2009] are also consistent with trench-perpendicular subslab splitting, but these results are not included in the plot.) The trench-oblique points correspond to Japan and to a segment of the Indonesia-Sumatra subduction zone.

![Figure 4](image-url)  
**Figure 4.** Sketch of the model explored in this paper for subslab anisotropy controlled by trench migration and slab decoupling [after Long and Silver, 2008].
of surrounding mantle material. Because the surrounding mantle cannot penetrate this barrier, it is forced to flow horizontally by the motion of the trench rather than underneath the slab. The presence of such a barrier to entrained flow has important implications for global mantle circulation and mass transfer between the lower and upper mantle; these are discussed in section 7. Third, in order to force mantle material to flow parallel to the trench, a distant barrier to horizontal flow is also required (in the absence of such a barrier, a moving trench would simply force mantle to translate along with it in the trench migration direction, rather than in a trench-parallel direction). This is demonstrated by numerical and analog box modeling studies; the boundaries of the box are required to induce trench-parallel flow [e.g., Buttles and Olson, 1998; Kincaid and Griffiths, 2003]. Such a barrier may be provided by the global system of subduction zones. Indeed, in the Pacific, where most contemporary subduction is taking place, the entire Pacific Rim of subduction zones likely represents a closed system [Husson et al., 2008], acting as distant boundaries that act as barriers to trench-normal motion.

4. Comparison of Splitting Parameters With Other Tectonic Parameters

[30] Comparisons among different subduction parameters have shown to be useful both in characterizing relationships between sets of parameters and in elucidating the physical processes that produce those relationships [e.g., Jarrard, 1986; Carlson, 1995; Heuret and Lallemand, 2005; Cruciani et al., 2005; Syracuse and Abers, 2006]. One promising approach, therefore, to understanding the processes that control seismic anisotropy is to undertake comparisons between shear wave splitting parameters and parameters that describe subduction. Here we present comparisons between average subslab splitting parameters, compiled as described in section 2, and a wide variety of parameters that describe the kinematics, behavior, and morphology of subduction zones and may be relevant to the subslab flow. Our previous work [Long and Silver, 2008] focused on the relationship between splitting parameters and trench migration velocity. Here, we present a more complete assessment of any relationship between subslab shear wave splitting and parameters that describes subduction. As discussed in section 2, the vast majority of subduction zones worldwide exhibit predominantly trench-parallel fast directions, with one major exception (Cascadia) and a few that also exhibit regions of oblique (Japan, Sumatra) or trench-perpendicular (South America) fast directions. Because the global subslab splitting signal is overwhelmingly dominated by trench-parallel fast directions, we focus on the subslab delay time $\tau_s$ as the splitting parameter that we compare to subduction parameters. We compare subslab $\tau_s$ ranges to a total of 11 parameters (in addition to the trench migration velocity, described in section 3) that describe subduction: total convergence rate, normal convergence rate, convergence obliquity angle, age of subducting lithosphere, slab dip, arc curvature, arc length, maximum depth of seismicity, maximum depth of slab penetration (from tomographic models), descent rate of the slab into the mantle, and the so-called thermal parameter, described below. In our companion study of global patterns of mantle wedge anisotropy, currently in progress, we also examine subduction parameters that may be related to processes in the wedge, such as overriding plate strain, overriding plate motion, volcanic production, and the depth to the slab beneath active volcanoes.

[31] Several recent compilations of subduction zone parameters have appeared, including those by Heuret and Lallemand [2005], Lallemand et al. [2005, 2008], and Syracuse and Abers [2006]. We used values of total
convergence rate, normal convergence rate, convergence obliquity, subducting plate age, descent rate, and the thermal parameter from Syracuse and Abers [2006]. The thermal parameter $\phi$ is defined as the product of the plate age and the descent rate of the slab [ Kirby et al., 1996; Syracuse and Abers, 2006]. We relied on estimates of arc curvature from Tovish and Schubert [1978]. Our estimates for the maximum depth of seismicity in each subduction zone were obtained from the slab contours described by Guðmundsson and Sambridge [1998] and Syracuse and Abers [2006]. For estimates of the arc length, we used a compilation of “slab width” values from Schellart et al. [2007] (roughly equivalent, in their definition, to the length of the arc); however, we have separated the Izu-Bonin-Marianas system and the Japan-Kurile-Kamchatka system into two separate arcs for the purpose of calculating arc length. We also examined the maximum depth of slab penetration as compiled by Lallemand et al. [2005] from various tomographic models; although there is, certainly, a great deal of subjectivity in the interpretation of tomographic images of slabs, many studies have distinguished between slabs that are stagnant in the transition zone and slabs that penetrate into the lower mantle [e.g., Fukao et al., 2001], and we regard this as an important descriptive parameter for the slab.

[32] Plots of subslab delay times ($\delta t$) against each of the subduction parameters described above are presented in Figure 6; the relationship between $\delta t$ and trench migration velocity in a hot spot reference frame, as compiled by Heuret and Lallemand [2005], is shown in Figure 5 and was discussed in section 3 and by Long and Silver [2008]. As with Figure 5, we have computed the correlation coefficient for each relationship shown in Figure 6. For each of the plots presented in Figure 6, we have not used the split time estimates for a given subduction zone if there is not a corresponding estimate for the subduction parameter under study. As Figure 6 demonstrates, we have not identified any other striking global correlations between subslab split time and parameters that describe subduction. None of the correlations tested here is significant at the 95% confidence level, in contrast to the correlation with $|V_p|$, which is significant at the 99%+ confidence level. The most striking relationship that emerges from our comparisons, therefore, is that between subslab split time and trench migration velocity in a Pacific hot spot reference frame.

[33] A key question, as discussed in section 2, is whether the subwedge delay time correlates with slab age (a proxy for slab thickness), which could potentially indicate a primary contribution from anisotropy in the slab itself. As shown in Figure 6d, there is no correlation between subwedge delay time and the age of the subducting lithosphere. This simple comparison does not take into account variations in path length through the slab due to differences in slab dip, however. Assuming nearly vertical SKS propagation, the range of slab dips contained in our data set (~10°–65°) would cause a variation in slab path length up to a factor of ~2.3. In order to more fully investigate this effect, we have carried out an additional test (similar to that shown in Figure 6d) in which the age of the slab was “corrected” by dividing by a path length factor calculated from the slab dip before testing for a statistical correlation. This test yielded a correlation coefficient of $R = 0.18$, which indicates that even when the age of the subducting lithosphere is corrected for the effect of different path lengths due to slab dip, there is no statistically significant correlation between slab age and subwedge delay time.

