LETTERS

Shallow fault-zone dilatancy recovery after the 2003 Bam earthquake in Iran

Eric J. Fielding[1], Paul R. Lundgren[1], Roland Bürgmann[2] & Gareth J. Funning[3]

Earthquakes radiate from slip on discrete faults, but also commonly involve distributed deformation within a broader fault zone, especially near the surface. Variations in rock strain during an earthquake are caused by heterogeneity in the elastic stress before the earthquake, by variable material properties and geometry of the fault zones, and by dynamic processes during the rupture[1,2]. Stress changes due to the earthquake slip, both dynamic and static, have long been thought to cause dilatancy in the fault zone that recovers after the earthquake[3–5]. Decreases in the velocity of seismic waves passing through the fault zone due to coseismic dilatancy have been observed[6] followed by postseismic seismic velocity increases during healing[5,7,8]. Dilatancy and its recovery have not previously been observed geodetically. Here we use interferometric analysis of synthetic aperture radar images to measure postseismic surface deformation after the 2003 Bam, Iran, earthquake and show reversal of coseismic dilatancy in the shallow fault zone that causes subsidence of the surface. This compaction of the fault zone is directly above the patch of greatest coseismic slip at depth. The dilatancy and compaction probably reflects distributed shear and damage to the material during the earthquake that heals afterwards. Coseismic and postseismic deformation spread through a fault zone volume may resolve the paradox of shallow slip deficits for some strike-slip fault ruptures[9].

Several seismic studies have found that the ratio of the velocities of S waves to P waves decreases with time after an earthquake, in addition to an overall velocity increase, reflecting a postseismic increase in the Poisson's ratio of fault zone material; this increase in Poisson's ratio is interpreted as an increase in the fraction of pore space filled with liquid[6,7] as the fault zone contracts. These postseismic changes reverse the coseismic velocity changes due to dilatancy in the fault zone. Seismic waves are relatively insensitive to the depth of the anomalies, but the damaged fault zone may be confined to the top 3–4 km of the crust and possibly to very shallow (~200 m) depths[8].

Fault zone dilatancy is caused by pervasive damage where small cracks open up within the deformed rocks, increasing the porosity and volume[3–5]. It has been difficult to measure the predicted coseismic dilatancy and postseismic volume decrease in earthquake fault zones geodetically, because dilatancy occurs within a very narrow damage zone along the earthquake rupture and many other postseismic processes also cause deformation of the Earth's surface.

Geodetic measurements of far-reaching postseismic deformation for many large earthquakes have been modelled as a combination of viscous relaxation in the lower crust and upper mantle and with aseismic afterslip on a deep continuation of the fault that ruptured[10]. Postseismic deformation in the upper crust has been explained by pore-pressure transients[11,12] and by afterslip both immediately surrounding coseismic slip patches and up-dip from the coseismic rupture[13–15], which can overwhelm the deformation from dilatancy recovery. We exploit interferometric synthetic aperture radar (InSAR) to measure negative volume change in the fault zone after a large earthquake in Iran.

Here we analyse 3.5 years of postseismic deformation after the M_w (moment magnitude) 6.6 earthquake that devastated the city of Bam, Iran, on 26 December 2003, using synthetic aperture radar from the European Space Agency Envisat satellite. The earthquake primarily ruptured a buried strike-slip fault that extends under the city of Bam and beneath a broad alluvial plain to the south[16–21]. This fault is part of a broad system of north–south oriented, right-lateral shear in eastern Iran[19]. Because the first synthetic aperture radar image was acquired 12 days after the earthquake, we cannot separate the earliest phase of postseismic deformation from the coseismic deformation. The stable surface around Bam provides optimal InSAR coherence[18] for mapping deformation during and after the earthquake. The surface rupture from coherence and field mapping[16,18] propagated from the south towards Bam[19]. Inversion of the surface deformation measured by InSAR[9,17], cross-correlation of optical imagery, and levelling data[20,21] require slip >2 m at depths of 4–5 km on the main fault, with less near the surface and little slip north of Bam. Maximum offsets across the surface ruptures were 0.25 m (refs 16, 18, 19), but total shear across a zone about 1 km wide is >1 m (refs 17, 21). A less pronounced shallow slip deficit has also been found for some other large strike-slip earthquakes[9]. We concentrate on surface displacement features with spatial scales less than 10 km that can resolve deformation in roughly the upper 8 km of the crust, because atmospheric effects mask InSAR signals at larger scales[22].

