Seismotectonic, rupture process, and earthquake-hazard aspects of the 2003 December 26 Bam, Iran, earthquake


To cite this version:


HAL Id: insu-00267061
https://hal-insu.archives-ouvertes.fr/insu-00267061
Submitted on 10 Mar 2021
Seismotectonic, rupture process, and earthquake-hazard aspects of the 2003 December 26 Bam, Iran, earthquake

J. Jackson,1 M. Bouchon,2 E. Fielding,1,3 G. Funning,4,5 M. Ghorashi,6 D. Hatzfeld,2 H. Nazari,6 B. Parsons,4 K. Priestley,1 M. Talebian,6 M. Tatar,7 R. Walker4 and T. Wright4

1 COMET, Bullard Laboratories, Madingley Road, Cambridge, CB3 0EZ. E-mail: jackson@esc.cam.ac.uk
2 Laboratoire de Geophysique Interne et Tectonophysique, UJF-CNRS, Grenoble, France
3 Jet Propulsion Laboratory, California Institute of Technology, MS 300-233, 4800 Oak Grove Road, Pasadena, CA 91109, USA
4 COMET, Department of Earth Sciences, Parks Road, Oxford, OX1 3PR
5 Now at: Berkeley Seismological Laboratory, University of California, 215 McCone Hall, Berkeley, CA 94720, USA
6 Geological Survey of Iran, PO Box 13185-1494, Tehran, Iran
7 International Institute of Earthquake Engineering and Seismology, Tehran, Iran

Accepted 2006 April 28. Received 2006 April 28; in original form 2005 November 17

SUMMARY

The catastrophic 2003 $M_w$ 6.6 Bam earthquake in southern Iran attracted much attention, and has been studied with an abundance of observations from synthetic aperture radar, teleseismic seismology, aftershock studies, strong ground motion, geomorphology, remote sensing and surface field work. Many reports have focused on the details of one or other data type, producing interpretations that either conflict with other data or leave questions unanswered. This paper is an attempt to look at all the available data types together, to produce a coherent picture of the coseismic faulting in 2003 and to examine its consequences for active tectonics and continuing seismic hazard in the region. We conclude that more than 80 per cent of the moment release in the main shock occurred on a near-vertical right-lateral strike-slip fault extending from the city of Bam southwards for about 15 km, with slip of up to 2 m mostly restricted to the depth range 2–7 km. Analysis of the strong ground motion record at Bam is consistent with this view, and indicates that the extreme damage in the city can be attributed, at least in part, to the enhancement of ground motion in Bam because of its position at the end of the northward-propagating rupture. Little of the slip in the main shock reached the Earth’s surface and, more importantly, aftershocks reveal that $\sim$12 km vertical extent of a deeper part of the fault system remained unruptured beneath the coseismic rupture plane, at depths of 8–20 km. This may represent a substantial remaining seismic hazard to the reconstructed city of Bam. We believe that some oblique-reverse slip (up to 2 m, and less than 20 per cent of the released seismic moment) occurred at a restricted depth of 5–7 km on a blind west-dipping fault that projects to the surface at the Bam-Baravat escarpment, an asymmetric anticline ridge that is the most prominent geomorphological feature in the area. This fault did not rupture significantly at shallow levels in 2003, and it may also represent a continuing seismic hazard. Widespread distributed surface ruptures north of the city are apparently unrelated to substantial slip at depth, and may be the result of enhanced ground motion related to northward propagation of the rupture. The faulting at Bam may be in the early stages of a spatial separation (‘partitioning’) between the reverse and strike-slip components of an oblique convergence across the zone. Such a separation is common on the continents, though in this case the slip vectors between the two faults differ only by $\sim$20$^\circ$ as a substantial strike-slip component remains on the oblique-reverse fault. The Bam earthquake is one in a series of large earthquakes involving faulting along the western edge of the Lut desert. In addition to the unruptured parts of the faults near Bam itself, continuing and substantial hazard is represented by unruptured neighbouring faults, particularly blind thrusts along the Jebel Barez mountains to the south and strike-slip faulting at Sarvestan to the west.

Key words: active faulting, continental tectonics, earthquakes, InSAR, Iran, seismology.
1 INTRODUCTION

The Bam earthquake of 2003 December 26, in the Kerman province of south-central Iran, was a catastrophe. It effectively destroyed the ancient city of Bam, with a population of around 150 000. The number of deaths will perhaps never be known exactly, but is thought to be between 26 000 (the official figure) and 40 000 (Berberian 2005). Even in the long and terrible earthquake history of Iran, where events of this nature are not rare (the last comparable one, again killing ~40 000 people, occurred in Rudbar 1990; see Berberian et al. 1992), the Bam earthquake was especially destructive.

The earthquake also attracted much scientific attention. It produced a series of enigmatic coseismic surface fractures and cracks, which were mapped and recorded by several groups, but which were small for a shallow event of this size ($M_w 6.6$). It was the first major destructive earthquake for which both pre- and post-seismic Envisat ASAR (Advanced Synthetic Aperture Radar) data were available, and spectacularly coherent radar images were obtained in the virtually vegetation-free desert surrounding the destroyed city itself. These images were sufficiently clear to observe the coseismic surface ruptures themselves, through the decorrelation, or lack of coherence, between pre- and post-seismic images observed along the fractures. The deformation of the surface, observed in radar interferograms, was used to infer the location and distribution of the coseismic faulting at depth. The earthquake was recorded by many stations of the Global Digital Seismograph Network (GDSN) and the seismic waveforms were also used to infer fault and rupture geometry. More than one dense local network of seismographs was installed to obtain aftershock locations and focal mechanisms, and in Bam itself a strong ground motion instrument recorded local accelerations approaching 100 per cent on both horizontal and vertical components in the main shock itself, with a strong directivity effect related to rupture propagation. Most of these different studies have now been published in some form, but there has been no attempt to bring them all together to provide a coherent overview of the coseismic faulting in this event that uses all these sources of information, and considers the extent to which they are compatible. That is one of the goals of this paper, and is worth doing, not just because we owe it to the memory of those who perished to find out what happened, but because that knowledge contains lessons for seismic hazard evaluation in Bam and elsewhere. In particular, we suggest, though it is not certain, that a substantial seismic hazard may remain at Bam, associated with unruptured parts of the active fault system that moved in 2003.

In 2003, the Bam earthquake was also the latest in a sequence of destructive earthquakes, beginning in 1980, that occurred along a system of N–S right-lateral strike-slip faults bordering the west side of the Dasht-e-Lut desert in SE Iran. (Another has occurred subsequently, at Dahuiyeh near Zarand, 280 km NW of Bam; Talebian et al. 2006). A review of what is known of the historical seismicity of the region is given by Berberian (2005). The question of what other active faults are nearby, which may be part of the same system and may themselves be reactivated in the near future, is thus also pressing and relevant. The tectonic context of the earthquake faulting at Bam, and its relation to other active faults that surround it, is therefore of more than academic interest, and is also commented on in this paper.

2 TECTONIC SETTING

The Bam region is on the west side of the Dasht-e-Lut, a flat, low-lying, aseismic, and probably undeforming, desert in SE Iran (Fig. 1). The Lut is bounded on both east and west sides by systems of N–S right-lateral strike-slip faults that together accommodate ~13–16 mm yr$^{-1}$ of N–S right-lateral shear between central Iran and western Afghanistan, which is part of rigid Eurasia (Vernant et al. 2004; Walker & Jackson 2004; Regard et al. 2005). There is insufficient GPS coverage in Iran to assess the relative importance of the two strike-slip systems directly, but limited dating of Quaternary offsets, together with estimates of the total cumulative offsets suggests that the faults bounding the west side of the northern Lut account for a relatively small part of the total N–S shear; in the region of 1–2 mm yr$^{-1}$ (Walker & Jackson 2002, 2004).

The fault system bounding the west side of the Lut begins in the north at ~33°N, continuing south for about 250 km as the Nayband Fault (Fig. 1b, top), which is remarkable for its linearity and the small relief across it, both of which are thought to indicate a nearly pure strike-slip nature. In spite of numerous clear Late Quaternary offsets across the Nayband Fault, it is associated with no known historical or instrumentally recorded earthquakes, though it must be regarded as capable of generating events of $M_w \geq 7.5$ (Berberian & Yeats 1999; Walker & Jackson 2002). Limited dating of Quaternary offsets suggests that the slip rate on the Nayband Fault is in the region of 1–2 mm yr$^{-1}$ (Walker & Jackson 2002, 2004). Between 30.5°N and 29.5°N the fault system changes strike to NNW–SSE (Fig. 1b), following the Gowk valley between Chahar Farsakh and Golbagh and acquiring an overall component of convergence. Destructive earthquakes occurred along this section, known as the Gowk Fault, in 1981 ($M_w 6.6$ and 7.0; Berberian et al. 1984, 2001), 1989 ($M_w 5.8$; Berberian & Qorashi 1994) and 1998 ($M_w 6.6$; Berberian et al. 2001). The coseismic surface ruptures and focal mechanisms of these earthquakes, together with their associated geomorphology, indicate the kinematics of the active faulting in the Gowk Fault zone. The evidence, reviewed by Berberian et al. (2001), Walker & Jackson (2002) and Fielding et al. (2004), particularly from the coseismic InSAR interferograms of the 1998 earthquake, suggests that the oblique right-lateral convergence is spatially separated, or ‘partitioned’, into its orthogonal pure strike-slip and thrusting components, with the strike-slip part being accommodated in the Gowk valley and the thrusting in the Shadad fold-and-thrust belt to the NE, adjacent to the Lut. A small normal component in the Gowk valley itself is consistent with a ramp-and-flat geometry on the master thrust fault at depth.

