JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 105, NO. B12, PAGES 28,081–28,094, DECEMBER 10, 2000

Observation of diffusion processes in earthquake populations and implications for the predictability of seismicity systems

David Marsan and Christopher J. Bean

Geology Department, University College Dublin, Ireland

Sandy Steacy and John McCloskey

School of Environmental Studies, University of Ulster, Coleraine, Northern Ireland, United Kingdom

Abstract. Scale invariance, either in space or in time, has been shown in many papers to characterize earthquake distributions. Unfortunately, little work has been dedicated to looking at the general space-time scaling invariance of seismicity systems, even though a better understanding of how the two domains (spatial and temporal) link together could help the development of the stochastic dynamical modeling of earthquake populations. In this paper we report the observation of diffusion processes of temporally correlated seismic activity for three different data sets: a mine (Creighton Mine, Canada), the Long Valley Caldera in eastern California, and a 7-year period of recorded seismic activity in southern California. The observed subdiffusion processes are indicative of the general space-time scaling of the system, taking the form of a slow power law growth $R(t) \sim t^H$ of the mean distance $R(t)$ between the main event and the temporally correlated afterevents occuring after a delay t. H is found on average to be small (0.1 for Creighton Mine, 0.22 for the Long Valley Caldera, and 0.22 for the southern California main events with magnitude ≥ 1.5) but fluctuates significantly from one main event to the other: the diffusion is found to be intermittent (non-Gaussian) and multiscaling, and except for the Long Valley Caldera, a systematic correlation between the sizes of the main event and subsequent afterevents and the growth exponent H is observed. While classical viscous relaxation models (e.g., elastic listhosphere-plastic asthenosphere coupling, or fluid flow triggered by sudden changes in pore pressure) have been proposed to characterize this relaxation by homogeneous (i.e., nonintermittent) normal ($H = 0.5$) diffusion processes, the direct implication of the reported results is that seismicity systems, at spatial scales from meters to hundreds of kilometers and small (microearthquakes in a mine) to intermediate magnitudes, relax spatiotemporally in a nonelastic way, revealing the stochastic space-time scale-invariant nature of such systems. Since these diffusion processes correspond to a loss of information with time on the location of the main event, they can be used to investigate the limits of predictability, at all spatial scales, of seismicity systems in terms of the spatiotemporal clustering of temporally correlated earthquakes.

1. Introduction

Earthquakes are intermittent phenomena associated with the stress state of the crust. Interactions and correlations between pairs or sets of earthquakes partly reveal the complex dynamics that temporally drive the stress field in seismically active zones: the understanding of such dynamics can therefore be approached, or at least better constrained, by studying systematic spatiotemporal trends and patterns in earthquake populations. An a priori more satisfying approach would be to infer the temporal behavior of the stress field in its phase space. However, scale invariance and intermittency of seismicity systems, as is very often reported (e.g., *Main* [1996] for review; see also discussion and references below), seem to rule out this alternative and call instead for a stochastic modeling: Since no spatial nor temporal scale can be singled out in the scaling regime, one has to consider an infinite number of degrees of freedom in order to account for the observed variabil-

Paper number 2000JB900232.
0148-0227/00/2000JB900232$09.00

ity. This is particularly well illustrated by the so-called self-organized critical modeling [*Bak et al.,* 1987], for which perturbations at small scale can force the system to algebraically depart from its unperturbed trajectory. To impose an arbitrary cutoff in the scaling range would thus lead to unrealistic dynamical behaviors by ignoring the fact that scales smaller than the cutoff can strongly interfer with the well-resolved larger scales. Also, the stress field cannot be considered as a closed system, interactions with numerous other physical fields (e.g., fluids, possibly chemically active, temperature, and more generally various types of heterogeneities, defects, and discontinuities in the physical properties of the crust) being well recognized. Therefore one cannot hope to achieve a pertinent modeling of the dynamics of both the stress field and associated earthquake populations by assuming a deterministic, Markovian temporal evolution in a well-constrained phase space and must instead resort to stochastic non-Markovian scale-invariant dynamics.

Scale invariance of seismicity systems is clearly revealed by statistical laws: The moment release follows the Gutenberg-Richter law [*Gutenberg and Richter,* 1944], the seismicity rate exhibits fractal clustering (e.g., *Omori*'s [1894] law for aftershock sequences following large main shocks; temporal correlation measures have also been reported to be scaling [*Kagan and Jackson,* 1991], a result implying that the seismic gap hypothesis predicting an anticorrelation at intermediate timescales might not be valid), and hypo/epicenter distributions are also typically found to be fractally clustered [*Hirata and Imoto*, 1991; *Hirabayashi et al.*, 1992; *Lei et al.*, 1993; *Hooge et al.*, 1994; *Robertson and Sammis*, 1995; *Wang and Lee*, 1996]. Consequently, stochastic dynamical modeling must link the spatial and temporal domains in a self-similar framework. Two different types of dynamical scale-invariant processes can be thought of: one in which there exists no interactions between space and time and another for which such interactions cannot be neglected. In the former case, the spatiotemporal correlation $C(r,t)$ between pairs of earthquakes separated in space by a distance r and in time by an interval t is of the form $C(r,t) = f(r)\ g(t)$, implying no temporal dependence on the correlation spatial structure. This was reported by *Shaw* [1993], who analyzed the spatiotemporal correlation following preselected main shocks and compared his results with a model based on deterministic subcritical crack growth dynamics. While no significant broadening of the spatial structure of the aftershocks was observed in time [*Shaw*, 1993, Figure 2], it should be noted that given the slowness of the subdiffusive processes reported below, the analysis conducted by Shaw is perhaps not sensitive enough to detect such a diffusion (in particular, one should take a systematic account of the local "background" seismicity pattern defined as the spatial distribution of the temporally uncorrelated regime and determine the spatial structure for nonoverlapping temporal windows, as explained in section 3).

In the latter case, the two domains cannot be separated anymore, and the spatial correlation structure evolves with time. This behavior has been observed in many instances, and typical migrations of aftershocks away from the main shock have been reported [e.g., *Tajima and Kanamori*, 1985; *Ouchi and Uekawa*, 1986; *Eneva and Pavlis*, 1991]. *Marsan et al.* [1999] studied such migrations in a scale-invariant stochastic framework for a mining-induced system and showed that these dynamics take the form of a power law growth of the temporally correlated spatial distribution of afterevents following any microearthquake in the population, with the mean distance $R(t)$ between temporally correlated earthquakes occuring at a delay t after the main event, and the main event increasing on average as $R(t) \sim t^H$, H being the diffusion or "growth" exponent. This phenomenon can be understood as a nonelastic relaxation of an initially strongly clustered stress concentration, spreading toward large scales with time. Such a relaxation is to be distinguished from the purely temporal fractal clustering of earthquakes, as for example illustrated by Omori's law: the latter corresponds to a dissipation in time of the seismic rate and does not a priori imply a temporal change in the spatial structure; therefore no spatiotemporal relaxation, i.e., stress diffusion, can be methodologically inferred on the basis of purely spatial or temporal clustering analyses (contrary to what is claimed by *Godano et al.* [1997]).

Numerous instances of stress diffusion processes following large earthquakes, particularly in subduction zones, have been documented, starting with *Mogi* [1968]. A coupling between an elastic lithosphere and a fluid asthenosphere is typically proposed as the viscoelastic mechanism responsible for such a diffusion, as initially advocated by *Elsasser* [1969], whose model, since based on an homogeneous fluid relaxation, leads to a normal diffusion (i.e., similar to the heat diffusion), therefore with a growth exponent $H = 0.5$. While such a normal diffusion can be expected for earthquakes big enough to significantly trigger a relaxation from the viscous asthenosphere, the results of *Marsan et al.* [1999], however, show that stress diffusion (1) can be found in much smaller systems (e.g., a mine), (2) can characterize the whole earthquake population regardless of their respective magnitudes, and not only a small subset of large earthquakes, and (3) does not follow a normal behavior, an instance of subdiffusive process with $H = 0.18$ rather than 0.5 being observed.

One immediate question is whether anomalous diffusion processes can be found in other seismic systems, at different scales, and characterized by different local conditions; a further step would then be to analyze what are the physical nonelastic mechanisms at work in these processes. Among possible mechanisms, one can cite crustal fluid flow [*Nur and Booker*, 1972; *Li*

et al., 1987; *Hudnut et al.*, 1989; *Noir et al.*, 1997], as strong spatial variability and discontinuities in the permeability field can cause the diffusion to be anomalous and nonhomogeneous. Also, while localized afterslip on the ruptured fault might be regarded as an unlikely candidate for the triggering of a stress diffusion (since it reproduces the same spatial pattern of static stress changes in time), afterslip due to either creep flow of aseismic crustal structures or viscous upper mantle relaxation as evoked above [e.g., *Nur and Mavko*, 1974; *Savage and Prescott*, 1978], or both [*Li and Rice*, 1987], is expected to cause a diffusion of seismic activity as a consequence of the broadening with time of the strain distribution away from the ruptured fault. The state-and-rate friction model of *Dieterich* [1994] also reproduces a temporal growth of the postseismicity distribution: owing to the nonlinear relation between time to instability and the logarithm of the slip speed, a large stress perturbation, i.e., at short distances from the focus, of a nucleation site distribution will create more postseismic activity but will also tend to reach a dissipative regime quicker than for a small perturbation (i.e., at larger distances from the focus).

