Mantle flow distribution beneath the California margin

Sylvain Barbot

Although the surface deformation of tectonic plate boundaries is well determined by geological and geodetic measurements, the pattern of flow below the lithosphere remains poorly constrained. We use the crustal velocity field of the Plate Boundary Observatory to illuminate the distribution of horizontal flow beneath the California margin. At lower-crustal and upper-mantle depths, the boundary between the Pacific and North American plates is off-centered from the San Andreas fault, concentrated in a region that encompasses the trace of nearby active faults. A major step is associated with return flow below the Eastern California Shear Zone, leading to the extrusion of the Mojave block and a re-distribution of fault activity since the Pleistocene. Major earthquakes in California have occurred above the regions of current plastic strain accumulation. Deformation is mechanically coupled from the crust to the asthenosphere, with mantle flow overlaid by a kinematically consistent network of faults in the brittle crust.
Plate tectonics provide a paradigm to understand the shape and dynamics of Earth’s surface. Most plate boundaries involve either divergence, like at oceanic spreading centers and continental rifts, or convergence, such as subduction and collision zones, but a small fraction involves transform boundaries that facilitate plate kinematics on the global sphere. These boundaries accommodate plate-parallel relative plate displacement by strike-slip motion on vertical or steeply dipping faults. The San Andreas fault system in California consists of major strike-slip faults and subsidiary thrusts faults that transfer extension at the East Pacific Rise in the Gulf of California spreading center to the south and the relative motion between the Gorda and Pacific plate along the Mendocino fracture zone to the north. Contemporary surface deformation at the California margin is captured by a dense continuous geodetic observatory1–3 and the long-term slip accumulation on numerous faults is well established4, providing a unique opportunity to study the mechanics of strain accumulation at a continental transform. Seismological observations, in particular shear wave splitting6,7 and seismic anisotropy8, provide unique constraints on the orientation of mantle flow, but a cohesive picture of flow distribution below the margin consistent with surface observations is still missing.

The relative motion between plates is accommodated at intracontinental transforms by various mechanisms of deformation. In the brittle layer, fault slip is the dominant mode of strain accumulation and release, mediated by the complex evolution of frictional resistance that leads to stick-slip and the earthquake and slow-slip phenomena. The bottom of the seismogenic layer is thought to represent an isotherm for the frictional stability of quartz-rich fault gouge9. In California, the maximum depth of seismicity varies from 10 to 20 km depending on location10, inversely correlated with the surface heat flow. Faults are steadily creeping below a locking depth that varies depending on the fault segment12–15 and often matches the bottom depth of seismicity. Variations of locking depth across the fault system may be attributed to differences in fluid content, geothermal gradients, and stability of fault slip in the seismogenic zone, and may be non-stationary17. Immediately below the seismogenic layer, deformation occurs in localized shear zones that creep steadily down to the brittle-ductile transition, below which plastic strain is accumulated in the ductile substrate on an on-going basis.

We seek to image the deep deformation that accrues along the California margin compatible with our understanding of lithospheric rheology. We present a method to image state-wide horizontal mantle flow from geodetic data with minimal assumptions and tradeoffs. At long wavelengths, the tectonic contribution of the velocity field at Earth’s surface during the interseismic period is caused by fault creep in the brittle layer and distributed flow in the viscoelastic substrate. In a viscoelastic medium, any parcel that deforms plastically entrains the surrounding rocks by elastic coupling18–20. We exploit this effect to reconstruct the subsurface strain-rate distribution by inversion of the surface velocity field (see Methods section). We find that the San Andreas fault system represents the extension of a 50-km-wide trans-lithospheric shear zone in the brittle crust. Major earthquakes have occurred in regions with rapid plastic strain accumulation and the orientation of the principal strain axes is compatible with the earthquake source mechanisms. The numerous parallel dextral faults in Southern California are associated with a return flow below the Mojave section of the Eastern California Shear Zone.

**Results**

**Modeling assumptions.** We obtain the velocity field from the Plate Boundary Observatory through the UNAVCO (unavco.org) repository (Fig. 1). The velocity field is representative of the deformation that accrues for at least 2.5 years up to 15 September 2018, and may include a contribution from the continuing postseismic relaxation of the 1992 Landers, 1999 Hector Mine, and 2010 El Mayor-Cucapah earthquakes. We assume that the deformation is mainly caused by distributed plastic flow in the lower crust and upper mantle. We devise a mesh of 20 × 20 km semi-infinite volume elements that extend down from 20 km depth along a coordinate system aligned with the plate boundary axis, with $x_1$ and $x_2$ the parallel and perpendicular axes, respectively. We assume that the viscous flow is horizontal and incompressible. This implies that the vertical plastic strain-rate tensor components are zero and that $\dot{\epsilon}_{11} + \dot{\epsilon}_{22} = 0$. The model ignores vertical shear, which accommodates the absolute motion of the plate boundary in a plume-fixed reference system, as this displacement is undistinguishable from a rigid-body translation at the regional scale. The maximum shear strain-rate is presumably aligned with the plate boundary axis, but small rotations are allowed by considering both $\dot{\epsilon}_{12}$ and $\dot{\epsilon}_{11}$ free parameters, strictly enforcing $\dot{\epsilon}_{22} = -\dot{\epsilon}_{11}$ and $\dot{\epsilon}_{12} \leq 0$ to ensure overall right-lateral shear parallel to the plate boundary and left-lateral shear perpendicular to the plate boundary, and incompressibility. The mesh extends from Baja California, Mexico to north of the Mendocino triple junction, encompassing 74 × 33 = 2442 volume elements.
To enforce kinematic continuity from the locking depth in the crust to the asthenosphere, we allow strike slip at the deep extension of major faults in the brittle layer, from 15 to 20 km. These embedded surface elements include the entire length of the Garlock, San Jacinto, Elsinore faults, a large section of the San Andreas fault from Gilroy in the Bay Area down to the Salton trough, and the southern section of the Calaveras fault. In addition, we allow shallow fault slip in the creeping section of the San Andreas fault, north of Parkfield. Variation of slip velocity on these deep segments reflect changes in locking depth along a fault and the possibility of a gradual transition from locked to creeping, and do not always represent the long-term fault slip-rate. Large deformation occurs around the Long Valley caldera due to volcanic unrest\cite{51}. Instead of removing the surrounding GPS stations, we model this deformation explicitly with shallow volume elements representing shear and volume change associated with a network of dyke, magma chambers, and connected faults.

