Geophysical Journal International

Geophys. J. Int. (2011) **186,** 1389–1403 doi: 10.1111/j.1365-246X.2011.05117.x

Shallow slip deficit due to large strike-slip earthquakes in dynamic rupture simulations with elasto-plastic off-fault response

Y. Kaneko and Y. Fialko

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California - San Diego, La Jolla, CA 92093*, USA.*
E-mail: ykaneko@ucsd.edu

Accepted 2011 June 20. Received 2011 June 20; in original form 2011 February 28

SUMMARY

Slip inversions of geodetic data from several large (magnitude ∼7) strike-slip earthquakes point to coseismic slip deficit at shallow depths (<3–4 km), that is, coseismic slip appears to decrease towards the Earth surface. While the inferred slip distribution may be consistent with laboratory-derived rate and state friction laws suggesting that the uppermost brittle crust may be velocity strengthening, there remains a question of how the coseismic slip deficit is accommodated throughout the earthquake cycle. The consequence of velocity-strengthening fault friction at shallow depths is that the deficit of coseismic slip is relieved by post-seismic afterslip and interseismic creep. However, many seismic events with inferred shallow slip deficit were not associated with either resolvable shallow interseismic creep or robust shallow afterslip. Hence, the origin of shallow 'slip deficit' remains uncertain. In this study, we investigate whether inelastic failure in the shallow crust due to dynamic earthquake rupture can explain the inferred deficit of shallow slip. Evidence for such failure is emerging from geologic, seismic and geodetic observations. We find that the amount of shallow slip deficit is proportional to the amount of inelastic deformation near the Earth surface. Such deformation occurs under a wide range of parameters that characterize rock strength in the upper crust. However, the largest magnitude of slip deficit in models accounting for off-fault yielding is 2–4 times smaller than that inferred from kinematic inversions of geodetic data. To explain this discrepancy, we further explore to what extent assumptions in the kinematic inversions may bias the inferred slip distributions. Inelastic deformation in the shallow crust reduces coseismic strain near the fault, introducing an additional 'artificial' deficit of up to 10 per cent of the maximum slip in inversions of geodetic data that are based on purely elastic models. The largest magnitude of slip deficit in our models combined with the bias in inversions accounts for up to 25 per cent of shallow slip deficit, which is comparable, but still smaller than 30–60 per cent deficit inferred from kinematic inversions. We discuss potential mechanisms that may account for the remaining discrepancy between slip deficit predicted by elasto-plastic rupture models and that inferred from inversions of space geodetic data.

Key words: Earthquake dynamics; Earthquake source observations; Rheology and friction of fault zones; Transform faults; Dynamics and mechanics of faulting.

1 INTRODUCTION

Inversions of space geodetic (in particular, Interferometric Synthetic Aperture Radar and Global Positioning System) data from large (moment magnitude ∼7) strike-slip earthquakes indicate that coseismic slip in the middle of the seismogenic layer (at depth of 4–5 km) is systematically larger than slip at the Earth surface (e.g. Simons *et al.* 2002; Fialko *et al.* 2005; Bilham 2010; Fialko *et al.* 2010). Examples include the 1992 M7.3 Landers earthquake, the 1999 M7.1 Hector Mine earthquake, the 2003 M6.5 Bam earthquake, the 2010 M7.0 Haiti earthquake and the 2010 M7.2 Sierra El Mayor (Mexico) earthquake (Fig. 1). In the case of the 2003 Bam and the 2010 Haiti earthquakes, observations showed that the rupture failed to propagate to the surface (Fialko *et al.* 2005; Bilham 2010). Slip inversions of Interferometric Synthetic Aperture Radar (InSAR) data from the recent Sierra El Mayor (Mexico) earthquake point to the overall deficit of shallow slip consistent with results for other M7 events (Fig. 1; Fialko *et al.* 2010). Determining the origin of the deficit of shallow slip is important both for understanding physics of earthquakes and for estimating seismic hazard, as suppression of shallow rupture could greatly influence strong ground motion in the vicinity of active faults (e.g. Somerville 2003;

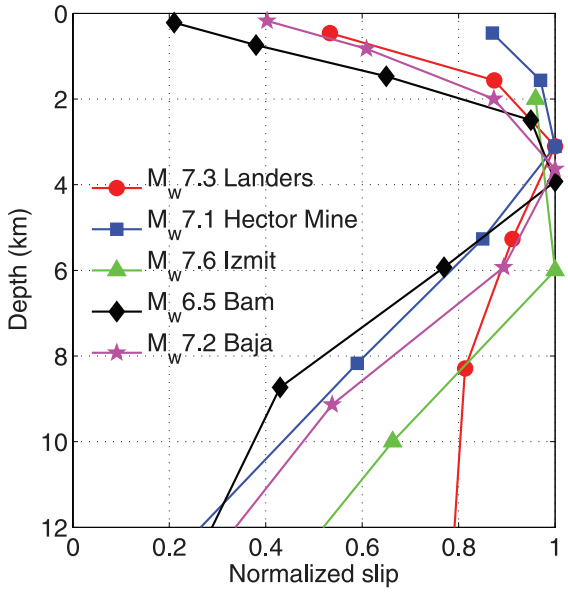

Figure 1. Distributions of coseismic slip for several ∼*M*7 strike-slip earthquakes averaged along the rupture length (Fialko *et al.* 2005, 2010). These slip distributions were obtained from the inversions of near-field InSAR data or of a combination of InSAR and GPS data. The coseismic slip sharply decreases towards the Earth surface from a maximum value in the middle of the seismogenic zone.

Kaneko *et al.* 2008; Pitarka *et al.* 2009). Note that we focus on slip distributions obtained from inversions of near-field InSAR data from more than one look direction (Fig. 1) because inversions based primarily on seismic data suffer from poor resolution at shallow depths and trade-offs between the slip distribution and rupture history (e.g. Fialko 2004a; Mai *et al.* 2007).

The inferred decrease in coseismic slip toward the Earth surface appears to be consistent with laboratory data indicating that the frictional behaviour of the uppermost brittle layer is velocity strengthening (Marone *et al.* 1991; Scholz 1998). Yet, there remains a question of how the coseismic slip deficit is accommodated throughout the earthquake cycle. The consequence of velocity-strengthening fault friction at shallow depths is that the deficit of coseismic slip is compensated as afterslip and interseismic creep (Marone *et al.* 1991; Rice 1993). However, none of the events included in Fig. 1 was associated with either shallow interseismic creep or robust shallow afterslip in the amount sufficient to remove the coseismic slip deficit in the shallow crust (Table 1) (Jacobs *et al.* 2002; Fialko 2004b; Fialko *et al.* 2005; Fielding *et al.* 2009). Geodetic observations also indicate that the occurrence of interseismic creep and afterslip at shallow depths is rather uncommon, except in certain locations near major creeping segments of mature faults and/or in areas with thick sedimentary covers with possibly overpressurized pore fluids (e.g. Wei *et al.* 2009, 2011). Hence, the inferred shallow slip deficit shown in Fig. 1 calls for other explanations.

Table 1. The magnitude of afterslip that occurred following the strike–slip earthquakes shown in Fig. 1. Since the rate of afterslip strongly decreases with time months after the earthquake, the afterslip over a longer time period would be smaller than twice the indicated value.

Event	Afterslip[a]	Duration of the data	Reference
1992 *M*7.3 Landers	∼0.15 m	6 yr	Fialko (2004b)
1999 *M*7.1 Hector Mine	∼0.10 m	1 yr	Jacobs *et al.* (2002)
1999 *M*7.6 Izmit	∼0.20 m	0.2 yr	Hearn *et al.* (2002)
2003 *M*6.5 Bam	∼0.10 m	3.5 yr	Fielding *et al.* (2009)
2010 *M*7.2 Baja	N/A	N/A	N/A

[a]The amount of afterslip in a region of the maximum afterslip in the top 2–3 km is reported.

Several mechanisms may be invoked to explain the inferred shallow slip deficit. Based on analytical models of quasi-static antiplane deformation, Rybicki (1992) and Rybicki & Yamashita (1998) suggested that low initial stress in low-rigidity shallow crust resulting from tectonic loading can lead to the reduction of coseismic slip at shallow depths. However, numerical studies of spontaneous earthquake sequences show that low-rigidity shallow layers alone do not lead to the reduction of coseismic slip (Kaneko 2009). Fialko *et al.* (2005) proposed that the shallow slip deficit may be caused by the bulk inelastic yielding of the host rocks in the shallow part of the brittle crust. In addition to accommodating some fraction of coseismic strain, inelastic deformation adjacent to the fault plane may introduce a systematic bias in inversions of geodetic data. In particular, inelastic deformation in the shallow crust reduces coseismic strain near the fault, which may introduce an artificial deficit in inversions of geodetic data that are based on purely elastic models. In this paper, we explore these ideas quantitatively and investigate whether the occurrence of inelastic deformation can account for the inferred shallow slip deficit.

There is growing evidence for extensive inelastic failure of the shallow crust around active faults. Geological observations of the structure of mature faults indicate that fault zones often consist of finely granulated gouge layer surrounded by wider zones of damaged, fractured or sometimes pulverized host rocks (e.g. Chester *et al.* 1993; Chester & Chester 1998; Dor *et al.* 2006). These observations suggest the occurrence of pervasive irrecoverable deformation off of the fault on spatial scales of tens to hundreds of metres or more. Persistent low-velocity zones near active faults have been recognized seismologically, for example, using fault-zone trapped waves (e.g. Li *et al.* 1998; Spudich & Olsen 2001; Peng *et al.* 2003; Vidale & Li 2003; Cochran *et al.* 2009). A kilometre-wide fault zones that are less rigid compared to the surrounding medium have also been inferred geodetically (Fialko *et al.* 2002; Fialko 2004b; Hamiel & Fialko 2007; Cochran *et al.* 2009). Although the structure of low-rigidity zones may be highly variable along the fault strike (Lewis & Ben-Zion 2010), widespread reduction in rigidity generally inferred around major earthquake faults is likely to be associated with accumulated damage from past seismic events or the associated post- or interseismic deformation (including aftershock activity).

Recent studies using numerical simulations of dynamic rupture with off-fault plasticity suggest that inelastic response of the bulk material can limit the peak slip velocity at the rupture front and diminish the corresponding coseismic slip and strong ground motion (e.g. Yamashita 2000; Andrews 2005; Ben-Zion & Shi 2005; Templeton & Rice 2008; Ma 2008). Since the confining pressure is relatively low near the surface, damage zones become progressively larger near the Earth surface than at seismogenic depths, resulting in so-called 'flower structures' (e.g. Ben-Zion & Shi 2005; Ma 2008; Finzi *et al.* 2009). These results imply that the occurrence of inelastic deformation might account for at least some fraction of the inferred slip deficit (Fig. 1) by reducing the amount of coseismic slip on the main fault at shallow depths.

In this work, we investigate whether inelastic failure of the shallow crust can explain the inferred deficit of shallow coseismic slip. We address this problem using numerical models of spontaneous dynamic rupture with off-fault plasticity. Our models incorporate

full rate and state rheology on the slip interface, elastic deformation, as well as yielding in the host rocks (Section 2). Inelastic deformation is represented as inelastic strain, a deviation from the linear elastic stress–strain relationship. We find that the shallow slip deficit due to coseismic inelastic deformation occurs under a wide range of parameters that characterize rock strength (Sections 3 and 4). However, the largest magnitude of slip deficit in our models is smaller than that inferred from kinematic inversions of geodetic data (Fig. 1). To explain this discrepancy, we explore to what extent assumptions in the kinematic inversions may bias the inferred slip distributions (Section 5). Finally, we discuss other potentially important causes of shallow slip deficit that are not included in our models (Section 6).

2 MODEL DESCRIPTION

We consider a vertical right-lateral strike-slip fault in a homogeneous half-space (Fig. 2a). For simplicity, antiplane (2-D) deformation is assumed such that the only non-zero component of the displacement is the one along the y direction and is given by $u(x, z, t)$, where t denotes time. The model consists of a 30 km by 48 km rectangular domain (Fig. 2a). The symmetries of the problem allow us to restrict the computational domain to the medium on one side of the fault ($x \geq 0$). The material properties are $V_s = 3.46$ km s^{-1}, $\rho = 2670$ kg m^{-3} and $G = 32$ GPa, where V_s, ρ and G are S wave velocity, density and shear modulus, respectively.

