GEOPHYSICAL RESEARCH LETTERS, VOL. 35, L24301, doi:10.1029/2008GL036249, 2008

Depth variation of coseismic stress drop explains bimodal earthquake magnitude-frequency distribution

O. Zielke[1] and J. R. Arrowsmith[1]

Received 7 October 2008; revised 11 November 2008; accepted 13 November 2008; published 16 December 2008.

[1] An essential part of seismic hazard analysis is the earthquake size-frequency relationship, used to estimate earthquake recurrence time, and thus probability. A key feature of those distributions is their bimodal character: small to moderate magnitude earthquakes follow the Gutenberg-Richter (GR) inverse power-law relation while large magnitude (characteristic) earthquakes are more frequent than anticipated from GR, following approximately a gaussian distribution around the maximum magnitude limited by fault geometry. Using a numerical earthquake simulator, we show that the temperature dependence of friction behavior and therefore the depth-variation of coseismic stress drop, derived from laboratory friction experiments, presents a simple and comprehensive explanation for the observed bimodal seismicity distribution. Characteristic earthquakes are the result of an abrupt increase in rupture width at the transition from small to large earthquakes. **Citation:** Zielke, O., and J. R. Arrowsmith (2008), Depth variation of coseismic stress drop explains bimodal earthquake magnitude-frequency distribution, *Geophys. Res. Lett.*, *35*, L24301, doi:10.1029/2008GL036249.

1. Introduction

[2] Regional and global seismicity is well described by the Gutenberg-Richter (GR) relation. Paleoseismic studies, on the other hand, have led to the formulation of the characteristic earthquake model (CEM) [*Schwartz and Coppersmith*, 1984], which postulates that individual faults and fault segments tend to generate essentially same size or characteristic earthquakes at or near the maximum magnitude. Estimates of earthquake recurrence time and probability may vary significantly depending on whether the GR or CEM is utilized, thus bearing major implications for the seismic hazard posed by a fault. A study by *Wesnousky* [1994] addressing the issue combined seismic and paleoseismic data for individual faults, revealing a bimodal earthquake size-frequency distribution: while small-moderate magnitude earthquakes follow the GR relation, large magnitude earthquakes follow a Gaussian distribution as described by CEM. However, the question of what causes this bimodal behavior remained unanswered. Because the quality of seismic hazard assessment depends on our understanding of the earthquake cycle and the processes involved, addressing this question is of major importance.

[3] Earthquakes are the result of stick-slip frictional instabilities that occur when shear stress τ along a fault surface exceeds the product of effective normal stress σ_n^{eff} and coefficient of friction μ. The amount of (shear) stress released during an earthquake is proportional to the coseismic change of μ and σ_n^{eff}. In the classic friction model, $\Delta\mu$ equals the difference between static and dynamic coefficient of friction, μ_s and μ_d. Normal stress can vary during an earthquake due for instance to pore fluid variation potentially present in the seismogenic region [e.g., *Bizzarri and Cocco*, 2006] or different material properties of sliding rocks [*Harris and Day*, 1997]. These parameters experience potentially high spatial and/or temporal variability across the fault. Considering a general, simplified case (no variation of pore fluid or rock type) we make the assumption that σ_n^{eff} remains constant. Then, the amount of (shear) stress released during an earthquake i.e. the static stress drop is proportional only to the coseismic change of μ: $\Delta\tau = (\mu_d - \mu_s)\ \sigma_n^{eff}$.