**Inversion method and resolution.** Imaging of localized and distributed deformation can be cast as an inverse problem (see Methods section) where plastic strain-rate in volume elements and slip velocity on surface elements are related to surface GPS velocity vectors by Green's functions\textsuperscript{52–56}. The formulation enforces that all the strain tensor components derive from a single-valued velocity field with conservation of angular and linear momentum in a half space\textsuperscript{49,50}. We also simultaneously fit for a rotation and a translation such that the results are independent of the reference frame of the velocity field. The inversion is under-determined with three components for each of the 937 GPS stations, summing up to 2811 data points and 5578 model parameters, among which 588 are dedicated to resolving deformation below Long Valley caldera. Stable inversion results can be obtained with first-order Tikhonov regularization with or without positivity constraints, but all results shown derive from a non-linear inversion that imposes a negative shear strain-rate component $\epsilon_{12}$. 

---

**Fig. 1 Distribution of the GPS velocity around the San Andreas fault.** The GPS velocity field is available at ftp://data-out.unavco.org/pub/products/velocity/. The red triangles correspond to quaternary volcanos. The trace of major faults is from the USGS catalog of quaternary faults https://www.usgs.gov/natural-hazards/earthquake-hazards/faults. The NUVEL1A velocity is for the Pacific plate relative to the North American plate\textsuperscript{76}. The dashed contour indicates the approximate extent of the Eastern California Shear Zone.
The imaging power provided by the Plate Boundary Observatory is shown in a checkerboard test in Fig. 2, indicating that regions of strain accumulation in the lower crust and upper mantle roughly 50 km wide can be resolved appropriately. Loss of resolution in the Sierra Nevada centered on the Long Valley caldera is due to the sparsity of the Plate Boundary Observatory geodetic network in the region and the challenging tradeoffs associated with resolving volcanic unrest and deep mantle flow. Inspection of the resolution matrix reveals a spatial pattern similar to the checkerboard test. As the data offer limited resolution at depth, we do not image how the flow changes across the Moho or how the shear zone may broaden in the asthenosphere. The model therefore represents an average strain accumulation, perhaps most representative of the asthenosphere due to its greater depths and dimensions leaving a broad footprint in the velocity field.

**Flow distribution.** The inversion of the Plate Boundary Observatory velocity field reveals the pattern of horizontal flow in the lower crust and upper mantle (Fig. 3). Plastic strain accumulation during the interseismic period is heterogeneous, with large strain-rates of the order of $1 \times 10^{-14}/s$ localized below major faults, but considerably smaller far from the plate-boundary faults. Deep shear zones are concentrated underneath major faults, spreading across a 50-km-wide shear zone at the minimum, implying substantial variations of ambient stress or rheological properties across the margin. Strain may be even more localized than currently resolved, but filtering by the elastic crust eventually limits the recovery of finer details. We surmise that the shear zone is narrower in the lower crust, broadening to a wider region in the asthenosphere, following theoretical predictions. The deep regions of strain concentration may reflect strain softening due to grain-size evolution, the development of plastic anisotropy, or thermo-mechanical feedbacks with the overlaying brittle section. Mantle flow is more localized throughout northern and central California, down to the Garlock fault, and in the Salton trough, from the San Bernardino mountains to the East Pacific Rise, with a 120-km-wide restraining step-over separating these two regions. The intervening space corresponds to a broad, spatially coherent shear zone, widely distributed from west to east from the Channel Islands offshore to the Eastern California Shear Zone in the...
Mojave block, extending as far north as the White Wolf fault and the Little Lake fault zone, across the Garlock fault. Across the central section of the San Andreas fault, the deformation resembles Couette flow on a horizontal plane.