Two features stand out in the postseismic deformation, with different spatial scales and temporal behaviour. We examine the hypotheses that these are caused by (1) afterslip, (2) localized poroelastic rebound at a restraining bend of the coseismic rupture, and (3) localized compaction of material that dilated during the earthquake. The first major feature in the postseismic pattern is two lobes of line-of-sight (LOS) change near the south end of the main coseismic rupture with positive LOS motion (towards the satellite locations) to the east of the rupture (A1 and A2 on Fig. 1 and Supplementary Fig. 5) and alternating zones of positive and negative LOS motion that have much lower amplitude, but similar spatial extent, about 10 km further north. The descending track (no. 120) is dominated by strong positive LOS motion (towards the satellite locations) to the southeast of the rupture (A1), whereas the ascending tracks (nos 156 and 385) have a smaller area of positive amplitude to the east combined with negative LOS motion to the west (A2), requiring a substantial eastward component of displacement. The combination of these features is consistent with afterslip above and to the south of the main coseismic asperity that ruptured during the 2003 earthquake. The time series shows that the afterslip rate decayed with a time constant of ~335 days if an exponential time function is assumed (Fig. 2a).

[1]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA. [2]Department of Earth and Planetary Science, University of California, Berkeley, California 94720, USA. [3]Department of Earth Sciences, University of California, Riverside, California 92521, USA.

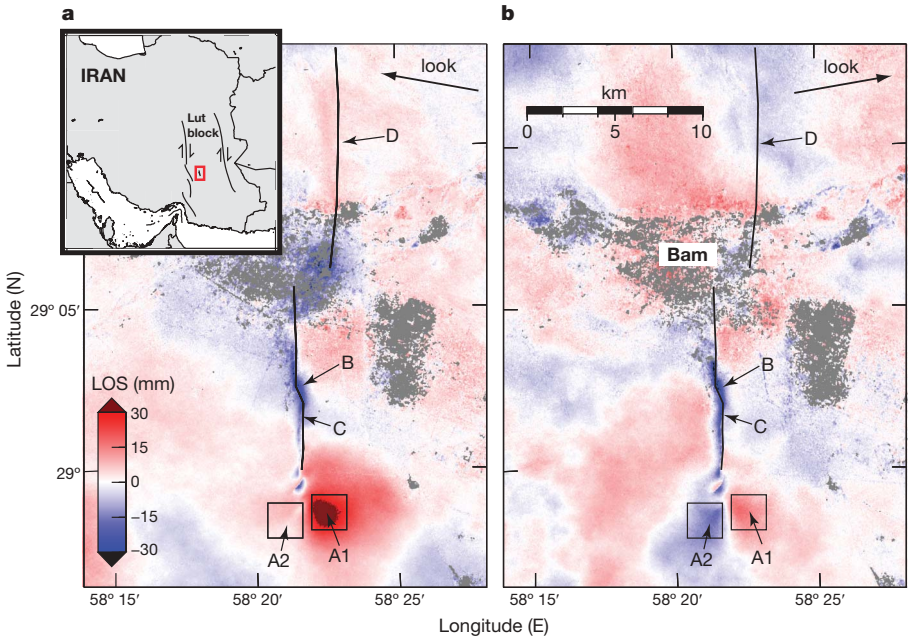
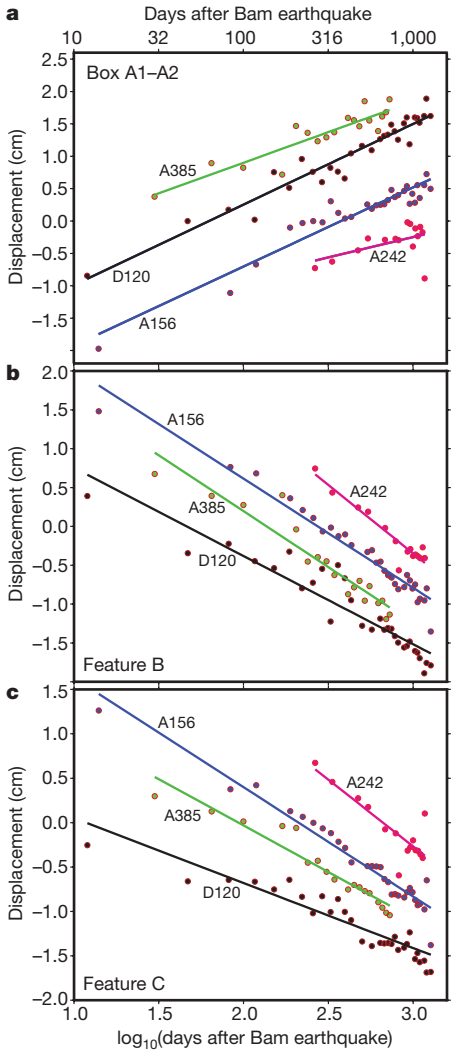