[4] Laboratory-derived rate-and-state friction (RSF) theory, states that μ depends on slip history as well as slip rate [e.g., *Marone*, 1998, and references therein]. The change of μ during a slip event (increase in slip velocity, $v_1 \rightarrow v_2$) is expressed with an $[a - b]$ term (Figure 1a). A positive slip rate dependence ($[a - b] > 0$, velocity-strengthening) prohibits slip instabilities, resulting in stable sliding, whereas a negative slip rate dependence ($[a - b] < 0$, velocity-weakening) permits initiation of slip instabilities (earthquakes). Assuming that v_1, v_2 are constant and slip is sufficiently large to reach steady state, we can formulate $\mu_s = \mu_{ss}(v1)$, $\mu_d = \mu_{ss}(v2)$ and $\Delta\mu = [a - b] = (\mu_d - \mu_s)$. The assumption of a constant coseismic slip velocity v_2 is a simplification: numerical earthquake models [e.g., *Dunham*, 2007] have shown that v_2 and thus $[a - b]$ evolves during the coseismic phase. We do not implement it here to keep the model complexity at a minimum level and define v_2 as the average coseismic slip velocity. Also, RSF laws have been derived under some restrictive conditions (low sliding velocity and low effective normal stress), which are far from seismological conditions. However, considering that RSF theory is capable of explaining virtually the entire range of observed seismic and interseismic fault behaviors [e.g., *Marone*, 1998, and references therein], extrapolation of laboratory results to real-world scale appears valid. With regards to earthquake rupture and under consideration of the assumptions made above, the product of $[a - b]$ and effective normal (i.e., lithostatic) stress along a fault surface at depth z defines its coseismic stress drop $\Delta\tau(z) = [a - b](z)\ \sigma_n^{eff}(z)$.

[5] Friction experiments on different rock types have shown that temperature [e.g., *Blanpied et al.*, 1991] as well as heterogeneity of rock material properties [e.g., *Boatwright and Cocco*, 1996] affect $[a - b]$. For granite, velocity-weakening is observed above 50°C and below 300°C, with

[1]School of Earth and Space Exploration, Arizona State University, Tempe, Arizona, USA.

0094-8276/08/2008GL036249$05.00

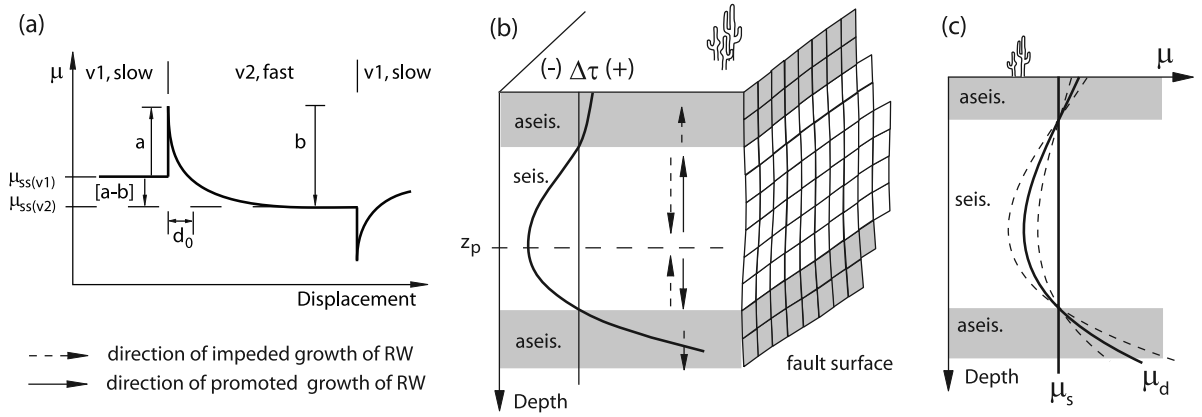

Figure 1. (a) Evolution of friction coefficient μ during a velocity stepping experiment, $v_1 \rightarrow v_2 \rightarrow v_1$). The slip rate dependent change of μ is expressed by $[a - b]$. (b) Depth variation of coseismic stress drop $\Delta\tau$ and inferred non-uniform evolution of rupture width. Shaded areas highlight aseismic zones, where velocity strengthening prohibits earthquake initiation. Also shown is an example of a non-planar self-similar fault plane as it is used in our simulations. (c) Depth evolution of μ_s and μ_d, the latter derived from depth evolution of $[a - b]$. Solid lines indicate average values for μ_s and μ_d. Dashed lines indicate the range of possible dynamic friction values.