The northern termination of the California margin exhibits a strong horizontal return flow around the Gorda slab, which may overprint an unresolved vertical corner flow due to plunging of the slab below the North American plate. Another return flow in southern California is more widely distributed due to the large extent of the restraining bend in the plate-parallel and plate-perpendicular directions. Flow is more uniform away from the San Andreas fault system. Smaller patches of strain accumulation can be found along Owens Valley, Panamint Valley, and Death Valley, all the way to the Walker Lane in Nevada, but low resolution around Long Valley caldera affects the imaging result in this region. A patchy, low-amplitude, background plastic strain-rate persists throughout the California margin from the Pacific plate to the Basin and Range, presumably reflecting the ambient tectonic stress or small-scale mantle convection driven by density anomalies that are not isostatically compensated.

Throughout California, the plate boundary at depth is not exactly aligned with the surface trace of the San Andreas fault, except in the Central Valley, as previously established from the distribution of surficial geodetic strain-rate. To the north, the shear zone is firmly east of the San Andreas fault, spreading across the Maacama and Bartlett Springs fault systems in Mendocino County and centered on the Hayward, Calaveras, and Greenville fault system underneath the Bay Area. In the Salton trough, a zone of high strain concentration spreads across the Coachella segment of the San Andreas fault and the San Jacinto fault, contiguous to a broad secondary zone of more distributed strain accumulation.

**Residual velocity.** The modeling assumptions are compatible with observations at the California margin, with a sum of squares of residuals of $(0.3 \text{ mm/yr})^2$ on average for each station and an
overall variance reduction of 95.5%. The vertical component of velocity is typically smaller and has less influence on the result (Supplementary Fig. 1).

The residuals with the Plate Boundary Observatory velocity field in the horizontal direction are shown in Fig. 4. Systematic and spatially coherent residuals in the horizontal components can be found near the Mendocino triple junction, near the Parkfield segment of the San Andreas fault, in the Transverse Range, and around the near-field of the recent 1999 Hector Mine and 2010 El Mayor-Cucapah earthquakes, shown in more detail in Fig. 5. The red triangles correspond to quaternary volcanos.

In northern California, the tectonic regime changes from strike-slip faulting on the San Andreas fault system to subduction of the Gorda plate underneath north America at the Cascadia trench. The strain associated with the Gorda down-going slab creates trench-perpendicular residuals around the Klamath mountains in Humboldt County. Other residuals in the Northern Coast Range may be due to hydrological activity. Misfits near Parkfield correspond to small scale variations of creep between the Cholame and creeping segments modulated by postseismic relaxation of the 2004 Parkfield earthquake. The shortening centered on the Oak Ridge fault system and the San Gabriel fault in the Los Angeles Basin is due to slip partitioning around the big bend of the San Andreas fault involving shallow thrusts splaying off a deeper sub-horizontal décollement.

Deformation in the Transverse Range therefore represents an example of thin tectonics, i.e., faults confined to the crust without mantle roots, instead of deformation driven by mantle downwelling. Other large residuals in the Los Angeles Basin are due to extraction and injection of groundwater and oil. Finally, the forward model leaves out residuals around recent large earthquakes in Southern California due to afterslip and other postseismic processes that were still ongoing in 2018. The postseismic deformation of large strike-slip earthquakes includes vertical shear and short-wavelength features that are not captured in the model. These are most evident in the near-field of the 1999

---

**Fig. 4 GPS residuals velocity.** Residuals are plotted at the same scale as in Fig. 2. The largest spatially coherent residuals occur off the Cascadia trench, around Parkfield, in the Transverse Range, and around the recent 1999 Hector Mine and 2010 El Mayor-Cucapah earthquakes, shown in more detail in Fig. 5. The red triangles correspond to quaternary volcanos.
Hector Mine earthquake and in the Salton trough, north of the El Mayor-Cucapah earthquake rupture, where slip on many shallow faults in the Imperial Valley was accelerated after the mainshock. The pattern of strain accumulation provides information about seismo-tectonic processes (Figs. 3 and 6), including the background stress, as the principal directions of strain-rate are identical to those of the driving deviatoric stress in the ductile region, and the direction of mantle flow at asthenosphere depths. In Mendocino county, the San Andreas fault is poorly oriented with the principal strain direction, contrarily to other faults inland. The strain principal directions are compatible with right-lateral strike-slip faulting on the San Andreas fault from Marin County, north of the Bay Area, to the triple junction at the Garlock fault and, further south, along the Coachella segment near the Salton Sea. Along the big bend of the San Andreas fault, the plastic strain-rate is smaller and less favorably oriented for strike-slip faulting, potentially compensated for by a weaker frictional resistance. The plastic strain-rate below the Garlock fault is optimally oriented for left-lateral strike-slip faulting from the junction with the San Andreas fault to Ridgecrest and Searles Valley, after which a restraining bend of the Garlock fault makes it less favorably oriented, compatible with the slip-rate of the Garlock decreasing to the east. In the Mojave desert, plastic strain-rate is optimally oriented for right-lateral slip for all of the Helendale, Lenwood, Emerson, Camp Rock, Blackwater, Calico, Pisgaw, and Ludlow faults that together give rise to bookshelf faulting.
Fig. 6 Principal strains, faults, seismicity, and seismic anisotropy. The top panel shows the principal strain direction of plastic strain-rate in relation to the orientation of strike-slip and thrust faults, and in relation with the compression axis of seismicity (Yang and Hauksson 2013). The bottom panel shows the direction of maximum shear of the plastic strain-rate and flow streamlines in relation to the orientation of seismic anisotropy (Lin et al. 2014). Red lines indicate strike-slip faults. Black lines with chevrons indicate major thrust faults. Red triangles indicate quaternary volcanos. Grey lines indicate quaternary faults. The dashed contour indicates the approximative extent of the Eastern California Shear Zone.
The background stress is also compatible with the source mechanisms of recent large Southern California earthquakes, including left-lateral and right-lateral slip for the 2019 Ridgecrest Mw 6.4 foreshock and Mw 7.1 mainshock, respectively, left-lateral and right-lateral slip for the 1992 Big Bear and Landers earthquakes, respectively, and right-lateral slip for the 1999 Hector Mine earthquake. The compressive axis of crustal seismicity is perpendicular to major thrust faults in the Transverse Range and oblique to strike-slip faults in the Mojave block, but generally aligned with the compressive axis of plastic strain-rate. An even better alignment is found with the surface geologic strain, indicating a small rotation of the stress tensor towards the surface, presumably due to the short-range influence of faulting in the brittle layer. These results indicate that deep deformation dictates the large-scale stress patterns in the brittle layer that can be relaxed by different types of faults of various orientations, each producing earthquakes with a specific focal mechanism.