Figure 1 | Postseismic surface deformation after the Bam, Iran, earthquake. **a**, **b**, Maps from two Envisat tracks in radar lines-of-sight (LOS); fits to each time series have been converted to estimates of total displacement between 12 and 1,097 days after the 26 December 2003 earthquake. Black lines, locations of coseismic fault ruptures mapped from InSAR[18]. Difference of boxes A1 and A2, used for Fig. 2a, measures afterslip at 2–3 km depth. Subsidence at B is poroelastic and at C is recovery of dilatancy. D is minor uplift on northern fault. **a**, Descending track no. 120. **b**, Ascending track no. 156. Horizontal components of look vectors shown with arrows. Inset, location of Bam in southeast Iran.



Second, the spatial pattern (see Fig. 1) includes a narrow zone of >20 mm subsidence combined with horizontal motion towards the coseismic surface ruptures observed south of the city of Bam[16,18]. Subsidence is strongest near a left bend of the right-lateral fault (feature B in Fig. 1) and continues along a roughly 500-m-wide zone north and south of the bend (feature C). This zone continued to subside for at least 3.5 years after the earthquake (Fig. 2b, c). We advocate that this subsidence has both poroelastic (feature B) and dilatancy recovery (feature C) components.

Readjustment of pore fluid pressure (poroelastic rebound) in the crust can cause volume changes following earthquakes, resulting in postseismic subsidence over areas of coseismic pressure increase and vice versa. The widest and strongest subsidence south of Bam is where the left bend caused compression (feature B on Figs 1 and 3) and coseismic uplift of at least 60 mm (Supplementary Fig. 6). A large part of the postseismic subsidence there is probably due to a poroelastic response reflecting partial deflation of the coseismic uplift due to fluid flow away from the compressed fault bend. Early relaxation in the first 12 days after the earthquake is included in the coseismic interferogram and cannot be measured. The Bam fault-bend subsidence has a nearly log-linear rate out to the end of the 3.5 years we analysed; the relaxation time for an exponential decay fit is 1.7 years.

The postseismic subsidence along the rupture zone south of Bam, however, extends >2 km to the north and >4 km south of the left bend in the fault (C in Fig. 3); this requires another process in addition to poroelastic rebound in the restraining bend, as the rest of the zone experienced no measurable coseismic compression. To the south of the bend, the surface ruptures mapped in the field[16,18]

Figure 2 | Postseismic deformation of features as function of time. **a**, Time series of difference between range change of boxes A1 and A2 (see Fig. 1) plotted versus the logarithm of the time in days since the earthquake, for four Envisat tracks. Lines show least squares fits with the log(t) function; line colour is keyed to data. Prefixes A and D denote respectively ascending and descending tracks. **b**, Subsidence in fault bend due to poroelastic rebound from difference between polygon B and boxes to east and west (see Fig. 3). **c**, Subsidence south of fault bend (difference between polygon C and two adjacent boxes, see Fig. 3) interpreted as recovery of coseismic dilatancy in the shallow fault zone.