a peak in weakening behavior around 200°C (see auxiliary material[1]). Considering the geothermal gradient, $[a - b]$ thus changes with depth and defines a seismogenic layer (typically extending between 2–15 km depth) that is bound by an upper and lower aseismic zone, satisfying our basic geologic model for faulting [e.g., *Marone and Scholz*, 1988; *Sibson*, 1977]. While normal stress increases gradually with depth, the $[a - b]$ profile contains a peak around the $200°C$ isotherm, defining the depth of maximum coseismic stress drop z_p (Figure 1b).

2. Numerical Model Formulation

[6] We investigate how the temperature dependence of $[a - b]$ and the resulting peak in coseismic stress drop within the seismogenic zone affects the seismic behavior of a fault by utilizing a quasi-static numerical earthquake simulator "*FIMozFric*" that is built on analytical expressions for internal displacements and strains due to slip along rectangular faults [*Okada*, 1992]. Faults are defined in a homogeneous, elastic half-space and each fault is subdivided into a large number of equal-sized, 1 by 1 km fault patches. Position and orientation of each fault patch is adjusted to reproduce the self-similar roughness of natural faults [e.g., *Power et al.*, 1987] (Figure 1b and auxiliary material).

[7] Each fault patch is assigned a static and dynamic coefficient of friction, μ_s and μ_d. While μ_s is taken constant with depth, μ_d is a function of geothermal gradient and $[a - b]$ depth profile. The experiments by *Blanpied et al.* [1991] to determine the temperature-dependence of $[a - b]$ were made using velocity stepping tests, periodically changing v from 0.1 to 1.0 μm/s. Because coseismic slip velocities (v_2) are approximately six magnitudes higher (~m/s), the $[a - b]$ values from these experiments cannot be used directly to determine μ_d. For that, we scale the $[a - b]$ profile so that its mean within the seismic zone ($mean[a - b]_{seis}$) becomes equal to $\Delta\mu$ = 2 to 10%μ_s (typical range for $\Delta\tau$ [*Scholz*, 2002]):

$$\mu_d(z) = \mu_s - \left(\frac{2 \text{ to } 10\%\mu_s}{mean[a-b]_{seis}}\right)[a-b](z) \quad (1)$$

Variability of $\Delta\mu$ is introduced as a bulk parameter to account for the large number of potentially interacting earthquake rupture mechanisms (e.g., thermal pressurization of pore fluid, rock melting, chemical environment changes, mechanical lubrication, gouge evolution) that may be of varying significance depending on specific site conditions (Figure 1c). Fault patches with $\mu_s > \mu_d$ lie in the seismogenic zone and vice versa. The numerical simulation is divided into interseismic and coseismic phases. During the interseismic phase, τ along seismic fault patches is incrementally increased during each time step (e.g., day). The amount of stress added varies slightly from patch to patch, depending on its orientation and position (horizontal distance) relative to the plate boundary (see auxiliary material). τ along aseismic fault patches is held constant at $\tau = \mu_s \sigma_n^{eff}$, assuming that exceeding shear stresses will be released constantly by stable slip or non-brittle, off-fault deformation. σ_n^{eff} along all patches is equal to the lithostatic pressure and held constant ($\sigma_n^{eff} = \sigma_n$).

[8] An earthquake is initiated if at least one fault patch becomes activated ($\tau > \mu_s \sigma_n$). Then the coseismic phase is entered. The frictional coefficient of activated patches instantaneously changes to its dynamic equivalent, $\mu_s \rightarrow \mu_d$ (in actuality an instantaneous stress release is unphysical because it causes an unbounded energy flux at the crack tip; change from $\mu_{ss}(v1)$ to $\mu_{ss}(v2)$ occurs over critical slip distance d_0, Figure 1a). For seismic fault patches, this results in a lowering of μ associated with fault slip that releases the excess stress so that $\tau = \mu_d \sigma_n$. The corresponding stress drop equals $\tau_0 - (\mu_d \sigma_n)$, where τ_0 is the pre-earthquake stress. Activated aseismic fault patches, on the other hand, experience an increase in friction coefficient

[1]Auxiliary materials are available in the HTML. doi:10.1029/2008GL036249.