South of the Gowk Valley, the faulting again renews a N–S trend, continuing through Sarvestan (Fig. 1b), west of Bam, to ~29°N, where it joins a series of thrusts along the northern flank of the Jebel Barez mountains. South of the Jebel Barez, at ~28.75°N, the N–S faulting reappears as the Sabzevaran fault system, running south through Jiroft and eventually stepping right to the N–S Minab fault system on the east side of the Gulf of Oman, which separates the Makran from the Zagros (Berberian & Yeats 1999; Regard et al. 2004, 2005). The slip rate on the southern part of the Sabzevaran–Jiroft fault system is estimated to be $6.2 \pm 0.7$ mm yr$^{-1}$ (Regard et al. 2005), but how much of this continues north into the Sarvestan–Gowk fault system is unknown. No large historical earthquakes are known from the Sarvestan or Sabzevaran faults, or from the faults near Bam discussed below, but, as in the case of the Nayband fault, the historical data are of questionable quality and almost certainly incomplete in those remote regions (Ambraseys & Melville 1982; Berberian & Yeats 1999; Berberian 2005).

As we will show, the system of faults near Bam also involves both N–S strike-slip and a component of thrusting, but occurs 50 km east of the main Gowk–Sarvestan–Sabzevaran faults, which are much clearer, longer and better developed. Walker & Jackson
Figure 1. (a) Yellow circles are earthquakes in the period 1964–2002 from the catalogue of Engdahl et al. (1998) and subsequent updates. Red arrows are velocities (in mm yr$^{-1}$) of points in Iran relative to stable Eurasia, measured by GPS from Vernant et al. 2004). The green circle is the epicentre of the 2003 Bam earthquake. (b) A mosaic of LANDSAT-7 TM images (bands 531-RGB) of the region around Bam with known active faults in black. The dotted black line running SSE from Bam and Baravat is the possible SSE extension of the Bam fault zone, discussed in Section 7.3. The red focal mechanism of the 2003 Bam earthquake is located at the teleseismically determined location, and is probably misplaced 15 km to the SW. Focal mechanisms further north are from earlier destructive earthquakes on the Gowk fault system, discussed in the text. The thick white line is the coseismic rupture trace of the 1998 March 4 Fandoqa earthquake.

Figure 2. (a) Shaded topographic image, from 90m SRTM data, of the Bam-Sarvestan valley. (b) LANDSAT-7 image (bands 742-RGB) of the same region. Note the principal drainage systems, merging in the middle of the valley, and then flowing east into the Dasht-e-Lut. The green colour of the Bam and Baravat oases is mostly from date-palm plantations.

(2002) provide evidence for a total strike-slip offset of $\sim$12 km on the Gowk-Sarvestan faults, which is probably much greater than that at Bam. Bam itself lies in a valley between two massifs of predominantly Eocene volcanic rocks; the Kafut mountains to the north and the Jebel Barez to the south (Fig. 2). Drainage into the valley from these two massifs is diverted east to a low playa SE of Bam, or into the Posht river (Fig. 3), which flows through Bam, ultimately to end in the southern Dasht-e-Lut. The surface of the Bam-Sarvestan valley consists of mixed fluvial–alluvial gravels, which can be quite coarse. Bam itself is a desert oasis, famous for the cultivation of date palms, most of which grow in the adjoining region of Baravat to the east (Fig. 3). Running for 10 km SSE from Bam, and separating it from Baravat, is a low ($\sim$30 m high) ridge of generally finer-grained silts, marls and sands of presumed Quaternary age, uplifted through the Recent gravels by a blind fault with a reverse component (discussed later).

There is little direct evidence for the age of initiation of these various faults. Allen et al. (2004) and Walker & Jackson (2004) point out that the main fault systems both sides of the Lut are capable of achieving their observed total offsets in about $5 \pm 2$ Ma at their present-day inferred slip rates (1–2 mm yr$^{-1}$ on the western and 10–15 mm yr$^{-1}$ on the eastern sides).

3 COSEISMIC SURFACE RUPTURES

Coseismic cracks and fissures were observed following the earthquake in several areas north and south of the city of Bam. They
are described briefly here, as their significance and relative importance only became clear after analysis of the InSAR interferograms and seismic waveforms. The radar data are discussed in detail in the next section, but in this section we use one aspect of the radar data, the decorrelation (or correlation decrease), to display the location of the observed ground ruptures on maps (Figs 4a and c). Two radar images, one taken before and one taken after the earthquake, are used to produce interferograms. At the same time another quantity, the radar correlation or coherence between the images, is derived, which provides a measure of changes in the ground surface (e.g. Fielding et al. 2005). High coherence (correlation values close to 1 in Figs 4a and c) indicate that there has been little modification to the ground surface in the time interval between the acquisitions of the two images. Poor coherence (small values of correlation in Figs 4a and c) suggest there have been significant changes to the scattering properties of the ground at these locations.

The main point to emphasize, which was recognized in the immediate aftermath of the earthquake, is that none of the ruptures described below indicate surface faulting sufficient to have generated the earthquake. Through the normal scaling relations (e.g. Scholz 1982), the observed seismic moment of $9 \times 10^{18}$ N m ($M_w 6.6$) should be associated with a fault 10–20 km long, with an average slip of about one metre. It is clear that the bulk of the faulting did not reach the surface.

To aid future trenching or palaeoseismological investigations, the GPS locations of all the coseismic ruptures observed in the field, and marked by red circles in Figs 4(a) and (c), are given in Table A1 in the Appendix.

3.1 North of Bam

The area north of Bam and the Posht river (Figs 3, 4a and b) consists of low-lying volcanic rocks and alluvial fans, the surfaces of which form a hard, pebbly, encrusted desert pavement, cemented by salt and carbonate. For about 2 km north of the river, and distributed over about 2 km E–W, were numerous minor cracks and fissures, mostly discontinuous hair-cracks with openings of 1–2 cm at most, and little direct evidence of the sense of lateral slip. Individual fractures could typically be followed for about 50–100 m with a N–S trend (e.g. at 29° 08.289′N 58° 21.997′E; location A in Fig. 4a). A little further east in the same area, cracks ran for about 400 m with a trend of 300° NW from 29° 08.482′N 58° 22.731′E (location E in Fig. 4a), and were clearly parallel to bedding in steeply dipping red shales exposed in places beneath the encrusted desert surface. These particular NW-trending fractures consisted of ‘pop-ups’ in the cemented pavement, separated by N–S trending, left-stepping en-echelon tension cracks, consistent with oblique right-lateral bedding-plane slip in the underlying shales. In the same general area, polygonal patterns of cracks were observed on the 2–3 m scale, similar to mud desiccation cracks but on larger scale and in the cemented pebbly desert surface. The most prominent set of fractures, continuous for 2 km
with a 170° trend, ran south from 29° 09.242′ N 58° 23.142′ E, and are marked B in Fig. 4(a). At their northern end, they consisted of small pop-up ridges a few cm high, generally in a right-stepping en echelon pattern (Figs 5a and b), consistent with the right-lateral slip. Adjacent to these fractures (location C in Fig. 4b) was a stone that had been ejected about 10 cm from its socket in the cemented gravel (Fig. 5c); a possible indication of the high surface accelerations that were recorded in Bam itself (discussed later).

An obvious N–S fault is visible in the ASTER satellite image, marked D in Fig. 4(a), following an escarpment bounding black volcanic rocks to the west. Along the base of this scarp were small, discontinuous hair-cracks between 29° 09.50′ and 29° 10.53′ N, but no signs of reactivation further north.

Although the fractures described in this section were among the first to be discovered after the earthquake, it is clear from the InSAR data, discussed later, that they were not the site of the principal coseismic slip at depth.

3.2 South of Bam

Following the acquisition of the coseismic interferogram, discussed in Section 4, it became clear that the principal fault slip at depth must
Figure 5. Field photographs of coseismic surface ruptures. (a) North of Bam at 29° 8.93′N 58° 23.16′E; location B in Fig. 4(a). View is north. The rupture trace cuts the cemented desert pavement. (b) Detail looking south at the same location as (a), showing minor pop-up features in the buckled, cemented desert crust. (c) A stone dislodged from the cemented surface and thrown about 10 cm, at location C in Fig. 4a, north of Bam. (d) View south towards the Jebel Barez mountains along the main rupture trace south of Bam, at 29° 01.390′N 58° 21.565′E, on segment G in Fig. 4(c). (e) Detail of the rupture trace at 29° 01.623′N 58° 21.628′E showing the left-stepping tension fractures indicative of right-lateral slip. (f) A vertical offset on segment I of Fig. 4(c), showing the cemented upper few cm of the desert pavement.

have occurred south of Bam (Talebian et al. 2004) in a region that was unvisited in the immediate aftermath of the earthquake (Figs 3, 4c and d). The region is an almost flat, featureless cemented bajada, sloping gently east (Fig. 5d). Guided by the decorrelation images from radar (Fig. 4c; Fielding et al. 2005), three of us (MT, HN, MG) mapped 8 km of ruptures along a N–S trend consisting, in detail, of four left-stepping en echelon segments, with clear evidence of right-lateral slip (Figs 5d and e). The northern segments F–H showed right-lateral horizontal offsets of typically less than 10 cm, reaching a maximum of 20 cm in one place, with no systematic vertical offset. The southern segment (I) had a different strike of 050–060°, a right-lateral offset of 5 cm and a systematic vertical component of ~12 cm down to the SE (Fig. 5f). Localized surface fractures comparable to those in segments F–I could not be traced north of 29° 03.50′N, where cracks became distributed, smaller, and discontinuous, before being lost in the disturbed outskirts of the built-up area.

As the InSAR data (discussed later) show, these ruptures south of Bam are the surface expression of the principal rupture surface in the earthquake, which was mostly blind (i.e. confined below ground). They are remarkable because they occur in a place where there is no apparent indication of earlier movement at the surface, or even the existence of a fault all, in either the topography, the geomorphology,
the soil development or vegetation. With effectively no vertical component of slip, we suspect all traces of earlier movement are simply removed by outwash flowing over the bajada surface in occasional floods, leaving the subdued drainage channels visible in Fig. 5(d). It is difficult to see how the location of an important seismogenic fault could have been discovered here before the earthquake, other than by some chance excavation.