2.1 Constitutive response of fault: rate and state friction

The fault is governed by rate and state friction with the aging form of state variable evolution. For time-independent effective normal stress σ_n (taken to be positive in tension), the shear strength τ on the fault is expressed as

$$\tau(z,t) = -\sigma_n(z)\left[f_o + a(z)\ln\frac{V(z,t)}{V_o} + b(z)\ln\frac{V_o\theta(z,t)}{L}\right]$$
$$\frac{d\theta(z,t)}{dt} = 1 - \frac{V(z,t)\theta(z,t)}{L}, \quad (1)$$

where a and b are rate and state constitutive parameters, V is slip rate, f_o is the reference friction coefficient corresponding to the reference slip rate V_o, θ is a state variable, which can be interpreted as the average age of contacts between two surfaces, and L is the characteristic slip for state evolution (Dieterich 1978, 1979, Ruina 1983). The parameter combination $a - b < 0$ corresponds to steady-state velocity-weakening friction and can lead to unstable slip, whereas $a - b > 0$ corresponds to steady-state velocity-strengthening and leads to stable sliding (Rice 1983; Ruina 1983). Throughout this paper, we omit the words 'steady-state' and simply refer to velocity weakening/strengthening.

The actual fault resistance to sliding in our model is given by rate and state friction regularized at zero slip velocity, as described in Appendix A. The response of constitutive laws (1), when extrapolated to coseismic slip rates, becomes qualitatively similar to the one given by linear slip-weakening friction (Cocco & Bizzarri 2002) widely used in dynamic rupture models (e.g. Ida 1972; Day *et al.* 2005). For simplicity, we use the standard rate and state friction framework (1), without the inclusion of enhanced dynamic weakening at high slip rates (e.g. Di Toro *et al.* 2003; Rice 2006; Noda *et al.* 2009).

2.2 Constitutive response of adjacent material: Drucker–Prager plasticity

The constitutive response of the materials surrounding the fault is governed by the Drucker–Prager plasticity (Drucker & Prager 1952), which has been widely applied to model deformation of rocks, soils and other materials with pressure-dependent failure criteria. In the Drucker–Prager plasticity, the yield stress depends on the mean normal stress, and such effect is recognized as an inherent property of rocks and soils (e.g. Davis & Selvadurai 2002). The Drucker–Prager plasticity can be described by the yield stress τ^y and the yield criterion τ^{yc}

$$\tau^y = -\sigma_m^o \sin(\phi) + c\ \cos(\phi)$$
$$\tau^{yc} = \sqrt{(1/2)s_{ij}s_{ij}}$$
$$\tau^{yc} \leq \tau^y, \quad (2)$$

where σ_m is the mean stress (compressive stress is negative), c is the rock cohesion, $\phi = \tan^{-1}\mu$ is the internal friction angle, μ is the internal rock friction and s_{ij} is the deviatoric stress (Fig. 2b). In the formulation (2), the yielding occurs in shear only, and there is no inelastic volumetric deformation. We assume that c and μ are uniform over the space and time and explore the effects of

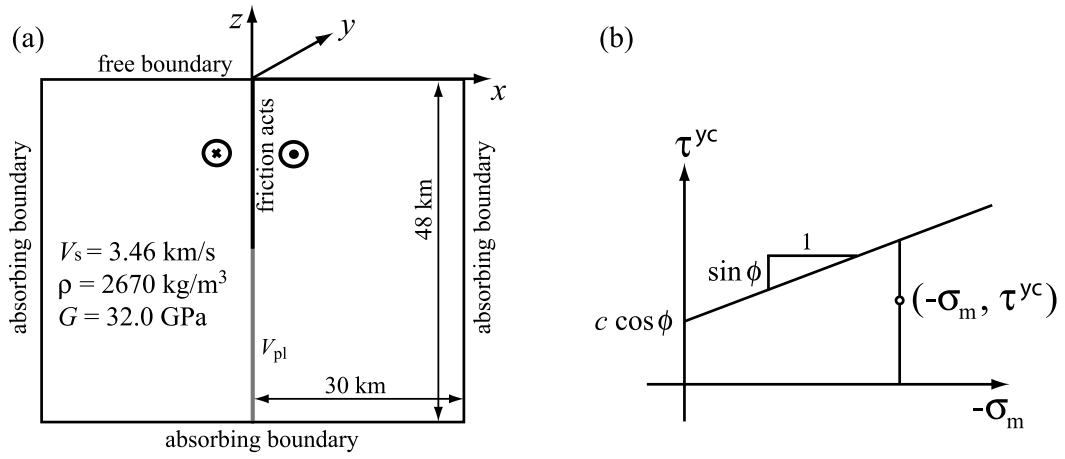

Figure 2. (a) A 2-D model of a vertical strike-slip fault. The fault segment governed by rate and state friction is marked by the red line. The fault is loaded by the region below 24-km depth steadily moving with a prescribed slip rate $V_{pl} = 35$ mm yr^{-1}. (b) Drucker–Prager yield criterion for plastic deformation. The vertical line shows the closeness of the stress state to failure envelop (CF) defined in the text.

different values of c and μ on the slip distributions of the simulated seismic events. Friction acting on a fault $\tau/(-\sigma_n)$ may be equal to the internal rock friction μ (e.g. Savage *et al.* 1996). However, we allow fault friction to be different from the internal friction of ambient rocks because well-developed faults, such as the ones we focus on in this study, may be weaker than the ambient crust due to the presence of gouge, different mineral or chemical phases within a fault zone, dynamic effects, etc.

To quantify the amount of inelastic deformation in our simulations, we compute the magnitude of plastic strain at each time step, which is given by

$$\gamma^{\mathrm{p}} = \int_0^t \sqrt{(1/2)\left(d\varepsilon_{ij}^{\mathrm{p}} - d\varepsilon_{kk}^{\mathrm{p}}/3\right)\left(d\varepsilon_{ij}^{\mathrm{p}} - d\varepsilon_{mm}^{\mathrm{p}}/3\right)}\,, \tag{3}$$

where $d\varepsilon_{ij}^{\mathrm{p}}$ is the plastic strain increment at one time step. Note that we assume no inelastic volumetric strain and hence $d\varepsilon_{kk}^{\mathrm{p}} = 0$ in eq. (3).

2.3 Numerical model and parameter selection

We use a spectral element method (SEM) to simulate elasto-plastic response. Off-fault plasticity described above is implemented into the existing SEM models of spontaneous dynamic rupture (Kaneko *et al.* 2008; Kaneko & Lapusta 2010). We discretize the computational domain into spectral elements with an average node spacing of 31 m, which is small enough to resolve the dynamic rupture on the fault (Appendix B). Absorbing conditions (Clayton & Engquist 1977) are used on all boundaries of the model except for the free surface and the fault boundary, to simulate a semi-infinite elastic half-space (Fig. 2a). The dynamic rupture code we use has been verified for a similar problem through the Southern California Earthquake Center Dynamic Earthquake-Rupture Code-Validation Exercise (Harris *et al.* 2009).

As eq. (1) indicates, σ_n, a and b vary with depth but not with time. We prescribe three normal stresses σ_{xx}, σ_{yy} and σ_{zz} as the lithostatic stress minus the hydrostatic pore pressure: $\sigma_{xx}^{\mathrm{o}} = \sigma_{yy}^{\mathrm{o}} = \sigma_{zz}^{\mathrm{o}} = -(\rho - \rho_w)gz = -[1.0 + 10.0\ z\,\mathrm{km}^{-1}]$ MPa (z is in kilometres), where ρ_w is the density of water and g is the gravitational acceleration. As a result, the effective normal stress on the fault is equal to the mean stress $\sigma_m = \sigma_n$ (Fig. 3a). Shear stresses in our models are given by $\sigma_{xy}^{\mathrm{o}} = f_o\sigma^{\mathrm{o}}{}_n + \Delta\sigma_{xy}^{\mathrm{el}}$, $\sigma_{xz}^{\mathrm{o}} = 0$, and $\sigma_{yz}^{\mathrm{o}} = \Delta\sigma_{yz}^{\mathrm{el}}$, where $\Delta\sigma_{xy}^{\mathrm{el}}$ and $\Delta\sigma_{yz}^{\mathrm{el}}$ are obtained from simulations of an earthquake sequence in an elastic model discussed below. For $\Delta\sigma_{yz}^{\mathrm{el}} = 0$, the resulting stress

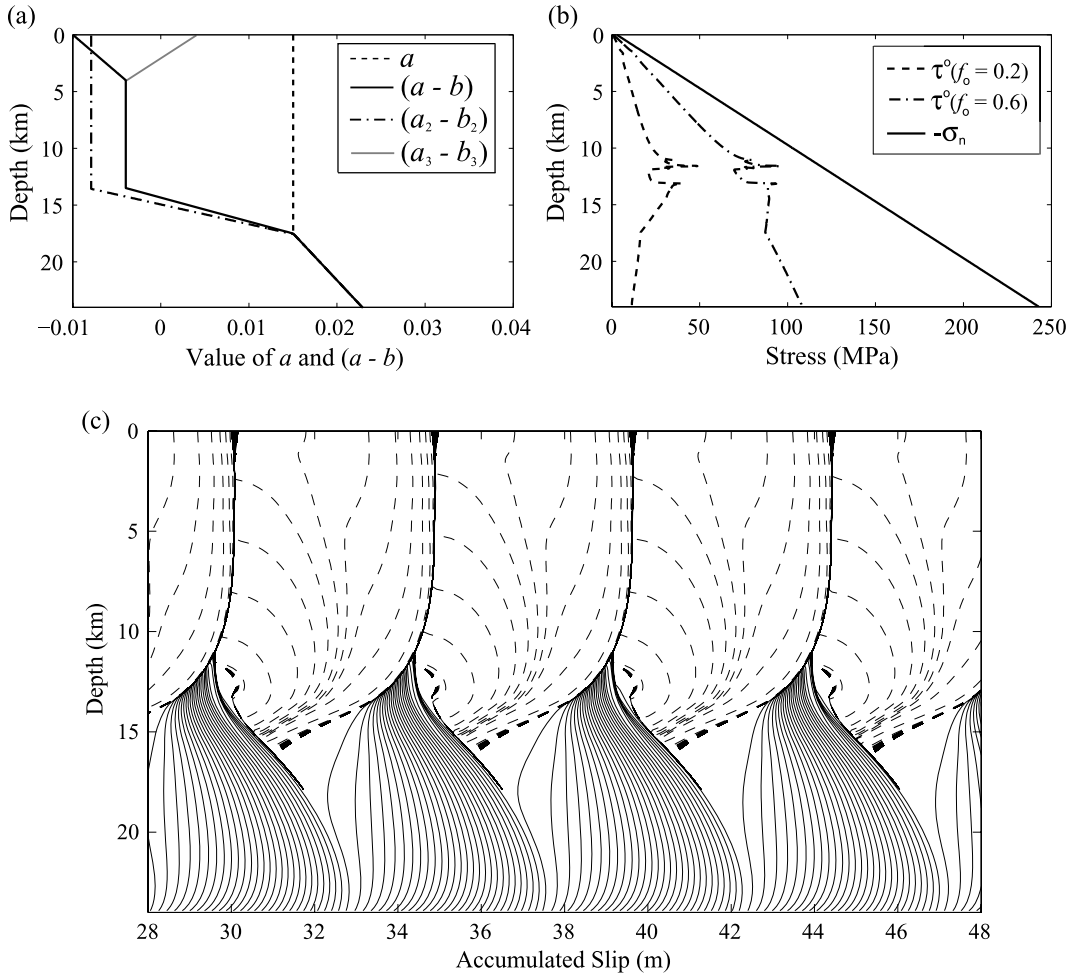

Figure 3. (a) Depth-variable distributions of the rate and state fault constitutive parameters $(a - b)$ and a. Unless noted otherwise, the indicated distributions of $(a - b)$ and a are assumed in our models; the distributions of $(a_2 - b_2)$ and $(a_3 - b_3)$ are discussed in Sections 4 and 6, respectively. (b) Distributions of the effective normal stress $(-\sigma_n)$ and prestress (τ^{o}). The distributions of τ^{o} are obtained from the simulations of earthquake sequences with $f_o = 0.2$ and 0.6. (c) An earthquake sequence in an elastic model. Accumulation of slip versus depth is shown. Solid lines show slip accumulation every 5 years, whereas dashed lines are intended to capture dynamic events and are plotted above 18-km depth every second during the simulated earthquakes.

state implies that the maximum compressive stress σ_1 is at 45° with respect to the fault plane.