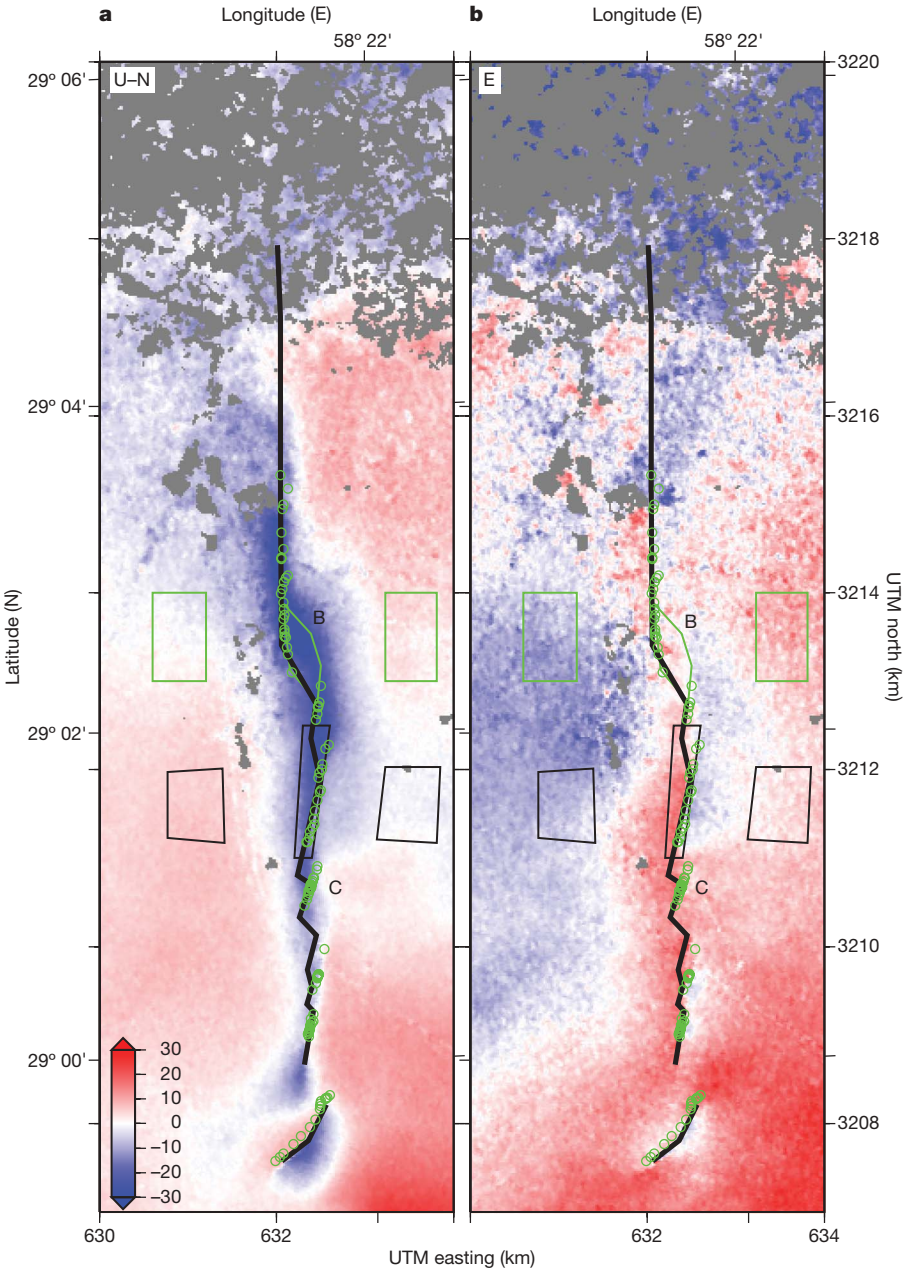


Figure 3 | Transformed components of total displacement from tracks 120 and 156. a, **b**, Components in mm (colour scale) for the area of the main coseismic rupture south of Bam, showing the narrow zone of subsidence along the fault: **a**, 'up' minus 0.1×'north' (U − N); **b**, 'east' (E). Black lines, locations of coseismic rupture from InSAR; green circles, locations of coseismic rupture mapped in the field[18]. Boxes B and C used for time series plots of Fig. 2b and c, respectively. On left, labelled and unlabelled tickmarks show respectively latitude and UTM (universal transverse Mercator projection, zone 40) north; and vice versa for tickmarks on right.

included a small opening component, indicating east–west coseismic expansion in the surface layer, which is also reflected in the coseismic InSAR[17,18]. The postseismic InSAR shows that there is a net motion of the surface on both sides of the fault towards the fault south of the bend (see Fig. 3b), which suggests that there was coseismic dilation or damage in the fault zone that subsequently relaxed. The material that compacted after the earthquake apparently did not have enough fluid flowing inward to fill the dilated pore space created during the earthquake to counteract the compaction and subsidence. There is no indication of dilatancy recovery along the rupture segment north of Bam (feature D in Fig. 1), where a small amount of postseismic uplift indicates different material properties or stress conditions of the shallow fault zone.