Modulations of such causal mechanisms by the scale-invariant cascading process of afterevents generating their own local diffusion of seismic activity, and so on, will then affect the quantitative nature of the relaxation. Also, *Marsan et al.* [1999] proposed that the observed dependence of H on the magnitude of the main event might reflect variations in the fractal geometry (hence a multifractal distribution) of faults, for the mining system studied.

In this paper, we investigate the existence of diffusion mechanisms for three different systems: a mine (Creighton Mine, Ontario), 13 years (1984-1997) of seismicity at and around the Long Valley Caldera (eastern California), and a 7-year period (1983-1990) in southern California. By considering three systems with different spatial extents (~0.5, 20, and 500 km, respectively) and characterized by different geophysical conditions, we hope to help in distinguishing between different mechanisms responsible for the observed diffusions. Further characterizations of the diffusion processes are conducted in order to quantify the degree of intermittency reflected by temporal and spatial variations of the H exponent; multiscaling symmetries are investigated. Also, the correlation between the magnitudes/energy releases associated with both the main event and its afterevents and the growth exponent H is analyzed. This paper extends the initial work of *Marsan et al.* [1999] by proposing a more robust and precise characterization of the diffusion, which is then systematically applied to the mentioned systems. The main finding is that the subdiffusion of temporally correlated seismicity following any earthquake appears to consistently characterize all three systems.

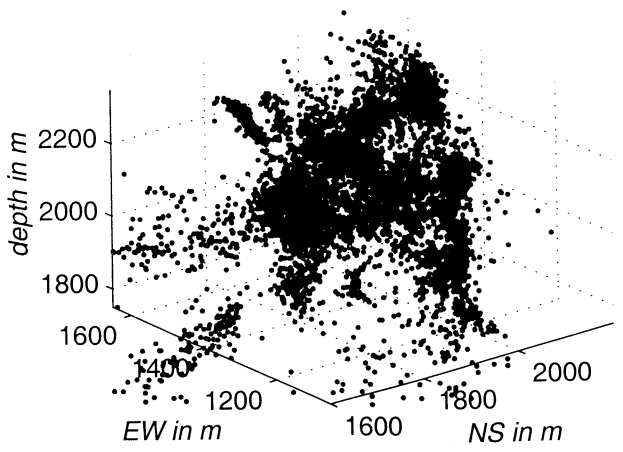

Figure 1. Spatial distribution of the microearthquakes for the Creighton Mine data set.

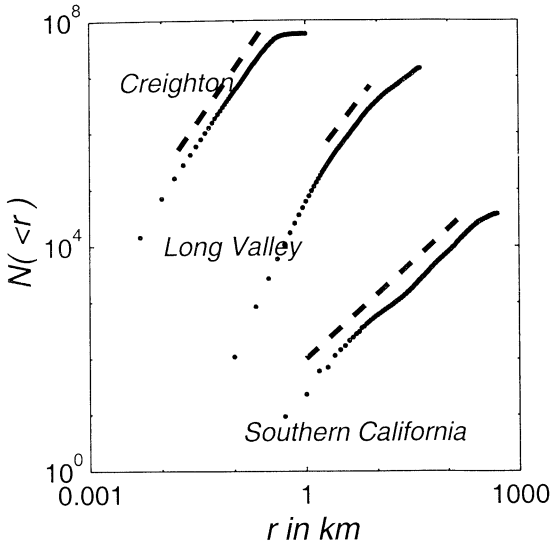

Figure 2. Pair integrals, i.e., number $N(< r)$ of pairs separated by distances shorter than r, for the three data sets. The power law best fits computed on the indicated scaling range (extent of the line) are shown by dashed line and yield fractal dimensions of 1.85 (17-250 m), 1.65 (2-8 km) and 1.15 (1-200 km) for Creighton Mine, the Long Valley Caldera, and southern California, respectively. The curves for the Long Valley Caldera and the southern California data have been arbitrarly shifted in the y direction for clarity.

2. Data

We study three different data sets with different spatial extents. The first set is the mining-induced microearthquake population recorded in Creighton Mine, Ontario, between October 1, 1997, and March 31, 1998, and was already partly analyzed by *Marsan et al.* [1999]. The 10,733 microearthquakes were recorded by the local Queen's Microseismic System (QMS) full waveform data network located in the deeper, active levels of the mine, at depth ranging between 2000 and 2200 m. The

seismically active volume, as recorded by the QMS network, has a size of $\sim 500^3$ m^3, and the estimated average error on the measured earthquake locations is 17 m. Temporal resolution is 1 s. Owing to incomplete magnitude information (cutoffs at small and large magnitudes, leaving 28% of the events with no magnitude estimate), we do not conduct systematic studies of the dependence of the diffusion on magnitude, though in section 3.4 we distinguish the two subsets of events that (1) saturated the QMS system (146 events) or (2) were too small to be given a magnitude (2909 events).

Figure 1 shows the spatial distribution of the microearthquakes in the mining volume. Figure 2 gives the results of the pair integral analysis [*Grassberger and Procaccia*, 1983] applied to this set of earthquakes, showing a scaling regime with fractal dimension $D = 1.85$ developing from ~17 m (error scale) up to ~200 m, at which scale the limited extent of the seismically active volume is felt. Note the break of scaling at ~17 m, with a steeper ($D \simeq 2.5$) scaling regime developing at small scales, which is an artefact of the "randomization" of the locations of the events at scales smaller than the error scale. Locations and times of occurence of the 1504 excavation and production blasts conducted during the studied period are also known.

The second set corresponds to the 36,032 earthquakes occuring between January 1, 1984, and November 18, 1997, at and around the Long Valley Caldera, in the area delimited by longitude -119.1 to -118.52, latitude 37.43 to 37.72, depth smaller than 15 km; see Figure 3. The Long Valley Caldera is a seismically active zone hydrothermally driven. A recent tomographic study by *Trygvasson* [1998] shows the Long Valley Caldera as possessing a complex velocity structure, believed to be strongly influenced by high temperatures, fluids, and hydrothermal alterations, with highly disrupted rocks that have been chemically and thermally altered by migrating fluids in the upper crust [*O'Doherty*, 1999]. The largest earthquake in the data set is the $M6.1$ Round Valley earthquake that occured on November 23, 1984,

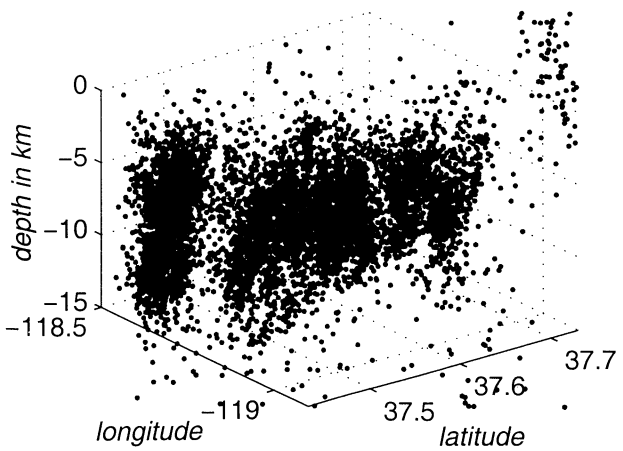

Figure 3. Spatial distribution of earthquakes for the Long Valley Caldera data set.

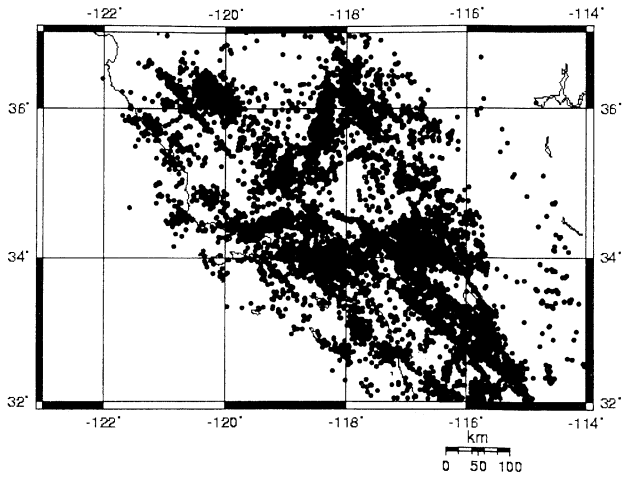

Figure 4. Locations of the earthquakes of the southern California data set.

at ~20 km southeast of the caldera. The magnitude range from 1.2 to 6.1 appears to be recorded without any truncations (no break in the magnitude-frequency distribution, giving a b value close to 1, within this range). As the spatial correlation shows (pair integral of Figure 2), the error scale on the locations of the earthquakes might here be as large as 1 km.

The third set is made of the 37,777 earthquakes of magnitude between 1.5 and 6.6 listed in the Southern California Seismic Network hypocenter catalog for the period extending from August 1, 1983, to March 1, 1990, see Figure 4. Only events of quality A, B, and C are kept in the data set, the typical error on the location being of the order of 1, 2, and 5 km for events of quality A, B, and C, respectively. As the surface extent of the set of earthquakes is much larger than its extension in depth, and in order to simplify the analyses by not having to account for such an anisotropy, only the epicenters were considered (depth was not taken into account). Completness of the distribution is ensured with a cutoff magnitude of 1.5. Several well-known and studied earthquakes are part of this set, among them the $M6.2$ Elmore Ranch and $M6.6$ Superstition Hill earthquakes (November 23, 1987, and November 24, 1987), the $M5.9$ Whittier Narrow earthquake (October 1, 1987), the $M5.6$ North Palm Springs earthquake (July 8, 1987), and the $M5.4$ Oceanside earthquake (July 13, 1987).