[34] The comparisons between subslab split time and subduction zone parameters shown in Figures 5 and 6 support the Long and Silver [2008] model for trench-parallel flow beneath subducting slabs controlled by trench migration. The new subduction zones (Izu-Bonin, Marianas, South Sandwich, and New Zealand) that we have added to our compilation fit the linear trend in subslab split time with trench migration velocity well (Figure 5), and the absence of other notable global correlations (Figure 6) between subslab split time and parameters that describe subduction lends support to this model. There are, as we have noted, a few exceptions to the global trend of trench-parallel fast directions, such as the consistent trench-perpendicular fast directions observed in Cascadia; these exceptions are discussed in the context of our model in sections 6 and 8. We also note that although the global subslab flow field appears to be controlled by trench migration, other effects likely modify this flow field locally. For example, complex slab morphology likely has an effect on the 3-D subduction flow field. This has been explored by Kneller and van Keken [2007, 2008], although they restricted their investigation to the mantle wedge region. Locally complex slab morphology likely has a significant effect on local fast directions; for example, the trench-perpendicular fast directions observed in a localized region in South America [Polet et al., 2000] coincide geographically with a change in the arc morphology and probably a change in slab morphology as well. With these caveats, however, the comparisons presented in Figures 5 and 6 appear to be consistent with our previously proposed subslab flow hypothesis [Long and Silver, 2008].

5. Reference Frames for Trench Migration and for the Convecting Mantle

[35] Our model invokes the migration of trenches relative to the ambient surrounding mantle to induce a three-dimensional flow field that results in dominantly trench-parallel flow beneath slabs. A correct characterization of the relative migration thus requires the specification of this mantle motion. If the motion is common to all subduction zones, then it would correspond to a single reference frame. The characteristics of this mantle motion, however, remain poorly known. Until now, we have used trench migration velocities in a Pacific hot spot reference frame [Gripp and Gordon, 2002] to describe trench motion, but the relationship shown in Figure 5 should depend on reference frame. The magnitude and, in some cases, the direction of trench migration depends heavily on the plate motion model and reference frame used [e.g., Stegman et al., 2006; Schellart et al., 2007, 2008; Lallemand et al., 2008]. For example, the central part of the South America subduction zone is retreating, with an average velocity of ~35 mm yr$^{-1}$, in a Pacific hot spot reference frame, while it is nearly stationary in an Indo-Atlantic hot spot reference frame and is advancing slowly (~10 mm yr$^{-1}$) in a no-net-rotation reference frame [Schellart et al., 2008]. The central and northern parts of the Ryukyu subduction zone are nearly stationary in a Pacific hot spot reference frame, but are
Figure 6. Subslab split time plotted against a variety of parameters describing subduction: (a) total convergence velocity, (b) trench-normal convergence velocity, (c) convergence obliquity angle, (d) age of subducting lithosphere, (e) slab dip, (f) arc curvature, (g) arc length, (h) maximum depth of seismicity, (i) maximum depth of slab penetration from tomography, (j) slab descent rate, and (k) thermal parameter. Values for each subduction parameter were compiled from the sources described in the text. Symbols represent different subduction systems according to the legend in Figure 5. The calculated correlation coefficients are $R = -0.28$ for Figure 6a, $R = -0.27$ for Figure 6b, $R = 0.20$ for Figure 6c, $R = -0.01$ for Figure 6d, $R = 0.06$ for Figure 6e, $R = 0.23$ for Figure 6f, $R = -0.02$ for Figure 6g, $R = -0.04$ for Figure 6h, $R = 0.39$ for Figure 6i, $R = -0.26$ for Figure 6j, and $R = -0.10$ for Figure 6k.
retreating with velocities of $\sim 20$ mm yr$^{-1}$ in the Indo-Atlantic hot spot and no-net-rotation reference frames [Schellart et al., 2008]. It is, therefore, imperative to test the comparisons between trench migration rate and the strength of subslab anisotropy established here in different reference frames. This will allow us to answer several questions: first, does the relationship we identified between subslab $\delta t$ and $|V_t|$ (the absolute value of the trench migration velocity) in the Heuret and Lallemand [2005] compilation hold if we use trench migration estimates for a variety of reference frames and plate motion models? Can we identify a different reference frame that improves the fit
of the relationship we observe in the Pacific hot spot reference frame? And might this, in principle, allow us to identify the best reference frame for the convecting mantle, at least as it relates to subduction zones (which are located mainly around the Pacific basin)?

[36] In order to answer these questions, we have reproduced the plot of \( \delta t \) versus \( |V_t| \) shown in Figure 5 for a variety of different plate motion models and reference frames (Figure 7). We used the estimates of trench migration velocities compiled by Schellart et al. [2008] for eight different combinations of reference frames and plate motion models: (1) the plate motion model and Pacific hot spot reference frame of Gripp and Gordon [2002], which was also used in the compilation of Heuret and Lallemand...
[2005] and by Long and Silver [2008] (Figure 7a), (2) the plate motion model and fixed hot spot reference frame of Gordon and Jurdy [1986] (Figure 7b), (3) the no-net-rotation model of Argus and Gordon [1991] (Figure 7c), (4) the no-net-rotation model of Kreemer et al. [2003] (Figure 7d), (5) the plate motion model of DeMets et al. [1994] in the Antarctic plate reference frame of Hamilton [2003] (Figure 7e), (6) the plate motion model of DeMets et al. [1994] with the Wessel et al. [2006] Pacific hot spot reference frame (Figure 7f), (7) the plate motion model of DeMets et al. [1994] in the Indo-Atlantic hot spot reference frame of O’Neill et al. [2005] (Figure 7g), and (8) the plate motion model of Kreemer et al. [2003] in the O’Neill et al. [2005] Indo-Atlantic hot spot reference frame (Figure 7h). Essentially, the differences among these eight sets of trench migration velocity values can be attributed either to differences in the reference frame, or differences in the model for back-arc deformation. Further details on this compilation are given by Schellart et al. [2008].

[37] We have also reproduced our δt versus |Vt| comparisons using the three different compilations of trench migration velocities given by Lallemand et al. [2008]: the HS3 Pacific hot spot reference frame [Gripp and Gordon, 2002] used by Heuret and Lallemand [2005], the no-net-rotation frame of DeMets et al. [1994], and the reference frame proposed by Steinberger et al. [2004] based on an improved reconstruction of the Emperor-Hawaii seamount chain using a mantle convection model. The Steinberger et al. [2004] reference frame and plate motion model incorporates an updated hot spot frame that accounts for moving hot spots. These three comparisons are shown in Figure 8.