The variation of friction parameters a and b with depth is shown in Fig. 3(a). This variation is the same as in Rice (1993) and Lapusta *et al.* (2000); it is derived from laboratory experiments (Blanpied *et al.* 1991, 1995), except that the region in the top 4 km has velocity-weakening conditions. In laboratory experiments, rock friction at low normal stress typically exhibits velocity-strengthening behaviour due to unconsolidated fault gouge (e.g. Marone *et al.* 1991; Marone 1998). For simplicity, we focus on effects of coseismic inelastic deformation and suppress interseismic creep by imposing velocity-weakening conditions throughout the upper crust (Fig. 3a). In an earthquake-cycle model discussed below, the interseismic creep rate at the surface resulting from the distribution of $a - b$ shown in Fig. 3(a) is $\sim$1/50 of the tectonic loading rate.

Dimensional arguments indicate that the rate of shallow creep may scale with the long-term fault slip rate (e.g. Savage & Lisowski 1993). Such a scaling is also predicted by numerical models of spontaneous earthquake sequences that incorporate a velocity-strengthening shallow layer; for example, the surface creep rates are $\sim$3.5 mm yr^{-1} when the tectonic loading rate is 35 mm yr^{-1} (Lapusta *et al.* 2000; Kaneko 2009). However, a number of mature fast-slipping strike-slip faults are not associated with resolvable interseismic shallow creep. Faults with the inferred shallow slip deficit (Fig. 1) are generally characterized by low slip rates and large recurrence intervals, so that the rates of shallow creep predicted by numerical models with a shallow velocity-strengthening layer would be below the geodetic detection limit (a fraction of mm yr^{-1}). The consequence of coseismic inelastic deformation in such scenario is discussed in Section 6.

2.4 Initial conditions for elasto-plastic dynamic rupture simulations

In the simulations of a single dynamic rupture, a stress field before the dynamic rupture and the nucleation location can be imposed as initial conditions. It is not clear whether those initial conditions would be compatible with parameterization of the model (e.g. the assumed friction law). To obtain initial conditions more appropriate for the parameterization of the model, we first simulate an earthquake sequence based on an elastic model (Fig. 3c). The earthquake-sequence simulation is done using a spectral boundary integral method (BIM) (Lapusta *et al.* 2000). The fault is driven below depth $z = -24$ km with a loading rate of $V_{pl} = 35$ mm yr^{-1} (Fig. 2a). Then we use values of stress, slip, slip rate and state variable at the onset of one of the earthquakes as an initial condition for an elasto-plastic simulation (Fig. 3b).

Since nucleation of dynamic ruptures occurs spontaneously in the earthquake-cycle model, we need to define the onset of the simulated seismic event. Following Kaneko *et al.* (2010), we define the beginning of an earthquake as the time when maximum slip rate $V_{\max}$ on the fault reaches 1 cm s^{-1}. The maximum slip rate of 1 cm s^{-1} during a self-accelerating nucleation ensures that, at the time of the switch from elastic to elasto-plastic simulations, the inertial effect is still negligible and a seismic rupture will nucleate within a timescale of seconds in the elasto-plastic simulation. Note that the initial conditions are independent of the value of a loading rate V_{pl}; different values of V_{pl} do not change the distribution of pre-stress and stress drop because the recurrence interval changes in proportion.

Since BIM renders a solution only on the fault nodes, we numerically solve the static equilibrium equation for a given stress distribution on the fault and obtain stress tensor everywhere in the medium. The resulting stress state in equilibrium implies that the maximum compressive stress σ_1 is depth-dependent but is approximately at 45° with respect to the fault plane over the depth range. This is because in this study the shear stress $\sigma^{\rm o}_{yz} = \Delta\sigma^{\rm el}_{yz}$ at the onset of a seismic event is small compared to $\sigma^{\rm o}_{xy} = f_{\rm o}\sigma^{\rm o}_{n} + \Delta\sigma^{\rm el}_{xy}$ over most of the seismogenic depth. We do not explore the effect of different angles of the maximum compressive stress on the model response as there are no apparent correlations between the angles of the maximum compressive stress and the inferred shallow slip deficit (Fig. 1). By assigning initial conditions obtained from the elastic model of an earthquake sequence, we implicitly make an assumption that the interseismic inelastic deformation is negligible and does not influence the overall slip budget. The validity of this assumption will be tested in future work.

3 ELASTO-PLASTIC RESPONSE AND SIMULATED SHALLOW SLIP DEFICIT

Figs 4(a)–(c) show a simulated seismic event and the resulting inelastic deformation for a cohesionless material ($c = 0$). The case with zero cohesion may apply, for example, to highly damaged or granulated host rocks. The seismic rupture nucleates at about 13-km depth and propagates updip. In the bottom part of the seismogenic zone at depth ($\gtrsim$7 km), the yield stress, which increases with the mean stress $\sigma_{\rm m}$, is higher than the stress at the tip of the propagating rupture, and hence the material responds elastically. As the rupture propagates towards the Earth surface, stress concentration at the tip of propagating dynamic rupture brings the stress state of the nearby host rocks to the failure envelope, and consequently, inelastic strain starts to accumulate (Figs 4a and b).

The distribution of the magnitude of the accumulated plastic strain γ^p at the end of the seismic event shows a broader zone of extensive inelastic deformation at shallower depths. Since confining pressure is relatively low near the subsurface and the particle velocity is amplified by the presence of the free surface (Fig. 4a), the yielding zone broadens at shallow depths and forms a 'flower-like' zone of accumulated inelastic strain (Fig. 4c), consistent with previous theoretical studies (Ben-Zion & Shi 2005; Ma 2008) and field observations (e.g. Sylvester 1988). The lateral extent of damage predicted by this simulation is similar to inferences of compliant fault zones from geodetic studies (e.g. Fialko *et al.* 2002; Fialko 2004a; Hamiel & Fialko 2007). It is worthwhile to point out that accumulated plastic strain at $z = 0$ is highest at $\sim$1 km away from the fault because the maximum value of the dynamic stress due to propagating mode III rupture with rupture speed $>0.7V_s$ is off of the fault plane (Poliakov *et al.* 2002).

The coseismic slip at the Earth surface in the numerical example shown in Figs 4(a), (b) and (c) is about 15 per cent smaller than the maximum slip at greater depths (Fig. 4e). In this study, we calculate the amount of shallow slip deficit as the ratio of the coseismic slip at $z = 0$ to the maximum coseismic slip. Note that there is 2 per cent shallow slip deficit even for the 'elastic' case in Fig. 4(e) because about 2 per cent of the total slip at the surface is accommodated through interseimic creep (Fig. 3c). In the case of $c = 0$ MPa shown in Fig. 4(e), there is 15 per cent shallow slip deficit. Lower values of rock cohesion (e.g. in pre-damaged rock) promote plastic strain accumulation near the Earth surface (Figs 4c and d) and increase the amount of shallow slip deficit (Fig. 4e). For example, in the case

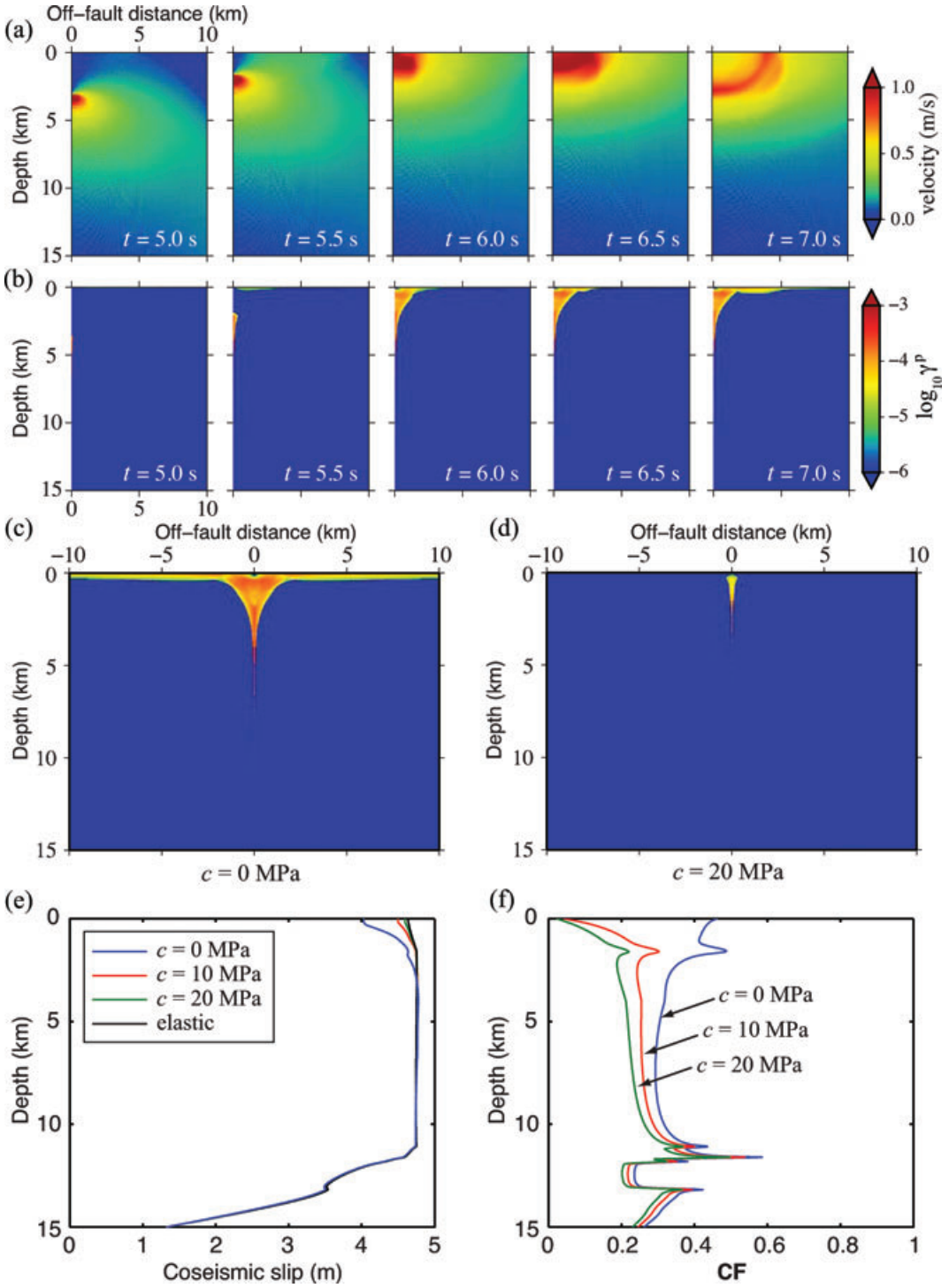

Figure 4. Snapshots of (a) the particle velocity field and (b) the accumulated plastic strain γ^p [eq. (3)] during the seismic event. The case with no cohesion ($c = 0$ MPa) and internal rock friction $\mu = 0.98$ is shown. By the symmetry consideration, the medium across the fault boundary has equal and opposite motion. (c) Distribution of the accumulated plastic strain γ^p at the end of the seismic event for the case shown in panels a and b. (d) Distribution of γ^p for the case with $c = 20$ MPa and $\mu = 0.98$. (e) Coseismic slip distributions for the cases with different values of rock cohesion c. (f) Distributions of the closeness of the stress state to the failure envelop (CF) at the onset of the seismic events. Larger values of CF near the Earth surface lead to a greater amount of shallow slip deficit.

with $c = 20$ MPa, the zone of the inelastic deformation (Fig. 4d) and the corresponding slip deficit (Fig. 4e) are smaller than those in the case with $c = 0$. Hence the slip deficit is larger for a larger amount of inelastic deformation near the Earth surface, as one might expect.