3. Diffusion of Temporally Correlated Seismic Activity

3.1. Methodology

We denote by (x_i, t_i) the position and time of occurrence of the ith event in the data set under study, which possesses a total of $\mathcal{N}$ events and covers a period of duration $\mathcal{T}$. We determine the space-time correlation existing between all the events of a data set by first com-

puting the average rate of events $N(r,t)$ occuring in a space-time window of size (L,T) at a distance r away and delay t after any event:

$$N(r,t) = \frac{1}{T\mathcal{N}_t} \sum_{i=1}^{\mathcal{N}_t} \sum_{j=i+1}^{\mathcal{N}} \Theta(t_j - t_i \in [t; t+T]) \cdot \Theta(|x_j - x_i| \in [r; r+L]), \quad (1)$$

where $\Theta(\mathcal{P})$ is 1 if $\mathcal{P}$ is true, 0 otherwise, and $\mathcal{N}_t$ is the largest index i such that $t_{\mathcal{N}} - t_i < t + T$. We call "main event" event i, and its "afterevents" the events $j > i$. $N(r,t)$ corresponds therefore to the stacking of the individual rates computed for all the events taken as main events. In our analyses we keep L constant while T varies, so that they correspond to time intervals separating an algebraically increasing set of times t. We here assume that the systems are (1) stationary, (2) homogeneous (invariant by translation), and (3) isotropic, so that $N(r,t)$ depends only on the relative space-time distances between events. An improvement of the method would consist in considering the focal mechanism of the main event and decomposing the distance $|x_i - x_j|$ in (1) into the two projections $r_{\|}$ and $r_{\perp}$ of $x_i - x_j$ onto the fault and the auxiliary planes.

For seismicity systems the total rate at time t $N(t) = \int dr\ N(r,t)$ is expected to decrease with t increasing, as it merely represents the average rate of afterevents at t. As already evoked in section 1, two cases can be distinguished concerning the dependence of $N(r,t)$ on t, i.e., either (1) $N(r,t)/N(t)$ changes with t, implying an interaction between the spatial and temporal domains or (2) $N(r,t)/N(t)$ does not significantly vary with t, implying a stationary spatial structure of correlation. In both cases, $N(r,t)$ is expected to converge with increasing t to $N(r)$, the spatial structure probed by the pair integral which, since computed for all pairs of events with temporal separations up to the maximum time period $\mathcal{T}$ covered by the data set, practically corresponds to the spatial structure of temporally uncorrelated events. $\bar{N}(r)$ is defined as the rate

$$\bar{N}(r) = \frac{1}{\mathcal{N}\mathcal{T}} \sum_{i}^{\mathcal{N}} \sum_{j}^{\mathcal{N}} \Theta(|x_j - x_i| \in [r; r+L]). \quad (2)$$

We define $C(r,t) = N(r,t) - \bar{N}(r)$. In order to determine the spatial structure of the temporally correlated afterevents, we compute the Green's function $G(r,t)$ corresponding to the temporally conserved version of $C(r,t)$:

$$G(r,t) = \frac{N(r,t) - \bar{N}(r)}{N(t) - \bar{N}} \quad (3)$$

with $\bar{N} = \int dr\ \bar{N}(r)$. $G(r,t)$ is the propagator, or Green's function, of the conserved temporally correlated seismic activity and gives the probability that knowing an earthquake is triggered by the main event after a delay t, it takes place at a distance r away from the main event.

3.2. Qualitative Evidence for Diffusion

We show in Figure 5 the average rates $N(r,t)$ and $\bar{N}(r)$ for the Creighton Mine data set, at two temporal windows $0 < t < 1$ mn and 10 h $< t <$ 11 h, with $L = 5$ m. The spatial structure $N(r,t)$ is clearly seen to converge toward $\bar{N}(r)$ as t increases, the rate of convergence being relatively faster at smaller scales. The temporally correlated structure $C(r,t) = N(r,t) - \bar{N}(r)$ extends significantly up to ~150 m for the first temporal window and up to ~200 m for the second window. Both effects (faster rate of convergence at smaller scales and extension of the temporally correlated spatial structure) are indicative of a diffusion process at work for the temporally correlated seismic activity following any event.

Figure 6 displays the Green's function $G(r,t)$ measuring the temporally correlated spatial structure after removal of the temporal dissipation for the three data sets at three different time windows. In all cases, the spreading of $G(r,t)$ toward large scales as t increases is evident. This diffusion is analogous to a decrease in the degree of clustering of the temporally correlated afterevents around the main event, that is, to a loss of information on the location of the main event as time increases.

3.3. Quantification of the Rate of Diffusion

Marsan et al. [1999] quantified the observed diffusion of the temporally correlated afterevents for Creighton

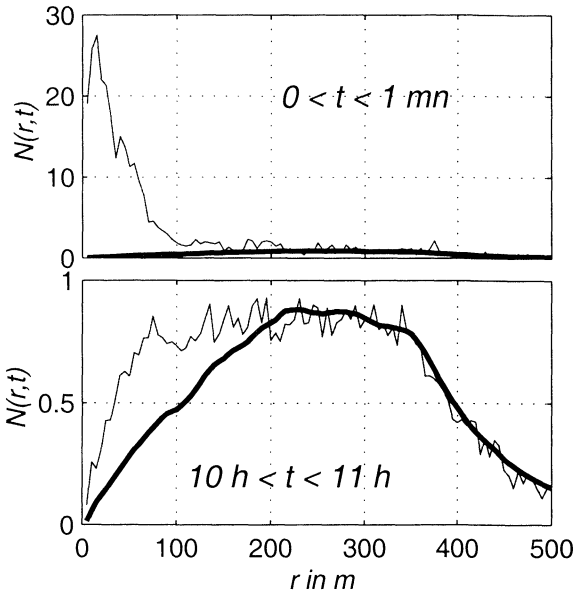

Figure 5. Rates of seismic activity at two different time windows: (top) $0 < t < 1$ mn and (bottom) 10 h $< t <$ 11 h, for the Creighton Mine data. The function $N(r,t)$ is shown along with the temporally uncorrelated $\bar{N}(r)$ (thick line). $N(r,t)$ and $\bar{N}(r)$ are in number of events per day, per 5^3 m^3.

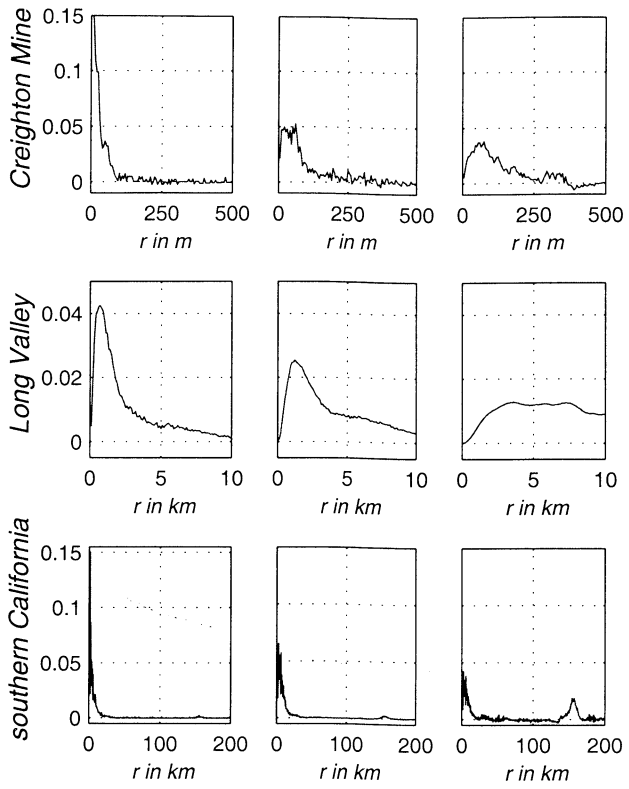

Figure 6. Green's function $G(r,t)$ for the three data sets: (top) Creighton Mine, (center) Long Valley Caldera, and (bottom) southern California. The time windows are from left to right 0 to 10 s, 12 to 16 min, 15 to 20 hours (Creighton Mine); 0 to 15 min, 18 hours to 1 day, 55 to 75 days (Long Valley Caldera); 0 to 5 min, 6 to 8 hours, 18 to 25 days (southern California). The time windows are chosen so that the second and third windows end at 96 times and 7200 times, respectively, the end of the first window.

Mine by defining the mean distance $R(t)$ between such temporally correlated afterevents and the main event, i.e., $R(t) = \int dr\, r\, G(r,t)$. While such a method directly estimates a growth exponent H such that $R(t) \sim t^H$, we will see that it is not unique, as other moments of $G(r,t)$ can also be used (e.g., second-order moment, corresponding to the width of $G(r,t)$).

