

# Fractal clustering of induced seismicity in The Geysers geothermal area, California

J. R. Henderson, D. J. Barton and G. R. Foulger

Department of Geological Sciences, University of Durham, Durham DH1 3LE, UK. E-mail: g.r.foulger@durham.ac.uk

Accepted 1999 June 2. Received 1999 May 27; in original form 1998 January 21

## SUMMARY

We analyse earthquakes recorded at The Geysers geothermal field in California, an area where industrial activity induces seismicity. The seismicity is characterized by the seismic  $b$ -value and  $D$ , the fractal dimension of earthquake hypocentres measured from sliding windows containing 200 events. We study a group of events strongly clustered around an injection well. Over most of the time period examined we find a positive correlation between  $b$  and  $D$ . However, during the initiation of injection into a new well we find instead a negative correlation. The differences in correlation are statistically significant at the  $1\sigma$  level but only marginally so at the  $2\sigma$  level. These results provide evidence for a transient change in the seismic mechanisms operating, and may be explained by a change from conditions of slow stress loading to rapid loading as a result of the build-up of the rate of water injection into the reservoir.

**Key words:** fractal, geothermal, seismicity.

## INTRODUCTION

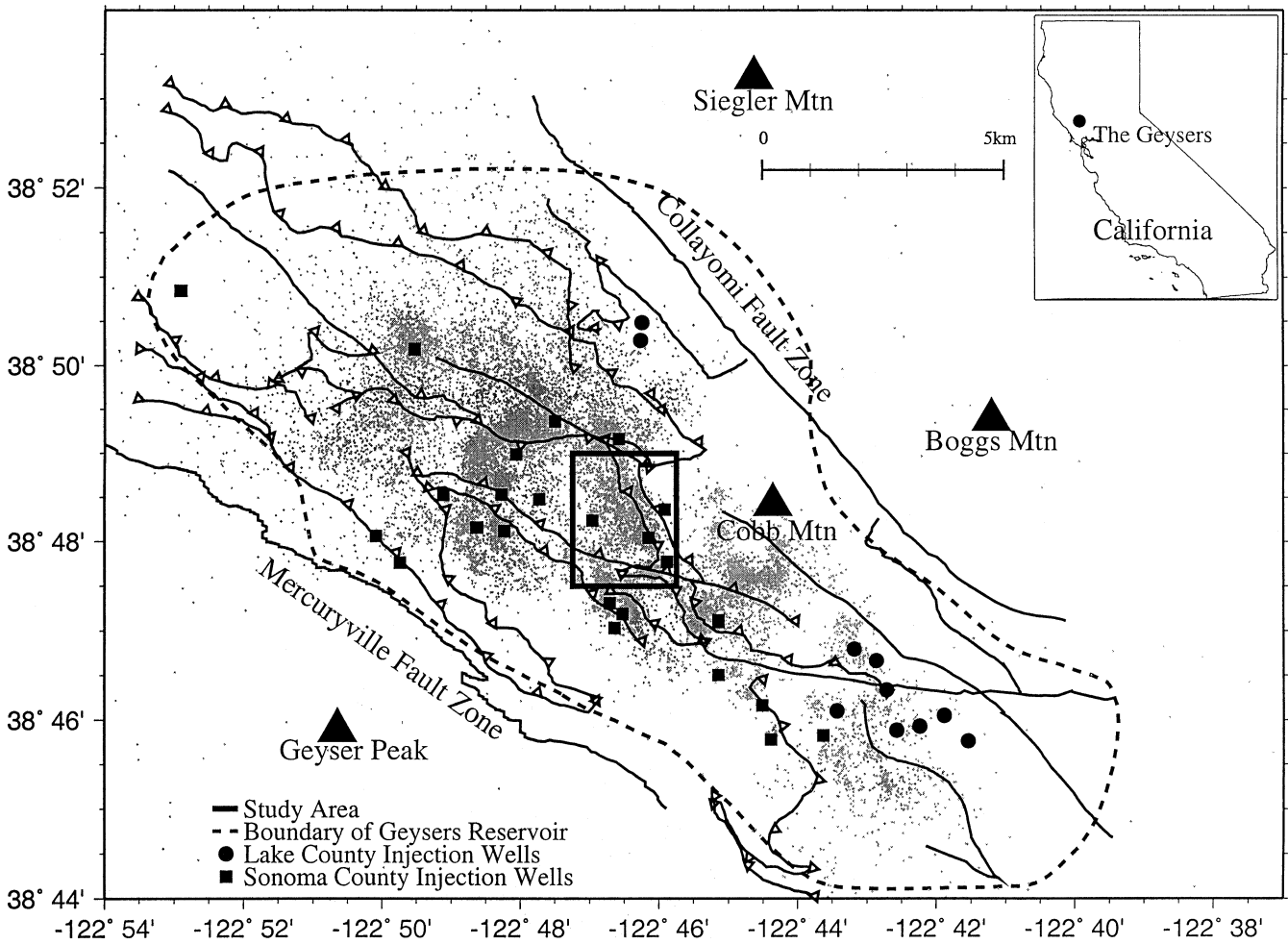
In recent years, fractal concepts have been widely used to characterize various aspects of seismicity. It has been shown that earthquake sizes have a power-law distribution, which is often expressed in terms of the Gutenberg–Richter relation. It has been claimed that these fractal measures, particularly the seismic  $b$ -value and the fractal dimension of hypocentres,  $D$ , vary in a systematic way related to the earthquake process. For example, it has been observed that the  $b$ -value may show a systematic variation in the period preceding a major earthquake (Smith 1981). Similar behaviour has been observed for  $D$  (e.g. DeRubeis *et al.* 1993; Legrand *et al.* 1996), and a correlation between  $b$  and  $D$  has been reported (e.g. Hirata 1989; Henderson *et al.* 1992; Oncel *et al.* 1995). Changes in  $b$  and  $D$  have also been observed during laboratory studies of rock deformation (Meredith *et al.* 1990; Sammonds *et al.* 1992; Lockner *et al.* 1991), and conceptual models have been produced to explain this (e.g. Main 1992; Henderson & Main 1992). However, since it is not clear what processes are operating during the seismic cycle, the origins of these changes in  $b$  and  $D$ , and hence the validity of the models, are not always obvious.