To explore the processes and kinematics of the postseismic transients that cause the surface deformation features, we utilize elastic half-space dislocation models. We model the afterslip using the two planar faults of the coseismic model: a near-vertical right-lateral strike-slip fault and an oblique-slip (fixed rake 151°) thrust dipping 64° to the west[17]. The preferred afterslip distribution on these faults has 100–136 mm of afterslip above and to the south of the large coseismic slip area on the main fault, and also has afterslip on the secondary fault just below the afterslip on the main fault (Supplementary Fig. 7). The peak afterslip is at about 3 km depth in the south and shallower than 2 km depth to the north. We explored alternative fault geometries, but few aftershocks were located in the area south of the main coseismic rupture to provide constraints on the fault geometry[23]. None of the afterslip models explain the narrow zone of subsidence along the main fault.

Broad-scale surface rebound due to pore-pressure relaxation after the coseismic stress change can be approximated by taking the difference

between surface deformation fields of undrained and drained elastic models[24]. The predicted deformation is opposite in sign from coseismic displacements and from postseismic deformation (see Fig. 1 and Supplementary Fig. 9). We cannot rule out a small amount (<10 mm) of broad-scale poroelastic rebound in the first two months after the Bam earthquake because longer-wavelength atmospheric effects are particularly strong in the two Envisat scenes acquired in early January 2004, obscuring deformation at scales greater than about 5 km.

We model the compaction of the fault zone with a grid of regularly spaced contracting Mogi point sources in an elastic half-space. This simple approximation allows efficient computations, but does not directly model the physical process that must involve a more continuous contracting volume. We performed an inversion with both the contracting Mogi source array and afterslip on the main rupture to determine an optimum joint distribution of deformation sources. The afterslip in the joint model is similar to the two-fault afterslip model, and the contraction sources are concentrated in a band that is deeper and stronger in the northern part (see Fig. 4c). The vertical extent of the contraction band depends on the amount of smoothing applied in the inversion, but the depth is constrained by the width of the observed subsidence. This model fits the measured subsidence and horizontal contraction (Figs 3 and 4). Our data only resolve contraction in the upper kilometre of the fault zone (with a total volume decrease of 8.4×10^4 m^3); however, 2×10^5 m^3 of dilatancy recovery could have occurred on the remaining coseismic slip zone below 1 km depth without producing significant surface deformation. Because the magnitude of coseismic dilatancy and its recovery is likely to be reduced at greater depths and confining pressures[25], dilatancy processes may have occurred at greater depths but are not resolved by geodetic data.

The restricted spatial extent and the three years of increasing subsidence in the fault zone south of Bam is very different from the broad-scale elevation changes of 1–2 months duration interpreted as pore-pressure transients after a similar earthquake in Iceland[11]. The Bam poroelastic response has a much longer time constant (Fig. 2b) than the relaxation time of 0.75 years estimated for Landers[24] and about 2 months for the Iceland earthquake[11]. The >1.7 year relaxation time for the Bam fault-bend subsidence (feature B) implies that the poroelastic rock volume has lower permeability.

The zone of strong postseismic contraction in the fault zone (Fig. 3, feature C) is directly above the largest coseismic slip[9,17], where the coseismic surface ruptures were widest and had the largest offsets in InSAR and field mapping[18]. This is also where coseismic optical and synthetic aperture radar pixel offsets show ~1 m of right-lateral shear over about 1 km centred on the fault[17,21]. It is impossible to determine from the surface displacements alone what part of the shear is elastic versus inelastic deformation, but theoretical models of deformation of porous media predict that shear at shallow depths is likely to involve damage with a dilatancy component[25]. The postseismic contraction occurred in the area where coseismic fault zone shear and hence damage were strongest. The depth of the contraction probably reflects the distribution of coseismic dilatancy; this is controlled by material strength variations and by distribution of water in the shallow crust, both modulated by stress.