Figure 7 shows the evolution of $R(t)$ for the three systems, which all exhibit power law growths on 4 to nearly 6 (in the case of the Long Valley Caldera) decades in time. Note that the scaling for the southern California catalog is only approximate at first order, while it is of good quality for the two other systems. The obtained growth exponents H are found to be small (0.1, 0.08, and 0.22), indicating slow, subdiffusive processes. Indeed, for both the Creighton Mine and the Long Valley Caldera sets a logarithmic fit $R(t) \sim \log t$ could also be proposed; we will, however, see later on that the rate of diffusion is significantly increased either by considering subsets of events with large energy release (Creighton Mine and southern California, section 3.4), or by deconvolving $G(r,t)$ with $G(r, t \to 0)$ (Long Valley Caldera, section 4), so that an algebraic rather than a logarithmic growth needs indeed to be considered. At large t the estimate of $R(t)$ becomes subject to large fluctuations, owing to the fact that $G(r,t)$ has a denominator $N(t) - \bar{N}$ tending to zero. Negative values of $G(r,t)$ and of $R(t)$ are then frequently obtained. We therefore stop the computation of $R(t)$ when such fluctuations become evident. Finally, note that the growth exponent $H = 0.1$ found for the Creighton Mine data set is significantly lower that $H = 0.18$ reported by *Marsan et al.* [1999], which was obtained by considering that in the so-called "temporally correlated regime", the mean distance between the main event and all the afterevents is assumed to be close to $R(t)$. While this is certainly true at very short timescales (for which $N(t) \gg \bar{N}$), a faster growth is obtained using this assumption when $\bar{N}$ is not negligible anymore compared to $N(t)$.

We detail in Appendix A the expected behavior of $C(r,t)$ and $G(r,t)$ under the action of a space-time con-

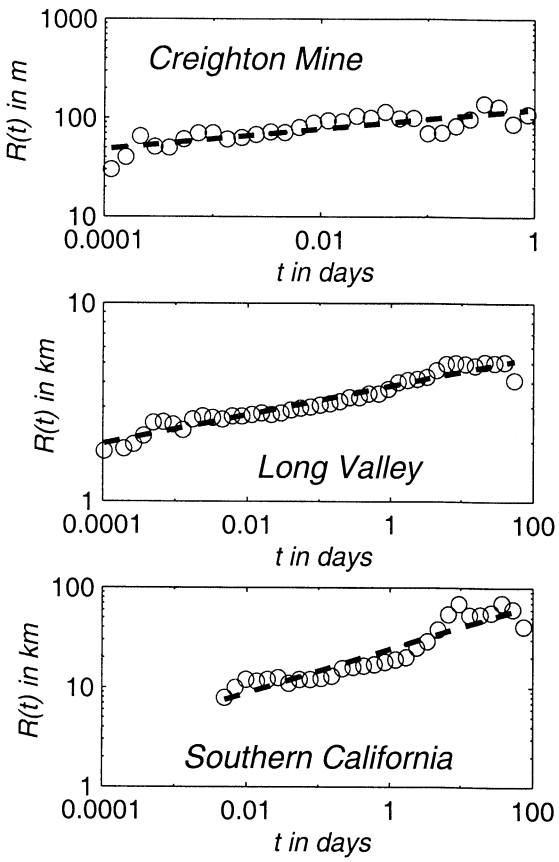

Figure 7. Mean distance $R(t)$ for the three data sets: (top) Creighton Mine, (middle) Long Valley Caldera, and (bottom) southern California. The power law best fits are shown by thick dashed lines, and correspond to growth exponents $H = 0.1$, 0.08, and 0.22, respectively. The Green's function $G(r,t)$ was used to compute $R(t)$, with the sampling scale L equals to 5 m (Creighton Mine), 0.1 km (Long Valley Caldera), 1 km (southern California), and r going up to 1 km (Creighton Mine), 25 km (Long Valley Caldera), and 200 km (southern California).

traction/dilatation operator for a system possessing (1) a fractal temporal clustering analog to a dissipation in $\sim t^{-p}$ of the temporally correlated seismic rate $N(t)-\bar{N}$ and (2) a diffusion given by $R(t) \sim t^H$ for the temporally correlated spatial structure.

Other moments of $G(r,t)$, rather than the mean distance, can be used in order to quantify the diffusion. For example, the second-order moment is associated with the temporal variation of the width of the temporally correlated spatial structure of the afterevents. We can generalize the method by determining any moment of order q:

$$\begin{aligned} R^{(q)}(t) &= \langle r_i^q(t) \rangle = \frac{1}{\mathcal{N}} \sum_i^{\mathcal{N}} r_i^q(t) \\ &= \frac{1}{\mathcal{N}} \sum_i^{\mathcal{N}} \int dr\, r^q \delta(r_i(t) - r) \\ &= \int dr\, r^q G(r,t) \\ &\sim t^{\zeta(q)}, \end{aligned} \tag{4}$$

where $r_i(t)$ corresponds to the individual diffusion originated by the ith main event. This defines the structure function $\zeta(q)$ related to the generalized growth exponent $H(q) = \zeta(q)/q$. An homogeneous diffusion would be characterized by a single value $H(q) = H$, therefore a linear structure function. Heterogeneous diffusions give nonlinear structure functions, indicative of a multifractal behavior, hence a whole hierarchy of growth exponents $H(q)$ and associated orders of singularity $\gamma = -d\zeta(q)/dq$ [*Schertzer and Lovejoy*, 1987]. In the latter case, $G(r,t)$ is locally scaling under the action of the local space-time contraction/dilatation operator characterized by a local exponent H varying in space and time.

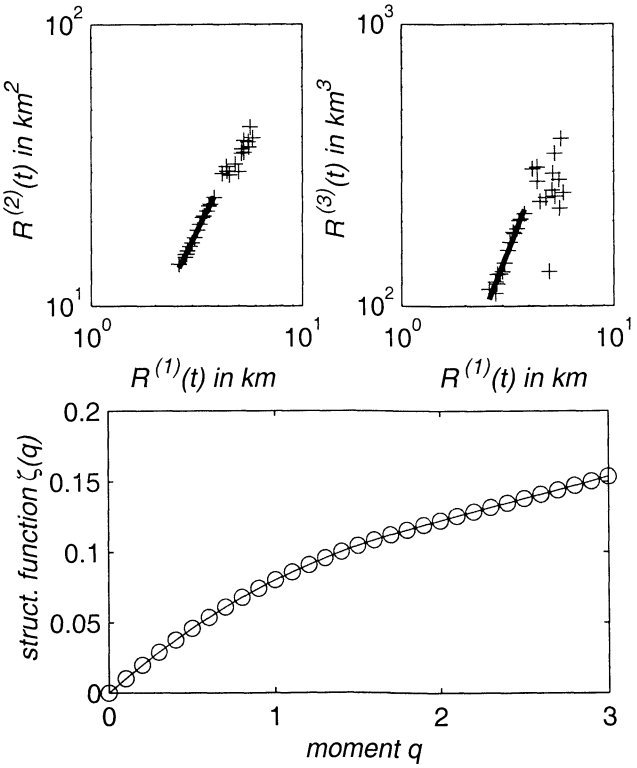

Figure 8. Multiscaling properties of the Long Valley Caldera data. (top) $R^{(2)}(t)$ and $R^{(3)}(t)$ versus $R^{(1)}(t)$ for t between 10^{-2} and 74.6 days. The best power law fits are shown by thick lines, with t ranging from 10^{-2} to 1 day, and give $\zeta(2)/\zeta(1) = 1.54$ and $\zeta(3)/\zeta(1) = 1.97$. (bottom) Structure function $\zeta(q)$ for q varying between 0 and 3.

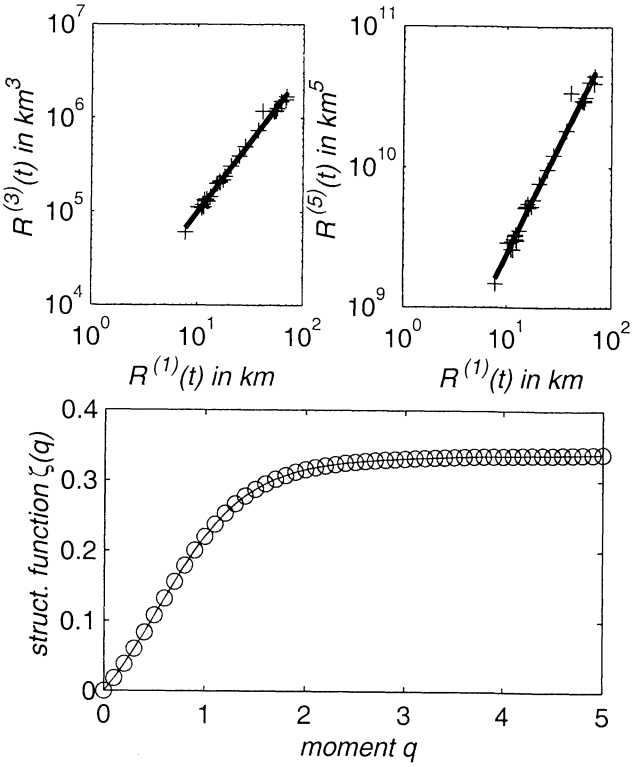

Figure 9. Multiscaling properties of the southern California earthquakes. (top) $R^{(3)}(t)$ and $R^{(5)}(t)$ versus $R^{(1)}(t)$ for t between 5 min and 74.4 days. The best power law fits are shown by thick lines, giving $\zeta(3)/\zeta(1) = 1.53$ and $\zeta(5)/\zeta(1) = 1.56$. (bottom) Structure function $\zeta(q)$ for q varying between 0 and 5.