3.3 The Bam-Baravat escarpment

The most obvious fault in the area is the escarpment running for \(\sim 12\) km south from the Posht river, between Bam and Baravat. It is clearly visible in satellite imagery (Fig. 3) and in the field (Fig. 6), and is mentioned in several earlier publications on the area (e.g. Berberian 1976; Walker & Jackson 2002). In form, the escarpment is an asymmetric fold with a steep eastern side and an almost-flat top (Fig. 6b), uplifting and exposing relatively fine-grained alluvium, marls and sands through the encrusted desert pavement.

The relief at the eastern edge of the escarpment is almost constant at \(\sim 30\) m along much of its length, dying away at both ends (Fig. 7). It has all the characteristics of a fold above a blind, west-dipping reverse fault (e.g. Berberian et al. 2000), with incised drainage cutting through the escarpment (Figs 6a and c), or abandoned as dry valleys on the ridge crest (Fig. 6d), or diverted around its southern end (Fig. 8). It is probable, however, that the fault also has a substantial right-lateral strike-slip component (discussed later).

After the earthquake, hairline cracks could be followed discontinuously along most of the length of the escarpment, close to its base. These were mostly in the form of open fissures less than one cm wide, with no consistent or reliable indication of strike-slip motion. Some cracks and fissures were also found in the steeper topography of the east-facing slope of the escarpment, many of which were caused or accentuated by landsliding.

The fold ridge is crossed by numerous W-to-E flowing underground water channels, called qanats (see Section 7.4), marked at the surface by lines of vertical shafts typically 10–20 m apart. These qanats tap an uplifted aquifer beneath the fold, in the hanging wall of the blind oblique-reverse fault, to bring water to the date-growing region of Baravat in the east. Indeed, the fold, and its uplifted aquifer, provide the main source of water for agricultural purposes, and together indicate the reason for the location of the Bam oasis and its agriculture (see Section 7.5). The lines of the qanats are shown in Fig. 8. Many of them follow incised ephemeral or dry stream valleys (as this requires less digging for the vertical shafts). Some of these streams, and hence qanats, are deflected south a few metres west of the frontal escarpment. This was noticed by Hessami et al. (2004), who interpreted this to indicate right-lateral offsets of originally straight qanats. On a larger scale, however, much of the drainage sweeping E or SE across the bajada west of the escarpment is deflected south, particularly at the southern end of the escarpment (Fig. 8), where the relief dies out. This deflection is characteristic of the propagating ends of hanging wall anticlines (e.g. at Sefidabeh in eastern Iran; Berberian et al. 2000), and several streams are bent round to the south even within the fold itself, perhaps testifying to a tip that has steadily propagated south. For this reason we are doubtful whether the deflections described by Hessami et al. (2004) are the result of historic strike-slip faulting. By contrast, there is little doubt that the main morphology of the fold escarpment is caused by a reverse faulting component at depth. The association between active reverse faulting, uplifted aquifers providing a water supply, and the locations of settlements, is common in the desert regions of Iran (Berberian et al. 2000; Jackson 2006), and is discussed later (Section 7.5).
Figure 7. (a) SRTM topography of the Bam-Baravat ridge illuminated from the NE. The steep eastern face of the ridge is marked by black arrows. The overall slope in this region is to the SE. (b) Three E–W profiles across the ridge along lines shown in (a), showing that the change in height across the ridge is constant for much of its length at about 15–20 m. (c) North-south profiles along the lines shown in (a), along the top of the ridge (west) and about 1 km east of the ridge at its base (east). The 15 m height difference between the profiles at the ends of the profiles is caused by the SE regional slope: the height difference increases rapidly from 15 to 30 m within 2 km at each end of the ridge.

4 RADAR INTERFEROMETRY

4.1 Dislocation modelling

The earthquake gave rise to exceptionally clear interferograms made from ENVISAT ASAR data, largely because of the excellent coherence between the pre- and post-earthquake images. Various groups have looked at interferograms from these images, and from them inferred source parameters of the coseismic faulting. The earliest study, by Talebian et al. (2004) used just the descending interferogram, whereas later ones by Wang et al. (2004), Funning et al. (2005), Fialko et al. (2005) and Motagh et al. (2006) used both ascending and descending tracks, thereby considerably reducing trade-offs between parameters. In addition, Funning et al. (2005) and Fialko et al. (2005) calculated azimuth offsets for both ascending and descending tracks, and when these were combined with ascending and descending interferograms, were able to recover all three components (east, north and vertical) of the 3-D displacement field of the surface around Bam. Source parameters of the faulting were then obtained by comparing the observed interferograms or ground displacements with those calculated from elastic dislocation models.

The studies of Wang et al. (2004), Funning et al. (2005), Fialko et al. (2005) and Motagh et al. (2006) differ in detail, but are agreed on the main features. All agree that the bulk (>80 per cent) of the moment was released by almost vertical, nearly pure strike-slip faulting beneath the ruptures described south of Bam in Section 3.2, with the main slip occurring over a distance of about 12 km running from the southern limit of the observed surface ruptures in Fig. 4(c) northwards into Bam itself. The maximum slip was about 2.7 m at a depth of about 5 km, dying away downwards at 10–12 km depth and upwards at 1–2 km beneath the surface. The lack of slip shallower than 1–2 km depth on the fault is also consistent with the terrestrial levelling data of Motagh et al. (2006) along the main Bam-Baravat road south of Bam. Wang et al. (2004) infer a small amount of slip on faults through and north of Bam (roughly beneath the faulting described in Section 3.1), which was not thought necessary by Funning et al. (2005) or Fialko et al. (2005).

More significantly, Funning et al. (2005) claimed that a substantial (~25 cm) eastward and upward component of motion they had detected in the SE quadrant of the main strike-slip rupture could only be modelled by additional slip with a reverse component of motion on a fault dipping west. They were able to reproduce this signal with oblique-reverse slip on a fault dipping 64°W and striking parallel to, and east of, the main strike-slip fault. They also show that the rms misfit of observed-to-calculated ground displacements reduce from 2.5 cm (for uniform slip) and 1.9 cm (for variable slip/rake) in their best one-fault models to 1.7 cm (uniform slip) and 1.3 cm (variable slip) in their best two-fault models; in other words, that the misfit is reduced by a third in both uniform- and variable-slip models. A test to argue that this improvement is statistically significant is given in the Appendix. The feature in the interferograms that requires something other than a simple, pure strike-slip rupture is the teardrop-shaped pattern in the SE quadrant of the main strike-slip fault (Fig. 9a), which has the same sign in both ascending and descending interferograms, and must therefore be a vertical signal. Other published InSAR studies (e.g. Wang et al. 2004; Fialko et al. 2005) also model this feature with a dip-slip component though not,
of their models is perhaps less obvious than in Funning et al. (2005) produced two models (Table 1): one in which the slip was uniform over the two fault surfaces, and another, with the same fault orientations and positions, that allowed slip to vary spatially over those surfaces. Their variable slip model and an observed and synthetic interferogram are shown in Fig. 9. Of particular importance, in the light of the aftershock data discussed below in Section 5.2, is the confidence with which we can say that the InSAR data rule out significant slip below 8–10 km depth. Although spatial resolution of slip on the fault plane degrades with depth, the detection of slip itself at depths of 8–10 km is good at the 15 cm one-sigma level on the main strike-slip fault and at 7 cm one-sigma level on the oblique reverse fault (Funning et al. 2005). There is thus little doubt that the majority of the slip responsible for the observed interferograms occurred above 8–10 km.

Several features of Fig. 9 need emphasizing. Both modelled faults are ‘blind’, in that nearly all the slip is beneath the surface, and the principal offset on both faults is strike-slip, with the rake values being 178° and 150°. The more oblique slip on the parallel west-dipping fault is so concentrated at its base, and so close to the main strike-slip fault, that it is unlikely the model presented here is unique. It is possible that a more complicated strike-slip fault geometry, involving a twisted surface that acquires a westward dip and a change in slip vector (to acquire a reverse-slip component) at its southern end, would also produce a satisfactory fit to the interferograms and surface displacement field. Wang et al. (2004) and Fialko et al. (2005) attempted to model the vertical component of ground motion in the SE quadrant of the main strike-slip fault by varying the rake on a single fault of constant orientation, but Funning et al. (2005) argue that the misfit from such a model is worse than for a two-fault model (see also the Appendix to this paper). The virtues of Funning et al.’s two-fault model in Fig. 9(c) are:

1. it is simple, in that there is no change of rake on a fault or fault curvature;
2. there is no doubt that a parallel west-dipping fault with a reverse component does exist beneath the Bam-Baravat anticline, and the cracking along the scarp suggests it may have moved at depth;
3. it is therefore consistent with the observed geology and
4. it is also consistent with the teleseismic waveform data (Section 5.1).

Although Funning et al. (2005) allow their fault surfaces to continue south of the observed ruptures south of Bam and south of the southern end of the Bam-Baravat escarpment, the patches of significant slip in their model on both surfaces are confined beneath the observed surface ruptures or project to the topographic expression of the ridge. As we will see in the seismic waveform analysis, the seismograms cannot be reproduced by a single strike-slip rupture alone, and also require a subevent with a reverse-faulting component. The slip of the oblique-reverse fault in Funning et al.’s model contributes only 15 per cent to the total moment release (Table 1).