[38] The plots of subslab δt versus |Vt| (Figures 7 and 8) demonstrate that the Pacific hot spot based reference frames described by Gripp and Gordon [2002], Wessel et al. [2006], and Steinberger et al. [2004] yield the strongest correlations between subslab anisotropy and trench migration. The correlations between subslab δt and |Vt| for the Gripp and Gordon [2002] reference frame are significant at the 99% confidence level (using both the Heuret and Lallemand [2005] and the Lallemand et al. [2008] compilations) and the 93% confidence level (using the Schellart et al. [2008] compilation), while the correlation for the Wessel et al. [2006] reference frame (Figure 7b) is significant at the 91% confidence level and that for the Steinberger et al. [2004] reference frame is at 88%. All of the correlations that are significant above the 88% confidence level, therefore, are associated with Pacific hot spot reference frames. All of the other plate motion models and reference frames (no-net-rotation, Indo-Atlantic hot spot, and Antarctic) yield no

Figure 7. Subslab split time versus absolute value of trench migration velocity for each of the plate motion model and reference frame combinations described by Schellart et al. [2008] and in section 5. (a) Plate motion model and Pacific hot spot reference frame of Gripp and Gordon [2002]. (b) Plate motion model and hot spot reference frame of Gordon and Jurdy [1986]. (c) The no-net-rotation reference frame and plate motion model of Argus and Gordon [1991]. (d) The no-net-rotation reference frame and plate motion model of Kreemer et al. [2003]. (e) The plate motion model of DeMets et al. [1994] in the Antarctic plate reference frame of Hamilton [2003] (f) The plate motion model of DeMets et al. [1994] in the Wessel et al. [2006] Pacific hot spot reference frame. (g) The plate motion model of DeMets et al. [1994] in the Indo-Atlantic hot spot reference frame of O’Neill et al. [2005]. (h) The plate motion model of Kreemer et al. [2003] in the O’Neill et al. [2005] reference frame. Symbols represent different subduction systems according to the legend in Figure 5. The calculated correlation coefficients are R = 0.40 for Figure 7a, R = 0.24 for Figure 7b, R = 0.19 for Figure 7c, R = 0.06 for Figure 7d, R = 0.28 for Figure 7e, R = 0.38 for Figure 7f, R = 0.23 for Figure 7g, and R = 0.04 for Figure 7h. The splitting estimate from GSN station PSI in the Tonga-Kermadec subduction zone is shown with a dashed line; accurate trench migration velocity estimates are difficult to make for this station because of the very large along-strike gradients in trench migration (the station is located near the transition from rapid trench retreat to the north to rapid trench advance to the south).
A discernable relationship between $\delta t$ and $|V_t|$. If our model for subslab flow is correct, the fact that the Pacific hot spot reference frames yield the clearest linear relationship between subslab $\delta t$ and $|V_t|$ suggests that such reference frames do the best job of describing a “rest” frame for Earth’s subduction zones relative to the upper mantle. This is perhaps unsurprising, as nearly all subduction zones are located either on the margins of the Pacific plate (e.g., Aleutians, Kuriles, Japan, Izu-Bonin-Marianas, Tonga) or in geographical proximity to it (e.g., Central America, Caribbean, Scotia, Ryukyu, Indonesia). The sole exception is the Calabria subduction zone in southern Italy. Because there is only one truly “non-Pacific” subduction zone in the compilation, it is difficult to evaluate whether the correlations with Pacific hot spot trench migration velocities would improve if subduction zones located far from the Pacific plate were removed from the plots shown in Figures 7 and 8.

In principle, the compilations of Lallemand et al. [2008] and Schellart et al. [2008] for the Gripp and Gordon [2002] Pacific hot spot reference frame should be similar, although in practice different estimates for back-arc deformation can result in incompatible estimates for $|V_t|$. In particular, we note that the correlation between subslab $\delta t$ and $|V_t|$ in the Pacific hot spot reference frame [Gripp and Gordon, 2002] is poorer using the Schellart et al. [2008] compilation than that of Lallemand et al. [2008] (Figure 7a versus Figure 8a), with a correlation coefficient of $R = 0.66$ for Lallemand et al. [2008] versus $R = 0.36$ for Schellart et al. [2008]. This is mainly due to the very different trench migration velocities obtained at the location of GSN station PSI in the Tonga-Kermadec subduction zone, at which splitting was measured by Long and Silver [2008], although a few other subduction systems (Izu-Bonin, Java-Sumatra, and Kamchatka) also yield dissimilar estimates using the different compilations. The disparity obtained for the
Kermadec subduction zone is mainly due to the unique kinematics of the Tonga-Kermadec subduction system; the northern part of the subduction zone is retreating rapidly in most reference frames, while the southern part is advancing; station PSI is located near the transition from trench retreat to trench advance. The different back-arc deformation estimates and along-strike spatial sampling used by Schellart et al. [2008] versus Lallemand et al. [2008] results in disparate estimates for $|V_t|$ at this station location. Physically, however, one would expect that the rapid trench movement to the north and south of this “hinge” point would result in considerable trench-parallel flow; taking this into account would improve the fit shown in Figure 7a.

[40] In general, dissimilar back-arc deformation estimates (which variously incorporate geological and geodetic data), treatments of tectonic erosion/accretion (which is ignored in the Lallemand et al. [2008] compilation), choices about which microplates and/or arc blocks to consider, and along-strike sampling variations between the two compilations can account for the differences between the two compilations, many of which are subtle (see Schellart et al. [2008] for a more complete discussion of the differences among trench migration velocity compilations). For example, Schellart et al. [2008] relied on geological estimates of back-arc deformation rates wherever possible in their Pacific hot spot reference frame compilations, while Lallemand et al. [2008] relied mainly on geodetic estimates. Despite the (often subtle) differences between $|V_t|$ estimates (with the exception of Tonga-Kermadec, all of the $|V_t|$ estimates for individual subduction zones shown in Figures 7a and 8a differ by less than $\sim 10–15 \text{ mm yr}^{-1}$), statistically significant correlations are observed for the Gripp and Gordon [2002] reference frame for both compilations that we examined.

