1 **Global large deep-focus earthquakes: source process and cascading**
2 **failure of shear instability as a unified physical mechanism**

3

4 **Authors:** Yu Chen^{1*}, Lianxing Wen^{1,2}

5

6 **Affiliations:**

7 ¹ Department of Geosciences, State University of New York at Stony Brook, Stony Brook,
8 New York, 11794, USA.

9 ² Laboratory of Seismology and Physics of Earth's Interior, University of Science and
10 Technology of China, Anhui, 230026, China.

11 *Corresponding author: Yu Chen, SUNY at Stony Brook, Stony Brook, New York,
12 11794, USA. (yu.chen@stonybrook.edu; esschenyu@gmail.com)

13

Abstract: We apply a multiple source inversion method to systematically study the source processes of 25 large deep-focus (depth > 400 km) earthquakes with $M_w > 7.0$ from 1994 to 2012, based on waveform modeling of P, pP, SH and sSH wave data. The earthquakes are classified into three categories based on spatial distributions and focal mechanisms of the inferred sub-events: 1) category one, with non-planar distribution and variable focal mechanisms of sub-events, represented by the 1994 M_w 8.2 Bolivia earthquake and the 2013 M_w 8.3 Okhotsk earthquake; 2) category two, with planar distribution but focal mechanisms inconsistent with the plane, including eighteen earthquakes; and 3) category three, with planar distribution and focal mechanisms consistent with the plane, including six earthquakes. We discuss possible physical mechanisms for earthquakes in each category in the context of plane rupture, transformational faulting and shear thermal instability. We suggest that the inferred source processes of large deep-focus earthquakes can be best interpreted by cascading failure of shear thermal instabilities in pre-existing weak zones, with the perturbation of stress generated by a shear instability triggering another and focal mechanisms of the sub-events controlled by orientations of the pre-existing weak zones. The proposed mechanism can also explain the observed great variability of focal mechanisms, presence of large values of CLVD (Compensated Linear Vector Dipole) and super-shear rupture of deep-focus earthquakes in the previous studies. In addition, our studies suggest existence of relationships of seismic moment $\sim$ (source duration)³ and moment $\sim$ (source dimension)³ in large deep-focus earthquakes.

Keywords: deep-focus earthquake, multiple source inversion, cascading failure, shear thermal instability, pre-existing weak zone

1. Introduction

Physics of deep-focus earthquakes remains enigmatic, since deep-focus earthquakes occur at the depth below 400 km where the pressure would strongly inhibit brittle failure and the temperature would result in ductile deformation (Scholz, 2002). Various models have been proposed to explain the unexpected occurrence of deep-focus earthquakes, including dehydration embrittlement, transformational faulting and shear thermal instability. Dehydration embrittlement suggests that deep earthquakes may be triggered by embrittlement accompanying dehydration of hydrous minerals (such as serpentine) (Jung et al., 2004; Meade and Jeanloz, 1991), although it is unclear if the water content below the 400 km depth can support this mechanism. Transformational faulting mechanism hypothesizes that a meta-stable olivine wedge exists in the mantle transition zone and the phase transformation can lead to catastrophic earthquake failure (Green, 2007; Green and Burnley, 1989; Green and Houston, 1995; Kirby, 1987; Wiens et al., 1993). And, shear thermal instability suggests that rapid deformation can be achieved (especially, in pre-existing rheological weak zones), when a stress or temperature perturbation promotes a positive feedback between shear heating and temperature-dependent rheology (Hobbs and Ord, 1988; Karato et al., 2001; Kelemen and Hirth, 2007; Ogawa, 1987).

Studies of source processes of large deep-focus earthquakes provided important constraints on deep-focus earthquake mechanisms. During the past 20 years, various methods have been applied to study source processes of deep-focus earthquakes. Rupture directivity is analyzed to determine fault plane orientations of deep-focus earthquakes occurring in the Tonga-Kermadec subduction zone (Warren et al., 2007). Both sub-

horizontal and sub-vertical fault planes are detected, corresponding to a complex rupture system with either a reactivated or new generated fault system (Warren et al., 2007). Super-shear rupture was also reported for an Mw 6.7 aftershock of the Okhotsk earthquake (Zhan et al., 2014b). A number of large deep earthquakes, especially the 1994 Mw 8.2 Bolivia and 2013 Mw 8.3 Okhotsk deep earthquakes, have also been studied by multiple source inversion (Chen et al., 2014; Silver et al., 1995; Zhan et al., 2014a) and finite fault inversion (Antolik et al., 1996; 1999; Estabrook, 1999; Wei et al., 2013; Ye et al., 2013). The source processes of these two largest ever-recorded deep earthquakes provided particularly important insights on the physical mechanisms of deep earthquakes. Both the Bolivia earthquake (Silver et al., 1995; Zhan et al., 2014a) and the Okhotsk earthquake (Wei et al., 2013; Ye et al., 2013) were determined to have a large source region that totally cut out of the assumed meta-stable olivine zone, seemingly excluding the transformation faulting as a possible mechanism. In a recent study, we showed that the Okhotsk earthquake consists of six sub-events that cannot spatially be fit into plane rupture and have variable focal mechanisms; we further suggested that the source process can be best explained by a cascading failure of shear thermal instabilities distributed at different depths, with one shear instability triggering another (Chen et al., 2014).

Though the past studies have provided important constraints on deep earthquake physics, several issues remain to be resolved. 1. Do deep-focus earthquakes rupture in a single fault plane? 2. Do deep-focus earthquakes behave the same between hot and cold subducted slabs? 3. Is there a unified mechanism to explain all the seismic results? In this study, we perform a systematical study of source processes of the global large deep-focus earthquakes with a moment magnitude greater than 7 and depth larger than 400 km,

based on waveform modeling of direct P and SH waves and near-surface reflected pP and sSH waves through multiple source inversion. Combining the direct and near-surface reflected phases places a tight constraint on the depth distribution of the seismic energy release during these earthquakes (Chen et al., 2014). The multiple source inversion also makes no *a priori* assumption on the focal mechanism of the sub-events in waveform modeling and is able to resolve varying focal mechanisms in the source process. In the following sections, we present event information and seismic data in section 2, multiple source inversion method in section 3, multiple source inversion results in section 4, a unified physical mechanism of shear thermal instability for explaining the seismic results in section 5, and implications to the observed focal mechanism variation and super-shear rupture of deep earthquakes in section 6.