Discussion
Geodetic constraints on lower-crustal and mantle strain accumulation illuminate how shallow brittle crust and the deeper plastic layer accommodate regional deformation consistently with significant interaction. Plastic strain accumulation in northern California spreads from west to east across the San Andreas, Maacama, and Bartlett Springs fault systems in the Coastal Range and across the San Andreas, Hayward, and Calaveras fault systems in the Bay Area. In the central segment, the San Andreas fault sits above a similarly wide shear zone that stretches from the Oceanic and Rinconada fault zones to the west to the Great Valley thrust system to the east. In southern California, the deep shear zone is distributed over a wide region that spans across the Newport-Inglewood-Rose Canyon, Elsinore, and San Jacinto fault systems and the Coachella section of the San Andreas fault, possibly extending farther offshore. These results are compatible with the view that a fault network coalesces to a broad zone of distributed flow in the ductile layer, with a separation distance for faults commensurate with the depth of the brittle layer. Theoretical considerations indicate that the thickness of parallel fault strands depends on the viscosity contrast between the lower and upper crust. However, the partitioning of fault slip to accommodate the underlying mantle flow can be accomplished by a set of parallel strike-slip faults, like in the northern and southern ends of the San Andreas fault system, or by a combination of strike-slip and en échelon folds and reverse faults, like in the central section. These results contrast with the simplistic view that major faults exist in isolation and extend from the surface to great sub-crustal depths with a similar degree of localization.

The broad deformation zone associated with a major restraining bend at asthenosphere depths impacts California tectonics from the Transverse Range to the Eastern California Shear Zone. A wide return flow surrounds the northern termination of the Salton trough shear zone, extending to north of the Garlock fault (Figs. 3 and 6b), where the flow direction rotates from SE-NW on the Pacific side, to W-E below the Mojave desert, and to NW-SE on the North American side. The flow pattern may shed light on the orientation of shear wave splitting in southern California. The development of seismic anisotropy in the upper mantle is thought to represent lattice preferred orientation of olivine minerals along the direction of maximum shear, which can be estimated assuming a stationary flow by integrating the plastic strain-rate along the streamlines. Considering a stable configuration of the San Andreas fault system in the last 4–6 Myr with particle velocity of 25 mm/yr and an average strain-rate of $5 \times 10^{-12}$/s corresponds to 90–120 km of particle motion along the current streamlines building up about 95% plastic strain, enough to eliminate any previous texture. Computing synthetic seismic anisotropy is outside the scope of this work, but if the plastic strain-rate is sufficiently uniform upstream of the streamline, the direction of maximum shear of the current plastic strain-rate may represent a good proxy for the fast axis of shear wave splitting. From the California Borderland and Peninsular Range to the Mojave desert, where these conditions are met, there is good agreement with seismic observations. However, shear wave splitting measurements have low depth resolution and may be sensitive to deformation below the asthenosphere, like a counter flow driven by the sinking of the Farallon slab.

The broad return flow below the most major step-over of the San Andreas fault also drives the eastward extrusion of the Mojave block. The advection of the eastern blocks of the San Andreas-Garlock unstable triple junction explains the formation of a new left-lateral shear zone along the White Wolf fault zone, which hosted the Mw = 7.3 Kern County earthquake in 1952, and may eventually supplant the role of the Garlock fault for the regional block kinematics (Fig. 7). The extrusion of the Mojave block perpendicular to plate motion at a rate of $7 \text{mm/yr}$ is responsible for the peculiar long-term evolution of fault activity in the Mojave desert with increasing cumulative slip from west to east, as faults move past a stable sub-crustal shear zone. The easternmost fault, the Ludlow fault, which simultaneously holds the largest cumulative offset and the smallest Holocene slip-rate, is no longer aligned with the Salton trough shear zone and is currently inactive. In contrast, the Calico and other faults to the west are currently still aligned with the Salton trough shear zone and have rapid slip-rates.