We conclude that the localized postseismic subsidence over the Bam fault zone is only partly due to poroelastic rebound in a prominent restraining bend in the rupture: this subsidence is dominated by healing and compaction of coseismic dilatation and damage in the upper ~1 km of the fault zone outside of the bend[25]. The postseismic compaction at

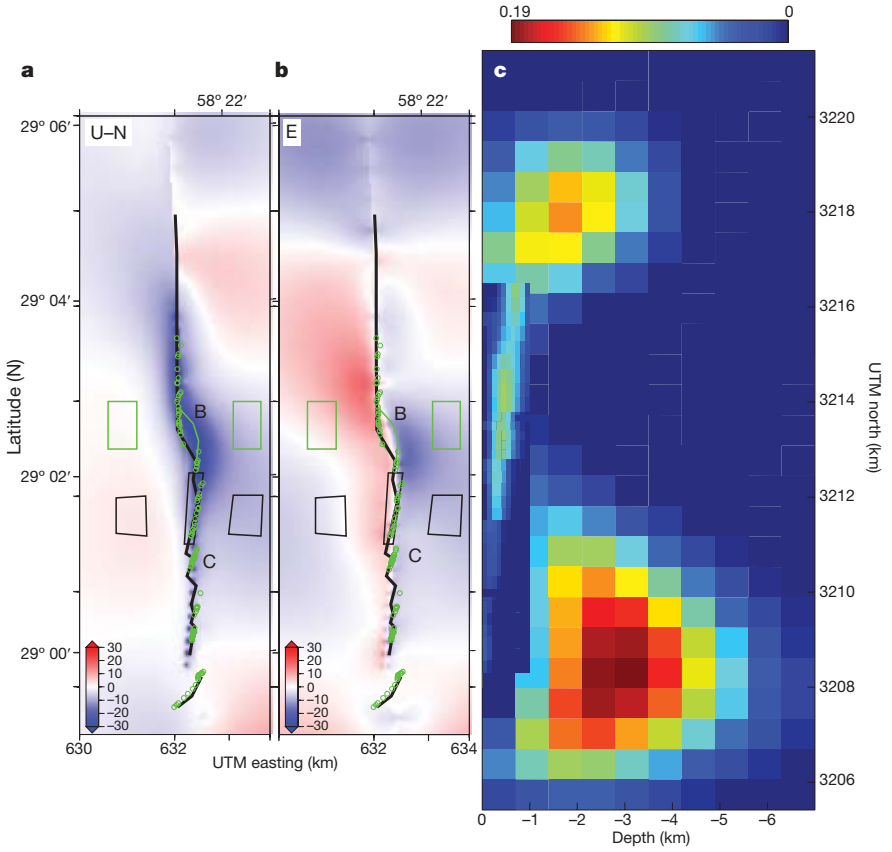


Figure 4 | Shallow fault zone contraction and main fault afterslip combined model results. Maps of predicted deformation in mm (colour scale) for the same area as Fig. 3; **a**, 'up' minus 0.1×'north' (U–N), and **b**, 'east' (E). **c**, Model source strength shown in north–south profile aligned with maps. Depth in km on horizontal axis. Shallow fault zone volume decrease (in 10^4 m^3) is shown as colour of small rectangles; main fault afterslip magnitude (in m) is shown as colour of large squares (same colour scale). Some afterslip patches are hidden behind the shallow fault zone.

Bam is approximately logarithmic in time, consistent with measurements of changes in the velocities of seismic waves in fault zones[5–8].

METHODS SUMMARY

We formed 342 full-resolution interferograms from 94 Envisat ASAR (advanced synthetic aperture radar) scenes on four different satellite tracks acquired between 7 January 2004 and 22 June 2007 (Supplementary Tables 1–4 and Supplementary Figs 1–4) using the JPL/Caltech ROI_pac[26] and Stanford SNAPHU[27] software. We calculated time series of line-of-sight (LOS) deformation for each track using a modified small baseline subset algorithm[28] (see Supplementary Information for full Methods description). The LOS vectors for the four tracks are nearly coplanar, so the north component of ground motion is poorly constrained (Supplementary Table 5). The temporal resolution is determined by the 35-day repeat cycle of the Envisat orbit and the spatial resolution is approximately 20×20 m on the ground.