2. Seismic data

We search the events with $M_w > 7.0$ and depth > 400 km from 1994 to 2012. There are a total of 28 earthquakes matching the criterion (Fig. 1 and Table S0). The studied events are well recorded by the Global Seismographic Network (GSN). The data consist of broadband compressional waves (P), near-surface reflected compressional waves (pP), transversally polarized shear waves (SH) and near-surface reflected transversally polarized shear waves (sSH) recorded at tele-seismic distances between 30° and 90° . The selected data constitute good azimuthal coverage for all the events (Figs. S1a – S25a). As we will mention later, the combination of direct P and SH waves with the near-surface reflected pP and sSH waves places a tight constraint on depth extent of seismic source. The displacement seismograms are deconvolved with their respective instrument responses and bandpass filtered between 0.01 and 4 Hz. Our source models

are constrained mostly by the seismic energy in a lower frequency band, but the use of the high-frequency data provides better arrival time reading of the P wave onsets while producing little artifacts in the inversions (supplementary information).

Fig. 1. Distribution of the global large deep-focus earthquakes ($M_w > 7$ and depth > 400 km) from 1994 to 2012, and their GCMT focal mechanisms, labelled with event numbers in Table S0. The earthquakes labelled with green, blue and red represent those in categories one to three as classified in Section 4. The events labelled with red numbers are the example events presented in the main text. Three earthquakes with black beach balls are not studied (see text for the reasons).

3. Multiple source inversion method

3.1. Method procedure

We apply a multiple point source inversion procedure to infer the source processes based on waveform fitting of direct phases P/SH and near-surface reflected phases pP/sSH (Chen et al., 2014). The method treats a large earthquake as a combination of multiple double couple point sources (sub-events) separated in space and time, and

resolves the spatio-temporal separation and focal mechanism of each sub-event by modeling the seismic data. Each sub-event is represented by nine parameters: seismic moment, seismic duration (the width of the source time function represented by a half sine function), strike, dip and slip for focal mechanism, dt for time separation of origin time from the initiation time, and de , dn and dz for distance separations from the initiation point in east-, north- and vertical-direction, respectively. Seismic Green's functions are computed by the Generalized Ray Theory method (Helmberg, 1968), based on the velocity and attenuation structures of PREM (Preliminary Reference Earth Model, downloaded from IRIS EMC-PREM) (Dziewonski and Anderson, 1981). The seismic response of each sub-event is obtained by convolving the Green's function with source time function (Ji et al., 2002). The $9 \times N$ (N , number of sub-events) model parameters are searched by the simulated heat annealing algorithm for the best-fitting solution that generates the least root mean square (RMS) residual between the seismic data and synthetic waveforms (Chen et al., 2014).

3.2. Importance for depth constraint by including the near-surface reflected waves

Most of previous studies inverted the source process using P waves only, which would have poor resolution in determining depth distribution of the seismic energy release during the earthquake, as the inferred depth of energy release trade-offs with the inferred occurring time of the energy release (Fig. 2). Combining pP wave with P wave data into the seismic inversion would place a tight constraint on the depth of seismic energy release, as pP waves take off from the opposite hemisphere of the source from the P waves (Fig. 2a).

We illustrate the necessity of combining P and pP waves to resolve the depth of propagating seismic energy release, using the following synthetic examples of a source consisting of two sub-events. We generate synthetics of sub-event 1 and a delayed sub-event 2 (black stars in Fig. 2b) at a group of hypothetical stations at different azimuths. We also generate synthetics for another sub-event 2 (red star in Fig. 2b) with a greater source depth and a delayed occurring time. These two sub-event 2s produce synthetics that exhibit only slight time offsets for P waves at various epicenter distances (Fig. 2c). These travel time differences are not resolvable in the real data inversion. In another word, strong trade-offs exist between source depth and event occurring time when only P waves are used in the inversion. However, the deeper sub-event 2 (red star in Fig. 2b) generates travel time delays of pP wave up to 3 s (red traces in Fig. 2d) in comparison with the shallower sub-event 2 (black star in Fig. 2b), since the deeper source has a longer path for pP wave. The above example indicates that combining direct waves and near-surface reflected waves is critical for resolving source distribution in depth.

Fig. 2 (a) Ray paths for P (green path) and pP (blue path) waves from earthquake source (star) to a seismic station (triangle). (b) Sub-event 1 (black star), sub-event 2s (red and black star) and stations (red triangles) used for synthetics calculations (sub-event 2 in red is located 20 km deeper,

but with a 1.5 s later occurring time than sub-event 2 in black). (c) P synthetics generated using a source that consists of sub-event 1 and sub-event 2 in black in (b) (black traces) and a source that consists of sub-event 1 and sub-event 2 in red in (b) (red traces). Synthetics are aligned along the maximal amplitudes of the seismic phase associated with sub-event 1. The pulses of two sub-events are labelled accordingly with their numbers. (d) Same as (c), except for pP synthetics.

4. Source processes of global large deep-focus earthquakes

We apply our multiple source inversion procedure to study the source processes of 28 deep-focus earthquakes. The observed waveforms of 25 of those earthquakes can be modeled reasonably well, using one to eight sub-events (green, blue, and red beach balls in Fig. 1). Three earthquakes (E26, E27 and E28, black beach ball in Fig. 1) are excluded in the present study. The seismic waves of event E27 are mixed with that of event E24 which occurred 24 minutes earlier, while E26 and E28 generated too complex waveform to be modeled by our method.

We align broadband P phases by the first arrival onsets, which can be clearly identified and picked in the high-pass filtered data above 1Hz. SH, pP and sSH phases are aligned iteratively in the inversion process: we first pick them based on IASP91 model (Kennett and Engdahl, 1991) and perform a preliminary inversion with P wave heavily weighted over other phases. The alignments of pP, SH and sSH phases are adjusted based on the best match between synthetics and the data during the inversions, until no further adjustment is observed from the previous inversion. The final inversion weights each phase equally. Synthetics tests indicate that such procedure of phase alignment is effective even in the presence of noise, large initial misalignment and waveform distortions due to attenuation or multi-path effects (supplementary information).

184 For each event, we also test whether the seismic data can be compromised with
185 rupture in a single fault plane. We test three planar models: the two GCMT (Global
186 Centroid Moment Tensor) fault planes and the plane that best geographically connects the
187 inverted best-fitting sub-events in the multiple source inversion. Obviously, the best-
188 fitting **sub-event** models of the multiple source inversion would always fit the data better
189 than any planar models, but we test if the better data fitting is significant. We define that
190 the earthquake can be modeled by a single fault plane if the misfit difference between the
191 best planar model and the best-fitting sub-event model is less than 5%, a threshold above
192 which better waveform fitting can be visually noticeable at stations around the nodal
193 planes. The synthetics (blue traces) of the best planar models are presented in Figs. S1 –
194 S25, along with the synthetics (red traces) predicted by the best-fitting sub-event models.