Even though a secondary zone of strain accumulation seems to connect the Eastern California Shear Zone and the Walker Lane with an average strain-rate of the order of $1 \times 10^{-12}$, the major plate-boundary shear zones sit beneath the Salton trough and the Central section of the San Andreas fault, separated by a 120-km-wide restraining bend. The alignment of Mojave block faults with the deep Salton trough shear zone appears to be controlling their slip history.

The rate of seismicity on mature faults is controlled, among others, by the rate of stress accumulation, the size of the unstable-weakening region, and the style of faulting, all of which challenging to unravel. Plastic strain accumulation at lower-crustal and upper-mantle depths contributes to loading the seismogenic layer by elastic coupling. In addition, kinematic boundary conditions require the long-term relative displacement of blocks bounding the shear zone to be the same in the brittle and ductile layers. We investigate this relationship by comparing the long-term fault slip-rates on vertical or sub-vertical strike-slip faults collected in Appendix B of the Uniform California Earthquake Rupture Forecast with the offsets of the underlyng volume elements in the fault-perpendicular direction. There is a broad agreement between geological slip rates and integrated plastic strain-rate (Fig. 8), even though mantle shear zones are diffuse, with no strict boundaries. Simple predictions based on underlying plastic strain-rate fall within the uncertainties of field observations for the Hayward, Calaveras, north San Andreas, south San Andreas, Newport-Inglewood, Elsinore, and San Jacinto faults. Large deviations occur for faults of the Eastern California Shear Zone, presumably due to off-fault deformation in the Mojave block and the variable rates of the Garlock fault in the late Pleistocene, early Holocene.

The overall agreement between shallow (geological) and deep deformation confirms the kinematic compatibility between the brittle and ductile layers, implying a strong mechanical coupling between upper-crust deformation and lower-crust and asthenosphere flow.

Assuming that plastic strain corresponds dominantly to simple shear, which holds true for most of California (Fig. 3), integration of
the plastic strain in the plate-perpendicular direction shows a relative velocity between the Pacific and the North American plates in agreement with global reconstructions of relative plate motion, i.e., 49.7 ± 2.7 mm/yr (Fig. 9). Local deviations within the uncertainties occur along the plate boundary along the extent of the 1906 Mw = 7.9 San Francisco earthquake rupture, the Mojave block, and the Salton trough, which may be attributed to transient acceleration of mantle flow following large earthquakes, and in the Eastern California Shear Zone, where deformation is not in simple shear, involving a significant component of plate-perpendicular motion.

Plastic strain accumulation below the California margin is compatible with the relative motion of the North American and Pacific plates at geological time scales, with the long-term slip-rate of major faults of the San Andreas fault system, and with the contemporaneous surface deformation during the interseismic period. In addition, the trace of major faults, background seismicity, and major earthquakes are situated above regions of high plastic strain-rate (Fig. 9). These results indicate that deformation at the plate boundary is mechanically coupled from the crust down to the upper mantle, despite the possible weak strengths of the lower-crust and the asthenosphere. Mantle flow and crustal faulting form a cohesive geometric assembly where deep shear zones are overlaid by a kinematically consistent network of faults in the brittle layer. Kinematic compatibility is accomplished by slip partitioning in the brittle layer and strain partitioning in the ductile substrate, with geographic overlap.

As models of crustal dynamics can now represent the coupling between brittle and ductile processes, imaging of deep deformation opens the door to more accurate models of crustal dynamics that assimilate a realistic distribution of mantle flow. Integration of brittle and ductile deformation processes into self-consistent mechanical representations will be key to better understand the geodynamics of the California margin.

Methods

Green’s functions. Any distribution of plastic strain in a viscoelastic medium causes deformation in its surrounding by elastic coupling. The details of the rheology in the ductile region is unspecified, but we assume that the flow is incompressible, driven by the deviatoric component of the stress tensor. With these assumptions, the plastic strain tensor \( \epsilon \) can be associated the moment density \( m = \)
Fig. 8 Anelastic strain accumulation and long-term slip-rate on California faults. a Spatial distribution of distributed strain and the location of paleo slip-rate on strike-slip fault segments cataloged in Appendix B of the Uniform California Earthquake Rupture Forecast (UCERF3; https://pubs.usgs.gov/of/2013/1165/). The squares indicate locations where the deep strain-rate immediately below the fault segment matches the long-term fault slip rate. Red triangles indicate quaternary volcanos. b Comparison of the UCERF3 slip rates with the integrated strain-rate below the corresponding fault segment.