In this study we use a large and relatively high-quality data set to examine the temporal changes in fractal clustering of earthquake hypocentres in a geothermal area where seismicity is induced by industrial activity. In such a situation it is possible to relate changes in the correlation between  $D$  and  $b$  to specific physical processes.

## DATA

The Geysers geothermal area lies within the San Andreas shear zone in northern California (Fig. 1). The steam-dominated geothermal reservoir has a surface area of about 75 km<sup>2</sup>, and extends from about 0.3 km above sea level to at least 3 km below sea level. Commercial development of the The Geysers commenced in the mid 1950s. The UNOCAL Corporation began large-scale development of the area in 1971, reaching a maximum installed capacity of 2043 MW in the mid-1980s.

The area is intensely seismically active, and generates many small earthquakes per day. Despite a dearth of pre-exploitation monitoring, a large majority of the earthquakes are attributed to the commercial extraction of steam (Hamilton & Muffler 1972; Ludwin & Bufe 1980; Majer & McEvilly 1979; Eberhart-Phillips & Oppenheimer 1984) and re-injection of condensate into the reservoir (Stark 1991). This is in agreement with case histories from other areas around the world where earthquake activity associated with fluid disposal and water impoundment suggests that fluid injection can induce earthquakes. It has been shown that injection activity does not always induce earthquakes (Stark 1991). Eberhart-Phillips & Oppenheimer (1984) state that injection under zero wellhead pressure, as practised at The Geysers, is unlikely to create the pore pressures required to cause earthquake activity by the Hubbert & Rubey (1959) mechanism, and that some of the seismic activity may be tectonic. However, later studies suggest that many of the earthquakes studied by Eberhart-Phillips & Oppenheimer (1984) were caused by proprietary injection not known to those authors.



**Figure 1.** Map of The Geysers geothermal area, showing the boundary of the area of geothermal production, principal injection wells, hypocentres of 30 000 automatically located events, and major faults. The box shows the boundaries of the area studied in detail, shown enlarged in Fig. 2. The inset shows the regional location of The Geysers.

Earthquakes at The Geysers are monitored by a permanent seismic network, operated until quite recently by the UNOCAL Corporation, which has comprised between 12 (1989) and 22 (1994) stations. Between 1989 and 1994 approximately 130 000 events were recorded. Of these events, we located 30 000 using *P* arrival times automatically picked from the digital seismograms. Samples of the autopicks were checked by hand and in most cases found to be accurate to 0.01 s. The earthquakes were located using a 3-D velocity model generated from seismic tomography (Julian *et al.* 1996)

Magnitudes for the events were calculated using the duration of the observed seismic signal, defined as the time from the first arrival to the time at which the signal level fell below some threshold determined with reference to the noise level. The relationship between magnitude and duration is (Bath 1981)