---

**Figure 8.** Geomorphic map of the southern end of the Bam-Baravat ridge. The red line is at the base of the steep eastern face of the escarpment. Blue lines and shading mark active drainage courses, thick dotted lines are dry valleys and drainage courses. Thin dotted lines are traces of underground irrigation tunnels (‘qanats’). The five black circles are areas of intense qanat activity. Southward deflection of rivers. The red line is at the base of the steep eastern face of the escarpment. Blue lines and shading mark active drainage courses, thick dotted lines are dry valleys and drainage courses. Thin dotted lines are traces of underground irrigation tunnels (‘qanats’). The five black circles are areas of intense qanat activity. Southward deflection of rivers. The red line is at the base of the steep eastern face of the escarpment. Blue lines and shading mark active drainage courses, thick dotted lines are dry valleys and drainage courses. Thin dotted lines are traces of underground irrigation tunnels (‘qanats’). The five black circles are areas of intense qanat activity. Southward deflection of rivers.
Figure 9. A summary of the coseismic interferogram dislocation modelling by Funning et al. (2005). (a) The observed wrapped ascending interferogram, with one colour cycle representing 28 mm line-of-sight change between the satellite and the ground. (b) Model ascending interferogram for the two-fault, variable-slip model of the Bam earthquake (Table 1). The thick black lines in (a) and (b) show the length of the modelled main strike-slip fault (west) and the reverse fault (right) projected to the surface. (c) The distribution of slip on the main strike-slip fault (left) and on both faults (right). Viewing position is from the NE. More details of the modelling and the interferograms are given in Funning et al. (2005).

Table 1. Source parameters of the Bam earthquake used for dislocation (Fig. 9) and teleseismic waveform modelling (Fig. 10); see Funning et al. (2005). Latitude and longitude are of the centre of the fault plane projected to the surface. Depths are centroid depths. The slip value is the peak slip. In the variable-slip model, the fault length and width are the dimensions of the area enclosed by the slip contour within which 95 per cent of the fault slip occurred. Funning et al. (2005) give, and discuss, values for the errors in the uniform-slip model, which was obtained by an inversion procedure. In the variable-slip model, only the spatial distribution and amount of slip on the surfaces were allowed to vary.

<table>
<thead>
<tr>
<th>Fault</th>
<th>Lat.</th>
<th>Long.</th>
<th>Depth</th>
<th>Strike</th>
<th>Dip °</th>
<th>Rake °</th>
<th>Slip m</th>
<th>Length km</th>
<th>Width km</th>
<th>Moment $\times 10^{18}$ Nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform slip</td>
<td>1</td>
<td>29.038</td>
<td>58.357</td>
<td>5.5</td>
<td>354.6</td>
<td>−177.8</td>
<td>2.14</td>
<td>12.0</td>
<td>8.6</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>28.988</td>
<td>58.406</td>
<td>6.7</td>
<td>180.0</td>
<td>63.9</td>
<td>149.5</td>
<td>2.04</td>
<td>14.8</td>
<td>1.4</td>
</tr>
<tr>
<td>Variable slip</td>
<td>1</td>
<td>29.040</td>
<td>58.356</td>
<td>6.4</td>
<td>354.6</td>
<td>−177.8</td>
<td>2.69</td>
<td>20.0</td>
<td>15.0</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>28.988</td>
<td>58.406</td>
<td>5.9</td>
<td>180.0</td>
<td>63.9</td>
<td>149.5</td>
<td>2.04</td>
<td>18.0</td>
<td>1.6</td>
</tr>
</tbody>
</table>

4.2 Decorrelation imaging

An extraordinary feature of the Envisat ASAR data at Bam was the contrast between the exceptional coherence of successive images in the surrounding desert and the decorrelation of those images along the ruptures south of Bam (Fig. 4c). This contrast was so extreme and clear that the ruptures themselves were visible in the correlation image (Fig. 4c) and this guided us to their previously undiscovered surface location a month after the earthquake (reported in Talebian et al. 2004). The high coherence is likely to result from the almost-rigid encrusted desert pavement, into which pebbles are cemented by carbonate or salt, as can be seen in Figs 5(c) and (f). Broad zones of decorrelation around fault ruptures have been observed elsewhere (Simons et al. 2002), but the level of detail seen at Bam is remarkable. This relation between the surface ruptures and the InSAR decorrelation has been studied in detail by Fielding et al. (2005). One result they highlight is the contrast between the clarity of the decorrelated discrete ruptures south of Bam (Fig. 4c) and the more dispersed decorrelation associated with the distributed ruptures north of Bam (Fig. 4a). They attribute this contrast to an E−W
shortening across the fault zone in the north, leading to a distributed pattern of cracking, and an E–W extension across the fault-zone in the south leading to a simpler, discrete set of fractures. We note that the ruptures in the north (Section 3.1) do, indeed, show numerous ‘pop-up’ features and en-echelon mole tracks, indicative of a shortening component, and that an E–W extension in the south is expected at the surface above a buried thrust. The linear decorrelation feature marked X in Fig. 4(a) corresponds to the location of the northern fault inferred from the InSAR data by Wang et al. (2004). Fielding et al. (2005) show that it coincides with a step of about 5–7 cm in E–W shortening, though the localized uplift component shows that any slip is very shallow. There may also be some strike-slip displacement at this feature, not resolved (or confused with the vertical component) in the InSAR analysis, due to its low sensitivity to N–S motion. No significant surface faulting was found at site X, though there were some minor ground cracks (Fielding et al. 2005).

The correlation images used to produce Figs 4(a) and (c) are from a descending track pair (2003 December–2004 February) with a much shorter baseline (roughly 10 m) than the pair used in Fielding et al. (2005), but the computational technique employed is the same as the phase-sigma technique of that paper. Fielding et al. (2005) were also able to use the change in correlation before and after the earthquake to map the damage distribution to buildings in the city of Bam itself.

5 SEISMOLOGICAL INFORMATION

5.1 Teleseismic waveform analysis

In the supplementary on-line material that accompanied their initial report on the Bam earthquake, Talebian et al. (2004) pointed out that the long-period P and SH waveforms of the main shock could not be explained by a single centroid source with a strike-slip mechanism. They suggested that the main shock consisted of two discrete pulses, a larger (Mw 6.5) N–S right-lateral strike-slip rupture, followed 9.5 s later by a smaller (Mw 6.0) thrust rupture on a N–S striking plane dipping 30° W. This analysis of the seismograms, and particularly the inclusion of the second smaller subevent, was influenced by (a) the minor surface ruptures observed along the Bam-Baravat escarpment and (b) the need for such a second pulse to explain the InSAR interferograms. With the much more detailed InSAR analysis now available (Funning et al. 2005; Fialko et al. 2005), including both ascending and descending tracks and azimuth-offset data, a re-analysis of the body waveforms is justified. Of particular interest is whether the waveforms can be explained by coseismic faulting with the same geometry as that inferred from the InSAR data by Funning et al. (2005) in Fig. 9.

We selected 25 P and 25 SH waveforms from the GDSN broadband network, which recorded the earthquake with good azimuthal coverage (Fig. 10). We first convolved the records from stations in the teleseismic distance range of 30–90° with a filter that reproduces the bandwidth of the old WWSSN 15–100 long-period instruments. For earthquakes of Mw 6.5 at these wavelengths, the source appears as a point source in space (the centroid) with a finite rupture time, and the resulting seismograms are sensitive to the source parameters of the centroid while relatively insensitive to the details of geological structure. We then used the MT5 version (Zwick et al. 1994) of McCaffrey & Abers’s (1988) and McCaffrey et al.’s (1991) algorithm, which inverts the P and SH waveform data to obtain the strike, dip, rake, centroid depth, seismic moment and the source time function, which is parametrized by a series of isosceles triangle elements of half-duration 1.0 s. The source was always constrained to be a double-couple. Details of the program, algorithm and approach we used are described in detail elsewhere (e.g. Nabélek 1984; McCaffrey & Nabélek 1987; Molnar & Lyon-Caen 1989; Taymaz et al. 1991).

We first generated synthetic seismograms for the main strike-slip rupture determined from InSAR data, using the same strike, dip, rake and moment as that in the uniform-slip model determined by InSAR (Table 1). We assumed a centroid depth of 5.5 km, close to the centre of slip distribution pattern in Fig. 9(c). We then allowed the inversion to determine the only remaining unknown source parameter, which is the source time function. The resulting waveforms are shown in Fig. 10(a), and reproduce the shape and amplitude of P and SH at most stations reasonably well. However, there are two obvious areas of the focal sphere where the fit is bad. The P waveforms at stations in the east and southeast (INCN, ENH, GUMO, MBWA, NWAO) are close to a P nodal plane and have small calculated waveforms, even though the observed waveforms are clear and impulsive. A similar effect is observed in the SH waveforms at southwest stations (LBTB, LSZ), which are also close to a nodal plane and have small calculated amplitudes, even though the observed onsets are large and clear. These deficiencies at these stations are inevitable given the orientation of the main strike-slip rupture, which is well constrained by the InSAR data.

The fits of synthetic to observed seismograms at these stations is improved by the addition of a second oblique-thrust subevent, with exactly the same strike, dip, rake and moment as used in the uniform-slip model determined by InSAR analysis (Table 1). We assumed a centroid depth of 6.7 km, close to the centre of the slip distribution for the second (oblique-reverse) rupture plane in Fig. 9(b), and thus the only parameters left undetermined were the source time function and its origin-time delay relative to the first subevent. We found the best fit occurred with a short time function delayed only 1 s after the onset of the first subevent (Fig. 10b). For this second subevent, those P stations in the east and SH stations in the southwest mentioned above are further from the nodal planes, and have substantial onsets. In particular, the increased dip-slip component of the rake in the second subevent introduces, through SV and SP, a relatively large amplitude to the P waves at stations in the east and southeast, even though the moment is five times smaller than in the first subevent. The comparison between waveforms from the one-source and two-source models at these stations is shown in Fig. 11. The fit at most other stations is much the same, or improved, by the addition of the second pulse. Only at DBIC (in the west) is the P waveform nodal for both subevents, and therefore small in the synthetics, even though it has an observed waveform of substantial amplitude. We suspect that this is because it is sensitive to small changes in the orientation of the nodal planes, and note the SH fit at this station is good.