Given the distribution of the accumulated inelastic strain shown in Fig. 4(c), the amount of the resulting slip deficit can be estimated. Shear strain ε is defined as $\varepsilon = u/x_o$, where u is the displacement in the along-strike direction and x_o is the length scale over which u occurs. We set x_o to be the width of the yield zone and ε to be the magnitude of the plastic strain γ^p. From Fig. 4(c), ε is about 10^{-4} over the width x_o of 3 km, and so the resulting slip ($2u$) associated with the inelastic deformation would be $2u = 2\varepsilon x_o = 0.6$ m, which

is in a good agreement with the simulated slip deficit (≈0.7 m) at $z = 0$ in the case of $c = 0$ (Fig. 4e).

To understand the dependence of the amount of shallow slip deficit on the plasticity parameters, we compute the closeness of the stress state to the failure envelop (CF), in terminology of Templeton & Rice (2008). The CF for the antiplane deformation is given by

$$\mathrm{CF} = \frac{r^o}{-\sigma_{\mathrm{m}}^{\mathrm{o}} \sin\phi + c\ \cos\phi}\ , \tag{4}$$

where $r^o = \sqrt{(\sigma_{xx} - \sigma_{yy})^2/4 + (\tau_{xy}^o)^2}$, $r^o = \tau_{xy}^o$ in this study, and τ_{xy}^o is the pre-stress on the fault (Fig. 2b). We compute CF versus depth on the fault plane at the onset of the seismic event (Fig. 4f). The computed shallow slip deficit is promoted when the value of CF near the Earth surface is closer to one (Fig. 4f). This is consistent with the findings by Templeton & Rice (2008) in that the accumulated inelastic deformation is larger for a larger value of CF.

The results discussed so far are obtained by assuming the reference friction $f_{\mathrm{o}} = 0.2$ in the rate and state formulation (1), the value corresponding to friction for clay minerals, fault gouge and serpentine in laboratory experiments (e.g. Reinen *et al.* 1994). The results of the elastic simulations are independent of the assumed value of f_{o} (as long as the resulting friction is non-negative), whereas those of dynamic rupture simulations with off-fault plasticity depend on the absolute level of the stress and hence f_{o}. Here, we also consider the cases with $f_{\mathrm{o}} = 0.6$, the representative value for most igneous and sedimentary rocks for low slip rates (10^{-9}–10^{-3} m s^{-1}) (Marone 1998). Following the same approach, we first simulate an earthquake sequence using the elastic model and then obtain the values of stress, slip, slip rate and state variable at the onset of one of the earthquakes as an initial condition for an elasto-plastic simulation (Fig. 3b).

Fig. 5(a) shows coseismic slip distributions for the cases with $f_{\mathrm{o}} = 0.6$. For a range of rock cohesion $0 \le c \le 30$ MPa, the seismic rupture fails to propagate. This behaviour can be understood from the distribution of CF (4). When $f_{\mathrm{o}} = 0.6$, the pre-stress τ_{xy}^o on the fault (approximately given by $-f_{\mathrm{o}}\sigma_{\mathrm{n}}$) becomes higher than that with $f_{\mathrm{o}} = 0.2$, and the stress concentrations at the tip of the nucleating rupture lead to CF > 1 locally (the cases with $c = 20$ and 30 MPa in Fig. 5b), meaning that the state of stress violates the yield criterion (2). This condition induces localized inelastic deformation at the nucleation depths immediately after the onset of the seismic event (Fig. 5b) and suppresses the propagation of the dynamic rupture. Even if the rupture is nucleated, the stress concentrations due to propagating dynamic rupture (Fig. 5c) can bring the nearby material to yield at greater depths. Hence localized inelastic deformation at greater depths inhibits earthquake nucleation and propagation when the pre-stress on the fault is comparable to the value given by the product of the Byerlee's friction ($f = 0.60$–0.85) and the assumed effective normal stress. It follows that faults may need to be permanently and considerably weakened compared to the ambient crust to enable nucleation of dynamic ruptures. Such scenario is also consistent with laboratory observations that various rock types including igneous and sedimentary rocks exhibit enhanced dynamic weakening at seismic slip rates (∼1 m s^{-1}) (e.g. Di Toro *et al.* 2003; Han *et al.* 2007). Even though we do not explicitly include enhanced dynamic weakening in our models, assuming $f_{\mathrm{o}} = 0.2$ gives rise to high peak friction ∼0.7 and low dynamic friction ∼0.2 at seismic slip rates (Fig. 5c), qualitatively similar to the laboratory observations.

In the cases with $f_{\mathrm{o}} = 0.6$, there is no shallow slip deficit in the seismic events that rupture the entire seismogenic section (the case with $c = 40$ MPa in Fig. 5a). For dynamic rupture to reach the Earth surface, rock cohesion c needs to be sufficiently high to overcome the effect of off-fault plasticity at nucleation depths. Larger rock cohesion also inhibits inelastic deformation at shallow depths, and hence there is no shallow slip deficit in the case with $c = 40$ MPa

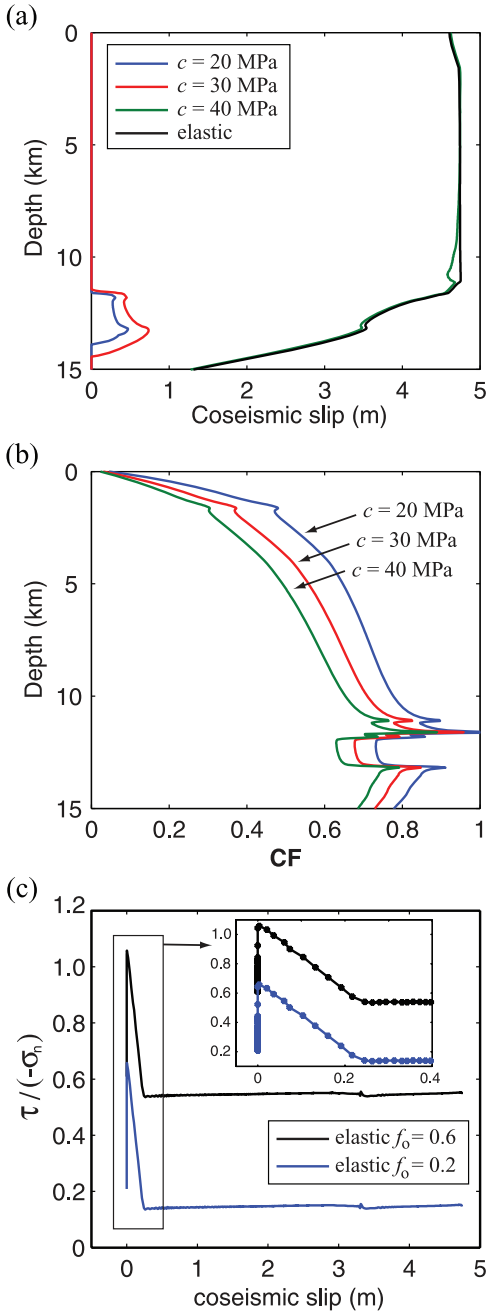

Figure 5. (a) Coseismic slip distributions for the cases with reference fault friction $f_{\mathrm{o}} = 0.6$, internal rock friction $\mu = 0.85$ and different values of rock cohesion c. For $f_{\mathrm{o}} = 0.6$ and lower values of rock cohesion $c \lesssim 30$ MPa, the earthquake rupture fails to propagate due to profound inelastic deformation at depths. (b) Distributions of the closeness of the stress state to the failure envelop (CF) at the onset of the seismic events. For the case with $f_o = 0.6$, the stress reaches the failure envelop at nucleation depths before the weakening commences, suppressing the propagation of dynamic rupture. (c) The effective slip dependence of rate and state friction at $z = -9$ km in the simulations with $f_o = 0.6$ and $f_o = 0.2$.

(the green curve in Fig. 5a). Therefore, to produce shallow slip deficit, not only the rock cohesion needs to be small, but also the value of CF needs to be sufficiently small (CF $\lesssim$ 0.5) at greater depths, which is not the case for $f_o = 0.6$ (Fig. 5b). In the next section, we analyse the dependence of the amount of slip deficit on these model parameters in more detail.

4 DEPENDENCE OF THE AMOUNT OF SLIP DEFICIT ON MODEL PARAMETERS

To understand the dependence of the amount of shallow slip deficit on the plasticity parameters, we compute shallow slip deficit for a number of cases with different values of rock cohesion c and internal rock friction μ (Fig. 6a). Larger values of rock cohesion decrease plastic strain accumulation near the Earth surface and hence decrease the amount of shallow slip deficit (Fig. 6a), consistent with the result discussed in Section 3. When the ratio of friction coefficient on the fault to internal rock friction is closer to one, inelastic deformation at depths delays/inhibits the earthquake nucleation. As a result, the rupture front does not reach the Earth surface or no dynamic rupture commences on a timescale of tens of seconds (cases with 'no rupture' in Fig. 6a).

Fig. 6(a) further shows that larger values of internal rock friction $\mu = \tan\phi$ result in larger shallow slip deficit. Let us consider a simplified situation with $c = 0$ and rewrite CF in eq. (4) as

$$\mathrm{CF}_{c=0} = \frac{\tau^o_{xy}}{-\sigma^o_m \sin\phi} \,. \tag{5}$$

The increase in internal friction μ results in lower values of CF, which should decrease shallow slip deficit if the latter increases with CF as in Section 3. However, the resulting shallow slip deficit is larger for larger values of μ (Fig. 6a).

To understand this dependence, we compare the distributions of CF and coseismic slip in the cases with $\mu = 0.73$ and 0.98 shown by black squares in Fig. 6(a). For $\mu = 0.73$, the largest coseismic slip is near the Earth surface, and there is no resulting shallow slip deficit (Fig. 7a) . The value of CF for $\mu = 0.73$ everywhere on the fault is larger than that for $\mu = 0.98$, resulting in larger plastic strain accumulation. Although the inelastic deformation for $\mu = 0.73$ is confined to a narrower zone at greater depths (Fig. 7c), the magnitude of the plastic strain is generally higher there than near the Earth surface. Consequently, the magnitude of the coseismic slip over much of the seismogenic depths is smaller than that in the elastic case (Fig. 7a). This is why lower internal friction suppresses the amount of shallow slip deficit even though the value of CF becomes larger. These results suggest that the amount of shallow slip deficit cannot be explained by a single value of CF, and is favoured by the conditions where the values of CF are larger at shallower depths ($\lesssim$3 km) and smaller at greater depths such that inelastic deformation preferentially occurs at shallower depths.