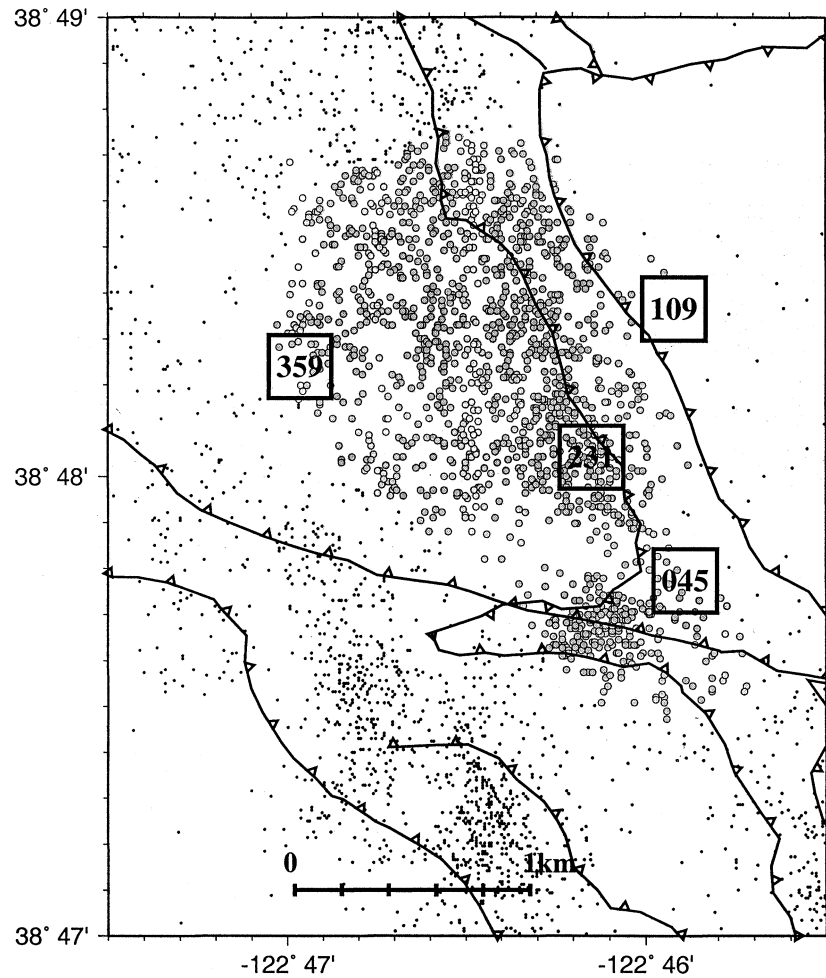
$$m_r = c_1 \log(L) + c_2 \Delta + c_3, \quad (1)$$

where *L* is the signal duration in seconds and  $\Delta$  is the epicentral distance in kilometres. In low-attenuation environments such as The Geysers, the coefficient  $c_2$  is typically negligibly small, and the coefficients  $c_1$  and  $c_3$  were determined by calibrating

the durations against magnitudes determined for some of the events from the NCSN seismograph network, operated by the US Geological Survey.

Geothermal well activity data were collected from various companies operating in the area. These data consist of production histories for steam and injection histories for water. Fig. 1 shows a map of the seismicity and injection-well locations.

There is a correlation between injection-well locations and distinct clusters of seismic activity, suggesting that the two are related. In order to investigate this phenomenon, a cluster of events was selected for more detailed study (Figs 1 and 2). This cluster was used because interactive examination of hypocentre locations in three dimensions revealed it to be clearly distinct from other areas of seismicity, and it contains a large number of events that occurred during a period when the configuration of the UNOCAL seismometer network was stable. Although there are several wells in the vicinity of the cluster, we focus on Well 045 because it is the only well where there was substantial variation in industrial activity, and because it lies adjacent to a large fault cutting across the study area. Well 231, which also lies on the fault, shows less variation in injection.



**Figure 2.** Detailed map showing the hypocentre locations of the events used in this study (large grey dots), other nearby events (small black dots), and injection wells in the area (numbered boxes).

Although it would be desirable to extend the analysis to the entire data set, in practice this is problematic because variations in the seismometer network and operating procedures reduce the homogeneity of the catalogue. Such deficiencies have been noted in other studies of seismicity (e.g. Zuñiga & Wyss 1995), and reduce the reliability of conclusions drawn from such studies. Instead we concentrate on a well-understood subset of the data.

#### METHOD OF ANALYSIS

We estimated the  $b$ -value using the formula of Page (1968):

$$\beta = \left\{ \bar{m} - \frac{m_{\min} - m_{\max} \exp[-\beta(m_{\max} - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \right\}^{-1}, \quad (2)$$

where  $\bar{m}$  is the average magnitude of events exceeding a threshold magnitude,  $m_{\max}$  and  $m_{\min}$  are the maximum and minimum magnitudes used, and  $b = \beta/\log_{10}e$ . The threshold magnitude  $m_{\min} = 0.5$ , was estimated by observing the deviation from linearity of the frequency–magnitude relationship for the area at low magnitudes, and did not vary over the time period of interest. A typical example of a frequency–magnitude plot for a window of 200 events is shown in the upper panel of Fig. 3(a).

The fractal dimension of earthquake hypocentres was estimated using the correlation dimension,  $D_c$  (Grassberger & Procaccia 1983):