We conclude that the seismograms can be fit reasonably well with a two-source model which is the same as that deduced from the InSAR analysis by Funning et al. (2005). The delay between the two pulses is only 1 s, and we could not resolve any spatial separation of the centroids. Clearly, given the long-period nature of the waveforms, we cannot distinguish between a model with two simple discrete subevents separated in time on surfaces of different orientations (as shown here) and a more complicated fault geometry involving propagating rupture on a single distorted rupture surface that acquires a more westward dip and a greater thrust component to its slip vector at one end. The same point was made regarding the interpretation of the InSAR interferograms in the previous section.
Figure 10. Long-period $P$ and $SH$ body wave modelling of the 2003 Bam mainshock, with the same fault parameters as those used in the uniform-slip InSAR modelling by Funning et al. (2005), shown in Fig. 9 and Table 1. (a) Seismograms for the strike-slip fault alone, with only the source time function (STF) free in the inversion. The event header shows the strike, dip, rake, centroid depth and scalar seismic moment (in Nm) of the minimum misfit solution. The top focal sphere shows the lower hemisphere stereographic projection of the $P$ waveform nodal planes, and the positions of the seismic stations used in the modelling routine. The lower focal sphere shows the $SH$ nodal planes. Capital letters next to the station codes correspond to the position on the focal sphere. These are ordered clockwise by azimuth, starting at north. The solid lines are the observed waveforms, and the dashed lines the synthetics. The inversion window is marked by vertical lines on each waveform. The source time function (STF) is shown, along with the timescale for the waveforms. The $P$ and $T$ axes within the $P$ waveform focal sphere are shown by a solid and an open circle respectively. Note that the fit of $P$ waves is particularly poor in the E and SE, and for $SH$ in the SW. Waveforms at some of these stations, highlighted in yellow, are shown in more detail in Fig. 11. (b) Seismograms for the two-source model, incorporating the second fault, with an oblique thrust mechanism (green focal sphere) used in the InSAR modelling.

Figure 11. A comparison between the one-source and two-source models at stations highlighted in yellow in Fig. 10(a). The top line shows observed (solid) and synthetic (dashed) $P$ waveforms at INCN, GUMO, MBWA and NWAO and $SH$ waveforms at LBTB and LSZ for the two-source model. Numbers beneath station codes are station azimuths from the epicentre. The bottom line shows waveforms for at the same stations for the single strike-slip source alone.

The important conclusions from this section are (a) that the seismograms can be interpreted in a way that is consistent with the InSAR interpretation, and (b) that both require a dip-slip component in addition to the nearly pure strike-slip component that is responsible for the bulk of the signal in both cases.

5.2 Aftershock locations and mechanisms

For the period 3–35 days after the earthquake a dense seismic network of 23 stations was operated in the epicentral region to record aftershocks, and the details of this experiment are reported by Tatar.
suggest that the aftershock zone has a steep (80◦ ∼ 85◦ south of Bam, broadening with depth (Fig. 12b). Tatar and (b).

north and does not show a concentration of aftershocks along the

the main InSAR interferogram signal (Talebian et al. 2005; Fialko et al. 2005). A number of important conclusions are worth noting here, to compare them with those from other data.

Firstly, the aftershock distribution delineates an intense N–S zone of activity (Fig. 12a) running through Bam and along the coseismic ruptures south of the city that were observed both in the radar decorrelation and on the ground (Talebian et al. 2004; Fielding et al. 2005) and which were clearly along the line of the fault responsible for the main InSAR interferogram signal (Talebian et al. 2004; Funning et al. 2005; Fialko et al. 2005). The pattern becomes diffuse in the north and does not show a concentration of aftershocks along the coseismic ruptures observed north of the city in Figs 3(a), (b), 4(a) and (b).

The aftershock zone is subvertical beneath the coseismic ruptures south of Bam, broadening with depth (Fig. 12b). Tatar et al. (2005) suggest that the aftershock zone has a steep (∼80◦) westward dip, compared to the steep (80–85◦) eastward dip required by the radar interferograms and teleseismic waveforms. Whether this discrepancy is real, or whether the aftershock zone simply becomes broader with depth, is not clear. What is remarkable about the aftershocks is not just their delineation of the N–S rupture, but their depth distribution (Fig. 13a). Nearly all the best-determined locations of Tatar et al. (2005) lie in the range 7–20 km, and thus almost entirely below the 2–8 km depth range in which most of the slip occurred in the main shock, and which produced the surface deformation revealed in the radar interferograms (Fig. 13). The depth range of maximum slip in the main shock is nearly completely free of aftershocks in the Tatar et al. (2005) study. This is a very significant observation as it suggests that the thickness of the seismogenic zone in the region is about 20 km, and that only half of this ruptured in the Bam main shock. An important question (discussed later) is whether the remaining, unruptured, half may still fail seismically in a future event.

With virtually no aftershocks shallower than 7 km, the aftershock distribution of Tatar et al. (2005) has little to say about the question of whether the main shock occurred on two separate faults or one twisted surface. Downward projections of the N–S strike-slip fault south of Bam and the fault with a reverse component beneath the Bam-Baravat escarpment would merge near the top of the aftershock zone, if we assume the dips for them that were estimated from the InSAR analysis (Fig. 12b).

After the period of Tatar et al.’s (2005) aftershock study (3–35 days after the main shock) ended, a second aftershock study was carried out by Nakamura et al. (2005) 41–70 days after the main shock. Their results were much the same, demonstrating a thickness to the seismogenic layer of ∼20 km, and a relative lack of earthquakes at shallow depths above the main coseismic rupture in the south. However, elsewhere, in contrast to the Tatar et al. (2005) study, they found much more aftershock activity at depths less than 5 km, and a more diffuse distribution of aftershocks, particularly in the north. Some of these differences may be attributable to the later period of activity, as aftershock distributions commonly broaden with time (e.g. Yielding et al. 1989). However, the contrast between the lack of shallow activity in the earlier study compared to that in the later may also be related to the different seismograph station distributions. Tatar et al. (2005) had a particularly dense station network, with at least 8 of their 23 stations less than 5 km apart above the central part of the aftershock zone in eastern Bam. This configuration is well able to resolve events shallower than 7 km, if they exist. By contrast, Nakamura et al.’s (2005) network had only 9 stations, with just a single station in Bam, and an inter-station spacing of typically ∼10 km; this configuration is much less able to resolve truly shallow depths of <5 km. We are therefore unable to evaluate Nakamura et al.’s (2005) suggestion that a band of aftershocks dips west at shallow (0–5 km) depths beneath the Bam-Baravat ridge; which is not a feature of the Tatar et al. (2005) aftershock distribution.

5.3 Strong ground motion
The ground motion produced by the earthquake was recorded within the city of Bam itself (Fig. 12a). This is a rare case where direct
recording of the ground motion was made in a place so massively destroyed. The instrument is a digital accelerometer installed by the Iranian Building and Housing Research Center (http://www.bhrc.gov.ir). Such an instrument records the ground acceleration and has a flat response in frequency so that the integration of the records yields the ground velocity (Fig. 14) while a second integration yields the ground displacement. The dominant feature of the ground motion is the high-amplitude E–W pulse, transverse to the fault (120 cm s\(^{-1}\) of peak velocity, 40 cm of peak displacement). This pulse is likely to have been a major cause of the catastrophic damage in Bam. During the 2–3 s duration of this pulse, the N–S and vertical motion are small. These characteristics are those theoretically expected near the fault for a N–S trending right-lateral strike-slip fault where rupture propagates northwards towards the station (Aki 1968; Somerville et al. 1997; Somerville 2003). The arrival of the hypocentral S wave corresponds to a polarity reversal on the transverse component, from the eastward motion of the near-field P wave to the westward motion of the SH wave. The records thus yield a clear S–P time of 1.9 s (Fig. 14), which, for the upper-crustal velocities inferred in the region (Tatar et al. 2005) places the zone of rupture initiation at about 14 km from the station. Because the event is shallow, this hypocentral distance places the point of initiation near the southern edge of the ruptured area. In such a configuration, the elastic strain energy released is strongly focused in the direction of the propagating rupture; that is, northwards. It is precisely there, near the northern edge of the slip patch, that Bam was located.

Synthetic seismograms calculated using the fault geometry and slip distribution determined by Funning et al. (2005) from the InSAR data (Fig. 9c), with the inferred hypocentral distance of 13.7 km, reproduce remarkably well the amplitude and shape of the transverse (E–W) ground-velocity pulse produced in Bam by the earthquake (Fig. 14). To fit the timing of the pulse requires a rupture velocity of 2.8 km s\(^{-1}\), which, within uncertainties, is the Rayleigh-wave velocity of the upper crust (Tatar et al. 2005) where faulting takes place. The width of the pulse constrains the rise time (the local duration of slipping) and implies that slip occurred at high slip velocity exceeding 2 m s\(^{-1}\) over a large part of the fault. The fit between data and synthetics is not very sensitive to the hypocentral depth itself provided that it lies between 3 and 10 km. The best match, however, is obtained for a depth close to 6 km.
About 1.5 s after the transverse E–W pulse, whose peak amplitude corresponds to the passage of the rupture front near the station, the Bam records show some complexity, particularly on the N–S component parallel to the fault, that is not explained by this main rupture. Its origin is unknown: it could relate to the later subevent, responsible for ∼20 per cent of the moment release, that was inferred from the teleseismic analysis (and indirectly from the InSAR), but we could not match it with the model for the second event used in Figs 9 and 10b (and Table 1), which is located too far from the station and produces too small an amplitude at Bam. With just one strong motion station, we cannot determine uniquely the source location and timing which caused this later disturbance. The aim was simply to show that the essential features of the observed ground motion in Bam can be reproduced by the same fault-slip model for the principal rupture that was obtained from InSAR analysis, which is also compatible with the observed teleseismic seismograms. The location of the hypocentre (rupture nucleation) inferred from the strong motion analysis is shown by a blue circle in Fig. 13(b).