The structure function $\zeta(q)$ is computed for the Long Valley Caldera (Figure 8) and the southern California set (Figure 9). We made use of the method developed for so-called extended self-similar systems [*Benzi et al.*, 1993], which determine $\zeta(q)$ as $\langle R^{(q)}(t) \rangle / \langle R^{(1)}(t) \rangle = t^{\zeta(q)/\zeta(1)}$, knowing $\zeta(1) = H$ previously determined. The observed nonlinearity of both structure functions reveals the heterogeneous character of the diffusion, which instead of being Gaussian occurs in an intermittent succession of "jumps" followed by calm periods during which only little diffusion happens. A physical picture is one of a seismic activity distributed as a set of space-time clusters; as time increases, new clusters of triggered afterevents are explored by the process, caus-

ing the diffusion to jump, while relatively calm periods develop between two successive jumps. For a hierarchy of space-time clusters interwoven at all scales, as expected for a scale-invariant system, the diffusion takes the form of a multifractional process possessing multiscaling properties.

3.4. Dependence of Diffusion Rate on Energy Release of Main and After Events

We have seen in the preceding sections that for the three studied systems, a main event triggers afterevents which tend to diffuse in an intermittent manner away from the focus of the main event, resulting in an heterogeneous loss of information on this location as time increases. It can be expected that the rate of this loss, measured by the structure function $\zeta(q)$, depends on the relative magnitudes or energy releases associated with the main event and its afterevents, so that bigger afterevents (relative to the size of the main event) will diffuse at a different rate compared to smaller afterevents. In a system with scale invariance in space, time, and energy the variation of $\zeta(q)$ with the energy of main and afterevents should only depend on the difference in magnitudes ΔM, amounting to an invariance of the laws by translation in magnitude.

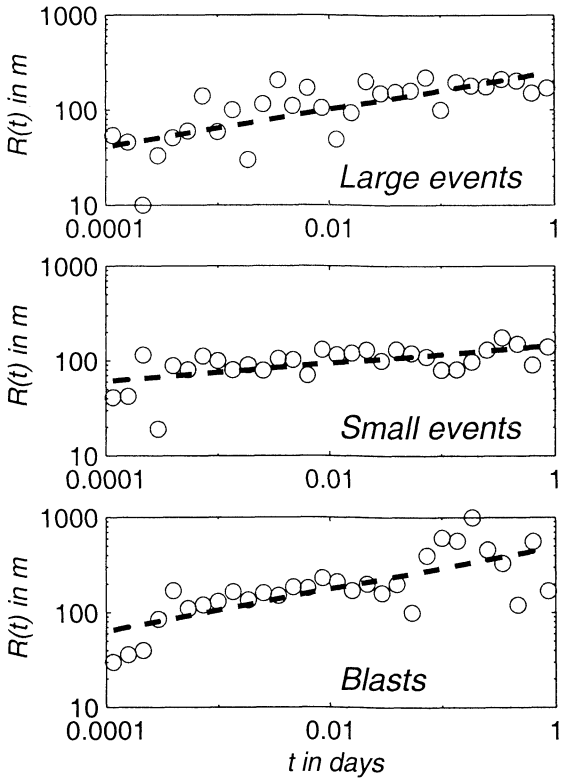

Figure 10. Mean distance $R(t)$ computed for the Creighton Mine data set for three subsets of main events (keeping all the 10,733 events as possible afterevents): (top) the 146 largest microearthquakes, (middle) the 2909 smallest microearthquakes, and (bottom) the 1504 blasts. The best power law fits are shown by thick dashed lines, yielding the growth exponents H=0.2, 0.09, and 0.23, respectively.

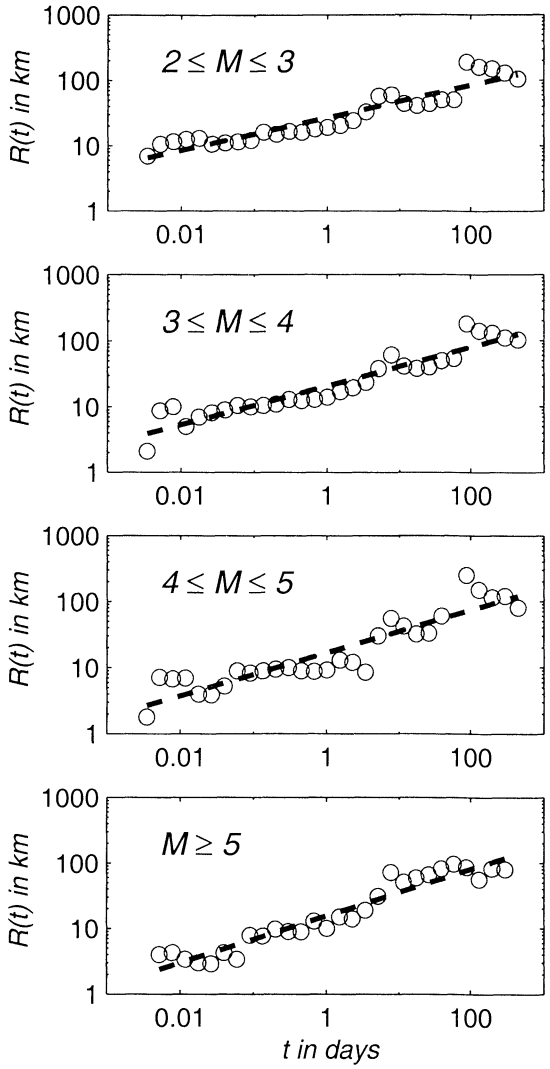

Figure 11. Mean distance $R(t)$ computed for the southern California earthquakes, for four subsets of main events and keeping all the 37,777 events as possible afterevents: (top) earthquakes with magnitude M between 2 and 3 (11,353 events), (second from top) earthquakes with magnitude between 3 and 4 (1137 events), (third from top) earthquakes with magnitude between 4 and 5 (116 events), and (bottom) earthquakes with magnitude greater or equal to 5 (12 events). The best power law fits are shown by thick dashed lines, yielding the growth exponents H=0.25, 0.29, 0.33, and 0.35, respectively.

For any actual data set, cutoffs at small magnitude (and sometimes also at large magnitudes, e.g., the Creighton Mine data set) break this symmetry, and an apparent dependence on the magnitude of the main event is observed. Such is the result given by *Marsan et al.* [1999], which shows that the subset of the 146 largest events in the Creighton Mine set triggered diffusions with a higher rate $H = \zeta(1)$ than observed when considering all the events as main events. We show in Figures 10 and 11 the growth of the mean distance $R(t)$ for various subsets of main events according to their magnitude/energy release for the Creighton Mine and

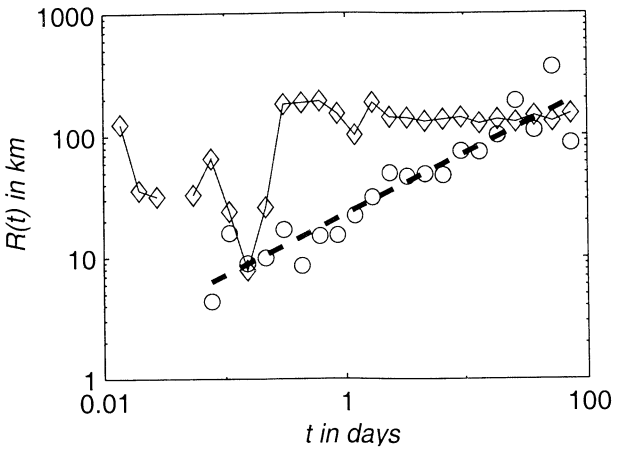

Figure 12. Mean distance $R(t)$ computed for the southern California earthquakes for two subsets of main events and possible afterevents: diamonds, main events with magnitude between 1.5 and 2 and afterevents with magnitude greater or equal to 5; and circles, main events with magnitude greater or equal to 5 and afterevents with magnitude between 1.5 and 2. The best power law fit for the second case (circles) is shown by dashed thick line and yields a growth exponent $H = 0.5$.

California data. For these two systems a clear increase in H with the size of the main event is obtained, while in the case of the Long Valley Caldera, no apparent significant changes in H are observed (not shown here). In the latter case, strong fluctuations in $R(t)$ are observed for magnitudes of main event larger than 4, limiting the observed scaling regime to less than 3 decades in time.

In the case of the southern California earthquakes (Figure 11) we note that at $t \simeq 100$ days the mean distance for all subsets is close to 100 km, which is the mean distance expected for a random distribution with distances ranging between 0 and 200 km. The actual mean distance computed for all pairs of earthquakes distant by 200 km or less is found to be equal to 110.8 km. We can thus interpret this as a complete loss of information, at this time and spatial scale, of the location of the main event, for the considered catalog. Note that it is not totally conclusive at this stage whether this maximum mean distance of ~100 km could be due to the actual maximum mean interaction range of the system for earthquakes of magnitude ranging between 1.5 and 6.6, rather than a size effect. The changes in the growth exponent H from one subset to the other result from changes in the initial clustering of the afterevents around the main event, with a stronger degree of clustering (i.e., a smaller mean distance) as the magnitude of the main event increases. This result is confirmed when determining the growth exponent H for (1) main events with magnitude between 1.5 and 2, and afterevents with magnitude greater or equal to 5 and (2) main events with magnitude greater or equal to 5 and afterevents with magnitude between 1.5 and 2 (Figure 12). In the first case, no diffusion ($H = 0$) is observed for t roughly between 0.1 and 100 days, and the constant mean distance ($\simeq$ 120 km) indicates a close-to-random distribution of the afterevents in space. In the second case, a normal diffusion $H = 0.5$ is obtained for 3 decades in time (0.1 to 100 days) before $R(t)$ reaches values around 100 km. This again can be explained by a weak initial clustering (if any) in the first case, which becomes insignificant at $t > 0.1$ day, implying that if small events can trigger bigger events (more than 3 orders of magnitude bigger), they, however, only weakly constrain their locations, while big events can trigger smaller afterevents and strongly constrain their locations, this constraint being then dissipated at an average rate increasing as one considers smaller and smaller afterevents.