Fig. 9 Strain partitioning across the San Andreas fault system. The top panel shows the plate parallel velocity from the integration of the shear component of the strain-rate tensor in the plate-perpendicular direction. The dashed lines indicate the average velocity and (1−σ) standard deviation of the Pacific plate relative to the North American plate from 18 different models, including REVEL-2000, NUVEL 1A, ITRF 2008, and MORVEL 2010 (refs. 76–79). The bottom panel shows the distribution of deep shear strain-rate in California from the Mendocino triple junction to the Salton trough, highlighting the strain partitioning primarily between the San Andreas fault and the Eastern California Shear Zone. The red stars and black circles show the hypocenters of historical earthquakes and earthquakes from moment magnitude 3 and above from 1990 to 2020 around California, respectively. The seismic catalog is obtained from the Northern California Earthquake Data Center (https://ncedc.org) and the Southern California Earthquake Data Center (https://scedc.caltech.edu). The gray dashed profiles correspond to the 5-km depth contours of the Cascadia megathrust. Red triangles indicate quaternary volcanos.
\( q \), where \( \mu \) is the local shear modulus, and the equivalent body-force \( f = -\nabla \cdot \mathbf{m} \). \( \nabla \cdot \mathbf{m} \) (1)

The resulting total displacement satisfies conservation of linear and angular momentum and can be obtained at any GPS point \( x \) of the domain by integration with the Green's functions \( G(x,y) \), where \( y \) is the location of the forcing term, to automatically conform with the boundary conditions, following

\[
\mathbf{u}(x) = \int_{\Omega} G(x,y) \cdot \mathbf{f}(y) \, dy,
\]

where \( \Omega \) is the region that deforms plastically, potentially the whole domain. To image the distribution of plastic strain accumulation in California compatible with an inverse problem, we discretize the domain into semi-infinite cuboids that extrude downwards from a burial depth of \( z_0 = 20 \) km. We approximate the total deformation field by superposition of the deformation caused by individual, non-overlapping volume elements of uniform plastic strain. Considering one volume element, the displacement field is given by

\[
u(x_1, x_2, x_3) = \int_{\Omega} G_e(x_1, x_2, x_3; y_1, y_2, y_3) m_{22} n_3 \, dy_1 dy_2 dy_3, \]

where \( \Omega \) represents the boundary of the volume element, and repeated indices are summed, following Einstein's convention. We use the Green's function for an elastic half-space and the coordinate system is aligned with the orientation of the semi-infinite cuboid, depth positive down. We consider cuboids of length \( L \) and thickness \( T \) in the \( x_1 \) and \( x_2 \) directions, respectively, with the origin centered in the middle of a top edge. Expanding the integration over the four vertical surfaces and the top horizontal one, we can write

\[
u(x_1, x_2, x_3) = \int_{T/2}^{T/2} \int_{-L/2}^{L/2} \int_{-L/2}^{L/2} G_1(x_1, x_2, x_3; y_1, y_2, y_3) \, dy_1 dy_2 dy_3
\]

\[+ \int_{L}^{L} \int_{-L/2}^{L/2} G_2(x_1, x_2, x_3; y_1, y_2, y_3) \, dy_1 dy_2 dy_3
\]

\[- \int_{-L}^{L} \int_{-L/2}^{L/2} G_3(x_1, x_2, x_3; y_1, y_2, y_3) \, dy_1 dy_2 dy_3.\]

We develop solutions for the displacement field using a numerical integration involving Gauss-Legendre quadrature along the \( x_1 \) and \( x_2 \) axes and double-exponential quadrature along the depth axis. For integration over depth, we use a change of variable that maps the semi-infinite interval into finite bounds. The displacement for semi-infinite cuboids of arbitrary strike angle from north at any coordinates in the reference system is obtained by translation and rotation of the displacement field. Surface displacement are associated with plastic strain; surface velocity with plastic strain-rate. Numerical codes that evaluate the displacement field are provided (see Code availability). The displacement fields for a buried volume element with non-zero plastic strain tensor components \( e_{11}, e_{22}, \) and \( e_{33} \) are shown in Supplementary Fig. 2. Each volume element produces similar displacement, accordingly shifted. Imaging the plastic strain accumulation can be thought of as a deconvolution of these displacement patterns. The diffusion of the surface velocity compared to the more localized extent of the plastic deformation represents the filtering effect of the elastic crust, which eventually limits the resolution of the inverse problem.

Geodetic inversion method. We design the inversion of the surface velocity field by constructing the displacement kernels for each volume element of the mesh, complemented by other deforming surface and volume elements described in the main text. The matrices \( L_{11}, L_{22}, \) and \( L_{33} \) represent the three-component surface velocity vectors at the coordinates of the GPS stations due to the plastic strain tensor components \( e_{11}, e_{22}, \) and \( e_{33} \), respectively, organized in a vector with one element per volume element in the mesh. For any distribution of horizontal plastic strain, the corresponding velocity field is given by the matrix-vector product

\[
L \cdot (e_{11}, e_{22}, e_{33}),
\]

where the Green's function matrix is defined as

\[
L = \begin{pmatrix}
L_{11} & 0 & L_{12} \\
L_{21} & L_{22} & 0 \\
L_{22} & L_{23} & L_{33}
\end{pmatrix},
\]

enforcing an equal and opposite contribution of east-west and north-south uniaxial strain accumulation. The second-order deviatoric tensor \( (e_{11} \otimes e_{22} - e_{12} \otimes e_{21}) \) is simply a 45° rotation of the tensor \( (e_{11} \otimes e_{22} - e_{12} \otimes e_{21}) / \sqrt{2} \), so inverting for these two orientations simultaneously provides a basis to capture horizontal pure shear in any orientation. The velocity field contains a rigid-body motion representing rotation about an unknown pole. However, rotation about any arbitrary pole can be decomposed into a rotation about the origin of the coordinate system and a residual translation, both of which can be represented by a linear operator that can be inverted. We represent the rotation about the origin by the spin vector \( \mathbf{w} = -w_0 \mathbf{E}_3 \) and the remaining translation by the vector \( \mathbf{u} \). The rigid-body component of the surface velocity is then given by the matrix-vector product \( \mathbf{R} \cdot (w_0, u_1, u_2)^T \), with

\[
\mathbf{R} = \begin{pmatrix}
-x_0 & 1 & 0 \\
x_0 & 0 & 1
\end{pmatrix}.
\]