Inferred shallow slip deficit for the 2003 Bam earthquake is the largest among all the events shown in Fig. 1. Interestingly, the Bam earthquake resulted in a relatively high stress drop of the order of 10 MPa (Fialko *et al.* 2005). Motivated by this observation, we consider the cases with higher stress-drop events (Fig. 6b). To simulate higher stress-drop events, we assume the distribution of the rate and state parameters $(a_2 - b_2)$ shown in Fig. 3(a), where $a_2 - b_2 = -0.008$ over the depth < 14 km, and set the value of characteristic slip $L =$ 16 mm. Fig. 6(b) shows that, for the higher stress-drop events, the slip deficit generally becomes larger, which may explain why the 2003 Bam earthquake produced the large slip deficit. Note the expanded 'no rupture' domain in Fig. 6(b). The seismic events with higher stress drops lead to larger dynamic stresses and larger plastic strain accumulation, resulting in larger slip deficit in some cases and 'no rupture' in other cases. Even though the slip deficit generally

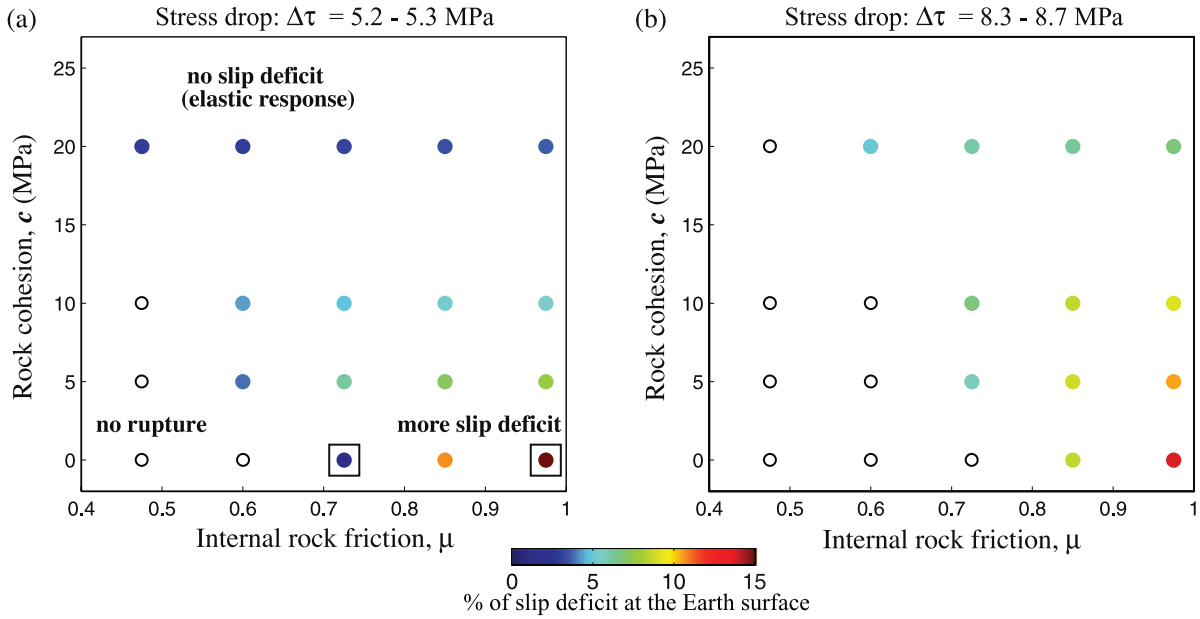

Figure 6. (a) Relation between internal friction, rock cohesion and the amount of simulated shallow slip deficit. Each colour dot represents the amount of shallow slip deficit for the case with $f_o = 0.2$. The range of stress drop for the seismic events (i.e. stress drop averaged over the region of positive stress drop) is indicated. The cases with black squares are discussed in Fig. 7. (b) Simulated slip deficit for the cases with higher stress-drop events. The higher stress-drop events are simulated by assuming the distribution of the rate and state parameters $(a_2 - b_2)$ shown in Fig. 3(a). The resulting slip deficit generally becomes higher in the higher stress-drop events, but the amount of the slip deficit increases by about a few percent.

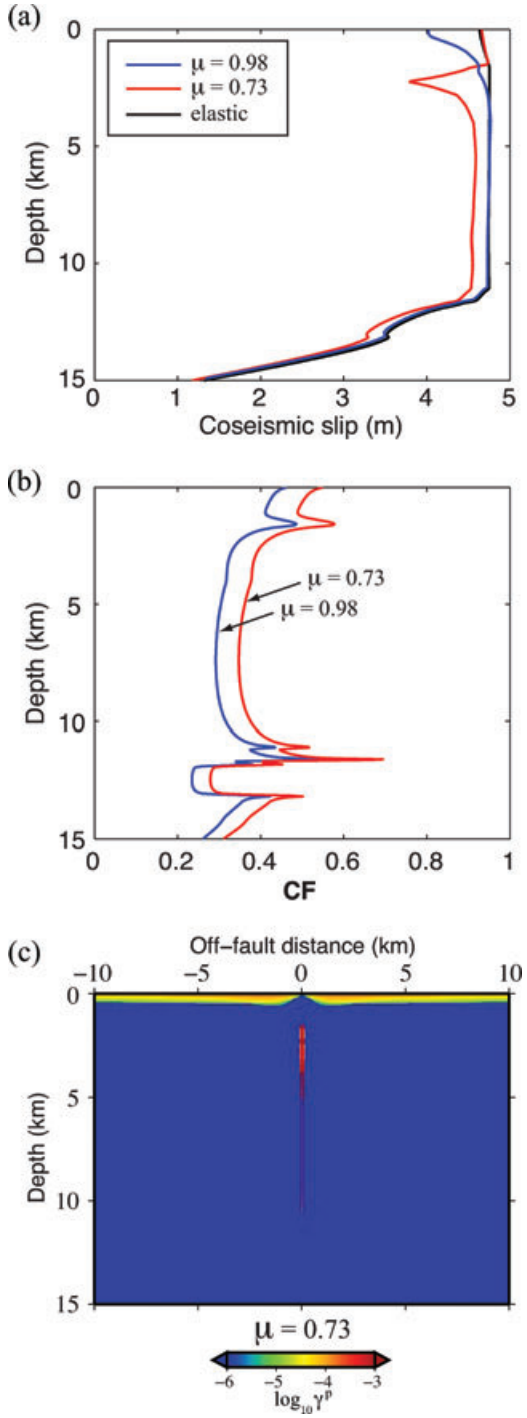

Figure 7. (a) Coseismic slip distributions for the cases indicated in black squares in Fig. 6(a) and the elastic case. For $\mu = 0.73$, the maximum coseismic slip occurs near the Earth surface. (b) Distributions of CF at the onset of the seismic events. (c) Distribution of the accumulated plastic strain γ^p at the end of the seismic event for $\mu = 0.73$. The narrow zone of large inelastic deformation suppresses the coseismic slip at depths, resulting in no shallow slip deficit.

becomes larger in the higher stress-drop events, the magnitude of the coseismic slip also becomes larger and hence the relative amount of the shallow slip deficit increases only slightly (by a few percent in most considered cases). Consequently, an increase in the shallow slip deficit is not linearly proportional to an increase in the static stress drop.

5 EVALUATION OF 'ARTIFICIAL' SLIP DEFICIT IN GEODETIC INVERSIONS THAT ARE BASED ON PURELY ELASTIC MODELS

As stated in Section 1, one explanation for the apparent slip deficit shown in Fig. 1 is a bias in inversions due to yielding of a bulk material. Inversions of geodetic data usually assume that the bulk material is linear elastic and do not account for the potential non-linear response of the medium surrounding the fault surface. We investigate this potential bias by inverting a displacement profile on the Earth surface obtained in our elasto-plastic model for a fault slip distribution at depth using a purely elastic forward model.

For a 2-D screw dislocation in a homogeneous elastic half-space, the analytical solution for the surface displacement $u(x)$ due to the fault slip δ within the depth interval $[z, z + \Delta z]$ is given by

$$u(x) = \frac{\delta}{\pi}\left(\tan^{-1}\frac{z+\Delta z}{x} - \tan^{-1}\frac{z}{x}\right). \quad (6)$$

Solution for an arbitrary distributed slip can be obtained by superpositions of eq. (6) for all depth intervals Δz on the fault segment. Conversely, one can invert a given surface displacement $u(x)$ for a slip distribution $\delta(z)$ using eq. (6) as a forward model.

We consider the surface displacement at the end of the seismic event shown in Figs 4(a) and (b). Our computational domain is finite, whereas the analytical solution (6) used for the inversion of the surface displacement assumes a semi-infinite plane. Due to the reflection of wave energy from the domain boundaries, the decay rate of the static displacement profile in our model is slower than that of the analytical solution. To reduce the influence of the reflection of waves from the boundaries, we increase the domain size by a factor of 2 and use the surface displacement only up to 15 km away from the fault ($x \leq \pm 15$ km) (Fig. 8a). Solid black curves in Figs 8(a) and (b) correspond to the simulated surface displacements and the slip distribution, respectively. The decrease in the displacement amplitude near the fault trace is caused by subsurface inelastic deformation that acts to reduce coseismic strain near the fault (the black curve in Fig. 8a).

To quantify the potential bias due to yielding of a bulk material, we perform two sets of inversions. We first obtain surface displacements predicted by the elastic solution (6) using the simulated slip distribution as an input. Then we invert these surface displacements (the blue curve in Fig. 8a) for a slip distribution at depths. As inversions of surface displacements are intrinsically non-unique, we impose the non-negativity and smoothness constraints to regularize the problem (e.g. Fialko 2004a; Barbot *et al.* 2008). The resulting slip distribution (the blue curve in Fig. 8b) is in a good agreement with the input slip distribution (the black curve in Fig. 8b), indicating robustness of the inversion procedure. We then invert the coseismic surface displacement obtained from the elasto-plastic simulation (the black curve in Fig. 8a) for a fault slip distribution at depth using eq. (6). As expected, the inferred slip distribution (the red curve in Fig. 8b) does not match well with the input model although the best-fitting surface displacements (the red curve in Fig. 8a) exhibit a good agreement with the synthetic data (the black curve).

The difference between the slip distributions obtained from these inversions (the red and blue curves in Fig. 8b) is a measure of bias in inversions that ignore yielding and inelastic deformation of the bulk material (Fig. 8c). The difference of the inverted slip distributions is about zero at the Earth surface and +0.5 m at a depth of the peak slip ($z \approx -4$ km) (Fig. 8c). This result suggests that the neglect of

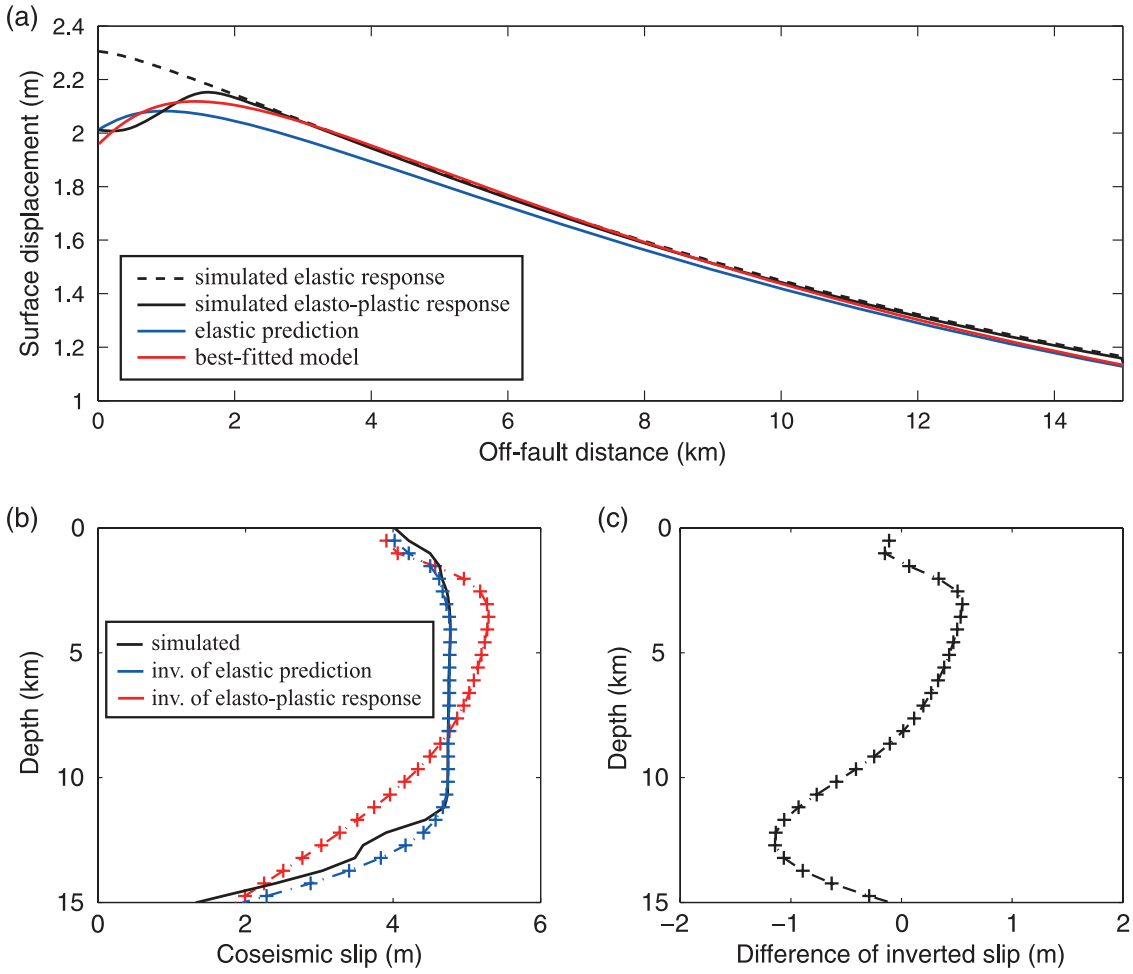

Figure 8. (a) The profiles of coseismic displacement on the Earth surface for the elastic simulation, an elasto-plastic simulation (with $\mu = 0.98$, $c = 0$ MPa), an elastic prediction for the simulated slip distribution shown in panel b, and the best-fitting model to the elasto-plastic response in the inversion. The inversions and forward predictions are done using the analytical solution (6). (b) The slip distribution in the elasto-plastic simulation and slip inversions of the surface displacements shown in panel a. The blue curve is the inverted slip of the surface displacement of the elastic prediction. The red curve is the inverted slip of the surface displacement of the simulated elasto-plastic response. (c) Difference between the inverted slip distributions shown in panel b. The difference shows that 10 per cent (0.5 m) of the coseismic slip at 3–4 km depths is due to an artefact of the inversion that is based on the elastic model.

inelastic effects may give rise to an overestimation of the shallow slip by as much as 10 per cent of the maximum coseismic slip at depth. In addition, inversions neglecting off-fault yielding tend to underestimate slip at the greater depths by as much as 20 per cent (Fig. 8c). The difference indicates that there would be an 'artificial deficit' in inversions that are based on a purely elastic model, if off-fault plastic yielding commonly occurs during large earthquakes. Our results motivate reinterpretation of existing data using forward models that explicitly include inelastic deformation in the near field of earthquake ruptures.