4. Implications of the Observed Diffusion for the Nature of the Dominant Postseismic Mechanisms

The temporally correlated spatial structure of afterevents has been shown to undergo a diffusion process for the three studied data sets, which cover different spatial extents and represent different geological and physical conditions. The processes are found (except in one case) to be subdiffusive ($H < 0.5$), accounting for a rather slow power law expansion of the afterevent area. *Tajima and Kanamori* [1985] reported expansions of aftershock areas following large earthquakes, mostly in subduction zones; even though they did not consider a scaling form for these expansions, a typical value for their "linear expansion ratio" $\eta_e(100) = 1.5$ would imply a H parameter equal to $\log \eta_e(100)/\log 100 \simeq 0.09$. The low values found for the Creighton Mine and the Long Valley Caldera data sets (see Figure 7) can be shown to be due either to the slow diffusion rate of the smaller afterevents (e.g., H significantly increasing in the case of the Creighton Mine when considering only the largest microearthquakes or the blasts as main events, see Figure 10) or to the broad initial distribution of afterevents in the case of the Long Valley Caldera. Indeed, seismicity in the Long Valley Caldera is often found to occur in swarms of earthquakes, that is, in clusters of activity with no "dominant" earthquake triggering them. Even though the b value of our data set is found to be very close to 1, the spatial distribution of temporally correlated afterevents at short timescales (seconds to tens of seconds) is found to be broad, extending to ~4 km after 20 s, as expected for a seismicity system experiencing diffuse swarms of activity. Given the spatial scale covered by the total distribution under study (~25 km), the assumption of an initial point-like source does not hold anymore. Instead, we need to deconvolve Green's function $G(r,t)$ by $G(r, t \to 0)$. We do this by taking $G(r, t \in [16.8\text{s}; 23\text{s}])$ as the "initial" Green's function (source); this time interval is chosen at short enough timescales to correctly model the initial distribution but at t large enough to count a sufficient

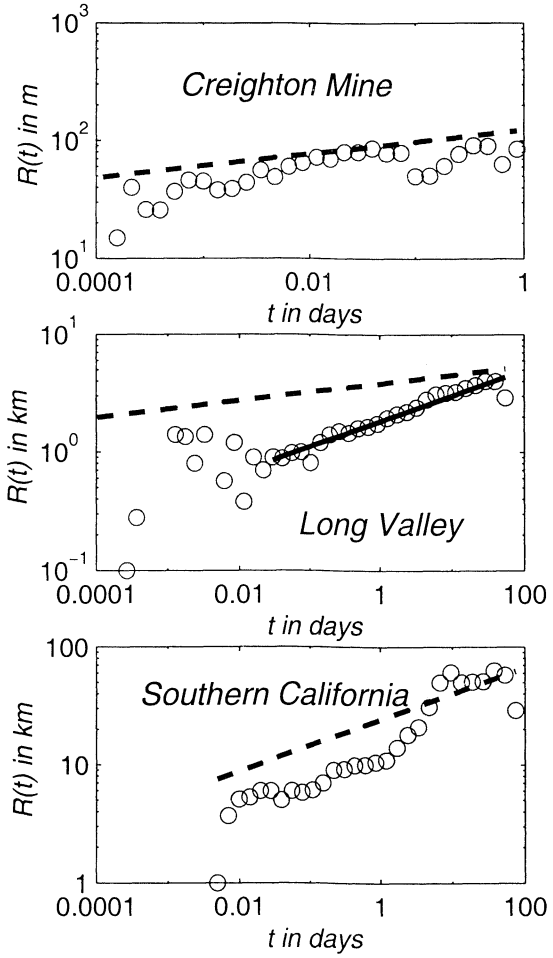

Figure 13. Mean distance $R(t)$ for the three data sets: (top) Creighton Mine, (middle) Long Valley Caldera, and (bottom) southern California after deconvolving the Green's function $G(r,t)$ by an initial source distribution given by $G(r,t_0)$, with t_0 0-10 s, 16.8-23 s, and 0-5 min, respectively. The power law best fit for the Long Valley Caldera is shown by the thick line and corresponds to a growth exponent $H = 0.22$. The thick dashed lines correspond to the best power law fits found for the mean distances obtained without such a deconvolution (see Figure 7).

enough number of events in this distribution. We determine the new distance $R(t)$ for t greater than the tens of seconds scale in Figure 13, along with the mean distances obtained similarly (by deconvolving $G(r,t)$ by an initial distribution taken at short timescale) for the two other data sets. The diffusion process, with a point-like source, characterizing the Long Valley Caldera region is then found to be faster ($H = 0.22$) for timescales between 10^{-2} and 100 days, while no significant change in the growth exponent H is obtained for the two other systems.

Interestingly, the largest growth exponent is found to be equal to 0.5 in the case of the southern California earthquakes when considering large main events (magnitude above 5) triggering small afterevents (magnitude between 1.5 and 2). Such a value is expected for diffusions due to the relaxation associated with viscoelastic structures at depth. In southern California, only earthquakes of magnitude greater than 6 are expected to rupture the whole thickness of the elastic lithosphere [*Scholz*, 1990], implying that smaller earthquakes are unlikely to trigger any significant response from the viscous lower crust. Only two events (the sequence of the Elmore Ranch-Superstition Hill earthquakes) are of magnitude greater than 6 in our data set, hence making it doubtful that the observed diffusion, obtained by averaging over the 12 earthquakes of magnitude greater or equal to 5, can be due to such coupling. Moreover, this mechanism would not explain why H decreases when considering all afterevents (at all magnitudes present in the data set) and still keeping only the 12 earthquakes of magnitude greater or equal to 5 as main events ($H = 0.35$, see Figure 11). $H = 0.5$ is also predicted for a fluid flow mechanism, assuming that (1) the flow occurs in an homogeneous medium and therefore is characterized by a normal diffusion and (2) it can trigger significant seismic activity. Both hypothesis are debatable; fluid flow in complex, heterogeneous media is expected to be subject to anomalous diffusion laws. Also, while numerous works have reported good agreement between fluid flow models and historical sequences of earthquakes [e.g., *Li et al.*, 1987; *Hudnut et al.*, 1989; *Noir et al.*, 1997; *Bosl*, 1999], *Kagan and Jackson* [1998] showed that at the statistical level the friction coefficient in the Coulomb law is close to 0 for southern California earthquakes, implying no significant dependence of the Coulomb stress on fluid pressure changes. Finally, this value $H = 0.5$ reported in the present study cannot be a priori taken as an upper limit, and it can be expected that, for example, by considering even smaller afterevents, i.e., with magnitudes smaller than 1.5, larger growth exponents might be retrieved. While our results do not agree with homogeneous fluid flow models, more complex fluid flow models, in particular involving a whole hierarchy of diffusion exponents H depending on the magnitude of the triggering earthquake, cannot be ruled out on the basis of these observations.

State-and-rate friction laws have been proposed to model the postseismic activity following an earthquake, and more particularly to explain Omori's law [*Dieterich*, 1994]. They also give a temporal growth of the postseismic active zone, as shown in Figure 6 of *Dieterich* [1994]. We therefore investigate whether such models can offer an explanation for the power law growth of the temporally correlated postseismicity zone as reported in this present paper. Using the notations and starting from equation (21) of *Dieterich* [1994], we assume (1) $t \ll t_a$, so that we are in a scaling temporal regime. Longer times t would probe the characteristic relaxation time t_a and hence give a nonscaling behavior for the postseismic activity diffusion; (2) we look at distances up to many times the size of the initial rupture zone c, so that the initial distribution of postseismic activity

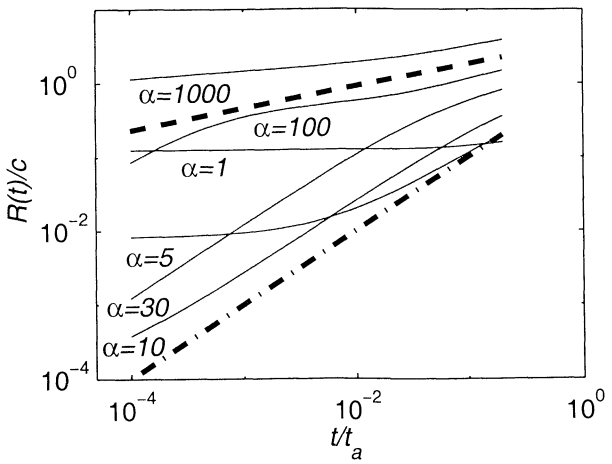

Figure 14. Mean distance $R(t)$ for the postseismic activity derived from the state-and-rate friction model of *Dieterich* [1994], for various $\alpha = -\Delta\tau_e/A\sigma$ values. The two scaling growths $R(t) \sim t$ and $R(t) \sim t^{0.3}$ are indicated by the dash-dotted and the dashed thick lines, respectively.