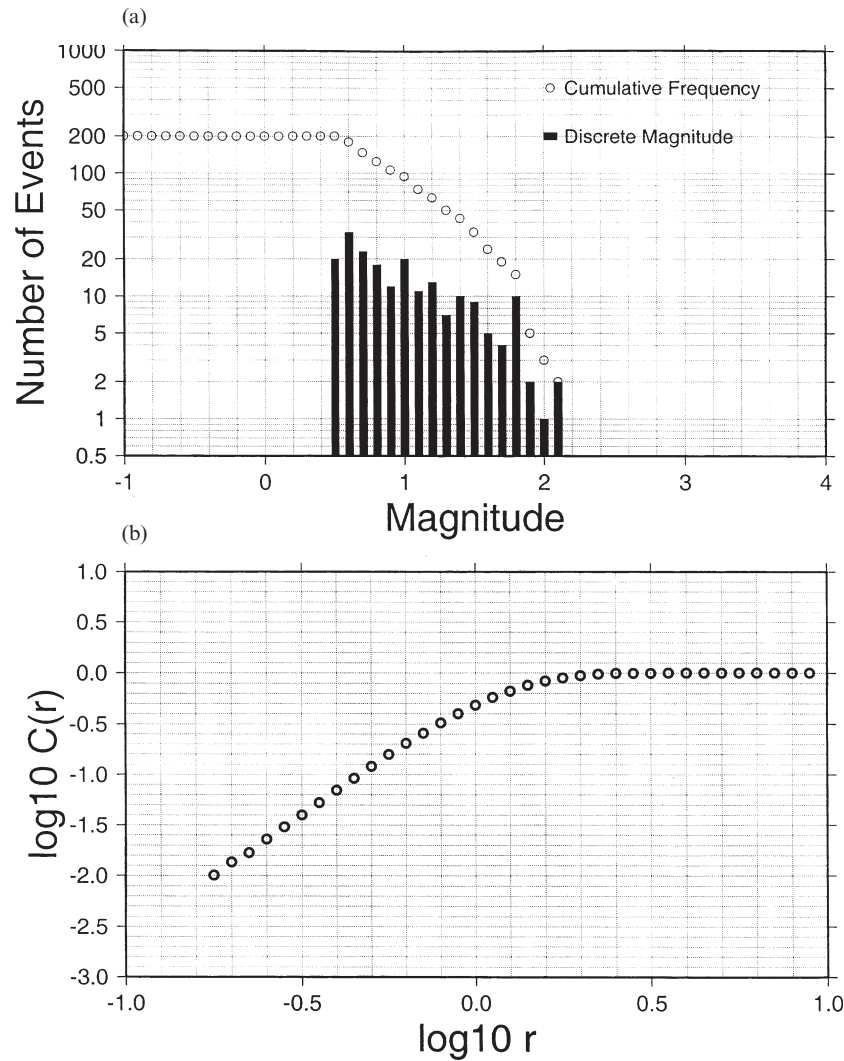
$$D_c = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r}, \quad (3)$$

where  $r$  is the radius of a sphere of investigation, and  $C(r)$  is the correlation integral:

$$C(r) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N H(r - |x_i - x_j|), \quad (4)$$

where  $N$  is the number of points in the analysis window, the  $x$  are the coordinates of the hypocentres, and  $H$  is the Heaviside step function  $H(x) = 0$  for  $x \leq 0$ ,  $H(x) = 1$  for  $x > 0$ . In simpler terms,  $C(r)$  is a function of the probability that two points will be separated by a distance less than  $r$ .

In the case of an infinite fractal distribution, the resulting plot of  $\log C(r)$  against  $\log r$  will be a straight line whose gradient is the fractal dimension. In practice, however, for large values of  $r$  the gradient is artificially low, whereas for small values of  $r$  the gradient is artificially high. These two conditions have been called ‘saturation’ and ‘depopulation’ (Nerenberg & Essex 1990). Whereas it is common for an estimate of the fractal dimension to be made by fitting a



**Figure 3.** Typical examples of (a) a frequency–magnitude plot and (b) a plot of  $\log C(r)$  against  $\log r$  for a 200-event window. The gradient of the frequency–magnitude plot above the threshold magnitude  $m_{\min}$  gives the  $b$ -value, and the gradient of the plot of  $\log C(r)$  against  $\log r$  at values of  $\log r < 0.35$  gives the fractal dimension.

straight line to a subjectively chosen straight part of the curve, Nerenberg & Essex (1990) provide formulae for determining the distances of depopulation and saturation,  $r_n$  and  $r_s$ :

$$r_n = 2R \left( \frac{1}{N} \right)^{1/d}, \quad r_s = \frac{R}{d+1}, \quad (5)$$

where  $d$  is the dimensionality of the data cluster, and  $2R$  is the approximate length of the side of the hypercube containing the data. As discussed by Eneva (1996) it is often safe to start the scaling range at values of  $r$  as low as  $r_n/3$ , but in the case studied here we choose the more conservative approach of measuring a gradient from  $r_n$ . In practice, the values used were  $r_n = 0.08$  km and  $r_s = 0.45$  km. An example of a typical plot of  $\log C(r)$  against  $\log r$  is shown in the lower panel of Fig. 3(b).