We mitigate atmospheric errors by computing the total LOS-change maps from a number of independent measurements. Long-wavelength errors due to atmospheric effects and imprecise orbit knowledge were reduced by fitting and removing a planar phase gradient from a 40×40 km subset of each date's solution (Fig. 1, Supplementary Fig. 5). This also removes any long-wavelength deformation signal, so we cannot resolve postseismic deformation due to viscoelastic relaxation or afterslip in the lower crust or mantle.

For efficient model parameter optimization, we selected about 1,000 samples from estimated total deformation for the target area (Fig. 1, Supplementary Fig. 5) from each of the four tracks using a source-dependent sampling method[29] with the main coseismic rupture planes[17] as the sources. Green's functions relating dislocations (shear or tensile) on fault patches to deformation at the surface were calculated with the Poly3D program[30]. We optimized the distribution of dislocation magnitudes on fixed-geometry surfaces with a non-negative least squares procedure including a smoothing constraint. In the joint source modelling, we used a shallow fault extending to 2 km depth along the main coseismic rupture to optimize the sampling. Details are in Supplementary Information.

Received 6 July 2008; accepted 9 January 2009.

1. Mai, P. M. & Beroza, G. C. A spatial random field model to characterize complexity in earthquake slip. *J. Geophys. Res.* **107,** 2308, doi:10.1029/2001JB000588 (2002).
2. Ben-Zion, Y. & Sammis, C. G. Characterization of fault zones. *Pure Appl. Geophys.* **160,** 677–715 (2003).
3. Nur, A. Effects of stress on velocity anisotropy in rocks with cracks. *J. Geophys. Res.* **76,** 2022–2034 (1971).
4. Scholz, C. H. Post-earthquake dilatancy recovery. *Geology* **2,** 551–554 (1974).
5. Vidale, J. E. & Li, Y.-G. Damage to the shallow Landers fault from the nearby Hector Mine earthquake. *Nature* **421,** 524–526 (2003).
6. Li, Y.-G., Chen, P., Cochran, E. S., Vidale, J. E. & Burdette, T. Seismic evidence for rock damage and healing on the San Andreas Fault associated with the 2004 M 6.0 Parkfield earthquake. *Bull. Seismol. Soc. Am.* **96,** S349–S363 (2006).
7. Schaff, D. P. & Beroza, G. C. Coseismic and postseismic velocity changes measured by repeating earthquakes. *J. Geophys. Res.* **109,** B10302, doi:10.1029/2004JB003011 (2004).
8. Peng, Z. G. & Ben-Zion, Y. Temporal changes of shallow seismic velocity around the Karadere-Düzce branch of the North Anatolian Fault and strong ground motion. *Pure Appl. Geophys.* **163,** 567–600 (2006).
9. Fialko, Y., Sandwell, D., Simons, M. & Rosen, P. A. Three-dimensional deformation caused by the Bam, Iran, earthquake and the origin of shallow slip deficit. *Nature* **435,** 295–299 (2005).
10. Bürgmann, R. & Dresen, G. Rheology of the lower crust and upper mantle: Evidence from rock mechanics, geodesy and field observations. *Annu. Rev. Earth Planet. Sci.* **36,** 531–567 (2008).
11. Jónsson, S., Segall, P., Pedersen, R. & Bjornsson, G. Post-earthquake ground movements correlated to pore-pressure transients. *Nature* **424,** 179–183 (2003).
12. Peltzer, G., Rosen, P., Rogez, F. & Hudnut, K. Postseismic rebound in fault step-overs caused by pore fluid-flow. *Science* **273,** 1202–1204 (1996).
13. Johanson, I. A., Fielding, E. J., Rolandone, F. & Burgmann, R. Coseismic and postseismic slip of the 2004 Parkfield earthquake from geodetic data. *Bull. Seismol. Soc. Am.* **96,** S269–S282 (2006).
14. Langbein, J., Murray, J. R. & Snyder, H. A. Coseismic and initial postseismic deformation from the 2004 Parkfield, California, earthquake, observed by global positioning system, electronic distance meter, creepmeters, and borehole strainmeters. *Bull. Seismol. Soc. Am.* **96,** S304–S320 (2006).