Fault creep between the seismogenic zone and the brittle-ductile transition of major California faults is captured with 55 embedded surface elements, 5-km-wide and of varying lengths, with strike-slip motion. Near-surface creep along a central segment of the San Andreas fault is captured with 48 shallow embedded surface elements, 20-km-long and 5-km-wide, with strike-slip motion. Deep and shallow creep are related to the surface displacement by the matrices \( F \) and \( C \), respectively. Additional volume elements with free plastic strain components \( e_{11}, e_{22}, \) and \( e_{33} \) organized in a \( 7 \times 7 \times 3 \) element mesh extending \( 70 \times 70 \times 15 \) km around Long Valley caldera are used to model volcanic unrest and \( M \) is the matrix used to compute the associated displacements. The forward model is then given by\( \hat{d} = Gm, \) where \( \hat{d} \) is the vector of observations and the design matrix is given by \( G = (L \cdot C \cdot M, R) \). The inversion is stabilized with first-order and second-order Tikhonov regularization. A non-linear inversion procedure is used to enforce right-lateral shear in the plate-perpendicular direction and left-lateral shear in the plate-parallel direction. The residuals with the PBO crustal velocity field is shown in Fig. 4 and Supplementary Fig. 1.

Data availability

All data used in this study are publicly available. A reference to each dataset can be found in the figure captions and in the main text. The source data underlying Figs. 1, 3, 4, 5, 8a, and 9 and Supplementary Fig. 2a are provided as a Source Data file. Source data are provided with this paper.

Code availability

The Matlab routines to compute Green's functions for plastic strain in a semi-infinite cuboid in an elastic half-space are available at bitbucket.org/sharbat/natocomm20. Source data are provided with this paper.

Received: 16 March 2020; Accepted: 13 August 2020;
Published online: 08 September 2020

References

1. Zeng, Y. & Shen, Z.-K. A fault-based model for crustal deformation, fault slip rates, and off-fault strain rate in California. Bull. Seismol. Soc. Am. 106, 766–784 (2016).
2. Klein, E. et al. Transient deformation in California from two decades of GPS displacements: Implications for a three-dimensional kinematic reference frame. J. Geophys. Res. 124, 12189–12223 (2019).
3. Field, E. H. et al. A spatiotemporal clustering model for the third Uniform California Earthquake Rupture Forecast (UCERF3-ETAS): toward an operational earthquake forecast. Bull. Seismol. Soc. Am. 107, 1049–1081 (2017).
4. Silver, P. & Holt, W. The mantle flow field beneath western North America. Science 295, 1054–1057 (2002).
5. Zandt, G. & Humphreys, E. Toroidal mantle flow through the western US slab window. Geology 36, 295–298 (2008).
6. Moschetti, M., Ritzwoller, M., Lin, F. & Yang, Y. Seismic evidence for widespread western-US deep-crustal deformation caused by extension. Nature 464, 885–889 (2010).
7. Barbot, S. Modulation of fault strength during the seismic cycle by grain-size evolution around contact junctions. Tectonophysics 765, 129–145 (2019).
8. Barbot, S. Slow-slip, slow earthquakes, period-two cycles, full and partial ruptures, and deterministic chaos in a single asperity fault. Tectonophysics 768, 228171 (2019).
9. Blanpied, M. L., Lockner, D. A. & Byerlee, J. D. Frictional slip of granite at hydrothermal conditions. J. Geophys. Res. 100, 13045–13064 (1995).
10. Nazareth, J. J. & Hauksson, E. The seismogenic thickness of the southern California crust. Bull. Seismol. Soc. Am. 94, 940–960 (2004).
11. Hauksson, E. Crustal geophysics and seismology in southern California. Geophys. J. Int. 186, 82–98 (2011).
12. Sauber, J., Thatcher, W., Solomon, S. C. & Lisowski, M. Geodetic slip rate for the eastern California shear zone and the recurrence time of Mojave Desert earthquakes. Nature 367, 264–266 (1994).
13. Smith, B. E. Sandwell, D. Coulomb stress accumulation along the San Andreas Fault system. J. Geophys. Res. 108, 10 (2003).
14. Platt, J. P. & Becker, T. W. Where is the real transform boundary in California? Geochim. Cosmochim. Acta 73, 1665–1677 (2009).
30. Pollitz, F. F. Transient rheology of the uppermost mantle beneath the Mojave. Nature Geosci. 10, 891–905 (1998).

29. Rollins, C., Barbot, S., & Avouac, J.-P. Postseismic deformation following the 1992 landers earthquake. J. Geophys. Res. 106, 9221–9240 (2001).

28. Bruhat, L., Barbot, S., & Avouac, J.-P. Evidence for postseismic deformation of the San Andreas fault zone during the 2004 rupture across the Salton Trough. Pure Appl. Geophys. 16, 1963–1985 (2013).