In summary, it is evident that the position of Bam at the northern end of the strike-slip rupture surface made it particularly vulnerable. The strong directivity effect, enhancing the E–W horizontal motion, was further amplified by the near-Rayleigh rupture velocity. The high slip velocity adds to these two factors which, all together, are responsible for the almost complete destruction of the part of Bam situated along the line of the main fault (e.g. NCC 2003; Fielding et al. 2005; Bouchon et al. 2006).

6 SYNTHESIS: COSEISMIC FAULTING IN THE BAM EARTHQUAKE

At this point we can attempt to resolve some of the uncertainties concerning the coseismic rupture surfaces in the main shock, many of which arise from ambiguities in any one type of data used on their own.

From the InSAR data, there is little doubt that the principal rupture surface during the main shock, responsible for the ground deformation pattern seen in the InSAR interferometry, was a near-vertical strike-slip fault extending from the city of Bam towards the south. The surface expression of this rupture was the relatively minor set of fractures described by Talebian et al. (2004) and observed in the radar decorrelation (Talebian et al. 2004; Fielding et al. 2005), shown in Figs 4c, 5d and e. However, it is clear from the amplitude of the signal in the InSAR interferograms that most of the slip, with peak values of more than 2 m at 5 km depth, failed to reach the surface. Although Talebian et al. (2004) reported a maximum surface offset of ∼20 cm on these fractures, Binet & Bollinger (2005), using subpixel correlation of SPOT-5 images, found a surface offset that was larger, with a maximum value of 1.2 m and a mean of ∼0.8 m. They attribute the difference to distributed surface deformation over a zone perhaps 500 m wide. This is consistent with the model of Funning et al. (2005), based on InSAR interferograms, which also has a maximum of 0.8 m slip in the upper 1 km, but which cannot resolve the spatial detail inherent in the SPOT images. Fialko et al. (2005) discuss the coseismic ‘slip-deficit’ at shallow (<4 km) depths, which must be taken up somehow in the interseismic period, perhaps by distributed, non-seismogenic processes.

One of the best-observed features of the deformation pattern seen in the InSAR interferograms is the wavelength of the signal away from the fault (Fig. 9a), which in turn is related to the depth extent of the rupture surface. Thus we know that the bulk of the slip on this rupture surface occurred above 8–10 km depth, yet the after-shock distribution (Fig. 13) strongly suggests that the fault zone extends, within the seismogenic layer, to about twice that depth. Dislocation modelling of the interferograms (Funning et al. 2005; Fialko et al. 2005) indicates a fault length for this segment of about 12–20 km (depending on whether uniform or variable slip models are used), slightly longer than the observed ruptures south of city and extending into the city itself. The aftershock zone associated with this segment over the month following the earthquake is somewhat longer (Fig. 12a), and may have broadened even more after that (Nakamura et al. 2005), but this is not unusual. The main strike-slip pulse seen in the seismograms (Fig. 10a) can be fit with the same moment value of 7.6 × 10^18 N m as that deduced for the strike-slip segment from the radar analysis (in the uniform-slip model), corresponding to an average slip of 2.1 m over a vertical fault of dimensions 12 × 9 km^2, suggesting that this seismic pulse was indeed caused by slip on the principal rupture surface identified by the radar imagery. This interpretation is also consistent with the observed and synthetic strong motion seismograms shown in Fig. 14.

The ruptures north of the Posht river and city of Bam (Figs 4a,b and 5a,b), which attracted such attention in the immediate aftermath of the earthquake (Talebain et al. 2004; Hessami et al. 2004) and which were discovered before the radar decorrelation, drew attention to the ruptures south of Bam, having a minor, barely resolvable effect on the InSAR interferograms (e.g. Wang et al. 2004). Slip on these fractures is small and, unlike those south of Bam, does not represent the surface expression of much more substantial movement at depth. They are distributed over an E–W width of several km (e.g. Hessami et al. 2004), and are not associated with any localized aftershock activity (Fig. 12a). Indications from the strong ground motion analysis are that the rupture on the main segment propagated north, and we suspect that the distributed fracturing and aftershock activity at the northern end of the main rupture may be related to the effects of this propagation; both from the enhanced ground accelerations and dynamic effects (e.g. Fig. 5c) and from the distributed minor faulting that is common at the end of main fault segments. A similar pattern of both more distributed fracturing and more distributed aftershocks was observed, for example, at the end of the main 1980 El Asnam earthquake fault rupture, in the direction of propagation (e.g. Yielding et al. 1989). Fielding et al. (2005) also point out, from the InSAR phase signal, that this northern region experienced a component of E–W shortening at the surface.

It remains to discuss the significance of the Bam-Baravat escarpment in this earthquake. There is no doubt that the antcline ridge is the surface expression of a fault that must dip west, and have a reverse component, and be largely ‘blind’ in that slip at depth is accommodated mostly by folding at the surface. The ridge is the most obvious geomorphological feature in the region (Figs 2 and 3) and was naturally the focus of attention following the earthquake. Minor coseismic cracks were observed along and close to the base of the scarp (Talebain et al. 2004; Hessami et al. 2004), but there is no doubt that the main ground deformation, as seen in the radar interferograms, was not caused by slip on this fault. Furthermore, the seismograms of the earthquake indicate that >80 per cent of the moment release in the earthquake occurred on a nearly vertical strike-slip fault with almost horizontal slip vector; which is not the characteristic of the fault beneath the Bam-Baravat ridge. On the other hand, neither the radar interferograms or the teleseismic seismograms can be explained by horizontal slip on a near-vertical strike-slip fault alone; both require an additional segment of slip with a reverse component. The radar data localize this extra component to a small patch with restricted depth near the base of the ...
main rupture (Fig. 9b). Again, the seismograms can be fit with a second subevent of the same moment as that deduced from the radar (Fig. 10b). The orientation of the second fault used in the radar analysis by Funning et al. (2005), shown in Fig. 9(b), and which is compatible with the seismic waveforms (Fig. 10b), projects to the surface at the base of the Bam-Baravat ridge. An obvious conclusion is that the oblique-reverse fault beneath the ridge did indeed slip a little in earthquake, but only at depths of 4–6 km (where the average slip may have been 1–2 m), and only minor cracks propagated to the surface. As we have pointed out above, in formal terms, given the uncertainties and trade-offs in the radar and seismic analysis that are inevitable with a complicated rupture, it is probably not possible to distinguish between slip on two distinct faults near their junction, and a single, more-complicated, warped surface with a changing slip vector. However, this seems to us to miss the point: there is no doubt that a vertical strike-slip fault exists beneath the main ruptures south of Bam and that a parallel-striking, west-dipping fault exists beneath the Bam-Baravat ridge. These faults must meet at depth, and the radar interpretation of Funning et al. (2005) in Fig. 9 suggests they meet near the base of the main rupture surface. The aftershock distribution (Fig. 12b) indicates that some sort of fault zone exists beneath this junction, but actual faulting near the junction itself must be more complicated than a simple intersection, as the reverse slip component on the dipping fault would offset a vertical strike-slip fault that reached from the surface to 20 km depth. It seems likely that the overall motion on the fault zone as a whole is oblique right-lateral and reverse in nature, and that it is partially separated ("partitioned") into pure strike-slip and oblique-thrust components on parallel-striking faults at shallow depths (less than ~8 km). The exact manner in which this separation happens is not clear, because of the severe strain compatibility problems, but as a phenomenon it is not unusual: it is, for example, a characteristic of the oblique left-lateral and reverse faulting associated with the great Gobi-Altay earthquake (M ~ 8) in Mongolia in 1957 (e.g. Kurushin et al. 1997; Bayasgalan et al. 1999).

If our interpretation is correct, and the small oblique-reverse subevent in the Bam earthquake represents slip at the base of the fault beneath the Bam-Baravat ridge, then the radar and teleseismic waveform interpretation indicate that this fault also has a strike-slip component of motion, as Hessami et al. (2004) suggest, though its geometrical expression is clearly dominated by the reverse component. In this case the ‘partitioning’ has not separated a pure reverse component, but this is also not unusual, and is also a characteristic of the "forebergs" on the Gobi-Altay earthquake fault, which are antichines that also had a strike-slip component of coseismic motion (Bayasgalan et al. 1999).

Thus, in summary, the Bam fault system seems to involve the branching of a fault zone within the seismogenic layer. At 10–20 km depth that fault zone probably involves localized strike-slip with a minor reverse component. At 8–10 km depth the fault zone appears to branch into a near-vertical, almost pure strike-slip fault (the one responsible for most of the moment release in the main shock), and into a west-dipping oblique-reverse fault that projects to the surface at the Bam-Baravat ridge. Tatar et al. (2005) found evidence for a change in velocity structure at about 8 km depth, from $P$-wave velocities of ~5.3 km s$^{-1}$ to ~6.2 km s$^{-1}$. Such a contrast may represent a change in material properties, and have limited the depth extent of rupture in the main shock. It may also be related to the branching of the fault that seems to occur near this depth. However, with coseismic slip in main shock above 8 km and aftershocks below that depth, the entire layer of 20 km thickness must be regarded as seismogenic, in principle.