[41] The dissimilarities in local trench migration velocities between different compilations and the unique kinematics of the Tonga-Kermadec subduction system both bring up a larger point about the utility of localized trench migration velocity estimates as an indicator of subslab mantle flow. Particularly for subduction systems for which the local trench migration velocity changes rapidly along strike, such as Tonga, the local value of $|V_t|$ may not actually be the best indicator of subslab flow—rapid trench retreat or advance on adjacent sections of the trench may induce considerable trench-parallel flow below trench segments that are themselves nearly stationary. Future modeling work should help to answer questions about how along-strike gradients in $|V_t|$ affect the subslab flow field and incorporating information about such along-strike gradients into plots such as those shown in Figures 7 and 8 may help to reconcile discrepancies between different trench migration compilations.

[42] There have been other attempts to define a “best” reference frame for mantle flow, and it is constructive to compare our results to them. The study closest to our own is that of Becker [2008], who compared the global constraints on azimuthal anisotropy in the upper mantle from surface wave models to the asthenospheric anisotropy predicted by numerical models of mantle flow. He concluded that the fit between predicted and observed asthenospheric anisotropy was best for a no-net-rotation (or minimal-net-rotation) reference frame. While this result appears to be at odds with our preference for the Pacific hot spot reference frame, it is possible that these two results reflect real hemispherical variability in flow in the upper mantle. Our own criterion for the “best” reference frame is heavily biased to the Pacific basin, since the subduction zones in our compilation are nearly all located on or near the margins of the Pacific Ocean. In contrast, the approach taken by Becker [2008] likely effectively downweights the Pacific. Because the Pacific is a fast moving plate, the orientation of asthenospheric shear (representing the difference between plate velocity and subasthenospheric mantle velocity) and therefore seismic anisotropy, is insensitive to the motion of the underlying subasthenospheric mantle [Behn et al., 2004]. For slow moving plates, such as the Atlantic, the anisotropy is, in contrast, much more sensitive to the subasthenospheric velocity, and these regions would consequently dominate any global fit.

[43] We therefore hypothesize that these two studies, taken together, represent actual hemispheric-scale variability in upper mantle flow. This may be located to structures in the lowermost mantle; the existence of prominent, large-scale low-velocity anomalies beneath the Pacific (the “Pacific Anomaly”) and beneath Africa (the “African Anomaly”) has been firmly established [e.g., Garnero and McNamara, 2008; Ni et al., 2002; Wang and Wen, 2007]. Such structures may have a dramatic effect on hemispheric-scale mantle convection; such an effect has been studied by, e.g., Gaboret et al. [2003]. If a Pacific hot spot reference frame is, indeed, the most appropriate “rest” frame for Pacific mantle flow, and either the no-net-rotation or the Indo-Atlantic hot spot reference frame is most appropriate for the Atlantic-African hemisphere, this implies a relative motion between the two hemispheres (the net rotation of the lithosphere in the HS3-Nuvel-1A model is $0.436^\circ \text{ Myr}^{-1}$ [Gripp and Gordon, 2002; Becker, 2008]). Our model does not constrain a mechanism for this hypothesized relative motion, but our suggestion seems to be consistent with previous studies that have suggested that Pacific hemisphere mantle convection is largely driven by a hemispheric-scale large upwelling anchored beneath the South Pacific super-swell [Gaboret et al., 2003] or that large-scale lower mantle return flow has driven net motion between the Hawaii-Emperor hot spot and the Indo-Atlantic hot spots [e.g., DiVenere and Kent, 1999; Tarduno et al., 2003, 2009]. The need to invoke different reference frames for the Pacific Ocean basin and the Indo-Atlantic oceans in order to explain observations of seismic anisotropy beneath ocean basins and beneath subducting slabs is, in fact, similar to the need to invoke different reference frames for Pacific versus Indo-Atlantic hot spots.

[44] Another approach to determining a mantle rest frame has been provided by Schellart et al. [2008], who established constraints based on a set of assumptions about trench migration behavior. In particular, they sought to minimize the amount of trench migration in the center of wide subduction zones, to maximize trench migration at subduction zone edges, and to maximize the number of trench segments that retreat (as opposed to advance). Schellart et al. [2008] concluded that the Indo-Atlantic hot spot reference frame best fits this set of assumptions. The approach taken by Schellart et al. [2008] differs markedly from that taken in this study, as ours relies on measurements of the subslab mantle flow field (which we
interpret as being due to trench migration) rather than assumptions about trench migration behavior.

6. Dynamics of the Subslab Mantle During Subduction: Models and Implications

[45] The observation that globally, subslab fast directions are nearly always parallel to the trench appears contrary to the expectations of a simple, 2-D model of entrained flow beneath downgoing slabs. This lack of evidence for entrained flow beneath slabs (with a few exceptions, such as Cascadia) has led us to consider a variety of models that could result in dominantly trench-parallel subslab flow directions. A closely related issue is the fate of suboceanic asthenosphere as a subducting plate begins its descent into the mantle, as we will show. Trench wave splitting behavior beneath oceanic plates far away from trenches or ridges is relatively well understood as being due to simple shear in the asthenosphere, with the geometry of the fast directions being controlled by the vectorial difference between sub-asthenospheric motion and plate motion [e.g., Behn et al., 2004; Conrad et al., 2007]. It is not at all clear, however, what the fate of the asthenosphere is at subduction zones, and this is a major consideration in the models we consider to explain the apparent lack of entrained flow beneath slabs.

[46] First, we consider if there is any way to reconcile the preponderance of trench-parallel fast directions with an entrained flow model. One possibility would be a model that invokes perfectly entrained flow, that is, in which downgoing slabs do entrain material, but there is no strain involved. In this scenario, there is a translation of mantle material down with the slab, along with a small component of trench-parallel strain that gives rise to the observed trench-parallel fast directions. However, a perfect translation of mantle material beneath the slab, that is, entrainment without deformation, would imply stress-free boundary conditions, and the fact that there is significant anisotropy (and therefore deformation) in many subslab regions contradicts this implication. Therefore, a model that invokes strain-free entrainment of mantle material in combination with a significant trench-parallel component of deformation seems physically unworkable.