27. Takeuchi, C. S., & Fialko, Y. Dynamic models of interseismic deformation and postseismic relaxation mechanisms. Pure Appl. Geophys. 171, 355–381 (2014).

26. Freed, A. M., & Bürgmann, R. Evidence of power-law strain accumulation on faults beneath Los Angeles, California. Geophys. Res. Lett. 39, B09401 (2012).

25. McCaffrey, R. Block kinematics of the Pacific–North America plate boundary in the southwestern United States from inversion of GPS, seismological, and geologic data. J. Geophys. Res. 114, B11 (2009).

24. Zhang, X., & Sagiya, T. Intraplate strike-slip faulting, stress accumulation, and shear localization of a crust-upper mantle system with nonlinear viscoelastic material. J. Geophys. Res. 113, B10319 (2008).

23. Takeuchi, C. S., & Fialko, Y. Evidence of power-law flow beneath a strong crust. Geophys. Res. Lett. 32, L20306 (2005).

22. Freed, A. M., & Bürgmann, R. Evidence of power-law flow in the Mojave desert mantle. Nature 430, 548–551 (2004).

21. Takeuchi, C. S., & Fialko, Y. Dynamic models of interseismic deformation and postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

20. Barbot, S., Moore, J. D., & Lambert, V. Displacement and stress associated with distributed anelastic deformation in a half-space. Bull. Seismol. Soc. Am. 97, 821–855 (2007).

19. Hirth, G. & Kohlstedt, D. L. In Rheology of the upper mantle and the mantle wedge: a view from the experimentalists. (ed. J. E. E.). Vol. 138. pp. 83–105 (Russell Publisher Ltd. Washington, D. C., 2003).

18. England, P., Housen, G., & Sonder, L. Length scales for continental deformation in convergent, divergent, and strike-slip environments: analytical and approximate solutions for a thin viscous sheet model. J. Geophys. Res. 90, 3551–3557 (1985).

17. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

16. Barbot, S., Moore, J. D., & Lambert, V. Displacement and stress associated with distributed anelastic deformation in a half-space. Bull. Seismol. Soc. Am. 107, 821–855 (2017).

15. Barbot, S. Deformation of a half-space from anelastic strain confined in a tetrahedral volume. Bull. Seismol. Soc. Am. 108, 2687 (2018).

14. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

13. Barbot, S., Moore, J. D., & Lambert, V. Displacement and stress associated with distributed anelastic deformation in a half-space. Bull. Seismol. Soc. Am. 107, 821–855 (2017).

12. Barbot, S. & Fialko, Y. Fourier-domain Green’s function for an elastic semi-infinite solid under gravity, with applications to earthquake and volcano inflation. Geophys. J. Int. 182, 568–582 (2010).

11. Barbot, S., Fialko, Y. A unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

10. Barbot, S., Moore, J. D., & Lambert, V. Displacement and stress associated with distributed anelastic deformation in a half-space. Bull. Seismol. Soc. Am. 107, 821–855 (2017).

9. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

8. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

7. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

6. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

5. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

4. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

3. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

2. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).

1. Takeuchi, C. S., & Fialko, Y. Unified continuum representation of postseismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. Geophys. J. Int. 182, 1124–1140 (2010).
73. Dixon, T. H. & Xie, S. A kinematic model for the evolution of the Eastern California Shear Zone and Garlock Fault, Mojave Desert, California. Earth Planet. Sci. Lett. 494, 60–68 (2018).
74. Lambert, V. & Barbot, S. Contribution of viscoelastic flow in earthquake cycles within the lithosphere-asthenosphere system. Geophys. Res. Lett. 43, 142–154 (2016).
75. Barbot, S. Asthenosphere flow modulated by megathrust earthquake cycles. Geophys. Res. Lett. 45, 6018–6031 (2018).
76. DeMets, C. & Dixon, T. H. New kinematic models for Pacific-North America motion from 3 Ma to present. Evidence for steady motion and biases in the NUVEL-1A model. Geophys. Res. Lett. 26, 1921–1924 (1999).
77. Sella, G. F., Dixon, T. H. & Mao, A. REVEL: a model for recent plate velocities from space geodesy. J. Geophys. Res. 107, ETG-11 (2002).
78. DeMets, C., Gordon, R. G. & Argus, D. F. Geologically current plate motions. Geophys. J. Int. 181, 1–80 (2010).
79. Altamimi, Z., Métivier, L. & Collilieux, X. ITRF2008 plate motion model. J. Geophys. Res. 117, B7 (2012).

Acknowledgements
The study benefited from funding from the National Science Foundation, under award number EAR-1848192. We thank UNAVCO for providing the GPS velocity field for the Plate Boundary Observatory. We thank the US Geological Survey and the Southern California Earthquake Center for providing the quaternary faults and southern California active faults databases, respectively. We thank the Northern California Earthquake Data Center and the Southern California Earthquake Data Center for the seismicity data.

Author contributions
The author conducted all the analysis and wrote the manuscript.

Competing interests
The author declares no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-18260-8.

Correspondence and requests for materials should be addressed to S.B.

Peer review information Nature Communications thanks Takeshi Sagiya and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020