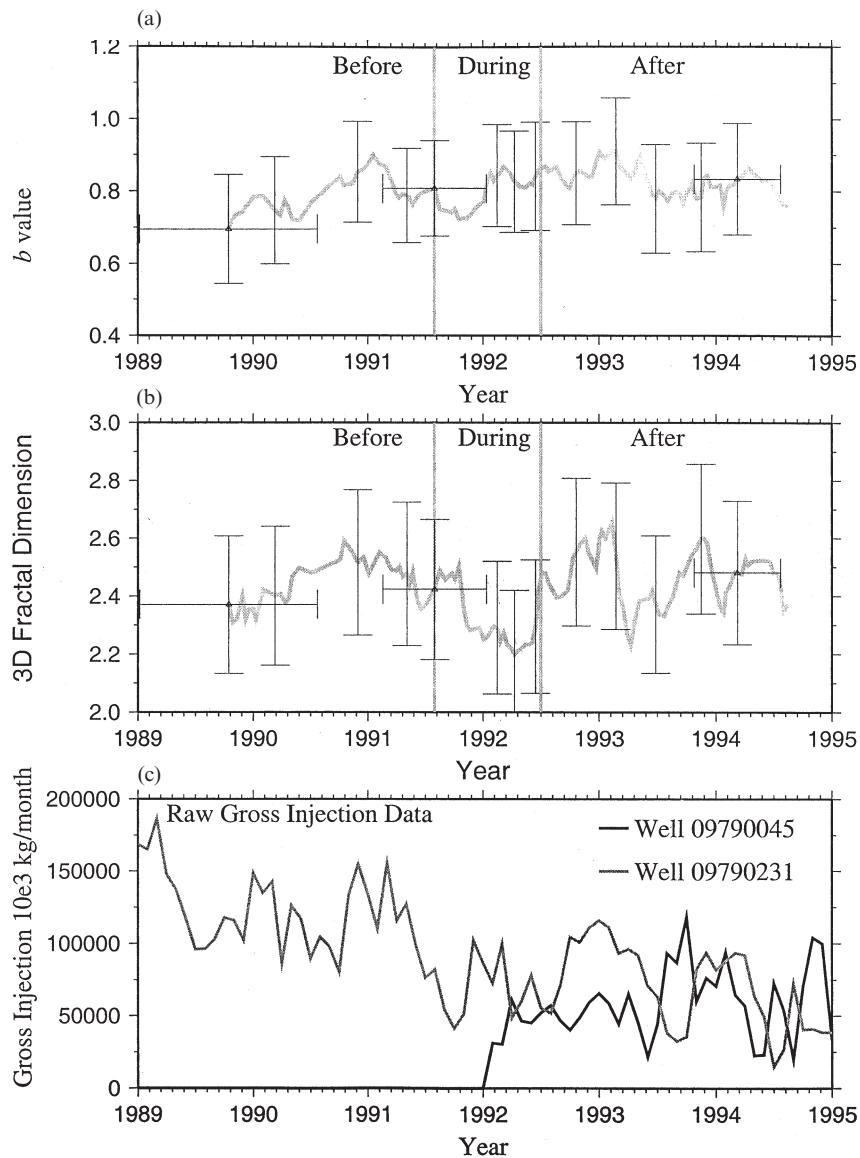
Another area of controversy in the estimation of fractal dimensions concerns the size of the data set used. It has been suggested that very large data sets are necessary for an accurate determination of  $D$  (Smith 1988); however, there is evidence that much smaller data sets are adequate, particularly in situations where it is *changes* in fractal dimension that are of

interest, rather than their absolute values (Nerenberg & Essex 1990). This viewpoint is supported by the findings of Havstad & Ehlers (1989) and Eneva (1996). In this study we use sliding windows of 200 events, overlapping by 10 events, from which to estimate  $b$  and  $D$ .

## RESULTS

Figs 4 (a) and (b) show the values of  $b$  and  $D$  calculated as described above. The error bounds shown for  $b$  are the 95 per cent confidence limits ( $2\sigma$ ) calculated using the method of Page (1968). These limits are more conservative than those found using the maximum likelihood formula  $\delta b = 1.96b/\sqrt{N}$ . The error bounds shown for  $D$  are 10 per cent of the calculated value, and are approximately appropriate for the 95 per cent confidence limits (Havstad & Ehlers 1989). Fig. 4(c) shows the injection activity at wells 231 and 045.

There is very little variation in either  $b$  or  $D$ , relative to their respective error bars (between  $1\sigma$  and  $2\sigma$ ). There is, however, a decline and subsequent revival of  $D$ , marginally



**Figure 4.** Diagrams showing the evolution of  $b$  (a) and  $D$  (b) for the study area during the periods before, during, and after the onset of injection in well 042 in early 1992.  $2\sigma$  error estimates described in the text are indicated by the vertical bars. Horizontal bars show representative window lengths. (c) shows the rate of injection for wells 045 and 231.

significant at the  $2\sigma$  level, over the period late 1991 to early 1993. This corresponds to an increase in seismic activity (Fig. 5), and occurs at the same time as the onset of injection well 045 (Fig. 4c).

Over the whole period studied there is no obvious correlation between any pair of  $D$ ,  $b$  and the absolute level of injection in the area. However, if the data are divided into temporal subsets, systematic behaviour may be discerned. Fig. 6 shows the relationship between  $b$  and  $D$  for three temporal subsets: before the initiation of activity at well 045; during the period over which injection rapidly increased; and subsequently, when injection continued at a roughly steady rate. These periods are defined with reference to the period over which injection rose from zero to over  $5 \times 10^7$  kg per month, in early 1992. The period labelled 'during' refers to all those windows containing a datapoint from this period of rising injection. The change in event rate means that these windows are of various lengths, but

they are typically about six months long. Some representative window lengths are shown in Fig. 4.