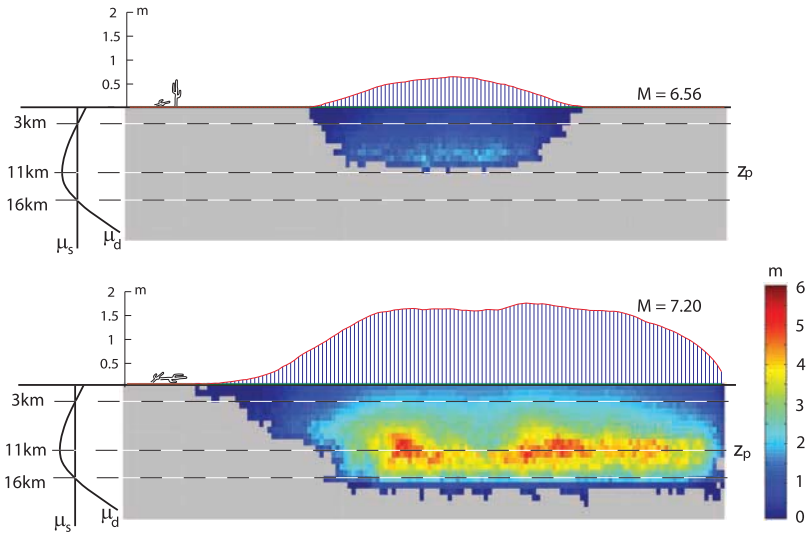

Figure 2. Color-coded amount of slip (22 km wide, 150 km long sub-vertical strike-slip fault with amplitude-wavelength ratio $\lambda = 0.003$) of a typical small and large earthquake. Gray fault patches did not slip in the respective events. Shown above is the respective surface-slip distribution. On the left is the average depth distribution of μ_s and μ_d and the location of z_p.

($\mu_d > \mu_s$) increasing their resistance to failure. Aseismic fault patches will only slip as part of an earthquake if coseismically induced stresses are sufficiently high so that $\tau > \mu_d \sigma_n$. Then, as with seismic fault patches, the excess stress will be released by coseismic slip, lowering the shear stress to $\tau = \mu_d \sigma_n$. An earthquake may grow because slip along one patch induces stress changes along the other patches. If these changes are sufficient to cause other patches to fail, they will do so as part of the same earthquake. The frictional coefficient of activated patches remains at the dynamic level for the duration of the event. Patches may slip more than once and continue doing so as long as $\tau > \mu_d \sigma_n$. The earthquake is over if τ along all fault patches is at or below a level corresponding to μ_d. Then the next interseismic phase is entered, the loading mechanism resumes, and all patches instantaneously regain their static frictional coefficient.

[9] The presented simulator does not incorporate dynamic aspects of the rupture process. However, based on results presented here, we argue that a quasi-static approach is capable of reproducing many aspects of fault behavior, including a bimodal earthquake magnitude-frequency distribution.

3. Results

[10] Using *FIMozFric*, we created 54 seismic catalogs for sub-vertical strike-slip faults, systematically changing fault geometry and strength (model parameters in Table S1). The combined, synthetic record is 540 kyrs long and contains approximately 900,000 earthquakes with rupture areas >5 km^2 ($M \approx 4.5$).