195 The inversion results of the 26 earthquakes can be divided into three categories
196 based on the focal mechanisms and spatial distributions of the inferred sub-events: 1)
197 category one, with non-planar spatial distribution and variable focal mechanisms of the
198 sub-events, represented by the Bolivia earthquake (E2, green beach ball in Fig. 1) and the
199 2013 Mw 8.3 Okhotsk deep earthquake (Chen et al., 2014), 2) category two, with planar
200 distribution but fault mechanisms of the sub-events inconsistent with the best-fitting
201 plane, including eighteen earthquakes (blue beach balls in Fig. 1), and 3) category three,
202 with planar distribution and focal mechanisms of the sub-events consistent with the plane,
203 including six earthquakes (red beach balls in Fig. 1). Earthquakes in categories one and
204 two have greater moment magnitudes, longer source durations and larger source regions.
205 We discuss the main features of the source process in each of the three categories, using
206 an earthquake in each category as example. The detailed results of all the earthquakes are

presented in the supplementary information with descriptions, tables of the best-fitting sub-event models (Tables S1 – S25) and the waveform fitting comparisons between the synthetics predicted by the best-fitting sub-event models (red traces) and the best planar models (blue traces) (Figs. S1 – S25). So are the results of our previous study of the 2013 Mw 8.3 Okhotsk earthquake (Chen et al., 2014), in Table S26 and Fig. S26.

4.1. Category 1: Non-planar spatial distribution of sub-events (Bolivia and Okhotsk)

The two largest ever recorded 2013 Okhotsk earthquake (Chen et al., 2014) and 1994 Bolivia earthquake exhibit large depth extensions of the sub-events. The sub-events of these two earthquakes are not distributed in a plane and exhibit variable focal mechanisms. We have reported the results related to the Okhotsk earthquake in our previous study (Chen et al., 2014). The Okhotsk earthquake consists of six sub-events with a horizontal extension of 64 km and a vertical extension of 35 km (Fig. S26b). The geographic distribution and focal mechanisms of the inferred sub-events and fore/aftershocks locations do not fit into plane rupture (Chen et al., 2014). We focus on discussing the results of the Bolivia earthquake here.

The Bolivia earthquake has been well studied (Antolik et al., 1996; Goes and Ritsema, 1995; Silver et al., 1995; Zhan et al., 2014a). All the previous studies have suggested that the Bolivia earthquake ruptured on the GCMT sub-horizontal plane (Silver et al., 1995; Zhan et al., 2014a). Seismic data used in those studies included mostly P waves, with exceptions of Goes and Ritsema (1995) where both P and pP waves were used at the stations around the azimuth of 310°, and Antolik et al (1996) where three pP observations were used. In this study, we revisit the issue by joint inversion of the P and

pP waves recorded globally. We find that the seismic P and pP waves of the Bolivia earthquake exhibit a vertical extension during the earthquake, inconsistent with rupture on the GCMT sub-horizontal plane (Fig. 3).

The earthquake begins with an ~10 s weak initiation phase, followed by strong moment release lasting ~35 s. As the present study deals with major source process of deep-focus earthquake, the initiation phase is ignored and ignorance of such phase does not affect the seismic results presented here (supplementary information). The observed P waveforms exhibit strong azimuthal variations with five major seismic pulses during the strong moment release, discernible with different move-outs (peak times labelled 1 to 5 in black, Fig. 3a). The pP waveforms exhibit similar azimuthal variations as the P waveforms (peak times label with 1 to 5 in red, Fig. 3a). However, pulses of sub-events 4 and 5 are delayed in pP waveforms than in P waveforms, and the duration of pP waveforms is about 7-8 s longer (Fig. 3a). We show five example pairs of P and pP seismograms observed at stations TAM, SUR, SPA, RAR and COR, to further illustrate this feature (Fig. 3b). The first two pulses present almost the same arrival times in P and pP waveforms, corresponding to the same source depths. The fourth and fifth pulses exhibit 3-4 s and 7-8 s delays in pP waveforms than in P waveforms, corresponding to event sources that are 15-20 km and 35-40 km deeper than the earthquake initiation point.

In the inversion, we adopt two sub-events 3a and 3b to better fit the observed waveform features in the time window labelled between 3 and 4 in Fig. 3. The Bolivia earthquake can be modeled by 6 sub-events (Figs. 4, Table S2), with a total duration of 44 s. The source has a horizontal extension of 44 km and a vertical extension of 38 km.

251 The inverted fault plane orientations of the sub-events are less than 10° different from the
 252 GCMT solution (Table S2). In comparison, the source of the Okhotsk earthquake has a
 253 duration of 28.5 s, a horizontal extension of 64 km and a vertical extension of 35 km
 254 (Chen et al., 2014). The sub-event durations of both earthquakes are about 10 s. We
 255 perform the planar tests for the Bolivia earthquake by compelling the sub-events on three
 256 planar rupture models. The sub-horizontal plane is determined to be the best planar model.
 257 However, the sub-horizontal plane model predicts synthetics arrival times that match the
 258 data in the early part of the seismograms (associated with sub-events 1, 2, 3a and 3b), but
 259 not in the later portion of the seismograms (associated with the last two sub-events) (Fig.
 260 4c). For example, in comparison with the observations (black traces), the last sub-event in
 261 the best planar model (synthetics in blue traces) predicts later P waves at stations TAM
 262 and HRV, but earlier pP waves at stations PAB, RAR and COR (Fig. 4c). Synthetics
 263 waveform amplitudes of the model also misfit the early portion of the seismograms
 264 associated with the first sub-event. (e.g., the P wave at station RAR, Fig. 4c). The GCMT
 265 sub-vertical plane model generates late P and pP arrival times around the azimuth of 0°
 266 and early arrival times around the azimuth of 180° . And, while the plane that best
 267 connects the spatial distribution of the inferred sub-events (with a strike of 281° and a dip
 268 of 45°) can predict correct arrival times for most of the sub-events except at the time
 269 window that is associated with sub-events 3a and 3b, it generates synthetics that do not fit
 270 the azimuthal variation of seismic amplitudes in the data. These modeling results suggest
 271 that the Bolivia earthquake, like the Okhotsk earthquake, is a combination of sub-events
 272 that are distributed at different depths and out of a plane, and have variable focal
 273 mechanisms.