6 DISCUSSION

In all the cases we have considered, the amount of 'true' shallow slip deficit due to coseismic inelastic off-fault deformation is, at most, of the order of 15 per cent (Section 4). An additional 10 per cent of 'artificial deficit' can be introduced in inversions of geodetic data that are based on purely elastic models due to reduction of near-field coseismic strain by off-fault yielding (Section 5). The largest magnitude of the slip deficit in our models (with a cohesionless material and large internal rock friction) combined with the bias in inversions gives shallow slip deficit of up to 25 per cent, which may be insufficient to explain 30–60 per cent deficit deduced from observations (Fig. 1). We discuss factors that are not included in our model but could be relevant for the remaining discrepancy (Sections 6.1–6.4). We also discuss the potential consequence of shallow slip deficit over a geologic timescale (Section 6.5).

6.1 Shallow interseismic fault creep below detection limit

For faults with low slip rates and large recurrence intervals such as those represented in Fig. 1, unresolvable shallow creep during the interseismic period may account for some of the inferred slip deficit. Since most geodetic measurements span only a few decays, interseismic creep rates smaller than $\sim$0.1 mm yr^{-1} may be unresolvable. One can estimate the contribution of potential interseismic creep to the overall slip budget. If we assume the fault slip rate of 1 mm yr^{-1}, the recurrence interval of $\sim$5000 years, and the surface creep rate of $\sim$0.1 mm yr^{-1}, the total shallow creep during the interseismic period is 0.5 m, accounting for up to $\sim$10 per cent of the average coseismic slip (Fig. 1). While this effect alone is insufficient to explain the inferred coseismic slip deficit, the occurrence of such

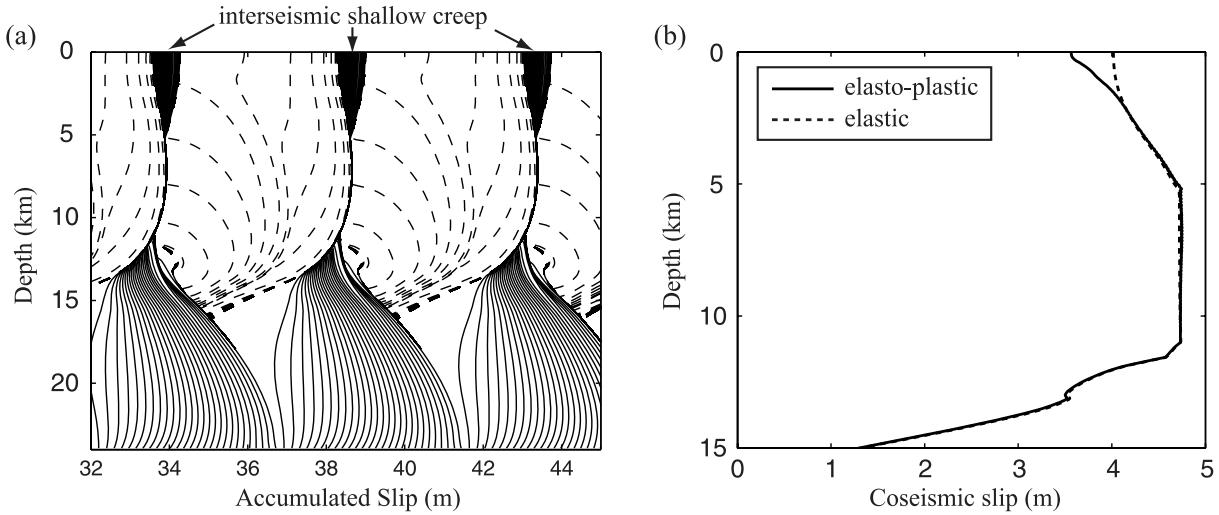

Figure 9. (a) An earthquake sequence in an elastic model with a velocity-strengthening shallow layer and a slower plate loading rate $V_{pl} = 1$ mm yr^{-1}. In this case, we use the distribution of the rate and state parameters $(a_3 - b_3)$ shown in Fig. 3(a). Solid lines show slip accumulation every 200 years, whereas dashed lines are plotted above 18-km depth every second during the simulated earthquakes. The surface creep rate during the interseismic period is $\approx$0.15 mm yr^{-1} and the recurrence interval is 4800 years. (b) Coseismic slip distributions for the elastic case shown in panel a and the corresponding elasto-plastic case with $c = 0$ MPa and $\mu = 0.98$. The amount of coseismic slip deficit in the elasto-plastic and elastic simulations is 25 per cent and 15 per cent, respectively. Shallow fault creep during interseismic periods leads to a larger coseismic slip deficit.

slow creep may increase the total amount of coseismic slip deficit in our elasto-plastic models.

Motivated by this argument, we further consider scenarios with a velocity-strengthening shallow layer that result in interseismic creep (Fig. 9a). To reproduce a fault with a low slip rate and a large recurrence interval, we assign a low loading rate $V_{pl} = 1$ mm yr^{-1} such that the resulting recurrence interval becomes 4800 years and the surface creep rate is $\approx$0.15 mm yr^{-1} (Fig. 9a). As expected, the occurrence of interseismic creep increases the amount of coseismic slip deficit in the elastic simulation (Fig. 9b). The slip deficit in the elasto-plastic simulation in this case is 25 per cent (Fig. 9b), larger than that in the case with a velocity-weakening shallow layer and with the same plasticity parameters (Fig. 4e). The difference in coseismic slip deficit between the elasto-plastic and elastic simulations is 10 per cent (Fig. 9b), smaller than that in the case with a velocity-weakening shallow layer (13 per cent in Fig. 4e). Since the existence of the shallow velocity-strengthening layer impedes the dynamic rupture, the resulting plastic strain accumulation is smaller than that without the velocity-strengthening layer. Nevertheless, up to 10 per cent of the slip deficit is due to coseismic inelastic deformation at shallow depths. These results suggest that the occurrence of slow interseismic creep could account for some of the remaining discrepancy between shallow slip deficit predicted by our models and that inferred from inversions of geodetic data.

6.2 Effects of three-dimensional (3-D) geometry

In this study, we use 2-D (antiplane) elasto-plastic models of spontaneous dynamic rupture that assume no variations of slip or material properties along the fault strike. Using 3-D simulations of spontaneous dynamic rupture, Ma (2008) showed that extensive inelastic deformation due to large strike-slip earthquakes occurs along strike and is mostly distributed on the extensional side of the fault. Based on dynamic rupture simulations in a 2-D in-plane setting, Templeton & Rice (2008) found that plastic deformation occurs primarily on the compressional side of the fault for shallow angles between the most compressive stress axis and the fault plane, and on the extensional side for higher angles of maximum compression axis. These results indicate that there may be more inelastic strain accumulation in 3-D models such that, when averaging the coseismic slip distribution along the fault strike, the amount of shallow slip deficit may become larger than that in the 2-D models considered in this study. We point out that 3-D effects are likely to be important for the $M7$ strike-slip earthquakes such as those included in Fig. 1.

6.3 Enhanced dynamic weakening at coseismic slip rates

There is growing evidence that friction may be much lower at seismic slip velocities than rate and state friction laws predict (e.g. Di Toro *et al.* 2003; Han *et al.* 2007; Rice 2006; Noda *et al.* 2009, and references therein). In this study, we assume that the fault constitutive response is represented by the rate and state friction formulation without accounting for the enhanced weakening at seismic slip rates of the order of 1 m s^{-1}. Our model reproduces the aseismic nucleation process in gradually varying zones of accelerating slip and the subsequent inertially controlled events with realistic slip velocities and rupture speeds. Yet, the dynamic stress fields and the amount of inelastic deformation may be smaller in our model than that with enhanced weakening, which causes abrupt and large changes in the stress field at the tip of the propagating rupture. A locally larger dynamic stress change due to such enhanced weakening may induce a larger amount of inelastic deformation near the fault surface and result in greater shallow slip deficit. What remains to be investigated is whether enhanced dynamic weakening could lead to the conditions where shallow slip deficit is mechanically favoured; that is, the stress state could be closer to the failure envelop (CF) at shallower depths ($\lesssim$3 km) than at greater depths such that inelastic deformation preferentially occurs at shallower depths. Other mechanisms that may increase damage generation at shallow depths include strong dynamic reduction of normal stress during seismic slip due to collision of rough fault surfaces (Lomnitz-Adler 1991; Brune *et al.* 1993) or a material contrast across a fault (Andrews & Ben-Zion 1997).

6.4 Inelastic deformation during interseismic periods

Following previous studies of dynamic rupture with off-fault plasticity, we have focused on single-rupture scenarios and assumed that interseismic inelastic deformation in the shallow crust is negligible in the overall slip budget. However, any interseismic inelastic deformation at shallow depths would reduce the ambient stress and therefore the amount of coseismic slip in the subsequent event. This may potentially account for the discrepancy between the simulated slip deficit (Fig. 6) and the observations (Fig. 1). It is also possible that the effects of interseismic inelastic deformation might be small because, even if the shallow crust next to a fault yields during the initial dynamic rupture events, it may respond elastically to subsequent events and secular loading. After multiple earthquake cycles, the model could evolve to self-consistent stress conditions that are compatible with parameterization of the model, while inelastic strain will keep accumulating over time. Quantifying interseismic inelastic deformation and its contribution to shallow slip deficit over multiple earthquake cycles needs to be addressed in future work.

6.5 Discrepancies between slip rates estimated from geology and geodetic methods

If shallow slip deficit due to off-fault inelastic deformation systematically occurred over many earthquake cycles, there would be a difference between the cumulative offset near the fault trace (within tens of metres) and at distances far from the fault (several kilometres away). This difference may bear on inferred discrepancies between slip rates estimated from geology and those estimated by geodetic methods (e.g. Oskin & Iriondo 2004; Titus *et al.* 2006). The presence of shallow slip deficit implies that the geologic rates would be smaller than the geodetic rates. Some studies (e.g. Matmon *et al.* 2005; Oskin *et al.* 2008) reported up to 200 per cent discrepancies between the estimates of geologic and geodetic fault slip rates, whereas other studies (e.g. Meade & Hager 2005; Chuang & Johnson 2011) argued that the discrepancies are much smaller ($\lesssim$10–20 per cent). An agreement between geologic and geodetic slip rates on mature faults (e.g. Fialko 2006) implies that shallow slip deficit may be restricted to young or infrequently slipping faults (Fig. 1). Estimates of shallow slip deficit for very long strike-slip ruptures on mature faults so far are scarce, in part due to poor correlation of InSAR data in the near field of recent strike-slip earthquakes with moment magnitude (M_w) greater than 7.5. Tong *et al.* (2010) found no shallow slip deficit for the 2008 M_w7.8 Wenchuan (China) earthquake. It remains to be seen if the amount of shallow slip deficit depends on the rupture size, fault slip rate, cumulative offset, recurrence interval, etc.