For the entire data set there is no significant correlation between  $b$  and  $D$ ; however, for each temporal subset, a weak correlation exists between  $b$  and  $D$ , and the nature of this correlation varies (Fig. 6). Before the initiation of well activity, there is a weak positive correlation between  $b$  and  $D$ . During initiation of injection, a negative correlation exists, and after this period the correlation is again positive. Errors were calculated assuming uncertainties in both  $b$  and  $D$ , and  $1\sigma$  values are shown in Fig. 6. For each subset the slopes calculated are significantly different from zero at the  $1\sigma$  level but not, or only marginally so, at the  $2\sigma$  level. The evidence presented here must thus be viewed as weak. This illustrates the notorious difficulty of detecting statistically significant variations in the fractal dimensions of earthquake activity: this difficulty arises from the very large numbers of high-quality data required to

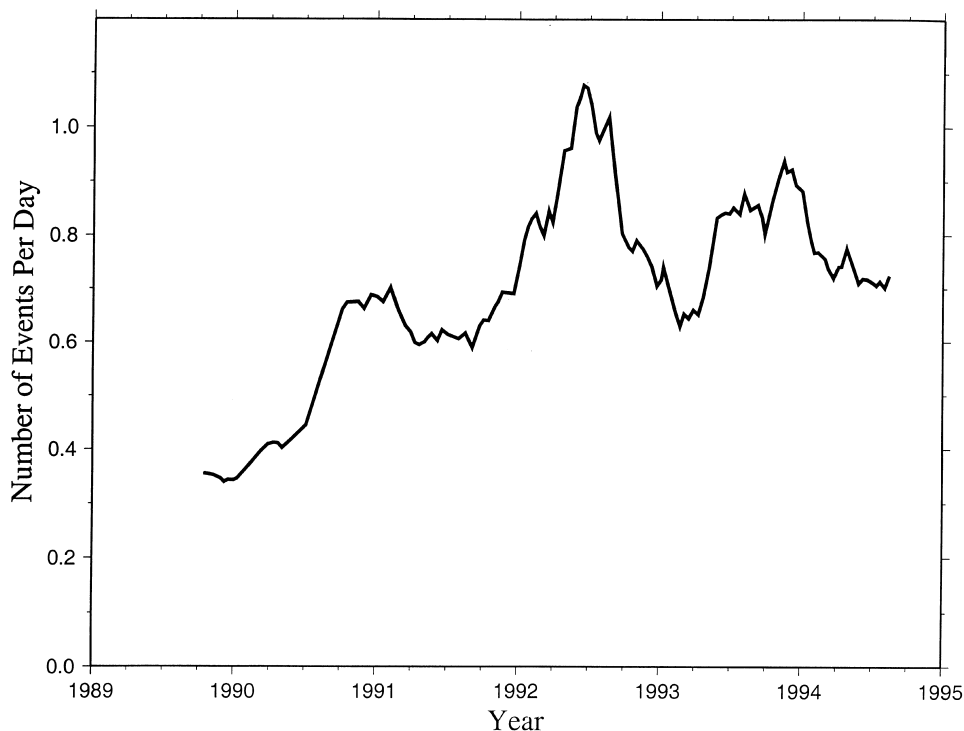


Figure 5. Rate of seismicity in the study area throughout the period of interest.

achieve high precisions and the small numbers that are typically available. Nevertheless, these results suggest that the nature of the seismicity, particularly the spatial clustering and the earthquake generation process, changes with the rate of injection of water into the geothermal reservoir.

## DISCUSSION

Whereas tectonic seismicity often shows significant changes in  $b$ -value associated with changes in event rate and clustering behaviour (e.g. Foulger 1988), the seismicity observed in the present study shows little change in  $b$ -value.  $D$  varies no more than  $b$  relative to the errors, but a weak negative anomaly and a surge of seismicity is associated with the onset of injection into well 045. A positive correlation between  $b$  and  $D$  was observed over the periods of time when well activity was fairly constant in the area, but this correlation became negative when the rate of injection was changing rapidly.

Henderson & Main (1992) presented a model for seismicity in which an initial phase where the proportion of small earthquakes is high (high  $b$ ) occurs in an anti-clustered manner (high  $D$ ). This is followed by low  $b$ -value activity in a strongly clustered geometry (low  $D$ ). This model, which is based on the notion that small, isolated earthquakes relieve local stresses, predicts a positive correlation between  $b$  and  $D$ , and is appropriate for a slowly loaded system. This pattern of behaviour is shown by the present data set during the periods before and after the rapid increase in injection rate. During these periods a preponderance of small earthquakes was associated with a relatively diffuse spatial distribution.

During the period when the injection rate increased rapidly, the opposite was the case. Then, the negative correlation between  $b$  and  $D$  shows that a predominance of small earthquakes (high  $b$ ) was associated with spatial clustering (low  $D$ ).