[11] Figure 2 shows example fault slip distributions for an arbitrarily chosen small and large earthquake. The distinction between small and large earthquake is based on rupture width *RW*: large earthquakes rupture the full seismogenic zone, and small earthquakes rupture less than that. The geothermal gradient of 20°C/km, used in our simulations, limits the seismic zone to a depth between 3–16 km, with the peak in $\Delta\mu$ at z_p = 11–12 km. While the large event shown in Figure 2 ruptured past z_p, the small one was not able to do so. Maximum slip of both events occurred at the depth of respective highest $\Delta\mu$. Simulated earthquakes with $M > 5.5$ were necessary to penetrate the upper aseismic zone and to create a surface rupture, consistent with earthquake surface rupture observations [*Bonilla*, 1988]. The average coseismic surface displacement of these events is lower than the respective displacement across the whole rupture.

[12] Figure 3a shows the probability distribution of earthquake magnitude for individual catalogs (blue) and the combined catalog (red). Although variation occurs because of different fault geometry and strength, a general trend can be recognized: small to moderate magnitude events follow the GR relation, while large events form a second probability peak around the maximum magnitude. This bimodal earthquake magnitude frequency distribution is consistent with observations for large, individual fault systems, occupying narrow and long spatial domains [e.g., *Wesnousky*, 1994; *Ben-Zion*, 2003].

[13] Figure 3b shows a contour-plot of a bivariate probability density function between earthquake rupture width and associated average coseismic stress drop. Earthquakes with $RW \leq 5$ km show a wide distribution of coseismic stress drop ranging from 0.1–15 MPa, with a probability peak at 1–2 MPa. These results are in very good agreement with empirically determined stress drops from measurements of radiated seismic waves of real small earthquakes [*Shearer et al.*, 2006, and references therein; *Hanks*, 1977]. As *RW* increases, the variability of $\Delta\tau$ decreases. Between 6 km $\leq RW \leq$ 12 km we find an approximately linear relationship of *RW* and $\Delta\tau$. While virtually no earthquakes with 13 km $\leq RW <$ 18 km occurred, we observe a second probability peak at 18 km $\leq RW \leq$ 20 km. The gap in rupture width distribution indicates that once an earthquake