[47] Another possibility is that there is a component of entrained flow beneath the slab, but the strain associated with this entrained flow does not lead to observable anisotropy and is thus “hidden.” This would be possible if strain in this region were accommodated by diffusion creep, which does not produce lattice preferred orientation (LPO) and therefore seismic anisotropy, instead of diffusion creep, which does produce LPO [e.g., Karato and Wu, 1993]. Again, however, constraints are provided by the observation of significant trench-parallel anisotropy in most subduction zones globally. A scenario in which entrained flow directly beneath the slab is accommodated by diffusion creep would have to accommodate a region of trench-parallel flow in the dislocation creep regime below the entrained flow layer, and a mechanism for a transition from diffusion creep to dislocation creep with increasing depth is difficult to envision. A more plausible “hidden” layer of entrained flow is one that is thin enough (on the order of ~10 km) so that any anisotropy is invisible to shear waves passing through it; as explained in more detail below, this is the scenario we favor. We note that either case (the “thick” layer of entrained asthenospheric flow accommodated by diffusion creep, or the “thin” layer of entrained asthenosphere) requires the advection of asthenospheric material that has distinct properties from the ambient mantle. If the asthenosphere merely represents the depth range over which the temperature is sufficiently close to the mantle solidus to cause significant weakening (that is, a reduction in viscosity), then it would be impossible for slabs to advect asthenospheric material deeper into the mantle; the physical properties of such an asthenosphere would be controlled only by depth (pressure) (Figure 9).

[48] Our preferred model for the subslab mantle is that downgoing slabs generally entrain a thin layer of asthenosphere that acts as a decoupling zone between the slab and the subslab mantle, which allows for the motion of the trench/slab system to induce trench-parallel flow (Figure 9). The viability of this model depends on the physical (and

Figure 9. Cartoons of two scenarios for the fate of the asthenosphere at subduction zones. (a) An asthenosphere that is controlled solely by pressure (depth); in this scenario, the low viscosity of the asthenosphere results from the proximity of the adiabatic temperature to the mantle solidus. If the depth extent of the asthenosphere is controlled only by pressure, then the asthenosphere will “pinch out” beneath the subducting slab and there is no mechanism for subslab decoupling at subasthenospheric depths. (b) A hot, buoyant asthenosphere that is formed by shear heating; a thin layer of such an asthenosphere is entrained beneath the slab, in a scenario similar to that envisioned by Phipps Morgan et al. [2007].
perhaps chemical) characteristics of the asthenosphere, but it requires an asthenosphere that is physically distinct from the surrounding ambient mantle. One possible mechanism is that suggested by Phipps Morgan et al. [2007]; they conducted numerical and laboratory experiments and showed that a thin (~10–30 km) layer of hot, buoyant asthenosphere may be entrained beneath downgoing slabs. This interpretation depends, however, on a particular view of the asthenosphere [Phipps Morgan et al., 1995]: namely, that is a consequence of positively buoyant upwelling plumes in the mantle that feed hot, chemically depleted material into the sublithospheric mantle. In this scenario, the asthenosphere is intrinsically buoyant, has a higher potential temperature and is much weaker than the mantle beneath it.

[40] The entrainment of a thin layer of buoyant asthenosphere beneath slabs does not, however, necessarily require a plume-fed, chemically depleted asthenosphere; here, we propose an alternate view of the asthenosphere that is consistent with such an entrainment scenario. For a model of the type presented by Phipps Morgan et al. [2007] invoking a thin entrained subslab asthenospheric layer to be viable, the asthenosphere must be buoyant. One possible mechanism for this is shear heating. We envision a scenario in which mantle material beneath newly created oceanic lithosphere, subjected to high-strain shear deformation induced by the motion of the overlying oceanic plate, undergoes shear heating as it moves progressively away from the ridge. Numerical models of corner flow in the vicinity of a spreading ridge [e.g., Nippress et al., 2007] indicate that the sublithospheric mantle is subjected to strains much larger than the mantle material directly beneath the ridge. If the viscosity of the starting material is large enough, the heating (and associated lowering of the viscosity) that would result from viscous dissipation in the sublithospheric mantle may be nonnegligible.

[40] The approximate effects of shear heating produced by asthenospheric flow can be estimated by obtaining the steady state temperature profile for Couette flow with both viscous dissipation and temperature-dependent viscosity. Here we briefly describe the equations derived by Turcotte and Schubert [1982] and discuss our estimate of the size of the shear heating effect using nominal values for mantle viscosity and temperature. Following Turcotte and Schubert [1982, pp. 320–324], the temperature, $T$, resulting from Couette flow in a channel of thickness $h$ driven by a stress $\tau$ applied to the upper boundary and with fixed lower boundary can be described by

$$
k \frac{d^2 T}{dx^2} + \frac{\tau^2}{\mu} = 0$$  \hspace{1cm} (1)

where $\mu$ is the temperature-dependent viscosity of the form $\mu = C e^{E_o/R T_o}$, $E_o$ is the activation energy, $R$ is the ideal gas constant, and $C$ is a constant. Boundary conditions are imposed such that $T = T_o$ on the upper boundary and $dT/dx = 0$ at the lower boundary (insulating). Assuming that the temperature perturbation, $T'$, due to shear heating is small compared to $T_o$, (1) can be rewritten as

$$
k \frac{d^2 \theta}{dx^2} + Br e^\theta = 0$$  \hspace{1cm} (2)

in terms of a dimensionless temperature perturbation $\theta = \gamma_o T'/T_o (\gamma_o = E_o/R T_o)$ and $\bar{x} = x/h$. $Br$ is the Brinkman number, defined as [Turcotte and Schubert, 1982]

$$Br = \frac{\tau^2 h^2 \gamma_o}{k \mu_o T_o}$$  \hspace{1cm} (3)

where $\mu_o$ is the reference viscosity at temperature $T_o$. The solution to (2) yields the temperature perturbation as a function of depth. This solution can be used to relate $Br$ to the maximum temperature perturbation in the layer, $\theta_t$ [see Turcotte and Schubert, 1982, Figure 7–17]. This temperature, however, is a double-valued function of $Br$, defining a high-temperature and low-temperature branch. The correct temperature can be calculated by specifying the velocity of the upper plate, $u$. Since $u$ is the integral of the strain rate, $\dot{\varepsilon}$, through the layer, i.e.,

$$u = \int_0^h \dot{\varepsilon}(x) dx = h \int_0^1 \dot{\varepsilon}(\bar{x}) d\bar{x},$$

we can use the relation for a Newtonian fluid, $\tau = \mu \dot{\varepsilon}$, to write $\dot{\varepsilon} = \tau \mu^{-1} = \tau \mu^{-1} e^\theta$ so that

$$u = \tau \mu^{-1} \int_0^1 e^\theta d\bar{x} \equiv \tau \mu^{-1} m.$$  

(Here $m$ denotes the integral expression in the previous equation.) Substituting $\tau = \mu_o u / h m$ in (3) yields $Br = \mu_o u^2 \gamma_o / k m^2 T_o$. Writing as $Br m^2 = \mu_o u^2 \gamma_o / k T_o$, the right-hand side consists of nominal parameters, and one can then determine the unique point on the $Br - \theta_t$ curve that satisfies this equation, which then defines the maximum temperature rise.