can be considered as being point-like; (3) no temporal changes in tectonic loading, i.e., $\dot{\tau} = \dot{\tau}_r$ (we also take the reference seismicity rate r equal to 1 and will therefore correct for a unit "background" activity rate when calculating the Green's function); and (4) seismicity can only take place on a fractal set of faults, with a dimension D. Denoting $\alpha = -\Delta\tau_e/A\sigma$ and by $\mathcal{A}(r,t)$ the postseismic activity rate at distance r and delay t from/after the earthquake, we have

$$\mathcal{A}(r,t) = \left\{ \frac{t}{t_a} + (1 - \frac{t}{t_a}) e^{-\alpha[(1-c^3/r^3)^{-1/2}-1]} \right\}^{-1}. \quad (5)$$

For $r \gg r_c = c(\alpha/2)^{1/3}$, $\mathcal{A}(r,t) \simeq 1/[1-(r_c/r)^3]$, and hence no significant temporal changes of the postseismic activity pattern are found at distances greater than r_c; for a typical $\alpha = 40$, $r_c/c \simeq 2.7$, and r_c can be interpreted as the typical distance for a time-invariant spatial pattern of remote triggering. Note also that the two limit cases $\alpha = 0$ and $\alpha \to \infty$ both give no diffusion, since $\mathcal{A}(r,t)$ then becomes independent of r. A maximum growth speed is expected for some value of α. The mean distance $R(t)$ is again computed with the integral $R(t) = \int dr\, r\, G(r,t)$, the Green's function $G(r,t)$ being

$$G(r,t) = \frac{r^{D-1}\,[\mathcal{A}(r,t)-1]}{\int dr'\, r'^{D-1}\,[\mathcal{A}(r',t)-1]}. \quad (6)$$

Figure 14 shows the resulting $R(t)$ obtained for a varying α, with $D = 1$ and r/c ranging from 1 to 100. The maximum growth speed is obtained at around $\alpha = 20$ and yields $R(t) \sim t$. A power law trend in $R(t) \sim t^{0.3}$ is also observed for values of α around 100. The results are found to depend weakly on the dimension D of the fault network. While this model can somewhat reproduce a scaling growth of the mean distance $R(t)$, the fact that these scaling regimes strongly depend on α and also that only the $R(t) \sim t^{0.3}$ trend is anywhere close to the typical growths observed for real systems, implies that this model would need to be modified in order to yield robust anomalous postseismic activity diffusions. In particular, the existence of the typical length r_c does not seem to be directly inferrable from seismicity data, which typically show scale-invariant clustering of earthquakes at this type of scale.

5. Loss of Information and Predictability of Seismicity Systems

The reported diffusions are equivalent to a loss of information on the location of the past main event, as time increases after the occurrence of this main event. Hence information on the future state of the system based on the known locations and times of occurrence of past earthquakes and exploiting two-point statistics becomes less and less spatially constrained as we want to predict further into the future, with a resolution scale growing as t^H on average. Thus one cannot expect to make significant forecasts of the future state of the system at spatial scales smaller than this resolution. This existence of scale-dependent predictability limits for seismicity systems is indeed expected for space-time scale invariant stochastic systems possessing an infinite number of degrees of freedom (see *Marsan et al.* [1996] and *Aurell et al.* [1996] for a similar discussion in the context of turbulence). A direct implication is that in the phase space the departure between the trajectory of the system with that of an initially perturbed version is algebraic, so that the perturbation spreads from the initial error scale ℓ_0 to a given larger-scale ℓ after, on average, a delay $t \sim (\ell/\ell_0)^{1/H}$. The intermittency of the diffusion processes, measured by the nonlinearity of the structure function, accounts for the existence of a hierarchy of space-time clusters of afterevents at all scales; space-time multifractality therefore implies that the actual departure between initially very close systems in the phase space can strongly fluctuate around the average behavior, that is, the perturbation can be much slower or much faster in spreading to scales larger than the initial error scale, and one needs to resort to a cascading, randomly variable hierarchy of Lyapunov exponents, a structure going well beyond classical deterministic chaotic models with only a few degrees of freedom. As a first approximation, the diffusion of temporally correlated postseismic activity can be directly incorporated into probabilistic seismic hazard assessment methods, for example, by providing an average, isotropic growth of the influence zone of a given earthquake. One can thus estimate a seismic hazard map for strong aftershocks following a major shock, for example, by calculating the scalar Coulomb criterion, and have it dynamically evolving through (1) a global dissipation of the activity in $\sim t^{-p}$ and (2) a dilatation of the spatial pattern in $\sim t^H$.

Finally, let us strengthen the analogy existing between seismicity and turbulence, as discussed by *Kagan* [1992]. As evoked above, space-time multiscaling is also present in turbulence [*Marsan et al.*, 1996]; denoting by δv_ℓ the velocity increment at scale ℓ, ϵ_ℓ the flux of energy cascading through ℓ, and τ_ℓ the lifetime of an eddy of size ℓ, one has that $\delta v_\ell \sim \ell/\tau_\ell$, and following *Kolmogorov*'s [1962] relation $\delta v_\ell \sim \epsilon_\ell^{1/3}\ell^{1/3}$, we have that $\ell \sim \tau_\ell^{3/2}\epsilon_\ell^{1/3}$. This is on average equivalent to a diffusion in $R(t) \sim t^{3/2}$, and the actual multiscaling is obtained from the multifractality of the energy flux $\langle \epsilon_\ell^q \rangle \sim \ell^{-K(q)}$, yielding, with our notations of section 3, that $R^{(q)}(t) \sim t^{\zeta(q)}$ with $\zeta(q) = 3q/2 - K(q/2)$. Contrary to seismicity, the diffusion associated with turbulence is superdiffusive; more remarkably, the velocity δv_ℓ increases with ℓ, while, following our results, the velocities implied by seismic diffusion processes decrease rapidly with increasing ℓ, i.e., $v_\ell \sim \ell^{1-1/H} \sim \ell^{-4}$ for a typical $H = 0.2$. The classical phenomenological understanding of turbulence [*Richardson*, 1922] is one of a cascade of structures (eddies) at all scales, with lifetimes scaling on average as $\tau_\ell \sim \ell^{2/3}$. Intermittency of turbulence implies that such lifetimes fluctuate strongly around this average. In the case of seismicity, still taking $H = 0.2$, the equivalent cascade would correspond to a hierarchy of structures (clusters of earthquakes) with associated average lifetimes $\tau_\ell \sim \ell^{1/H} \sim \ell^5$.

6. Conclusions

The dynamical behavior of a seismicity system is controlled by the interactions existing between earthquakes, for example, by advancing the time of occurrence of other, impeding earthquakes. By analyzing the correlations, both in space and in time, within an earthquake population, one can therefore measure how these interactions vary, on average, with the distance and the delay separating two earthquakes. Scaling in space (fractal/multifractal distributions of focii and faults) and in time (Omori's law and the temporal clustering of large magnitude events) implies that these interactions do not "see" any characteristic space or time scale, at least within wide enough scaling ranges.

Here we have examined how these two symmetries, scaling in space and scaling in time, relate to each other. A global space-time scaling symmetry is proposed to characterize the dynamics of seismicity systems and can be empirically probed by measuring the rate of dilatation of the "influence" zone of any earthquake in such systems, as indicated by the diffusion of temporally correlated postseismic activity, away from the earthquake focus. The characteristics of this diffusion process are analyzed; in particular, the diffusion exponent H measuring the mean scaling anisotropy between space and time, its variability which is manifested by the intermittency of the process, and its dependence on ΔM, the magnitude difference between the two earthquakes, are estimated. Three systems are studied, and all exhibit in a very robust way anomalous ($H < 1/2$) average diffusions for the available ΔM ranges. The variability of this process from one system to another is mainly seen by the variations in H; a better understanding of the dependence of the diffusion with the physical characteristics of the system (e.g., tectonic setting, size, rheology, etc.) should help in explaining this variability.

Appendix A: Space-Time Scaling Symmetry of $C(r,t)$ and $G(r,t)$

Here we derive the full space-time scaling symmetry characterizing $C(r,t)$ and the Green's function $G(r,t)$ defined in section 3.1, constrained by the two conditions of (1) fractal temporal clustering $C(r,t) \sim t^{-p}$ and (2) diffusion $R(t) \sim t^H$. $C(r,t)$ and $G(r,t)$ are assumed to be scale-invariant in space and time, that is, there exists a scale-changing operator $T_\lambda(x,t) = (\lambda^a x, \lambda^b t)$ for all $\lambda > 0$ such that $C(T_\lambda(r,t)) = \lambda^\mu C(r,t)$ and $G(T_\lambda(r,t)) = \lambda^\nu G(r,t)$. The two conditions given above can be rewritten as (1) the projection $\int dr^d\ C(r,t)$ scales as t^{-p} and (2) the anisotropy in scale invariance between space and time is given by $a/b = H$. Without any loss of generality, we can take $b = 1$; thus $T_\lambda(x,t) = (\lambda^H x, \lambda t)$. The first condition then leads to

$$\int dr\ C(r,\lambda t) = \int dr'\ \lambda^H\ C(\lambda^H r', \lambda t) \qquad \text{(A1)}$$

with the change of variable $r = \lambda^H r'$, and thus

$$\int dr\ C(r,\lambda t) = \lambda^{H+\mu} \int dr'\ C(r',t); \qquad \text{(A2)}$$