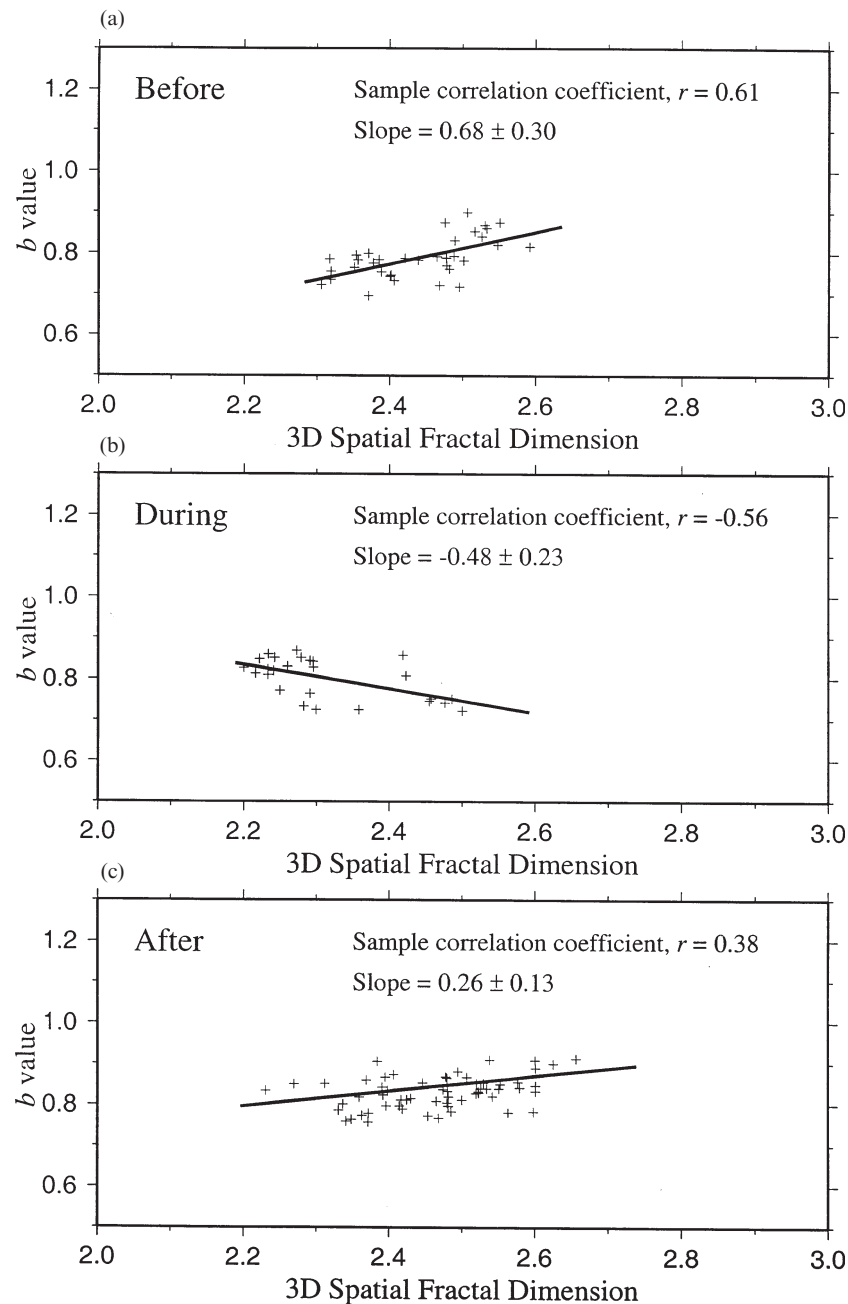
This behaviour occurred under conditions of rapid loading and, although only weakly statistically significant, has been observed in a number of other cases at The Geysers (Barton 1999).

A possible interpretation for these observations is as follows. Preceding and following the initiation of well activity, small fluctuations in pore fluid pressure led to seismic activity which had the effect of locally inhibiting further activity. The mechanism for this may have been dilatant hardening (Scholz *et al.* 1973). The rapid increase in the rate of injection overcame this and triggered numerous earthquakes by a process involving pore pressure diffusion. Such a style of activity is manifest in the spatial clustering of the seismicity ( $\log D$ ). Henderson & Maillot (1997) proposed that, where changes in pore fluid pressure are large compared with fault zone permeabilities, seismicity will be dominated by small events (high  $b$ -value), and that model seems to be applicable in this case.

These results suggest that there are fundamental differences in the style of seismicity associated with steady-state industrial activity at The Geysers and periods of rapid changes. They further suggest a relationship between the driving forces of seismicity and the nature of the correlation between  $b$  and  $D$ .

## ACKNOWLEDGMENTS

The seismic data used in this study were supplied by the IRIS-PASSCAL Data Center, Berkeley, and the well data were supplied by the California Division of Oil, Gas and Geothermal Resources and the Northern California Power Agency. DJB was supported by a NERC Research Studentship. This manuscript was considerably improved by the detailed comments of Ian Main, two anonymous reviewers and the Editor.



**Figure 6.** Graphs showing the correlation between  $b$  and  $D$ , and the associated best-fitting straight lines, for three time periods: (a) before the injection into well 045; (b) during the build-up of injection; and (c) subsequent to the build-up of injection. The sample correlation coefficients, slopes and  $1\sigma$  uncertainties are shown in each panel.