[51] A reasonable set of values for the upper mantle (a temperature of 1500 K at the base of the lithosphere, an initial (reference) viscosity of $10^{21}$ Pa s, a plate velocity of 60 mm yr$^{-1}$, $\gamma_o = 20$, and $k = 4$ W m$^{-1}$ K$^{-1}$) places us on the high-temperature branch, and we would predict a temperature increase of over 100 K and a viscosity decrease of ~1 order of magnitude. (For a more detailed derivation of the relationship between the Brinkman number and the steady state temperature profile in the channel, we refer the reader to Turcotte and Schubert [1982]. We are also currently undertaking further analytical and numerical modeling of asthenospheric shear heating.) In this model, therefore, newly formed suboceanic asthenosphere (after steady state is reached) would have a viscosity of about $10^{18}$–$20$ Pa s (depending on the values used in the calculation), consistent with estimates from geophysical observations such as postglacial rebound [e.g., Fjeldskaar, 1994] or triggered earthquakes [e.g., Rydelek and Sacks, 1988]. The thermal buoyancy provided by the ~100 K temperature rise would make the entrainment of a thin decoupling layer beneath the slab feasible.

[52] Our shear heating model for the origin of the asthenosphere makes testable predictions that should provide additional means for discriminating among different models for the mechanism which gives rise to subslab decoupling. Our model implies a starting viscosity value
for protoasthenosphere that is the same as that of ambient mantle material; the reduction in viscosity to $\sim 10^{19} - 10^{20}$ Pa s is accomplished by shear heating through progressive strain that results in a temperature rise of $\sim 100$ K. A prediction of our model, therefore, is that the asthenosphere will have an excess potential temperature (that is, the temperature relative to an adiabatic temperature profile) of $\sim 100$ K, and a key test of our model is whether this prediction can be reconciled with the one-dimensional seismic velocity profiles inferred for the oceanic mantle. Another test can be provided by examining lateral variations in seismic velocities at asthenospheric depths: our model predicts that “young” asthenosphere near spreading ridges that has not yet undergone sufficient strain to experience shear heating should be relatively cooler than fully developed asthenosphere away from spreading ridges. A key question is whether this is borne out by three-dimensional seismic velocity models of the mantle; however, it is questionable whether global models provide the resolution needed to examine such small-scale lateral variations, particularly since seismic velocities in the immediate vicinity of ridges are strongly affected by the presence of partial melt. For example, global surface wave models such as that of Nettles and Dziewonski [2008] show low velocities in the immediate vicinity of spreading ridges and relatively higher velocities away from the ridges. This is the opposite of the effect predicted by the shear heating model; however, the low velocities beneath ridges are very likely due to the effect of partial melt beneath the ridge. The poor lateral resolution afforded by such global studies does not allow for the separation of this partial melt effect from the temperature effect predicted by the shear heating model. The paucity of seismic stations in the oceans means that there are few, if any, previous regional seismic studies that have both the resolution and the lateral extent to look for lateral temperature variations in the vicinity of ridges, but this could represent an important scientific target for future studies.

While our hypothesized shear heating mechanism provides one possibility for asthenospheric entrainment, there are certainly other possibilities, depending on the physical state of the asthenosphere. For example, it has been suggested that the asthenosphere possesses a high volatile content [Karato and Jung, 1998] or, alternatively, that it corresponds to a minimum in water solubility in aluminous orthopyroxene [Mierdel et al., 2007]. Other possible mechanisms for creating a thin decoupling zone beneath downgoing slabs may be related to the power law rheology of the mantle, the possible presence of partial melt in the asthenosphere, or mechanical anisotropy, which has been shown to have a significant effect on mantle flow in numerical models [e.g., Lev and Hager, 2008]. We emphasize that the basic requirement of our slab mantle flow model is that a thin decoupling zone must be present; we remain open to alternative models that could accomplish this. The class of decoupling models that invoke the entrainment of a thin layer of weak asthenosphere beneath the slab includes the shear heating model explored here, but could also include models in which the asthenosphere is defined by the presence of partial melt or is rich in volatiles. The shear heating model is one possibility, but it is not the only one.

The suggestion that the asthenosphere beneath oceanic plates has characteristics (such as excess temperature or partial melt) that allow it to be entrained beneath subducting slabs is consistent with recent studies of the lithosphere-asthenosphere boundary (LAB) beneath oceans [Rychert and Shearer, 2009; Kawakatsu et al., 2009]. In particular, Kawakatsu et al. [2009] imaged a sharp LAB beneath the subducting Pacific slab beneath Japan using receiver functions; their preferred model to explain their observations invokes partial melt as a cause for the sharp reduction in shear wave velocity across the LAB. The proposed existence of partial melt in the asthenosphere is consistent with the excess temperatures predicted by our shear heating model, and an asthenosphere that contains partial melt would be able to be entrained beneath subducting slabs to provide subslab decoupling. In general, models for subslab asthenospheric entrainment, including the shear heating model proposed here, raise questions about the fate of the suboceanic asthenosphere at subduction zones, and, more broadly, about the nature of the asthenosphere itself. Many of the questions raised by the inference of a thin decoupling zone beneath subduction zones relate to larger questions about how the asthenosphere is defined, what controls its physical properties, and what is its fate when the overlying lithosphere begins to descend into the mantle at subduction zones. Many of these issues are poorly understood, and represent important avenues for further research.

7. A Constraint on Mass Transfer Between the Upper and Lower Mantle?

The model proposed here invokes a decoupling zone beneath the slab and a partial barrier to flow at depth, as first proposed by Russo and Silver [1994] for South America. Although there is ample evidence from global seismic tomography that some slabs penetrate the transition zone and sink into the lower mantle [e.g., van der Hilst et al., 1997; Li et al., 2008], the extent to which such slabs entrain the surrounding mantle, and therefore the amount of material flux across the transition zone and the degree of mixing between the upper and lower mantle, remain hotly debated [e.g., Tackley, 2008]. Our model for subslab flow in subduction zones suggests that in cases where the downgoing slab is decoupled from the underlying mantle, this partial barrier to flow forces upper mantle material to flow parallel to the trench if the trench is moving in the reference frame of the convecting mantle.