274 Although the horizontal distribution of the sub-events in our study is similar to
275 those previous studies using P waves (e.g., Silver, et al., 1995, Zhan et al., 2014a), our
276 results indicate that the sub-events 4 and 5 are 18 and 38 km deeper, and the spatial
277 distribution of the sub-events is inconsistent with rupture on the GCMT sub-horizontal
278 plane. This difference in the seismic results lies on that only the P waves were used in
279 those studies (Silver et al., 1995; Zhan et al., 2014a), and P waves alone do not have
280 source depth resolution (Fig. 2). We discuss the discrepancies between our results and
281 those of the two previous studies using P and pP wave observations below. In the study
282 of Antolik et al. (1996), they used three pP wave observations and only tested two plane
283 models from the GCMT solution. Their best-fitting plane model (the GCMT sub-
284 horizontal plane) predicted earlier pP wave than the data recorded at station ALE (Figure
285 4 in Antolik et al., 1996), which implies inaccuracy in the inferred source depths. In the
286 study of Goes and Ritsema (1995), they concluded that similar durations were observed
287 between the stacked P and pP waveforms around the azimuth of 310°. However, we
288 should note that the first sub-event is weak in the azimuth range between 310° and 70°
289 (Fig. 3a), which would mislead the comparison of the durations of P and pP waves. At
290 other azimuths, the observed pP waveforms clearly exhibit longer durations than P
291 waveforms (Fig. 3).

Fig. 3. (a) Comparison of observed P (black traces) and pP (red traces) displacement seismograms as a function of azimuth. The black and red lines indicate the peak times of the sub-events for P and pP, respectively. (b) Overlays of P and pP displacements observed at five example stations: COR, RAR, SPA, SUR and PAB. The black and red lines mark the maximal of the first, second, fourth and fifth sub-events labelled in (a). For the display purpose, the polarities of some records have been flipped.

Fig. 4. (a) Locations of the best-fitting sub-events (red points), the best planar model (blue points, the inferred sub-events compelled to the sub-horizontal plane), and the initiated location (black star) of the 1994 Bolivia earthquake (E2), along with slab depth contours (black traces, labelled with slab depths). Black beach ball describes GCMT solution of the event. Red and blue beach balls represent inferred focal mechanisms of the sub-events of the best-fitting sub-event model and the best planar model, labelled with the sub-event peak times/moment magnitudes/depths. (b) Three-dimensional spatial distribution and focal mechanisms of inferred sub-events of the best-fitting sub-event model (red beach balls) and the best planar model (blue beach balls) labelled with their occurring sequence. (c) Selected observed seismograms in Fig. S2 (black traces) comparing with synthetics (red traces) predicted by the best-fitting sub-event model and synthetics (blue traces) predicted by the best planar model. The red and blue lines label the peak times of the sub-events predicted by the best-fitting sub-event model and the best planar model.

4.2. Category 2: Planar spatial distribution of sub-events, but with focal mechanisms inconsistent with the plane (18 events)

The geographic distribution of the inferred sub-events of category two earthquakes (E1, E3, E5, E8, E9, E11, E12, E13, E15, E16, E17, E19, E20, E21, E22, E23, E24 and E25) can fit into a plane, but the focal mechanisms of the sub-events are inconsistent with the plane orientations. Most of the fault plane orientations of the sub-events deviate less than 20° from the fault plane orientations of the GCMT solutions, but some up to 50° (e.g., E1, Table S1). Synthetics of the best planar model exhibit large amplitude misfits to the seismic data (and even different polarities) at some stations (Fig. 5c). The hypocenter depths of this group of earthquakes range from 425 to 675 km, their moment magnitudes between 7.0 and 7.9, total source durations from 8 to 26 s, and the maximum distances among the sub-events between 15 and 70 km. The typical sub-event durations of these earthquakes are between 2 and 8 s.

We present the results of event E19 beneath Tonga region as example. This event consists of three sub-events, with inverted focal mechanisms (strike/dip/rake) of 46/71/262, 51/79/258 and 78/60/200 (Fig. 5a, Table S19). The best-fitting sub-events of the earthquake are distributed around the GCMT sub-vertical fault plane (Fig. 5, Table S19). The first two sub-events are roughly consistent with 49/76/252 of the GCMT solution, but the third one is $\sim 30^\circ$ different from the GCMT sub-vertical fault plane. We compare the best-fitting sub-event model (red traces) and the best planar model (blue traces) that compels the sub-events distributed on the GCMT sub-vertical plane (Fig. 5, S19). The blue traces can fit the early part of the data (corresponding to the prediction of the first two sub-events) at most of the stations, but significantly misfit the latter portion of the data that corresponds to the pulse associated with the third inferred sub-event. The azimuthal distribution of the latter portion of the seismograms would require a source that has a quite different radiation pattern from the early sub-events and from the GCMT solution (Fig. 5c).

Fig. 5. Similar as Fig. 4, except for event E19, with the best planar model being the GCMT sub-vertical plane. (c) Selected comparisons between data (black traces) and synthetics based on the best-fitting sub-event model (red traces) and the best planar model (blue traces) in Fig. S19.

4.3. Category 3: Planar spatial distribution and focal mechanisms of sub-events consistent with the plane (6 events)

The spatial distribution of the inferred sub-events of the category three earthquakes (E4, E6, E7, E10, E14, and E18) fits into one plane and the focal mechanisms of the inferred sub-events are consistent with the best-fitting plane. Naturally, synthetics of the best planar model can fit the data as well as that of the best-fitting sub-event model. The hypocenter depths of this group of earthquakes range from 495 km to 610 km, and their moment magnitudes between 7.0 and 7.4. The extent of the lateral

distribution of these events varies from 7 to 25 km, and the total durations of the earthquakes in this category from 4 to 14 s. The typical sub-event durations of these earthquakes are between 2 and 4 s.

We present the results of event E6 beneath Tonga region as example. The earthquake can be modeled by four sub-events with a horizontal extension of 25 km (Table S6, Fig. 6a). The first three sub-events are distributed on both the GCMT fault planes (along the interface of the two planes) with focal mechanisms consistent with the GCMT solution (Fig. 6b). The fourth sub-event is weak and its location is not well constrained. In the test of planar models, both GCMT fault planes could model the data reasonably well, with the sub-vertical plane slightly better (Figs. 6c, S6). Previous studies also reached different conclusions on the best-fitting plane, between the sub-vertical plane (Warren et al., 2007) and sub-horizontal plane (Antolik et al., 1999).

Fig. 6. Similar as Fig. 4, except for event E6, with the best planar model being the GCMT sub-vertical plane. (c) Selected comparisons between data (black traces) and synthetics based on the best-fitting sub-event model (red traces) and the best planar model (blue traces) in Fig. S6.