## REFERENCES

- Bath, M., 1981. Earthquake magnitude—Recent research and current trends, *Earth Sci. Rev.*, **17**, 315–398.
- Barton, D.J., 1999. Systematics of the frequency–magnitude distribution and spatial fractal dimension at The Geysers geothermal area and Long Valley caldera, California, *PhD thesis*, University of Durham, pp. xvi + 330.
- DeRubeis, V., Dimitriu, P., Papadimitriou, E. & Tosi, R., 1993. Recurrent patterns in the spatial behaviour of Italian seismicity revealed by the fractal approach, *Geophys. Res. Lett.*, **20**, 1911–1914.
- Eberhart-Phillips, D. & Oppenheimer, D.H., 1984. Induced seismicity in The Geysers geothermal area, California, *J. geophys. Res.*, **89**, 1191–1207.
- Eneva, M., 1996. Effect of limited data sets in evaluating the scaling properties of spatially distributed data: an example from mining-induced seismic activity, *Geophys. J. Int.*, **124**, 773–786.
- Foulger, G., 1988. The Hengill triple junction, SW Iceland: 1. Tectonic structure and the spatial and temporal distribution of local earthquakes, *J. geophys. Res.*, **93**, 13 493–13 506.
- Grassberger, P. & Procaccia, I., 1983. Measuring the strangeness of strange attractors, *Physica*, **9D**, 189–208.
- Hamilton, R.M. & Muffler, L.J.P., 1972. Microearthquakes at The Geysers geothermal area, California, *J. geophys. Res.*, **77**, 2081–2086.
- Havstad, J.W. & Ehlers, C.L., 1989. Attractor dimension of non-stationary dynamical systems from small datasets, *Phys. Rev. A*, **39**, 845–853.

- Henderson, J.R. & Maillot, B., 1997. The influence of fluid flow in the fault zone on patterns of seismicity: a numerical investigation, *J. geophys. Res.*, **102**, 2915–2924.
- Henderson, J.R. & Main, I.G., 1992. A simple fracture-mechanical model for the evolution of seismicity, *Geophys. Res. Lett.*, **19**, 365–368.
- Henderson, J.R., Main, I.G., Meredith, P.G. & Sammonds, P.R., 1992. The evolution of seismicity: observation, experiment and a fracture-mechanical interpretation, *J. struct. Geol.*, **14**, 905–914.
- Hirata, T., 1989. A correlation between the *b*-value and the fractal dimension of earthquakes, *J. geophys. Res.*, **94**, 7507–7514.
- Hubbert, M.K. & Rubey, W.W., 1959. Role of fluid pressure in the mechanics of overthrust faulting, *Geol. Soc. Am. Bull.*, **70**, 115–166.
- Julian, B.R., Ross, A., Foulger, G.R. & Evans, J.R., 1996. Three dimensional imaging of reservoir depletion at The Geysers geothermal area, California, using *V<sub>p</sub>/V<sub>s</sub>* ratios, *Geophys. Res. Lett.*, **23**, 685–688.
- Legrand, D., Cisternas, A. & Dorbath, L., 1996. Multifractal analysis of the 1992 Erzincan aftershock sequence, *Geophys. Res. Lett.*, **23**, 933–936.
- Lockner, D.A., Byerlee, J.D., Kuksenko, V., Ponomarev, A. & Sidorin, A., 1991. Quasi-static fault growth and shear fracture energy in granite, *Nature*, **350**, 39–43.
- Ludwin, R.S. & Bufe, C.G., 1980. Continued seismic monitoring of The Geysers, California geothermal area, *Geophysics*, **44**, 246–269.
- Main, I.G., 1992. Damage mechanics with long-range interactions: correlation between the seismic *b*-value and the two point correlation dimension, *Geophys. J. Int.*, **111**, 531–541.
- Majer, E.L. & McEvilly, T.V., 1979. Seismological investigations at The Geysers geothermal field, *Geophysics*, **44**, 246–268.
- Meredith, R.G., Main, I.G. & Jones, C., 1990. Temporal variations in seismicity during quasi-static and dynamic rock failure, *Tectonophysics*, **175**, 249–268.
- Nerenberg, M.A.H. & Essex, C., 1990. Correlation dimension and systematic geometric effects, *Phys. Rev. A*, **42**, 7065–7074.
- Oncel, A.O., Alptekin, O. & Main, I.G., 1995. Temporal variations of the fractal properties in the western part of the North Anatolian fault zone: possible artifacts due to improvements in station coverage, *Nonlin. Proc. Geophys.*, **2**, 147–157.
- Page, R., 1968. Aftershocks and microaftershocks of the great Alaska earthquake of 1964, *Bull. seism. Soc. Am.*, **58**, 1131–1168.
- Sammonds, R.R., Meredith, R.G. & Main, I.G., 1992. Role of pore fluids in the generation of seismic precursors to shear fracture, *Nature*, **359**, 228–230.
- Scholz, C.H., Sykes, I.R. & Aggarwal, Y.R., 1973. Earthquake prediction: A physical basis, *Science*, **181**, 803–810.
- Smith, L.A., 1981. Intrinsic limits in dimension calculations, *Phys. Lett. A*, **133**, 283–288.
- Smith, W.D., 1981. The *b*-value as an earthquake precursor, *Nature*, **289**, 136–139.
- Stark, M.A., 1991. Microearthquakes—a tool to track injected water in The Geysers reservoir, *Geothermal Resources Council, Monograph on The Geysers Geothermal Field, Special Report No. 17*.
- Zuñiga, F.R. & Wyss, M., 1995. Inadvertent changes in magnitude reported in earthquake catalogues: their evaluation through *b*-value estimates, *Bull. seism. Soc. Am.*, **85**, 1858–1866.