If our model correctly describes the subslab flow field then there are important implications for large-scale mantle dynamics, in particular for the amount of mass transfer across the transition zone and the extent to which mantle convection is well mixed. The qualitative nature of the convecting mantle continues to be vigorously debated; end-member models invoke either whole mantle convection and mixing or a nearly impenetrable barrier to flow at the base of the transition zone that separates the mantle into two distinct layers. Attempts to reconcile geochemical and geophysical observations have resulted in a range of intermediate models, many of which involve exchange between the upper and lower mantle which is limited to slab flux only [Silver et al., 1988] or invoke partial or intermittent layering [e.g., Tackley, 2008].
Constraints from seismic tomography have been incorporated into the mantle mixing debate, as the predictions made by mantle mixing models are compared to the “snapshot” of present-day mantle structure provided by tomographic inversions for seismic wave speed [e.g., Tanimoto and Lay, 2000]. However, while global tomographic techniques can image structures such as subducting slabs at depth, they cannot directly image active mantle flow. The constraints on mantle flow provided by observations of seismic anisotropy are more direct than those provided by tomography, and the constraints on the subslab mantle flow field provided by shear wave splitting observations imply that even though slabs may penetrate the lower mantle, they are unlikely to entrain much surrounding mantle material with them. This, in turn, suggests that the material flux associated with subductions slabs that penetrate into the lower mantle may simply represent the slab flux. The constraints on the character of mantle downwellings provided by the observation that the subslab flow field is dominated by horizontal rather than entrained flow may help to constrain models of mantle convection and mixing; the apparent restriction on mass transfer between the upper and lower mantle reservoirs suggested by our model could help to discriminate, for example, among numerical models of mantle convection [e.g., Tackley et al., 2005]. The barrier to entrained flow required by our model tends to support mantle convection models that have limited material flux across the 660-km discontinuity, such as models that invoke “leaky layering” (see the overview by Tackley [2008]). Similar support for the idea that mass flux across the 660-km discontinuity is limited has been offered by studies of the heterogeneity spectrum of tomographic models, which suggest that there may be an abrupt change in this spectrum across the discontinuity at the base of the transition zone [e.g., Dejewonski, 2005].

8. Discussion

The observation that shear wave splitting due to subslab anisotropy nearly always exhibits fast polarization directions that are parallel to the trench (Figure 3) is one of the most robust features of the global splitting data set in subduction zone regions. However, there are several important exceptions to this rule, and the significant nearly trench-parallel splitting attributed to subslab flow beneath the downgoing Juan de Fuca slab in the Cascadia subduction zone is the most notable and best constrained. Preliminary observations of strong, trench-parallel SKS splitting at stations overlying the Middle America subduction zone beneath Mexico [Subaio and Davis, 2007; Leon Soto et al., 2009], to the north of the region of the Middle America subduction zone that is included in our compilation, may also provide an example of trench-parallel fast directions beneath a subducting slab. An important question, then, is whether our model for subslab mantle flow can accommodate these exceptions. One way that subduction-parallel entrained flow beneath a subducting slab can be reconciled with our model is if the mechanism for decoupling the slab from the mantle material beneath it is not active. Both Cascadia and the northern and southern ends of the Middle America trench feature the subduction of very young lithosphere; the lithosphere entering both of these subduction zones is between 5 and 10 Myr in age, and these represent the most extreme examples of the subduction of young lithosphere on the present-day Earth. In this paper, we have proposed a mechanism for the decoupling of slabs and subslab mantle material that invokes shear heating in the highly deformed asthenosphere and subsequent entrainment of a thin layer of asthenosphere beneath the downgoing slab that serves as a weak, rapidly deforming, decoupling zone. This mechanism depends on the asthenosphere experiencing a large amount of shear strain; for the case of very young seafloor, we hypothesize that the amount of strain needed for the shear heating mechanism to reach steady state has not yet been reached, and that decoupling between the slab and the subslab mantle is not accomplished.

If this hypothesis is correct, it has several important implications. This model makes testable predictions about the character of subslab shear wave splitting in regions where the subducting lithosphere is young, and would also predict an along-strike transition from trench-perpendicular to trench-parallel subslab splitting in subduction zones where the subducting lithosphere is generally young and there is an along-strike transition in its age. Future studies of along-strike variations in subslab splitting in such systems provides a direct test of this hypothesis, which should provide further constraints on the viability of our proposed model for subslab decoupling. Particularly interesting test cases could be provided by, for example, the subduction of the Cocos Plate beneath southern Mexico and Central America. In this subduction system, the age of the downgoing lithosphere varies from just over 5 Myr to the northwest to over 20 Ma for the region beneath Guatemala, El Salvador, and northwestern Nicaragua; seafloor ages then decrease again to the southeast, with ages of ~10 Ma beneath southern Costa Rica. Splitting results from the TUCAN experiment [Abt et al., 2005, 2009] indicate that subslab anisotropy is trench-parallel beneath the relatively older lithosphere that underlies stations located in Nicaragua. Our hypothesis would predict that the subslab splitting signal would rotate to trench-perpendicular fast directions for the youngest lithosphere in the northwesternmost and southeasternmost parts of the Central American subduction zone. This prediction could be tested by collecting and analyzing data from stations located above the youngest parts of the subducting Cocos slab. Preliminary data from the Rivera subduction zone (where the young Rivera microplate, a fragment of the Cocos plate, is subducting beneath North America) indicate that subslab anisotropy beneath the young (~<10 Ma) subducting plate is dominantly trench-parallel. If there is, indeed, a transition in subslab fast directions related to the age of the subducting lithosphere, then this would place constraints on the “critical distance” from the ridge needed for significant shear heating of young asthenosphere that would produce subslab decoupling. If our model for subslab decoupling is correct, this has significant implications for our understanding of the nature and dynamics of the asthenosphere, which are currently poorly understood [e.g., Rychert et al., 2005].