5. Physical mechanism of deep-focus earthquakes

5.1. Possible physical mechanisms for earthquakes in each category

The 2013 Okhotsk earthquake and the 1994 Bolivia earthquake in category one consist of sub-events with non-planar distribution. Our previous study indicated that the geographic distribution and focal mechanisms of the inferred sub-events of the Okhotsk

earthquake and its fore/aftershocks locations do not fit into plane rupture (Chen et al., 2014). The source geometry of the event does not prefer phase transformational faulting, since it is not consistent with the meta-stable olivine wedge (Chen et al., 2014; Meng et al., 2014). Previous studies suggested that the Bolivia earthquake ruptured on the GCMT sub-horizontal plane and the inferred lateral dimension of the Bolivia event was larger than the hypothesized meta-stable olivine wedge, making the phase transformation faulting unlikely for the event (Silver et al., 1995). Our seismic results are inconsistent with a plane rupture, either on the GCMT sub-horizontal plane or on the plane that best connects the sub-event locations. The spatial distribution of the sub-events of the Bolivia earthquake is inconsistent with the likely slab geometry in the region (Fig. 4a) and the lateral extent of the source distribution of the event is estimated to be 50 km (maximum separation distance between the sub-events plus the rupture length of sub-events with 10 s duration and half of local shear wave velocity), exceeding the 10-20 km thick meta-stable olivine wedge estimated based on theoretical estimation of the temperature condition within the South American slab (Silver et al., 1995). The best-fitting sub-event model can be best explained by a cascading failure of shear instabilities with focal mechanisms similar as the GCMT solution.

Eighteen earthquakes in category two consist of sub-events with planar (the GCMT fault planes) distribution, but with focal mechanisms inconsistent with the plane. Those earthquakes cannot be interpreted as rupture in a fault plane. Spatial distributions of the sub-events of some of these earthquakes are also inconsistent with the geometry of the assumed meta-stable olivine wedge. For example, the spatial distribution of the inferred sub-events of E8 has a horizontal extension of ~50 km with an ~10° dip to the

east, while the slab geometry is $\sim 70^\circ$ dip to the west (Fig. S8); and the spatial distribution of the inferred sub-events of E15 possess a horizontal dimension of ~ 70 km along $\sim N340^\circ W$ and a vertical extension of ~ 40 km with $\sim 80^\circ$ dipping to the east, while the slab geometry $\sim 60^\circ$ dips to the west along approximately the north direction. We suggest that the source processes of the earthquakes in this category cannot be explained by transformational faulting, similar to the inference based on the aftershock distribution of the 1994 Tonga earthquake (E1), which suggested the aftershock distribution totally cut out of the possible meta-stable olivine wedge (McGuire et al., 1997; Wiens et al., 1994). We suggest that the source processes of this category of earthquakes can also be best explained by a cascading failure of shear instability, same as category one. It is just that the shear thermal instabilities during this category of the earthquakes are aligned approximately along a plane, while those of the 2013 Okhotsk and 1994 Bolivia earthquakes are not.

Six earthquakes in category three (Tables S0, S4, S6, S7, S10, S14 and S18) consist of sub-events with planar distribution and focal mechanisms consistent with the plane. These events have relatively smaller magnitudes, source geometry and fewer sub-events. The inferred source processes of these earthquakes can be interpreted in many ways: rupture in a fault plane, transformational faulting or cascading failure in close space and time (as for the earthquakes in the first and second categories). It is interesting to note that the source processes still exhibit complex spatial and temporal relationship: two earthquakes (E4, E7) exhibit unilateral propagations (Tables S4, S7), two earthquakes (E10, E18) bilateral propagation (Tables S10, S18), one earthquake (E14)

backward propagation (Table S14), and one earthquake (E6) a backward-and-forward propagation (Fig. 6a, Table S6).

5.2 Cascading failure of shear thermal instabilities as a unified physical mechanism for large deep-focus earthquakes.

While multiple mechanisms exist to explain the inferred source processes of some groups of deep-focus earthquakes, the seismic results of all large deep-focus earthquakes can be explained by a unified physical mechanism of cascading failure of shear thermal instabilities occurring within pre-existing weak zones. Shear thermal instability is a failure due to positive feedback between viscous heating and temperature dependent rheology (Kelemen and Hirth, 2007). Such feedback system is naturally influenced by perturbation of stress. An earthquake would generate static or dynamic stress change in the region, may trigger a cascading failure of shear instability in close space and time (Hill, 2008; Tibi et al., 2003). Shear thermal instability can occur in a viscous medium with a uniform rheological property (Ogawa, 1987) or a pre-existing zone of different rheological properties (Kelemen and Hirth, 2007) which is known to exist within the subducted slab, due to either past faulting before subduction (Silver et al., 1995) or existence of preexisting fine-grained shear zones (Kelemen and Hirth, 2007). The focal mechanism of a shear instability that occurs in a uniform viscous medium would be controlled by the state of the medium before failure and the triggering force, while the focal mechanism of a shear instability that occurs in a medium with pre-existing weak zones would be controlled by the orientations of the pre-existing weak zones, as the pre-existing weak zones would prescribe the instability to occur in the weak zones. We

suggest that the inferred source processes of these earthquakes are consistent with shear thermal instabilities in the pre-existing weak zones for the following reasons: 1) focal mechanism changes during many large deep-focus earthquakes (the 2013 Okhotsk earthquake in category one and the earthquakes in category two). We argue that the triggering force for a single large earthquake unlikely trigger instabilities with various principle stress axes in a homogeneous viscous medium, but may do so in the pre-existing weak zones with various orientations; 2) it appears that the focal mechanisms of the sub-events are related to the locations of the sub-events, regardless of their hosting large-events. For example, the inferred sub-events of E9 beneath the Kuril-Izu-Bonin subducted slab exhibit two types of focal mechanisms, with one type (the second sub-event) similar to the focal mechanisms of the inferred sub-events of E11 (Fig. 7e). The second sub-event of E9 is 10 km south of E11 and 5 km to the north of the first sub-event of E9. There appears existent of location control of focal mechanism between south and north regions of the source region of E9. Location control of focal mechanism of the sub-events regardless of the hosting large earthquakes favors shear thermal instabilities in the medium with pre-existing weak zones over those in a homogeneous medium. In short, the inferred results of our studied events can be best explained by a cascading failure of shear thermal instabilities in pre-existing weak zones, with the perturbation of stress generated by a shear instability triggering another and the orientations of the pre-existing weak zones controlling the focal mechanisms of these instabilities.