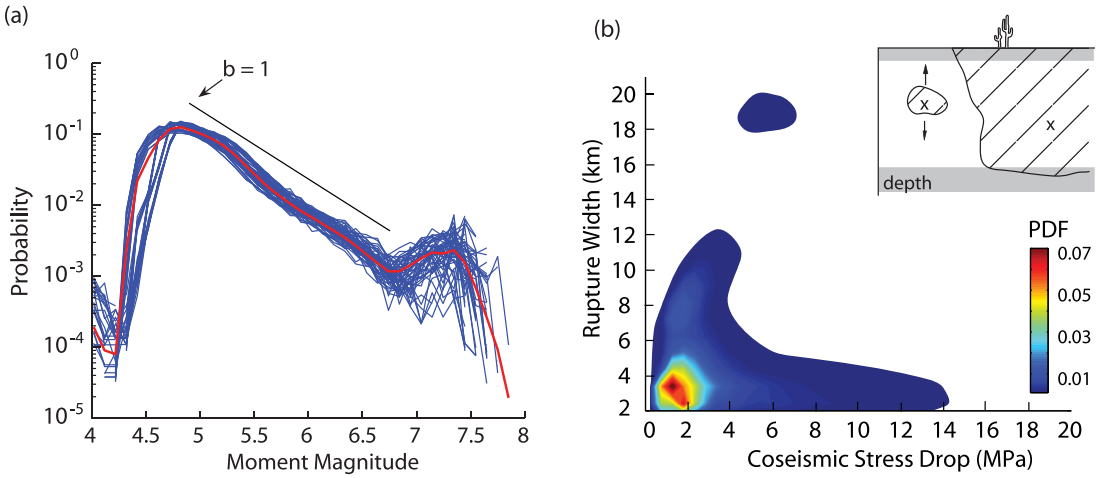

Figure 3. Results of numerical earthquake simulations. (a) Magnitude-frequency distribution follows a power-law relation for small events (M = 4.8–6.8) and has a second peak in event probability for large events (M ≈ 7.3). The drop-off at M < 4.8 is due to minimum earthquake size (only RA ≥ 5 km^2 in catalog). (b) Color-coded bivariate histogram of earthquake rupture width and associated average coseismic stress drop. Small earthquakes with $RW \leq 12$ km show a wide range of possible stress drops (between 0.1–15 MPa, with a peak at 1–2 MPa). Large events that rupture the full seismogenic zone have an approximately constant stress drop (≈5 MPa). Inset shows the potential variability of moment centroid depth (x) for small earthquakes which causes the observed variability in stress drop.

rupture overcomes z_p (at 11–12 km depth), its width increases abruptly, activating the full seismogenic zone and further extending into the lower aseismic zone, thus becoming a large, characteristic event (forming the secondary peak in Figure 3a).

[14] While small earthquakes experience significant stress drop variation, large, characteristic earthquakes with $RW \geq 18$ km, have an approximately constant coseismic stress drop of $\Delta\tau \approx 5$ MPa. As Figure 1b indicates, the coseismic stress drop changes with depth (as a function of $[a - b]$). The average stress drop of an earthquake is determined by its average depth (i.e., moment centroid depth). The observed range in coseismic stress drop of small earthquakes is the result of the variability of moment centroid depth (Figure 3b, insert) which decreases as RW increases. Large earthquakes, rupturing the full seismogenic zone, have an approximately constant moment centroid depth (around z_p) and therefore coseismic stress drop (≈5 MPa). We want to remind the reader that we define $\Delta\tau$ as the product of $(\mu_s - \mu_d)$ and σ_n^{eff}. A weak fault (low $\Delta\mu$ or low σ_n^{eff} due to high pore pressure) will experience lower stress drops and vice versa. Although the absolute value of $\Delta\tau$ changes with $\Delta\ \mu$ and σ_n^{eff}, the observed distribution (Figure 3b) remains the same. Further model results can be found in the auxiliary material.

4. Discussion

[15] Our simulations are capable of reproducing seismological observations of earthquake magnitude frequency distribution and coseismic stress drop, suggesting their general validity. Our results show that the rupture width of small earthquakes is limited to the distance from z_p, to the depths of top or bottom limits of the aseismic zones. This distance is always less than the width of the seismogenic zone itself. Widening an earthquake rupture towards z_p is impeded because it requires higher coseismically induced stresses than rupture away from z_p. Consequently, if the fault at depth z_p is activated, upward and downward widening of the earthquake rupture is promoted (Figure 1b). Once a rupture increases sufficiently past z_p it will cause failure of the whole seismogenic width, thus becoming a large and characteristic earthquake that ruptures the whole fault or fault segment. Because RW of small earthquakes is always less than the width of the seismogenic zone, the transition from small to large earthquake is associated with a sudden increase in rupture width. The moment magnitude of an earthquake is defined as the product of shear modulus G, average coseismic slip d, and rupture area A: $M_0 = G\,d\,A$. As the rupture width increases abruptly at the transition from small to large earthquakes, rupture area and therefore moment magnitude will increase in the same manner, producing the observed bimodal magnitude-frequency distribution (Figure 3a).

[16] The depth variation of coseismic stress drop and the resulting abrupt increase in rupture width present a simple explanation for the observed bimodal magnitude-frequency distribution. Characteristic earthquakes are an inherent feature of every fault or fault segment assuming finite fault width and depth-dependent coseismic stress drop. The time to recover shear stresses at the depth of maximum coseismic stress drop z_p gives a first order approximation of the recurrence time of large, characteristic earthquakes. A key problem that remains to be addressed is which physical properties or differences thereof define a fault segment. Structural complexity as well as variations of friction properties and rupture mechanisms along fault are probably the main contributors.