15. Hsu, Y. J. *et al.* Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science* **312,** 1921–1926 (2006).
16. Talebian, M. *et al.* The 2003 Bam (Iran) earthquake: Rupture of a blind fault. *Geophys. Res. Lett.* **31,** L11611, doi:10.1029/2004GL020058 (2004).
17. Funning, G. J., Parsons, B. E., Wright, T. J., Jackson, J. A. & Fielding, E. J. Surface displacements and source parameters of the 2003 Bam (Iran) earthquake from Envisat advanced synthetic aperture radar imagery. *J. Geophys. Res.* **110,** B09406, doi:10.1029/2004JB003338 (2005).
18. Fielding, E. J. *et al.* Surface ruptures and building damage of the 2003 Bam, Iran, earthquake mapped by satellite synthetic aperture radar interferometric correlation. *J. Geophys. Res.* **110,** B03302, doi:10.1029/2004JB003299 (2005).
19. Jackson, J. A. *et al.* Seismotectonic, rupture-process, and earthquake-hazard aspects of the 26 December 2003 Bam, Iran, earthquake. *Geophys. J. Int.* **166,** 1270–1292 (2006).
20. Motagh, M. *et al.* Combination of precise leveling and InSAR data to constrain source parameters of the Mw= 6.5, 26 December 2003 Bam earthquake. *Pure Appl. Geophys.* **163,** 1–18 (2006).
21. Peyret, M. *et al.* The source motion of 2003 Bam (Iran) earthquake constrained by satellite and ground-based geodetic data. *Geophys. J. Int.* **169,** 849–865 (2007).
22. Bürgmann, R., Rosen, P. A. & Fielding, E. J. Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation. *Annu. Rev. Earth Planet. Sci.* **28,** 169–209 (2000).
23. Sadeghi, H., Aghda, S. M. F., Suzuki, S. & Nakamura, T. 3-D velocity structure of the 2003 Bam earthquake area (SE Iran): Existence of a low-Poisson's ratio layer and its relation to heavy damage. *Tectonophysics* **417,** 269–283 (2006).
24. Peltzer, G., Rosen, P., Rogez, F. & Hudnut, K. Poroelastic rebound along the Landers 1992 earthquake surface rupture. *J. Geophys. Res.* **103,** 30131–30145 (1998).
25. Hamiel, Y., Lyakhovsky, V. & Agnon, A. Coupled evolution of damage and porosity in poroelastic media: Theory and applications to deformation of porous rocks. *Geophys. J. Int.* **156,** 701–713 (2004).
26. Rosen, P. A., Hensley, S., Peltzer, G. & Simons, M. Updated repeat orbit interferometry package released. *Eos* **85,** 47 (2004).
27. Chen, C. W. & Zebker, H. A. Network approaches to two-dimensional phase unwrapping: Intractability and two new algorithms. *J. Opt. Soc. Am. A* **17,** 401–414 (2000).
28. Berardino, P., Fornaro, G., Lanari, R. & Sansosti, E. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans. Geosci. Rem. Sens.* **40,** 2375–2383 (2002).
29. Lohman, R. B. & Simons, M. Some thoughts on the use of InSAR data to constrain models of surface deformation: Noise structure and data downsampling. *Geochem. Geophys. Geosyst.* **6,** Q01007, doi:10.1029/2004GC000841 (2005).
30. Thomas, A. L. *Poly3D: A Three-Dimensional, Polygonal Element, Displacement Discontinuity Boundary Element Computer Program with Applications to Fractures, Faults, and Cavities in the Earth's Crust*. M.Sc. Thesis, Stanford Univ. (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Envisat data are copyright 2004–07 European Space Agency and were provided under AOE project 668. We thank R. Lohman for the use of data resampling programs and discussions. We thank J. Jackson, M. Talebian, D. McKenzie, J.-P. Avouac, Z. Liu, Z.H. Li, E. Cochran, F. Horowitz, G. Peltzer, Y. Ben-Zion and E. Hearn for discussions. Part of this research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Author Contributions E.J.F. performed the data analysis, interpretation and modelling. P.R.L. wrote the time series analysis and modelling framework programs, and contributed to interpretation. R.B. and G.J.F. contributed to modelling and interpretation. E.J.F. and R.B. wrote the manuscript with contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.J.F. (Eric.J.Fielding@jpl.nasa.gov).