In the context of shear instability in pre-existing weak zones as the physical mechanism of deep-focus earthquakes, we suggest that caution is needed to relate the focal mechanism of deep-focus earthquakes to the stress state in the subducted slabs. The

468 observed great variability of focal mechanisms of small and large deep-focus earthquakes
469 may just reflect existence of pre-existing weak zones with various orientations and may
470 not necessarily be related to the ambient stress in the region of the earthquake sources.

Fig. 7 Locations (points) and focal mechanisms (beach balls) of the inferred sub-events of the 25 deep earthquakes (labelled with event IDs as in Table S0 and inferred source propagating velocity ranges in parentheses) in four subducted slabs and two doublet-event areas, including (a) South American, (b) Sumatra-Java-Philippine, (c) Tonga-Kermadec, (d) Kuril-Izu-Bonin, (e) E9 and E11, and (f) E12 and E25, along with slab depth contours (black traces, labelled with slab depth). Sub-events are color coded with their hosting large earthquakes and labelled by the occurring sequence in the earthquake.

6. Discussion

Evaluating the uniqueness of our multiple-CMT solutions is difficult, given the non-linear nature of our inversion approach, and we cannot exclude the possibility that other solutions exist that may fit that data equally well. In addition, we have chosen point sources to explain various energy packets, which **could** have been resulted from finite rupture. However, it is nonetheless useful to provide a set of solutions computed in a uniform way that fit the main features seen in the direct and depth phases for a large number of deep earthquakes. Base on the seismic result, we suggest that our results can be explained by a unified mechanism of cascading failure of shear thermal instabilities in pre-existing weak zones.

The mechanism **of cascading failure of shear thermal instabilities** can also explain previously **noted** features of **some deep earthquakes: including** great variability **of focal mechanism in small distances, large** CLVD (Compensated Linear Vector Dipole) component **of the inferred moment tensor**, and super-shear rupture. It also provides new insights as well.

Previous studies suggested that focal mechanisms of deep-focus earthquakes exhibit great variability **between** different slabs or in one slab (e.g. Frohlich, 2006; Myhill, 2013). Yu and Wen (2012) also reported that the focal mechanism of the deep-focus

earthquakes can vary significantly within a small distance. It is further noted in this study that focal mechanism also varies significantly during the occurrence of most large earthquakes, as manifested in the source processes of the events in categories one and two. In addition, focal mechanism also varies significantly between large earthquakes occurring in a small distance. We illustrate this feature with the results of earthquake pairs that occurred in a close distance but with complex sub-event focal mechanisms (Fig. 7). We determine the relative location of the initiation points of these earthquake pairs based on differential onset times and a master relocation program (Wen, 2006). E25 is 20 km west to and 20 km deeper than E12 in the Tonga-Kermadec subducted slab (Fig. 7f). E12 occurred within normal faults, while E25 is dominated by strike-slip faulting (Tables S12, S25). E23 is 15 km west to and 20 km shallow than E22 in the Sumatra-Java-Philippine subducted slab (Fig. 7b). The two earthquakes occurred with slightly different focal mechanisms (Tables S23, S22). Cascading failure of shear thermal instabilities in pre-existing weak zones would explain these observed features well. In the context of this mechanism, the strong variability of the observed focal mechanism reflects complex orientations of the pre-existing weak zones.

The great variability of focal mechanisms during large deep earthquakes can also be the reason of the existence of CLVD component in the GCMT moment tensor solutions of these earthquakes. Kuge and Kawakatsu (1990) reported that the CLVD component of a deep earthquake can be interpreted by the summation of two different double couple sub-events. Later studies further suggested that the CLVD components of shallow and deep earthquakes can be explained by source irregularities (Kuge and Kawakatsu, 1992, 1993; Kuge and Lay, 1994). We compare the GCMT CLVDs and the

CLVDs in the summed solution over the inferred focal mechanisms of the sub-events (Fig. 8a). CLVDs are similarly present in the summed solution for the events with large GMT CLVDS, although CLVDs from the summed solutions are general smaller. The differences probably lie on the different ways to represent the earthquake source and different frequency contents of the data used between the studies. Such comparison suggests that GCMT CLVDs may just reflect the variability of focal mechanisms during a large deep-focus earthquake.

Previous studies also suggested that source behavior and physical mechanisms may depend on slab temperature (Wiens, 2001; Wiens and Gilbert, 1996; Wiens and McGuire, 1995; Zhan et al., 2014a). Earthquakes within a warmer slab may be related to a lower rupture propagating velocity and some earthquakes may experience super-shear rupture. However, we find that the source propagating velocities are not correlative with slab temperature with super-shear rupture pervasive during many large earthquakes (Tables. S1 – S26). For each inferred sub-event, we define its source propagating velocity as the ratio between distance separation of the sub-event from the earthquake initiation point, and the time difference between the peak time of the inferred sub-event and the earthquake initiation time. The source propagating velocities are mostly from 0.2 to 1.2 times of the S wave velocity in the source region, consistent with the previous studies (e.g., Frohlich, 2006). The inferred propagating velocity higher than S wave velocity corroborates well with the finding of super-shear rupture in a deep earthquake (Zhan et al., 2014b). Our results further indicate that high propagating velocity is widespread during the occurrence of large deep-focus earthquakes and occurs in various subduction zones. Propagation velocity is also heterogeneous during an earthquake and between the

earthquakes in a particular region (Fig. 7). For example, beneath the South American subduction where slabs are warm, the 1994 Bolivia earthquake has propagating velocities of 0.9 – 1.6 km/s (Fig. 7a, Table S2), while event E15 occurring in the similar region in 2003 has propagating velocities from 1.8 to 5.6 km/s (Fig. 7a, Table S15). Beneath the Tonga-Kermadec subduction where slabs are cold, event E6 has propagating velocities between 1.1 and 3.0 km/s (Fig. 7c, Table S6), but event E19 has propagating velocities varying from 3.1 to 5.4 km/s (Fig. 7c, Table S19). We argue that this observed variability of propagating velocity of the sub-events and super-shear rupture do not require additional physical mechanisms for deep earthquakes, as the static or dynamic stress generated by the preceding sub-event(s) could trigger sub-events with various time delays (starting from 0 s by static stress triggering). In fact, some of the high source propagating velocity is not consistent with super-shear rupture when focal mechanism variation is considered. For example, the third sub-event of E19 has a source propagating velocity of 5.4 km/s, higher than S wave velocity (Table S19). However, its fault plane orientation is $\sim 30^\circ$ different from those of the first and second sub-events (Fig. 5, Table S19), making it unlikely a super-shear rupture on a fault plane.