7 TECTONIC AND SEISMIC HAZARD ISSUES

The 2003 Bam earthquake was one of a series of $M_w > 6$ events on the western side of the Dasht-e-Lut since 1981 (Fig. 1b), all of them associated with right-lateral shear and shortening between the Lut and central Iran. This activity has continued since 2003, with the 2005 $M_{w}$ 6.4 Dahuiyeh (Zarand) earthquake further north of the region in Fig. 1b (Talebian et al. 2006). Iran is a country with a long and well-studied historical earthquake record (Ambraseys & Melville 1982; Berberian 1994) which shows that periods of enhanced activity along fault zones, separated by periods of relative quiescence, are not unusual. Such behaviour is also expected in some dynamic models of fault systems (e.g. Ben-Zion et al. 1999; Kenner & Simons 2005). The faulting on the western side of the Dasht-e-Lut is apparently going through such an enhanced seismically active period at this time. The issue of the seismic hazard represented by other faults that have yet to be ruptured is therefore important. Several of the earthquakes in this active period since 1981 have involved surface rupture, and the faulting in the Bam earthquake, in this more general context of faulting within the fault system west of the Lut, raises questions of general tectonic significance, beyond just the context of Iran. These are issues we discuss in this section.

7.1 Spatial 'partitioning' of slip on faults

In general, the motion between the Dasht-e-Lut and central Iran is thought to be nearly pure N–S right-lateral shear (Fig. 1a; Vernant et al. 2004), parallel to the Nayband and Sarvestan strike-slip faults (Fig. 1b) which have almost no associated topography that might represent a shortening component (Walker & Jackson 2002, 2004). However, in the region of the Gowk Valley, between Chahar Farsakh and Golbaf (Fig. 1b) the fault system is oblique to this trend, and the overall motion is spatially separated into strike-slip faulting in the Gowk valley and shortening on the Shahdad thrust system. This separation was dramatically revealed by the coseismic interferogram of the 1998 Fandoqa earthquake (Berberian et al. 2001; Fielding et al. 2004) showing that both fault systems had moved in a time window covering that event. The folds associated with the Shahdad thrust system are much larger, with ridge crests 100 m or more above the plain to the NE, than the Bam-Baravat ridge. Nevertheless, we argued above that a similar kind of spatial separation, or slip ‘partitioning’ occurs on the faults near Bam, and in the 2003 Bam earthquake itself, though the geomorphology and structure at Bam is much less developed at Bam than at Shahdad-Gowk. The Nayband-Golbaf-Sarvestan fault system is clearly a much more significant regional structure than the Bam faults, and almost certainly has a larger horizontal offset (estimated as ~13 km by Walker & Jackson 2002). Near Golbaf, the available evidence suggests the system has evolved into an underlying ramp-and-flat thrust geometry with the strike-slip fault as a higher-level, steeper splay off a structural ramp (Walker & Jackson 2002). At Bam, the strike-slip and reverse-component faults are still closer together and still evidently merge at quite shallow depth (~8 km). Moreover, the shortening component occurs on a fault that is probably still substantially strike-slip in character, based on the InSAR and teleseismic waveform analysis of the 2003 main shock. It is possible that the Bam faults represent an earlier stage of the evolution of such a partitioned system.

If the Bam fault system is in an early stage of development, it might explain the peculiar observation of Fielding et al. (2005), from the InSAR phase signal, that the region of northern fractures...
(Figs 4a,b and 5a,b) experienced about 100 mm of E–W shortening at the surface, although this was not concentrated on any one coseismic rupture. The contrast between this area and the region south of Bam may be that, in the north, the shortening component has not yet concentrated onto a discrete structure, as it has on the Bam-Baravat ridge to the south.

7.2 Partitioning with depth and future seismic hazard at Bam

The comparison may also be made between how the Bam and Golbaf-Gowk fault systems behave in earthquakes. The 1981 earthquakes, of $M_w$ 6.6 and 7.0, in the Gowk valley produced very small surface displacements. In particular the $M_w$ 7.0 event in 1981 produced a maximum surface offset of only 40 cm, whereas the smaller ($M_w$ 6.6) event in 1998 produced up to 300 cm offset along the same surface fault (Berberian et al. 2001). The explanation is almost certainly that the 1998 earthquake ruptured the shallow part of the fault system (InSAR constrains the depth extent to be shallower than 7 km), whereas the 1981 event ruptured the deeper part of the fault system (the seismic waveform analysis suggests a centroid depth of 15–20 km, but is not well constrained). The comparison with the 2003 Bam earthquake is now instructive. The aftershock distribution (Figs 12b and 13) clearly indicates that there is a substantial thickness (>10 km) of the fault zone within the seismogenic layer that was below the 2003 coseismic rupture in the main shock, and which did not rupture in 2003.

The important question is whether the unruptured parts of the Bam fault system represent a significant future hazard. It is not unusual for aftershocks to concentrate round the edges of fault patches that slipped in main shocks (e.g. Mendoza & Hartzell 1988; Bakun et al. 2005). The issue at Bam is whether the substantial thickness of the seismogenic layer at depths of 10–20 km, i.e. below the patch that slipped in the main shock, will eventually fail seismically or by creep. The fact that this layer is seismogenic in aftershocks (Fig. 13b) is not necessarily a guide, as they could represent either a response to sudden loading by massive shallow slip, or loading of the deeper region in preparation for a second massive slip event. These issues have been most studied in California. Following the 2004 Parkfield, California, earthquake, post-seismic afterslip in the first four months accounted for roughly half the coseismic moment release and occurred in the same place as substantial aftershock activity (Bakun et al. 2005; Johanson et al. 2006). On the Hayward Fault, in northern California, there is abundant repeating microseismicity in the same areas as steady fault creep (Burgmann et al. 2000; Schmidt et al. 2005). No measurements of fault creep are available at Bam and the expected long-term slip rates are so low (~1 mm yr$^{-1}$; see Section 2) that it would be very difficult to detect. Two years of post-seismic InSAR observations at Bam studied by one of us (EF) have so far mapped an absence of any significant post-seismic strike-slip motion on the main 2003 rupture surface, in contrast to that observed following the 2004 Parkfield earthquake. However, the faults at Hayward and Parkfield have long-term slip rates (~10–30 mm yr$^{-1}$) that are much faster than that expected at Bam, much larger total offsets, and thinner seismogenic layers (~12 km vs. 20 km). These factors, together with the relative rarity of confirmed creeping faults in the seismogenic layer elsewhere in the world, give us little confidence that the behaviour of those faults provide a reliable guide to the Bam fault system. Indeed, it is suggested (e.g. Scholz 2002) that faults in plate boundary systems, such as those in California, behave fundamentally differently from faults in intraplate regimes such as eastern Iran, where the stress drops and slip-to-length ratios estimated for individual earthquakes tend to be larger, earthquakes less frequent, and long-term slip rates lower, indicating perhaps that such intraplate faults are better able to heal between earthquakes.

It is, therefore, possible that the main Bam fault slides aseismically below the 2–8 km depth that ruptured in the main shock, in a way similar to parts of the Hayward fault and the Parkfield section of the San Andreas fault. However, given the lack of post-seismic strike-slip afterslip, and the behaviour of the earthquakes in the nearby Gowk Valley, it would be prudent to conclude that the unruptured, deeper part of the Bam fault system is capable of moving in a future separate earthquake. The dimensions of the unruptured fault plane indicated by the aftershocks suggest that such an earthquake, if it occurred, could be of similar size, and in a similar location (though slightly deeper) than that in 2003. The consequences of such an event for Bam are very severe, and should be taken into consideration in the reconstruction strategy.

Since it is also clear that the shallow part of the oblique-reverse fault beneath the Bam-Baravat ridge did not move a significant amount in 2003, it too could represent a future hazard, if it were to break in an earthquake. However, in this context its role is less clear. The 2003 coseismic slip at depth, estimated to be as much as 2 m by Funning et al. (2005) must be released at shallower levels somehow, though not necessarily by seismic slip in earthquakes. Aseismic shallow slip by creep may also be possible, as is suspected to have occurred at Shahdad in 1998 (Berberian et al. 2001; Fielding et al. 2004). Post-seismic InSAR observations at Bam studied by one of us (EF) do show a decrease in line-of-sight distance on both ascending and descending tracks that requires uplift east of the southern end of the main 2003 rupture, similar to the ‘teardrop’ uplift pattern in Fig. 9(a) discussed in Section 4.1 and in the same place. This can best be explained as continuing dip-slip motion on a fault in the same area, and is not compatible with strike-slip motion on the main 2003 rupture plane.

7.3 Southern extent of the Bam fault system

The approximate southern limit of the rupture in the 2003 Bam earthquake is well determined by both the radar interferograms (Fig. 9a) and the aftershock distribution (Fig. 12a). It corresponds, in latitude, with the southern end of the Bam-Baravat ridge, whose relief dies out to the south (Fig. 7) and around whose end the drainage is deflected (Fig. 8). Many seismotectonic studies and maps (e.g. Berberian 1976; Fu et al. 2004), including some of our own (Walker & Jackson 2002, 2004) mark the Bam fault system as continuing further to the SSE (dotted line in Fig. 1b), following a prominent lineation picked out by vegetation on the satellite imagery (Fig. 3). If the fault system does indeed continue along this line, its unruptured state after the 2003 earthquake represents a serious hazard for Bam and the surrounding settlements, particularly as Bam is along-strike from this lineation and vulnerable to enhanced propagation and directivity effects.

After examination in the field, and of the satellite imagery and digital topography, we now doubt whether this lineation is really an active fault. It is clearly picked out by the vegetation, which in turn is sensitive to springs that occur along this lineation. Springs are common along active faults in Iran (see Sections 7.4.2 and 7.5) and the association of a spring line with an active fault is, in general, a reasonable assumption. In this particular case, however, the line occurs precisely at the junction between major fan systems coming off the Jebel Barez mountains to the south. Fig. 3 shows a major
7.4 Neighbouring unruptured seismogenic faults

The Bam faults themselves are not the only ones that pose a seismic hazard to the Bam valley. Given the apparently active phase of the fault systems bounding the western side of the Dasht-e-Lut, two other fault systems merit immediate attention. Not much is known about previous earthquake history on either one, as historical coverage is poor in this region and there has been very little modern instrumentally recorded activity (Berberian 2005). However, both of them have clear geomorphological signals that indicate late Quaternary activity, and are described briefly below.