7 CONCLUSIONS

Using simulations of spontaneous dynamic rupture with off-fault yielding, we have investigated whether the occurrence of inelastic deformation can account for the shallow slip deficit inferred from slip inversions of several well-documented M7 earthquakes. Our results suggest that shallow slip deficit due to coseismic inelastic deformation occurs under a wide range of parameters that characterize strength of the uppermost crust. The main results can be summarized as follows.

The amount of shallow slip deficit scales with the amount of inelastic deformation near the Earth surface. Lower values of rock cohesion (e.g. in pre-damaged rock) promote plastic strain accumulation near the Earth surface and hence increase the amount of slip deficit. When the ratio of friction coefficient on the fault to internal rock friction is closer to one, inelastic deformation at depths suppresses dynamic rupture propagation and, in some cases, even inhibits earthquake nucleation. Shallow slip deficit is often suppressed when there is strong inelastic deformation at depths greater than 3–4 km. Hence shallow slip deficit is favoured when the stress state is closer to the failure envelop (CF) at shallower depths ($\lesssim$3 km) and further away at greater depths such that inelastic deformation preferentially occurs at shallower depths. Earthquakes with higher stress drop would lead to larger dynamic stresses and larger plastic strain accumulation, leading to larger slip deficit in some situations.

We have found that the amount of 'true' shallow slip deficit due to inelastic off-fault deformation is, at most, of the order of 15 per cent in the cases that we have considered. Inelastic deformation, however, reduces coseismic strain near the fault, which can introduce an additional 'artificial' deficit of up to 10 per cent of the maximum slip in inversions of geodetic data that are based on purely elastic models. The combined effect is close to, but still smaller than the magnitudes of the inferred shallow slip deficit in inversions of geodetic data, although potentially important factors that are not included in this study may account for the discrepancy. Our results further indicate that off-fault yielding during dynamic rupture propagation may explain kilometre-wide fault damage zones imaged by geodetic and seismic studies.

ACKNOWLEDGMENTS

This study was supported by the National Science Foundation (Grant EAR 1114019), Scripps Institution of Oceanography Postdoctoral Fellowship, and the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 1475. We thank Sylvain Barbot for providing us the inversion code. We also thank Jean-Paul Ampuero for helpful discussion. The paper benefited from comments by two anonymous reviewers and Editor Yehuda Ben-Zion.

REFERENCES

Andrews, D.J., 2005. Rupture dynamics with energy loss outside the slip zone, *J. geophys. Res.,* **110,** doi:10.1029/2004JB003191.

Andrews, D.J. & Ben-Zion, Y., 1997. Wrinkle-like slip pulse on a fault between different materials, *J. geophys. Res.,* **102,** 553–571, doi:10.1029/96JB02856.

Barbot, S., Fialko, Y. & Sandwell, D., 2008. Effect of a compliant fault zone on the inferred earthquake slip distribution, *J. geophys. Res.,* **113,** B06404, doi:10.1029/2007JB005256.

Ben-Zion, Y. & Shi, Z., 2005. Dynamic rupture on a material interface with spontaneous generation of plastic strain in the bulk, *Earth planet. Sci. Lett.,* **236,** 486–496, doi:10.1016/j.epsl.2005.03.025.

Bilham, R., 2010. Lessons from the Haiti earthquake, *Nature,* **463,** 878–879, doi:10.1038/463878a.

Blanpied, M.L., Lockner, D.A. & Byerlee, J.D., 1991. Fault stability inferred from granite sliding experiments at hydrothermal conditions, *Geophys. Res. Lett.,* **18,** 609–612, doi:10.1029/91GL00469.

Blanpied, M.L., Lockner, D.A. & Byerlee, J.D., 1995. Frictional slip of granite at hydrothermal conditions, *J. geophys. Res.,* **100,** 13 045–13 064, doi:10.1029/95JB00862.

Brune, J.N., Brownb, S. & Johnsona, P.A., 1993. Rupture mechanism and interface separation in foam rubber models of earthquakes: a possible solution to the heat flow paradox and the paradox of large overthrusts, *Tectonophysics,* **218,** 59–67, doi:10.1016/0040-1951(93)90259-M.

Chester, F.M. & Chester, J.S., 1998. Ultracataclasite structure and friction processes of the Punchbowl fault, San Andreas system, California, *Tectonophysics,* **295,** 199–221, doi:10.1016/S0040-1951(98)00121-8.

Chester, F.M., Evans, J.P. & Biegel, R.L., 1993. Internal structure and weakening mechanisms of the San Andreas Fault, *J. geophys. Res.,* **98,** 771–786, doi:10.1029/92JB01866.

Chuang, R.Y. & Johnson, K.M., 2011. Reconciling geologic and geodetic model fault slip-rate discrepancies in Southern California: consideration of nonsteady mantle flow and lower crustal fault creep, *Geology,* **39,** 627–630, doi:10.1130/G32120.1.

Clayton, R. & Engquist, B., 1977. Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. seism. Soc. Am.,* **67,** 1529–1540.

Cocco, M. & Bizzarri, A., 2002. On the slip-weakening behavior of rate and state dependent constitutive laws, *Geophys. Res. Lett.,* **29**(11), doi:10.1029/2001GL013999.

Cochran, E.S., Li, Y.-G., Shearer, P.M., Barbot, S., Fialko, Y. & Vidale, J.E., 2009. Seismic and geodetic evidence for extensive, long-lived fault damage zones, *Geology,* **37,** 315–318, doi:10.1130/G25306A.1.

Davis, R.O. & Selvadurai, A.P.S., 2002. *Plasticity and Geomechanics,* Cambridge University Press, New York, NY.

Day, S.M., Dalguer, L.A., Lapusta, N. & Liu, Y., 2005. Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture, *J. geophys. Res.,* **110,** B12307, doi:10.1029/2005JB003813.

Di Toro, G., Goldsby, D.L. & Tullis, T.E., 2003. Friction falls towards zero in quartz rock as slip velocity approaches seismic rates, *Nature,* **427,** doi:10.1038/nature02249.

Dieterich, J.H., 1978. Time-dependent friction and the mechanics of stick-slip, *Pure appl. Geophys.,* **116,** 790–806, doi:10.1007/BF00876539.

Dieterich, J.H., 1979. Modeling of rock friction: 1. Experimental results and constitutive equations, *J. geophys. Res.,* **84,** 2161–2168, doi:10.1029/JB084iB05p02161.

Dor, O., Ben-Zion, Y., Rockwell, T.K. & Brune, J., 2006. Pulverized rocks in the Mojave section of the San Andreas Fault Zone, *Earth planet. Sci. Lett.,* **245,** 642–654, doi:10.1016/j.epsl.2006.03.034.

Drucker, D.C. & Prager, W., 1952. Soil mechanics and plastic analysis or limit design, *Q. J. Appl. Math.,* **10,** 157–165.

Fialko, Y., 2004a. Probing the mechanical properties of seismically active crust with space geodesy: study of the co-seismic deformation due to the 1992M_w7.3 Landers (southern California) earthquake, *J. geophys. Res.,* **109,** doi:10.1029/2003JB002756.

Fialko, Y., 2004b. Evidence of fluid-filled upper crust from observations of post-seismic deformation due to the 1992M_w7.3 Landers earthquake, *J. geophys. Res.,* **109,** doi:10.1029/2004JB002985.

Fialko, Y., 2006. Interseismic strain accumulation and the earthquake potential on the southern San Andreas fault system, *Nature,* **441**, 968–971, doi:10.1038/nature04797.

Fialko, Y., Sandwell, D., Agnew, D., Simons, M., Shearer, P. & Minster, B., 2002. Deformation on nearby faults induced by the 1999 Hector Mine earthquake, *Science,* **297,** 1858–1862, doi:10.1126/science.1074671.

Fialko, Y., Sandwell, D., Simons, M. & Rosen, P., 2005. Three-dimensional deformation caused by the Bam, Iran, earthquake and the origin of shallow slip deficit, *Nature,* **435,** 295–299, doi:10.1038/nature03425.

Fialko, Y., Gonzalez, A., Gonzalez, J., Barbot, S., Leprince, S., Sandwell, D. & Agnew, D., 2010. Static rupture model of the 2010 M7.2 El Mayor-Cucapah earthquake from ALOS, ENVISAT, SPOT and GPS data, *EOS, Trans. Am. geophys. Un.,* **91**(52), T53B–2125.

Fielding, E.J., Lundgren, P.R., Bürgmann, R. & Funning, G.J., 2009. Shallow fault-zone dilatancy recovery after the 2003 Bam earthquake in Iran, *Nature,* **458,** 64–68, doi:10.1038/nature07817.

Finzi, Y., Hearn, E.H., Zion, Y.B. & Lyakhovsky, V., 2009. Structural properties and deformation patterns of evolving strike-slip faults: numerical simulations incorporating damage rheology, *Pure appl. Geophys.,* **166,** 1537–1573, doi:10.1007/978-3-0346-0138-2_2.

Hamiel, Y. & Fialko, Y., 2007. Structure and mechanical properties of faults in the North Anatolian Fault system from InSAR observations of coseismic deformation due to the 1999 Izmit (Turkey) earthquake, *J. geophys. Res.,* **112,** B07412, doi:10.1029/2006JB004777.

Han, R., Shimamoto, T., Hirose, T., Ree, J.-H. & Ando, J., 2007. Ultralow friction of carbonate faults caused by thermal decomposition, *Science,* **316,** doi:10.1126/science.1139763.

Harris, R.A. *et al.*, 2009. The SCEC/USGS dynamic earthquake rupture code validation exercise, *Seismol. Res. Lett.,* **80**(1), 119–126, doi:10.1785/gssrl.80.1.119.

Hearn, E., Bürgmann, R. & Reilinger, R., 2002. Dynamics of Izmit earthquake postseismic deformation and loading of the Duzce earthquake hypocenter, *Bull. seism. Soc. Am.,* **92,** 172–193, doi:10.1785/0120000832.

Ida, Y., 1972. Cohesive force across the tip of a longitudinal-shear crack and griffith's specific surface energy, *J. geophys. Res.,* **77,** 3796–3805, doi:10.1029/JB077i020p03796.

Jacobs, A., Sandwell, D., Fialko, Y. & Sichoix, L., 2002. The 1999 (M_w7.1) Hector Mine, California, earthquake: near-field postseismic deformation from ERS interferometry, *Bull. seism. Soc. Am.,* **92,** 1433–1442, doi:10.1785/0120000908.

Kaneko, Y., 2009. Investigations of earthquake source processes based on fault models with variable friction rheology, *PhD thesis*, California Institute of Technology, Pasadena, CA.

Kaneko, Y. & Lapusta, N., 2010. Supershear transition due to a free surface in 3-D simulations of spontaneous dynamic rupture on vertical strike-slip faults, *Tectonophysics,* **493,** 272–284, doi:10.1016/j.tecto.2010.06.015.

Kaneko, Y., Lapusta, N. & Ampuero, J.-P., 2008. Spectral element modeling of spontaneous earthquake rupture on rate and state faults: effect of velocity-strengthening friction at shallow depths, *J. geophys. Res.,* **113,** B09317, doi:10.1029/2007JB005553.

Kaneko, Y., Avouac, J.-P. & Lapusta, N., 2010. Towards inferring earthquake patterns from geodetic observations of interseismic coupling, *Nat. Geosci.,* **3,** 363–369, doi:10.1038/NGEO843.

Lapusta, N., Rice, J., Ben-Zion, Y. & Zheng, G., 2000. Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction, *J. geophys. Res.,* **105**(B10), 23 765–23 789, doi:10.1029/2000JB900250.