We investigate the relationships of seismic moment with source duration and source dimension for these deep earthquakes (Figs. 8b and 8c). We define seismic moment (M) as the summed moments of the sub-events, source duration as peak time of the last sub-event from the initiation time plus its half duration, and source dimension as maximum distance separation among the sub-events. Source durations increase with seismic moments, roughly described by a relationship of $M \sim \text{duration}^3$ (Figs. 8b). The relationship is consistent with that of the previous studies (Figure 6.3 in Frohlich, 2006).

Source dimensions and seismic moments can also be roughly described by a relationship of $M \sim \text{dimension}^3$ (Fig. 8c).

Fig. 8 a) Comparison between GCMT CLVDs and CLVDs in the summed solution of the inferred focal mechanisms of the sub-events. (b-c) Relationships between seismic moment and (b) source duration and (c) source dimension. Circles represent the analyzed earthquakes (E1 – E25) and the Okhotsk earthquake (Table S0).

7. Conclusion

We apply a multiple source inversion method to study the source processes of 25 deep-focus earthquakes from 1994 to 2012, with $M_w > 7.0$ and depth > 400 km, based on

waveform modeling of P, SH, pP and sSH waves. The multiple source inversion method treats a large earthquake as a combination of smaller sub-events in space and time, and resolves the spatio-temporal separation and focal mechanism of each sub-event, based on waveform modeling of seismic data. We show two aspects of the scientific approach are of particular importance in resolving the source process of a large deep earthquake: 1) combining the direct and near-surface reflected phases as seismic constraints, which is needed to resolve the depth distribution of the sub-events during a large earthquake, and 2) adopting the multiple source inversion, which makes no *a priori* assumption on the focal mechanism of the sub-events in waveform modeling and is capable of resolving the focal mechanism change during the earthquake.

All the studied deep-focus earthquakes can be modeled using one to eight sub-events. We classify these earthquakes into three categories based on the focal mechanisms and spatial distributions of the inferred sub-events: 1) category one, with non-planar spatial distribution and variable focal mechanisms of sub-events, represented by the Bolivia earthquake and the 2013 Mw 8.3 Okhotsk deep earthquake, 2) category two, with planar distribution but fault mechanisms of the sub-events inconsistent with the best-fitting plane, including eighteen earthquakes, and 3) category three, with planar distribution and focal mechanisms of sub-events consistent with the plane, including six earthquakes.

We discuss possible physical mechanisms for earthquakes in each category, including transformational faulting, shear instability in a homogeneous medium and shear instability in pre-existing weak zone. We suggest that the inferred results of all the

studied events can be best explained by a cascading failure of shear thermal instabilities in pre-existing weak zones, with the perturbation of stress generated by a shear instability triggering another and the orientations of the pre-existing weak zones controlling the focal mechanisms of these instabilities. The proposed mechanism can also explain the observed great variability of focal mechanisms, large values of CLVD component and super-shear rupture of deep-focus earthquakes in the previous studies. We further suggest that caution is needed to relate the focal mechanism of deep-focus earthquakes to the stress state in the subducted slabs, as the observed great variability of focal mechanisms of small and large deep-focus earthquakes may just reflect existence of pre-existing weak zones with various orientations and may not necessarily be related to the ambient stress in the region of the earthquake sources. The relationships between seismic moment and source duration and dimension in large deep earthquakes can be roughly described as seismic moment $\sim$ (source duration)³ and moment $\sim$ (source dimension)³.

Acknowledgements

We thank editor Peter Shearer and two anonymous reviewers for insightful and constructive comments and suggestions. The seismic data are downloaded from the Incorporated Research Institutions for Seismology (IRIS). The work is supported by NSF EAR1214215.

618

619

620

621

622 **References**

623 Antolik, M., Dreger, D., Romanowicz, B., 1996. Finite fault source study of the great
624 1994 deep Bolivia earthquake. *Geophysical Research Letters* 23, 1589-1592.

625 Antolik, M., Dreger, D., Romanowicz, B., 1999. Rupture processes of large deep-focus
626 earthquakes from inversion of moment rate functions. *Journal of Geophysical Research-*
627 *Solid Earth* 104, 863-894.

628 Chen, Y., Wen, L., Ji, C., 2014. A cascading failure during the 24 May 2013 great
629 Okhotsk deep earthquake. *Journal of Geophysical Research-Solid Earth* 119, 3035-3049.

630 Dziewonski, A.M., Anderson, D.L., 1981. Preliminary reference Earth model. *Physics of*
631 *the Earth and Planetary Interiors* 25, 297-356.

632 Estabrook, C.H., 1999. Body wave inversion of the 1970 and 1963 South American large
633 deep-focus earthquakes. *Journal of Geophysical Research-Solid Earth* 104, 28751-28767.

634 Frohlich, C., 2006. *Deep earthquakes*. Cambridge university press.

635 Goes, S., Ritsema, J., 1995. A broad-band P-wave analysis of the large deep Fiji island
636 and Bolivia earthquakes OF 1994. *Geophysical Research Letters* 22, 2249-2252.

637 Green, H.W., 2007. Shearing instabilities accompanying high-pressure phase
638 transformations and the mechanics of deep earthquakes. *Proc. Natl. Acad. Sci. U. S. A.*
639 104, 9133-9138.

640 Green, H.W., Burnley, P.C., 1989. A new self-organizing mechanism for deep-focus
641 earthquakes. *Nature* 341, 733-737.

642 Green, H.W., Houston, H., 1995. The mechanisms of deep earthquakes. *Annual Review*
643 *of Earth and Planetary Sciences* 23, 169-213.

644 Helmberg, D.V., 1968. The crust-mantle transition in Bering sea. Bulletin of the
645 Seismological Society of America 58, 179-214.

646 Hill, D.P., 2008. Dynamic stresses, Coulomb failure, and remote triggering. Bulletin of
647 the Seismological Society of America 98, 66-92.

648 Hobbs, B.E., Ord, A., 1988. Plastic instabilities - implications for the origin of
649 intermedite and deep focus earthquakes. Journal of Geophysical Research-Solid Earth
650 and Planets 93, 10521-10540.

651 Ji, C., Wald, D.J., Helmberger, D.V., 2002. Source description of the 1999 Hector Mine,
652 California, earthquake, part I: Wavelet domain inversion theory and resolution analysis.
653 Bulletin of the Seismological Society of America 92, 1192-1207.