[17] **Acknowledgments.** We would like to thank the Numerical Earthquake Simulator initiative of the Southern California Earthquake Center, bringing together many researchers that work on earthquake simulations for the purpose of model comparison and testing. W. Thatcher, N. Beeler, and T. Hanks provided timely and helpful reviews. We also thank A. Bizzari, F. Pollitz, and S. Ward for their review of the manuscript and

associate editor F. Florindo. We would also like to thank the Fulton School of Engineering at Arizona State University for allowing us to use its high performance computing facilities.

References

Ben-Zion, Y. (2003), Key formulas in earthquake seismology, in *International Handbook of Earthquake and Engineering Seismology*, *Int. Geophys. Ser.*, vol. 81, edited by W. H. K. Lee, pp. 1857–1875, Academic, Amsterdam.

Bizzarri, A., and M. Cocco (2006), A thermal pressurization model for the spontaneous dynamic rupture propagation on a three-dimensional fault: 2. Traction evolution and dynamic parameters, *J. Geophys. Res.*, *111*, B05304, doi:10.1029/2005JB003864.

Blanpied, M. L., D. A. Lockner, and J. D. Byerlee (1991), Fault stability from granite sliding experiments at hydrothermal conditions, *Geophys. Res. Lett.*, *18*, 609–612.

Boatwright, J., and M. Cocco (1996), Frictional constraints on crustal faulting, *J. Geophys. Res.*, *101*, 13,895–13,909.

Bonilla, M. G. (1988), Minimum earthquake magnitude associated with coseismic surface faulting, *Bull. Assoc. Eng. Geol.*, *25*, 17–29.

Dunham, E. M. (2007), Conditions governing the occurrence of supershear ruptures under slip-weakening friction, *J. Geophys. Res.*, *112*, B07302, doi:10.1029/2006JB004717.

Hanks, T. C. (1977), Earthquake stress drops, ambient tectonic stresses and stresses that drive plate motions, *Pure Appl. Geophys.*, *115*, 441–458.

Harris, R. A., and S. M. Day (1997), Effects of a low-velocity zone on a dynamic rupture, *Bull. Seismol. Soc. Am.*, *87*, 1267–1280.

Marone, C. (1998), Laboratory-derived friction laws and their application to seismic faulting, *Annu. Rev. Earth Planet. Sci.*, *26*, 643–696.

Marone, C., and C. H. Scholz (1988), The depth of seismic faulting and the upper transition from stable to unstable slip regimes, *Geophys. Res. Lett.*, *15*, 621–624.

Okada, Y. (1992), Internal deformation due to shear and tensile faults in a half-space, *Bull. Seismol. Soc. Am.*, *82*, 1018–1040.

Power, W. L., T. E. Tullis, S. R. Brown, G. N. Boitnott, and C. H. Scholz (1987), Roughness of natural fault surfaces, *Geophys. Res. Lett.*, *14*, 29–32.

Scholz, C. H. (2002), *The Mechanics of Earthquakes and Faulting*, 2nd ed., 471 pp., Cambridge Univ. Press, Cambridge, U. K.

Schwartz, D., and K. J. Coppersmith (1984), Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res.*, *89*, 5681–5698.

Shearer, P. M., G. A. Prieto, and E. Hauksson (2006), Comprehensive analysis of earthquake source spectra in southern California, *J. Geophys. Res.*, *111*, B06303, doi:10.1029/2005JB003979.

Sibson, R. H. (1977), Fault rocks and fault mechanics, *J. Geol. Soc. London*, *133*, 191–213.

Wesnousky, S. G. (1994), The Gutenberg-Richter or characteristic earthquake distribution, which is it?, *Bull. Seismol. Soc. Am.*, *84*, 1940–1959.

J. R. Arrowsmith and O. Zielke, School of Earth and Space Exploration, Arizona State University, Bateman Physical Science Complex, 550 East Tyler Mall, Tempe, AZ 85287-1404, USA. (ramon.arrowsmith@asu.edu; olaf.zielke@asu.edu)