Lewis, M.A. & Ben-Zion, Y., 2010. Diversity of fault zone damage and trapping structures in the Parkfield section of the San Andreas Fault from comprehensive analysis of near fault seismograms, *Geophys. J. Int.,* **183,** 1579–1595, doi:10.1111/j.1365-246X.2010.04816.x.

Li, Y.G., Vidale, J.E., Aki, K., Xu, F. & Burdette, T., 1998. Evidence of shallow fault zone strengthening after the 1992 M7.5 Landers, California, earthquake, *Science,* **279,** 217–219, doi:10.1126/science.279.5348.217.

Lomnitz-Adler, J., 1991. Model for steady state friction, *J. geophys. Res.,* **96,** 6121–6131, doi:10.1029/90JB02536.

Ma, S., 2008. A physical model for widespread near-surface and fault zone damage induced by earthquakes, *Geochem. Geophys. Geosyst.,* **9,** doi:10.1029/2008GC002231.

Mai, P.M. *et al.*, 2007. Source-inversion blindtest: initial results and further developments, *Geophys. Res. Abstr.,* **9,** 07351, SRef-ID: 1607-7962/gra/EGU2007-A-07351.

Marone, C., 1998. Laboratory-derived friction laws and their application to seismic faulting, *Annu. Rev. Earth Planet. Sci.,* **26,** 643–696, doi:10.1146/annurev.earth.26.1.643.

Marone, C., Scholz, C.H. & Bilham, R., 1991. On the mechanics of earthquake afterslip, *J. geophys. Res.,* **96,** 8441–8452, doi:10.1029/91JB00275.

Matmon, A., Schwartz, D.P., Finkel, R., Clemmens, S. & Hanks, T., 2005. Dating offset fans along the Mojave section of the San Andreas fault using cosmogenic ^{26}Al and ^{10}Be, *Geol. Soc. Am. Bull,* **117,** 795–807, doi:10.1130/B25590.1.

Meade, B.J. & Hager, B.H., 2005. Block models of crustal motion in southern California constrained by GPS measurements, *J. geophys. Res.,* **110,** B03403, doi:10.1029/2004JB003209.

Noda, H., Dunham, E.M. & Rice, J.R., 2009. Earthquake ruptures with thermal weakening and the operation of major faults at low overall stress levels, *J. geophys. Res.,* **114,** B07302, doi:10.1029/2008JB006143.

Oskin, M. & Iriondo, A., 2004. Large-magnitude transient strain accumulation on the Blackwater fault, Eastern California shear zone, *Geology,* **32,** 313–316, doi:10.1130/G20223.1.

Oskin, M., Perg, L., Shelef, E., Strane, M., Gurney, E., Singer, B. & Zhang, Z., 2008. Elevated shear zone loading rate during an earthquake

cluster in eastern California, *Geology,* **36,** 507–510, doi:10.1130/G24814A.1.

Peng, Z., Ben-Zion, Y., Michael, A.J. & Zhu, L., 2003. Quantitative analysis of seismic trapped waves in the rupture zone of the 1992 Landers, California earthquake: evidence for a shallow trapping structure, *Geophys. J. Int.,* **155,** 1021–1041, doi:10.1111/j.1365-246X.2003.02109.x.

Pitarka, A., Dalguer, L.A., Day, S.M., Somerille, P.G. & Dan, K., 2009. Numerical study of ground-motion differences between buried-rupturing and surface-rupturing earthquakes, *Bull. seism. Soc. Am.,* **99,** 1521–1537, doi:10.1785/0120080193.

Poliakov, A.N.B., Dmowska, R. & Rice, J.R., 2002. Dynamic shear rupture interactions with fault bends and off-axis secondary faulting, *J. geophys. Res.,* **107**(B11), doi:10.1029/2001JB000572.

Reinen, L.A., Weeks, J.D. & Tullis, T.E., 1994. The frictional behavior of lizardite and antigorite serpentinites: experiments, constitutive models, and implications for natural faults, *Pure appl. Geophys.,* **143,** 317–358, doi:10.1007/BF00874334.

Rice, J.R., 1983. Constitutive relations for fault slip and earthquake instabilities, *Pure appl. Geophys.,* **121,** 443–475, doi:10.1007/BF02590151.

Rice, J.R., 1993. Spatio-temporal complexity of slip on a fault, *J. geophys. Res.,* **98**(B6), 9885–9907, doi:10.1029/93JB00191.

Rice, J.R., 2006. Heating and weakening of faults during earthquake slip, *J. geophys. Res.,* **111,** doi:10.1029/2005JB004006.

Rice, J.R. & Ben-Zion, Y., 1996. Slip complexity in earthquake fault models, *Proc. Nat. Acad. Sci. U.S.A.,* **93,** 3811–3818, doi:10.1073/pnas.93.9.3811.

Ruina, A.L., 1983. Slip instability and state variable friction laws, *J. geophys. Res.,* **88,** 10 359–10 370, doi:10.1029/JB088iB12p10359.

Rybicki, K.R., 1992. Strike-slip faulting in the presence of low-rigidity inhomogeneities, *Bull. seism. Soc. Am.,* **82**(5), 2170–2190.

Rybicki, K.R. & Yamashita, T., 1998. Faulting in vertically inhomogeneous media and its geophysical implications, *Geophys. Res. Lett.,* **25**(15), 2893–2896, doi:10.1029/2002GL014672.

Savage, J. & Lisowski, M., 1993. Inferred depth of creep on the Hayward fault, central California, *J. geophys. Res.,* **98,** 787–793, doi:10.1029/92JB01871.

Savage, J.C., Byerlee, J.D. & Lockner, D.A., 1996. Is internal friction friction?, *Geophys. Res. Lett.,* **23**(5), 487–490.

Scholz, C.H., 1998. Earthquakes and friction laws, *Nature,* **391,** 37–42, doi:10.1038/34097.

Simons, M., Fialko, Y. & Rivera, L., 2002. Coseismic deformation from the 1999 Mw 7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations, *Bull. seism. Soc. Am.,* **92,** 1390–1402, doi:10.1785/0120000933.

Somerville, P.G., 2003. Magnitude scaling of the near fault rupture directivity pulse, *Phys. Earth planet. Inter.,* **137,** 201–212, doi:10.1016/S0031-9201(03)00015-3.

Spudich, P. & Olsen, K.B., 2001. Fault zone amplified waves as a possible seismic hazard along the Calaveras Fault in central California, *Geophys. Res. Lett.,* **28**(13), 2533–2536, doi:10.1029/2000GL011902.

Sylvester, A., 1988. Strike-slip faults, *Geol. Soc. Am. Bull,* **100,** 1666–1703, doi:10.1130/0016-7606(1988)100<1666:SSF>2.3.CO;2.

Templeton, E.L. & Rice, J.R., 2008. Off-fault plasticity and earthquake rupture dynamics, 1. Dry materials or neglect of fluid pressure changes, *J. geophys. Res.,* **113,** doi:10.1029/2007JB005529.

Titus, S.J., DeMets, C. & Tikoff, B., 2006. Thirty-five-year creep rates for the creeping segment of the San Andreas fault and the effects of the 2004 Parkfield eart hquake: constraints from alignment arrays, continuous Global Positioning System, and creepmeters, *Bull. seism. Soc. Am.,* **96,** S250–S268, doi:10.1785/0120050811.

Tong, X., Sandwell, D. & Fialko, Y., 2010. Coseismic slip model of the 2008 wenchuan earthquake derived from joint inversion of interferometric synthetic aperture radar, GPS, and field data, *J. geophys. Res.,* **115,** doi:10.1029/2009JB006625.

Vidale, J.E. & Li, Y.-G., 2003. Damage to the shallow Landers fault from the nearby Hector Mine earthquake, *Nature,* **421,** 524–526, doi:10.1038/nature01354.

Wei, M., Sandwell, D. & Fialko, Y., 2009. A silent M4.8 slip event of October 3-6, 2006, on the Superstition Hills fault, Southern California, *J. geophys. Res.,* doi:10.1029/2008JB006135.

Wei, M., Sandwell, D., Fialko, Y. & Bilham, R., 2011. Slip on faults in the Imperial Valley triggered by the 4 April 2010 Mw 7.2 El Mayor-Cucapah earthquake revealed by INSAR, *Geophys. Res. Lett.,* **38,** doi:10.1029/2010GL045235.

Yamashita, T., 2000. Generation of microcracks by dynamic shear rupture and its effects on rupture growth and elastic wave radiation, *Geophys. J. Int.,* **143,** 395–406, doi:10.1046/j.1365-246X.2000.01238.x.

APPENDIX A: RATE AND STATE FRICTION REGULARIZED AT ZERO SLIP VELOCITY

In expression (1), shear frictional strength τ is undefined for slip velocities $V = 0$, which is unphysical. To regularize (1) near $V = 0$, we follow the approach of Rice & Ben-Zion (1996) and Lapusta *et al.* (2000) in using a thermally activated creep model of the direct effect term $a \ln(V/V_0)$ to obtain

$$\tau = -a\sigma_{\rm n} \operatorname{arcsinh}\left[\frac{V}{2V_0}\exp\left(\frac{f_0 + b\ln(V_0\theta/L)}{a}\right)\right]. \qquad \text{(A1)}$$

This regularization produces a negligible change from eq. (1) in the range of slip velocities explored by laboratory experiments; the difference in V at $V \sim V_0$ is of the order of $\exp(-2f_0/a)$ or less, and the typical value of f_0/a in this study is 40.

APPENDIX B: SELECTION OF A SPATIAL RESOLUTION FOR WELL-RESOLVED SIMULATIONS

For the rate-independent form of Drucker–Prager plasticity that we employ, shear localization in numerical solutions is known to be limited by a grid spacing used in the simulations (Templeton & Rice 2008). Since the degree of shear localization may affect the amount of slip deficit in our models, it is important to make sure that the solution of our interest (i.e. slip distribution after a seismic event) is independent of the employed grid spacing. We simulate the same scenario as shown in Fig. 4(e) with $c = 0$ using a twice-smaller node spacing and compare these results (Fig. B1). The simulated slip distribution and slip velocity for two different resolutions are nearly identical, indicating that the results are converged to reasonable accuracy. We use the average node spacing of $\Delta x = 31.2$ m for all the cases we consider in this study.

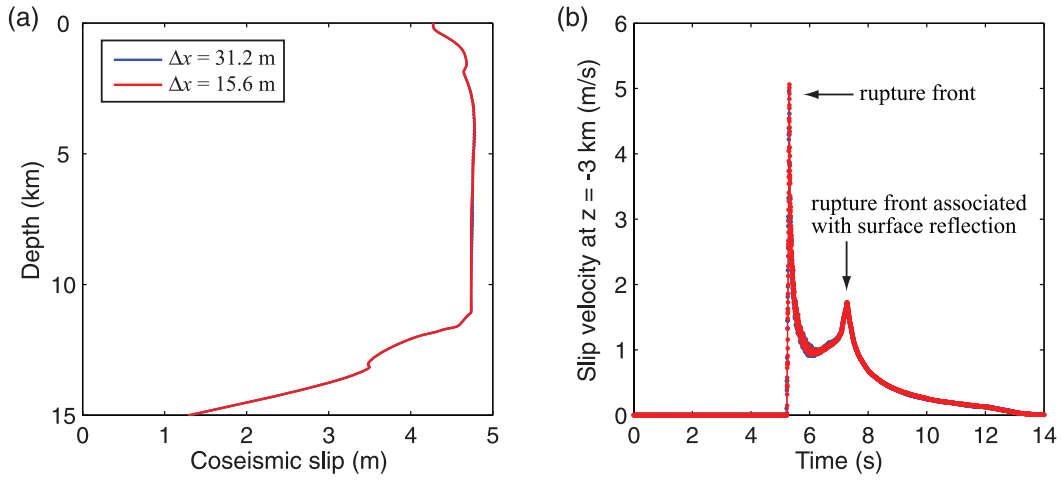

Figure B1. (a) Coseismic slip distributions for the case with $\mu = 0.85$ and $c = 0$ MPa and the same case with a twice-higher resolution. Indicated values of Δx are the average node spacing used in each case. (b) Slip velocity at $z = -3$ km as a function of time for the cases in panel a. The resolution $\Delta x = 31.6$ m gives essentially the same results as $\Delta x = 15.6$ m.