Another fruitful area for future work on the nature and implications of the subslab flow field is the integration of shear wave splitting constraints with numerical and analog models of subduction zone flow in the presence of trench migration and subslab decoupling. The model for
subslab mantle flow that we propose here is consistent with the global shear wave splitting data set, but it needs to be validated using geodynamical models to verify whether it is dynamically plausible and whether the relationships between trench migration and strength of subslab anisotropy can be reproduced. Laboratory models have been successful in producing trench-parallel flow above subducting slabs for systems with retreating trenches [e.g., Kincaid and Griffiths, 2003; Fuciniello et al., 2003] but such modeling studies have not emphasized the behavior of the subslab flow field. Issues such as the character of the flow field for models incorporating trench advance, as opposed to trench retreat, whether laboratory models can reproduce the splitting observations, and the effect of a subslab decoupling zone on mantle flow fields remain to be addressed. Another key question is to what extent trench-parallel strain induced by trench migration is localized directly beneath the slab, as opposed to being distributed over a larger region of the mantle. Some previous laboratory modeling work [Buttles and Olson, 1998] has indicated that mantle flow beneath subslab slabs is strong enough to locally produce trench-parallel anisotropy, but the character of this subslab flow field will depend on the viscosity structure of the mantle, the nature of the far-field boundary conditions, and other variables, and remains to be more thoroughly explored through laboratory and numerical modeling studies.

[61] As discussed in section 2, we believe that the subwedge splitting observations are most consistent with trench-parallel flow in the subslab mantle, rather than a primary contribution from the slab itself. However, alternative models that invoke anisotropy within the slab have been proposed; in particular, recent work by Faccenda et al. [2008] and Healy et al. [2009] has explored the feasibility of explaining SKS splitting data with a model for slab anisotropy that invokes the hydration and serpentinization of faults that penetrate the lithospheric slab. From an observational point of view, it is difficult to discriminate between slab versus subslab anisotropy in the subwedge splitting signal, so most of the arguments about which model better fits the observations are indirect. For several reasons, we believe that subslab flow does a better job of explaining the global subwedge splitting signal than slab anisotropy due to hydrated faults. As discussed in section 4, the lack of a correlation between the strength of subwedge splitting and slab age (which correlates with both slab thickness and, perhaps, with the extent of fault hydration) argues against a primary contribution to the global subwedge splitting signal from the slab. It is also difficult to explain the global variability in subwedge split times with a slab anisotropy model; the strength of subwedge splitting varies from little to none (e.g., Ryukyu or the Aleutians) to up to 1.5–2 s (Tonga-Kermadec or Calabria), and there is no obvious mechanism to explain this variability with the hydrated fault model. While Faccenda et al. [2008] argue that trench-parallel flow beneath subducting slabs is inconsistent with the results of analog and numerical modeling studies, we point out that in the studies cited, the decoupling between the slab and the subslab mantle required by our model was not taken into account. Additionally, if the subslab flow field is in fact dominated by entrained flow, as argued by Faccenda et al. [2008], then the anisotropy in the slab due to hydrated faults must be strong enough not only to explain the large subwedge splitting observed in regions such as Tonga-Kermadec, Calabria, or Middle America, but also to counteract (through destructive interference) the effect of trench-perpendicular splitting due to entrained flow beneath the slab. Finally, we point out that observations of azimuthal anisotropy beneath oceanic plates from both splitting observations and surface wave studies are very well explained by active flow in the upper mantle (that is, simple shear between the oceanic lithosphere and the underlying mantle) and that there is no obvious correlation with the geometry of (presumably hydrated) faults [e.g., Behn et al., 2004; Conrad et al., 2007; Becker, 2008]. This observation seems to be contrary to the predictions of a hydrated fault model: if hydrated faults in downgoing slabs indeed control the SKS splitting observed above subduction zones, then the SKS splitting signal within oceanic plates should presumably be controlled by faults as well. While hydrated faults in the upper 10 km of subducting slabs may make some contribution to shear wave splitting in some regions, this mechanism does not appear to be the primary control on SKS splitting patterns in subduction zones globally.

[62] On the basis of only the compilation of subwedge anisotropy parameters presented in this paper, we cannot entirely rule out alternatives to the model we propose such as the hydrated fault model of Faccenda et al. [2008] or the model proposed by Jung et al. [2009], in which B-type olivine fabric predominates beneath the slab and the observed trench-parallel fast directions are due to 2-D entrained flow. However, our trench-migration-controlled model seems to be the most consistent with the global subwedge splitting patterns documented here. Specifically, our model is able to explain the variations in subwedge delay time, which varies globally from ~0 s to ~1.5 s or more; the correlation we have documented between subwedge delay time and trench migration velocities in a Pacific hot spot reference frame (Figures 5 and 8) is significant at the 99%+ confidence level. The model proposed by Long and Silver [2008] and elaborated upon here can also explain the “exceptions” in the global data set where the subwedge splitting is apparently trench-perpendicular or oblique, such as Cascadia or Mexico.

9. Conclusions

[63] We have presented a global compilation of subslab splitting parameters and have tested for relationships between those parameters and parameters that describe subduction, with the aim of determining the dominant properties of the subslab flow field on a global scale. We have found that globally, the subslab region is nearly always dominated by trench-parallel fast directions, and we have identified a striking correlation between the strength of subslab anisotropy and the trench migration rate in a Pacific hot spot reference frame. We did not identify any other global correlations that may be identified with first-order contributions to the global subslab splitting signal. These comparisons support a model in which the subslab flow field is controlled by the migration of trenches with respect to the convection mantle, as previously proposed by Long...
and Silver [2008]. Such a model requires a thin decoupling zone beneath slabs; we hypothesize that this is accomplished by entraining a thin layer of hot, buoyant asthenosphere, which in our model represents ambient mantle that has been subjected to shear heating during deformation, resulting in a viscosity reduction and temperature increase. Our model also requires a partial barrier to flow, likely at the base of the transition zone, which permits downdragging slabs to penetrate into the lower mantle but which prevents the entrainment of the surrounding mantle material. Further tests of our model for the creation and demise of oceanic asthenosphere will include the investigation of splitting parameters in regions where young lithosphere is being subducted and characterizations of the mantle temperature beneath oceanic lithosphere from petrology and seismic constraints.

[64] Acknowledgments. We thank David Abt, Bill Holt, Chris Kincaid, Laurent Montesi, Jim Ni, Ray Russo, David Stegmam, Wouter Schellart, and Jessica Warren for helpful and thought-provoking discussions. Wouter Schellart kindly provided us with his compilation of trench anisotropy, without which this work would not be possible. We are grateful to the thoughtful and constructive comments provided by two anonymous reviewers and by the Associate Editor, which helped to improve the manuscript considerably. This work was supported by the Department of Terrestrial Magnetism, Carnegie Institution of Washington.

References