654 Jung, H., Green, H.W., Dobrzhinetskaya, L.F., 2004. Intermediate-depth earthquake
655 faulting by dehydration embrittlement with negative volume change. Nature 428, 545-
656 549.

657 Karato, S., Riedel, M.R., Yuen, D.A., 2001. Rheological structure and deformation of
658 subducted slabs in the mantle transition zone: implications for mantle circulation and
659 deep earthquakes. Physics of the Earth and Planetary Interiors 127, 83-108.

660 Kelemen, P.B., Hirth, G., 2007. A periodic shear-heating mechanism for intermediate-
661 depth earthquakes in the mantle. Nature 446, 787-790.

662 Kennett, B., Engdahl, E., 1991. Traveltimes for global earthquake location and phase
663 identification. Geophysical Journal International 105, 429-465.

664 Kirby, S.H., 1987. Localized polymorphic phase-transformation in high-pressure faults
665 and applications to the physical-mechanism of deep earthquakes. Journal of Geophysical
666 Research-Solid Earth and Planets 92, 13789-13800.

667 Kuge, K., Kawakatsu, H., 1990. Analysis of a deep “non double couple” earthquake
668 using very broadband data. Geophysical Research Letters 17, 227-230.

669 Kuge, K., Kawakatsu, H., 1992. Deep and intermediate-depth non-double couple
670 earthquakes - interpretation of moment tensor inversions using various passbands of very
671 broad-band seismic data. Geophysical Journal International 111, 589-606.

672 Kuge, K., Kawakatsu, H., 1993. Significance of non-double couple components of deep
673 and intermediate-depth earthquakes - implications from moment tensor inversions of
674 long-period seismic waves. *Physics of the Earth and Planetary Interiors* 75, 243-266.

675 Kuge, K., Lay, T., 1994. Systematic non-double-couple components of earthquakes
676 mechanisms : The role of fault zone irregularity. *Journal of Geophysical Research-Solid*
677 *Earth* 99, 15457-15467.

678 McGuire, J.J., Wiens, D.A., Shore, P.J., Bevis, M.G., 1997. The March 9, 1994 (M-w
679 7.6), deep Tonga earthquake: Rupture outside the seismically active slab. *Journal of*
680 *Geophysical Research-Solid Earth* 102, 15163-15182.

681 Meade, C., Jeanloz, R., 1991. Deep-focus earthquakes and recycling of water into the
682 earths mantle. *Science* 252, 68-72.

683 Meng, L., Ampuero, J.-P., Buergermann, R., 2014. The 2013 Okhotsk deep-focus
684 earthquake: Rupture beyond the metastable olivine wedge and thermally controlled rise
685 time near the edge of a slab. *Geophysical Research Letters* 41, 3779-3785.

686 Myhill, R., 2013. Slab buckling and its effect on the distributions and focal mechanisms
687 of deep-focus earthquakes. *Geophysical Journal International* 192, 837-853.

688 Ogawa, M., 1987. Shear instability in a viscoelastic material as the cause of deep-focus
689 earthquakes. *Journal of Geophysical Research-Solid Earth and Planets* 92, 13801-13810.

690 Scholz, C.H., 2002. *The mechanics of earthquakes and faulting*. Cambridge university
691 press.

692 Silver, P.G., Beck, S.L., Wallace, T.C., Meade, C., Myers, S.C., James, D.E., Kuehnel, R.,
693 1995. Rupture characteristics of the deep Bolivian earthquake of 9 June 1994 and the
694 mechanisms of deep-focus earthquakes. *Science* 268, 69-73.

695 Tibi, R., Wiens, D.A., Inoue, H., 2003. Remote triggering of deep earthquakes in the
696 2002 Tonga sequences. *Nature* 424, 921-925.

697 Warren, L.M., Hughes, A.N., Silver, P.G., 2007. Earthquake mechanics and deformation
698 in the Tonga-Kermadec subduction zone from fault plane orientations of intermediate-
699 and deep-focus earthquakes. *Journal of Geophysical Research-Solid Earth* 112.

Wei, S.J., Helmberger, D., Zhan, Z.W., Graves, R., 2013. Rupture complexity of the M-w 8.3 sea of okhotsk earthquake: Rapid triggering of complementary earthquakes? *Geophysical Research Letters* 40, 5034-5039.

Wen, L., 2006. Localized temporal change of the Earth's inner core boundary. *Science* 314, 967-970.

Wiens, D.A., 2001. Seismological constraints on the mechanism of deep earthquakes: temperature dependence of deep earthquake source properties. *Physics of the Earth and Planetary Interiors* 127, 145-163.

Wiens, D.A., Gilbert, H.J., 1996. Effect of slab temperature on deep-earthquake aftershock productivity and magnitude-frequency relations. *Nature* 384, 153-156.

Wiens, D.A., McGuire, J.J., 1995. The 1994 Bolivia and Tonga events - fundamental different types of deep earthquakes. *Geophysical Research Letters* 22, 2245-2248.

Wiens, D.A., McGuire, J.J., Shore, P.J., 1993. Evidence for transformational faulting from a deep double seismic zone in Tonga. *Nature* 364, 790-793.

Wiens, D.A., McGuire, J.J., Shore, P.J., Bevis, M.G., Draunidalo, K., Prasad, G., Helu, S.P., 1994. A deep earthquake aftershock sequence and implications for the rupture mechanism of deep earthquakes. *Nature* 372, 540-543.

Ye, L., Lay, T., Kanamori, H., Koper, K.D., 2013. Energy Release of the 2013 Mw 8.3 Sea of Okhotsk Earthquake and Deep Slab Stress Heterogeneity. *Science* 341, 1380-1384.

Yu, W.-c., Wen, L., 2012. Deep-Focus Repeating Earthquakes in the Tonga-Fiji Subduction Zone. *Bulletin of the Seismological Society of America* 102, 1829-1849.

Zhan, Z., Kanamori, H., Tsai, V.C., Helmberger, D.V., Wei, S., 2014a. Rupture complexity of the 1994 Bolivia and 2013 Sea of Okhotsk deep earthquakes. *Earth and Planetary Science Letters* 385, 89-96.

Zhan, Z.W., Helmberger, D.V., Kanamori, H., Shearer, P.M., 2014b. Supershear rupture in a M-w 6.7 aftershock of the 2013 Sea of Okhotsk earthquake. *Science* 345, 204-207.

Figureg

Figure 1

Figure 2

Figure 3

Figure 4a

Figure 4b

Figure 4c

Figure 5a

Figure 5b

Figure 5c

Figure 6a

Figure 6b

Figure 6c

Figure 7

Figure 8