7.4.1 Jebel Barez thrust system

The Nayband-Golbaf-Sarvestan strike-slip system ends in the south by turning east along the northern flank of the Jebel Barez mountains. With the change in strike, the motion on the fault system becomes dominantly thrusting, whose geomorphological expression is of uplifted anticlines along the front of the range. An example is shown in Fig. 16(a), which shows a series of uplifted and now inactive fan surfaces elevated above the currently active outwash surfaces. The inactive fan surfaces have an abrupt northern edge, forming an arcuate shape above a mostly blind thrust fault dipping south. Rivers cutting though the uplifted surfaces leave a step-like series of terraces (Fig. 16b), in a morphology reminiscent of many other similar situations in Iran (e.g. Walker et al. 2003, 2005). The fault segment in Fig. 16(a) is 15–20 km long, and capable of generating an earthquake of at least $M_w$ 6.5. Such an earthquake poses a hazard to the Bam valley as well as to settlements along the mountain.
front, though these are mostly occupied during the summer pasture months only. The village of Meej (Fig. 16a) was undamaged in the 2003 event, but it was also unoccupied at the time as all its inhabitants were in Bam for the winter, where about 300 of them perished in the earthquake.

7.4.2 Sarvestan fault

The Sarvestan fault is the southern part of the Nayband-Golbaf system, and has not ruptured in modern times. It has a very clear expression as a linear feature, both in satellite imagery and on the ground (Fig. 17). Spring lines along the fault are responsible for vegetation and the existence of the Sarvestan oasis itself. Deflected streams along its length (Figs 17a, b and c) and its subdued topography (Fig. 17f) attest to its nearly pure right-lateral strike-slip character. Its length of 30–40 km and its continuity make it a possible source of a substantial earthquake of $M_w$ 7.0 or more, which represents a serious hazard for the settlement of Sarvestan and the western Bam valley. As Fig. 17 shows, there are numerous sites along it with potential for palaeoseismological investigation.

7.5 Habitation and faulting in the desert

A final point to emphasize and explain is the close relation between habitation and active faulting in the desert regions of Iran (Jackson 2006). Several recent earthquakes have apparently ‘targeted’ isolated desert settlements with an accuracy that seems either vindictive or extreme bad luck. Examples include the earthquakes at Sepidabeh in 1994 (Berberian et al. 2000), Tabas in 1978 and Ferdows in 1968 (Walker et al. 2003), as well as Bam in 2003. In each case the destroyed town concerned was the only sizeable settlement for 50 km in any direction. The reason for this apparent targeting is clearly illustrated by the example of Bam. Bam and Baravat are neighbouring desert oases, famous for growing dates that are much prized throughout Iran. The reason for their existence is the availability of water, which is provided by the uplifted aquifer in the hanging wall of the oblique-reverse fault beneath the Bam-Baravat escarpment (Fig. 18). This structure traps the subsurface drainage flowing east beneath the fans of the western Bam valley (Figs 2 and 3), and is tapped by the numerous underground tunnels (‘qanats’) that cross the ridge to supply the date growing region of Baravat (Fig. 8). It is therefore no accident that Bam and Baravat were targeted in a ‘bull’s-eye’ fashion by the 2003 earthquake; they owe their existence to the presence of the fault system in the first place. It is the fault system that makes life in the desert possible, though it destroys life when it moves in earthquakes. This association applies equally to the earthquakes at Sepidabeh, Tabas and Ferdows and to numerous other historical earthquakes around the desert rims of Iran, as well as to the spring-lines that are exploited along strike-slip faults. It is explored further in Jackson (2006), and is crucial for understanding the vulnerability of human settlements in these arid regions.

8 CONCLUSIONS

We have tried to use the extraordinary wealth of diverse data from InSAR, seismology, geomorphology and surface observations to produce a coherent picture of the coseismic faulting in the 2003 Bam earthquake. Most of the coseismic rupture occurred on a near-vertical strike-slip fault within and south of the city, with an average slip of $\sim 2.0$ m mostly confined to depths between 2–8 km. Little of this slip reached the Earth’s surface and, more importantly, the same fault system remained unruptured within the seismogenic layer for at least a further 10 km beneath the coseismic slip surface, which may represent a significant seismic hazard for the future. Some slip occurred at depths of 5–7 km with a reverse component, consistent with partial activation at depth of a blind oblique-reverse fault beneath the Bam-Baravat escarpment, though radar data indicate no significant slip on this fault at shallower depths. This too may represent a future seismic hazard. The minor distributed coseismic ruptures north of the city of Bam appear to be unrelated to significant slip at depth, and are probably the result of enhanced ground motions and dynamic effects related to northward rupture propagation. The faulting in the Bam region may represent the early stages in the evolution of a spatially separated, or ‘partitioned’, fault system in which oblique strike-slip and shortening is eventually taken up on subparallel strike-slip and reverse faults. The Bam earthquake was one of the most recent in a series along the western margins of the Lut desert, and there are other unruptured active faults nearby that pose a significant remaining hazard to the region.

ACKNOWLEDGMENTS

This work was supported by an NERC grant to COMET (http://comet.nerc.ac.uk). Part of this research was carried out at...
Figure 17. The Sarvestan fault. (a) LANDSAT ETM image (bands 741-RGB sharpened with panchromatic band 8), showing the segmented fault at Sarvestan. A folded ridge of alluvial deposits (marked ‘push-up ridge’) occurs south of Sarvestan. The dotted yellow line is a stream apparently deflected 500 m in a right-lateral sense. Locations of photos in (c)–(f) are marked. (b) Southern continuation of (a). The fault race cuts through alluvial fans and river terraces, displacing streams (yellow dotted lines) in a right-lateral sense by ∼500 m. Green regions along the fault trace are springs. (c) View NE from the push-up ridge in (a) towards Sarvestan. The low scarp can be seen heading north (white arrows). (d) A 1 m high vertical face in cemented conglomerate, at the fault trace, possibly marking the fault itself. (e) View SE of a right-lateral offset of the river marked by the dotted line in (a). (f) View NW of the ∼15 m high scarp north of Sarvestan.

The Jet Propulsion Laboratory of the California Institute of Technology, under a contract with NASA. We are grateful to INSU-CNRS, the French Embassy, the Geological Survey of Iran and IIEES for support in Iran and France, to BHRC for making available the strong motion record at Bam, and to M. R. Bolourchi for help in producing Fig. 18. We thank the European Space Agency for supplying the copyrighted Envisat ASAR data under projects AOE-621 and AOE-668. We benefited from thoughtful reviews by two anonymous reviewers.
REFERENCES


NCC (National Cartographic Centre of Iran), 2003. Preliminary damage map for Bam, Iran earthquake of 26 December 2003.


© 2006 The Authors, GJI, 166, 1270–1292
Journal compilation © 2006 RAS


Vernant, Ph. et al., 2004. Contemporary crustal deformation and plate kinematics in Middle East constrained by GPS measurements in Iran and northern Oman, Geophys. J. Int., 157, 381–398.


APPENDIX A: STATISTICAL TEST OF MODEL COMPLEXITY

Funning et al. (2005) showed that the fit between modelled and observed InSAR data at Bam is improved when a second fault is added, compared to models that utilize a single fault plane. We present here a test to argue that the improvement is statistically significant.

Using the $F$ test, a standard statistical test used to compare how well two models with different numbers of free parameters fit the same set of data (e.g. Stein & Gordon 1984), we can evaluate the likelihood that the reduction in misfit that occurs when more free parameters are added is greater than would be expected through pure chance. We use the statistic

$$F = \frac{[L_{\text{rms}}(r^2) - L_{\text{rms}}(p^2)]/(p - r)}{L_{\text{rms}}(p)/(N - p)}, \quad (A1)$$

where $p$ and $r$ are the numbers of free parameters in the two models ($p > r$), $L_{\text{rms}}(p)$ and $L_{\text{rms}}(r)$ are the rms misfits of the two models, and $N$ is the number of data points. The distribution of $F$ is controlled by the degrees of freedom $v_1 = (p - r)$ and $v_2 = (N - p)$. The test evaluates the probability that the calculated $F$ value could exceed that computed for a random sample with the same degrees of freedom. Here we test at the 1 per cent level—that is to say that if the value of $F$ we calculate is larger than the tabled value (hereafter $F_{\alpha,\nu}$), the probability of the improvement in misfit being due to chance is 1 per cent. In the following calculations, we use the misfit values obtained by Funning et al. (2005) for a suite of different fault
models. Comparing the simplest case of a two-fault model—the uniform slip case, where \( L_{\text{rms}} = 0.017 \) and \( p = 23 \) (9 parameters describing the main fault, 8 describing the second fault and 6 nuisance parameters), with the single fault uniform slip case (\( L_{\text{rms}} = 0.025, r = 15 \)), where \( N = 4470 \), the computed \( F \) value of 646.3 is vastly greater than the tabled \( F_{0.01} \) value of 2.511 for \( v_1 = 8 \) and \( v_2 = \infty \) (Dixon & Massey 1969), indicating that the improvement in misfit is highly significant. In the case of the variable slip/rake models on a single fault plane, such as those favoured by Wang et al. (2004) and Fialko et al. (2005), application of the \( F \) test depends on the estimation of the number of effective free model parameters for the smoothed model—a number which, due to the smoothing, is likely to be different from the total number of model parameters (the number of fault patches multiplied by the number of slip basis functions for each fault patch), and is not trivial to estimate. However, as the misfit for such a model (\( L_{\text{rms}} = 0.019 \)) is higher than for the two-fault uniform-slip case, and because the number of effective free parameters for a model with several hundred fault patches is likely to be greater than 23, we believe that the two-fault model is again favoured here.