therefore $H + \mu = -p \Rightarrow \mu = -p - H$. The general solution of $C(r,t)$ is of the form

$$C(T_\lambda(r,t)) = \lambda^{-p-H}\ C(r,t) \qquad \text{(A3)}$$

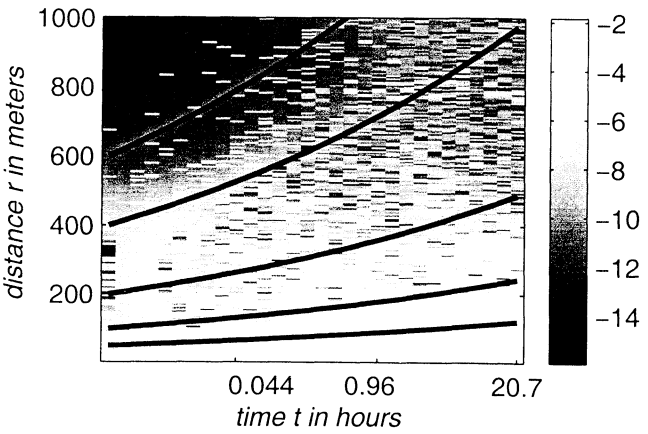

Figure 15. Logarithm of the Green's function $G(r,t)$ for r between 0 and 1 km and t between 0 and 20.7 hours, for the Creighton Mine data, along with an arbitrary set of six $r \sim t^H$ curves invariant under the space-time scale changing operator T_λ, with $H = 0.1$ for this system as computed in section 3.3.

with $T_\lambda(r,t) = (\lambda^H r, \lambda t)$. By definition of the Green's function $G(r,t)$, we have

$$G(T_\lambda(r,t)) = \frac{\lambda^{-p-H}\, C(r,t)}{\int dr'\, C(r',\lambda t)}, \qquad \text{(A4)}$$

and as we saw earlier (equation (A2)), the denominator scales as $\int dr'\, C(r',\lambda t) = \lambda^{-p} \int dr'\, C(r',t)$, so that

$$G(T_\lambda(r,t)) = \lambda^{-H}\, G(r,t). \qquad \text{(A5)}$$

The Green's function therefore only captures the changes in the correlation structure due to the diffusion and, by construction, is not sensitive to the temporal dissipation of seismic activity.

Note that the spatial integral $\mathcal{G}(r,t) = \int_0^r dr'\, G(r',t)$ of $G(r,t)$ up to r is found to scale as

$$\mathcal{G}(T_\lambda(r,t)) = \int_0^{\lambda^H r} dr'\, G(r',\lambda t) \qquad \text{(A6)}$$

$$\mathcal{G}(T_\lambda(r,t)) = \int_0^r dr''\, \lambda^H\, G(\lambda^H r'', \lambda t); \qquad \text{(A7)}$$

hence

$$\mathcal{G}(T_\lambda(r,t)) = \mathcal{G}(r,t) \qquad \text{(A8)}$$

i.e., $\mathcal{G}(r,t)$ is the invariant associated with the scale symmetry T_λ (Noether's theorem). Figure 15 shows $G(r,t)$ for the Creighton Mine data, along with curves invariant under T_λ with $H = 0.1$, as found in section 3.3 and figure 7 for these data.

Acknowledgments. We thank the International Nickel Company (INCO), owner of Creighton Mine, for their help in providing the data for the mine, the Northern California Earthquake Data Center and David Oppenheimer for the Long Valley Caldera data, and the Southern California Earthquake Center for the southern California data. Ian Main, Francesco Mulargia, and an anonymous referee helped in improving the clarity and the quality of this paper. S.S. was funded by INCO; D.M. was funded by INCO and Enterprise Ireland.

References

Aurell, E., G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani, Growth of noninfinitesimal perturbations in turbulence, *Phys. Rev. Lett.*, *77*, 1262, 1996.

Bak, P., C. Tang, and K. Wiesenfeld, Self-organized criticality: An explanation of $1/f$ noise, *Phys. Rev. Lett.*, *59*, 381, 1987.

Benzi, R., S. Ciliberto, C. Baudet, G. Ruiz Chavarria, and R. Tripiccione, Extended self-similarity in the dissipation range of fully developed turbulence, *Europhys. Lett.*, *24*, 275, 1993.

Bosl, W. J., Computational studies of crustal fluids and earthquakes, Ph.D. thesis, Stanford Univ., Stanford, Calif., 1999.

Dieterich, J., A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, *99*, 2601, 1994.

Elsasser, W. M., Convection and stress propagation in the upper mantle, in *The Application of Modern Physics to the Earth and Planetary Interiors*, edited by S. K. Runcorn, p. 223, Wiley-Interscience, New York, 1969.

Eneva, M., and G. L. Pavlis, Spatial distribution of aftershocks and background seismicity in central California, *Pure Appl. Geophys.*, *137*, 35, 1991.

Godano, C., M. L. Alonzo, and G. Vilardo, Multifractal approach to time clustering of earthquakes. Application to Mt. Vesuvio seismicity, *Pure Appl. Geophys.*, *149*, 375, 1997.

Grassberger, P., and I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.*, *50*, 346, 1983.

Gutenberg, B., and C. F. Richter, Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.*, *34*, 185, 1944.

Hirabayashi, T., K. Ito, and T. Yoshii, Multifractal analysis of earthquakes, *Pure Appl. Geophys.*, *138*, 591, 1992.

Hirata, T., and M. Imoto, Multifractal analysis of spatial distribution of microearthquakes in the Kanto region, *Geophys. J. Int.*, *107*, 155, 1991.

Hooge, C., S. Lovejoy, D. Schertzer, S. Pecknold, J.-F. Malouin, and F. Schmitt, Multifractal phase transitions: The origin of self-organized criticality in earthquakes, *Nonlinear Proc. Geophys.*, *1*, 191, 1994.

Hudnut, K. W., L. Seeber, and J. Pacheco, Cross-fault triggering in the November 1987 Superstition Hill earthquake sequence, southern California, *Geophys. Res. Lett.*, *16*, 199, 1989.

Kagan, Y. Y., Seismicity: 'Turbulence' of solids, *Nonlinear Sci. Today*, *2*, 1, 1992.

Kagan, Y. Y., and D. D. Jackson, Long-term earthquake clustering, *Geophys. J. Int.*, *104*, 117, 1991.

Kagan, Y. Y., and D. D. Jackson, Spatial aftershock distribution: Effect of normal stress, *J. Geophys. Res.*, *103*, 24.453, 1998.

Kolmogorov, A. N., A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number, *J. Fluid Mech.*, *13*, 82, 1962.

Lei, X., O. Nishizawa, and K. Kusunose, Band-limited heterogeneous fractal structure of earthquakes and acoustic-emission events, *Geophys. J. Int.*, *115*, 79, 1993.

Li, V. C., and J. R. Rice, Crustal deformation in great California earthquake cycles, *J. Geophys. Res.*, *92*, 11,533, 1987.

Li, V. C., S. H. Seale, and T. Cao, Postseismic stress and pore pressure readjustement and aftershock distributions, *Tectonophysics*, *144*, 37, 1987.

Main, I., Statistical physics, seismogenesis, and seismic hazard, *Rev. Geophys.*, *34*, 433, 1996.

Marsan, D., D. Schertzer, and S. Lovejoy, Causal space-time multifractal processes: Predictability and forecasting of rain fields, *J. Geophys. Res.*, *101*, 26,333, 1996.

Marsan, D., C. J. Bean, S. Steacy, and J. McCloskey, Spatio-temporal analysis of stress diffusion in a mining-induced seismicity systems, *Geophys. Res. Lett.*, *26*, 3697, 1999.

Mogi, K., Sequential occurences of recent great earthquakes, *J. Phys. Earth*, *16*, 30, 1968.

Noir, J., E. Jacques, S. Békri, P. M. Adler, P. Tapponier, and G. C. P. King, Fluid flow triggered migration of events in the 1989 Dobi earthquake sequence of central Afar, *Geophys. Res. Lett.*, *24*, 2335, 1997.

Nur, A., and J. R. Booker, Aftershocks caused by pore fluid flow?, *Science*, *175*, 885, 1972.

Nur, A., and G. Mavko, Postseismic viscoelastic rebound, *Science*, *183*, 204, 1974.

O'Doherty, K. B., Coda wave imaging of the Long Valley Caldera, California, Ph.D. thesis, Univ. College Dublin, Ireland, 1999.

Omori, F., On the aftershocks of earthquakes, *J. Coll. Sci. Imp. U. Tok.*, *7*, 111, 1894.

Ouchi, T., and T. Uekawa, Statistical analysis of the spatial distribution of earthquakes-Variation of the spatial distribution of earthquakes before and after large earthquakes, *Phys. Earth Planet. Inter.*, *44*, 211, 1986.

Richardson, L. F., *Weather Prediction by Numerical Processes*, 1922. (Reprinted, Dover, Mineola, N.Y., 1965).

Robertson, M. C., and C. G. Sammis, Fractal analysis of three-dimensional spatial distributions of earthquakes with a percolation interpretation, *J. Geophys. Res.*, *100*, 609, 1995.

Savage, J. C., and W. H. Prescott, Asthenosphere readjustement and the earthquake cycle, *J. Geophys. Res.*, *83*, 3369, 1978.

Schertzer, D., and S. Lovejoy, Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *J. Geophys. Res.*, *92*, 9693, 1987.

Scholz, C. H., *The Mechanics of Earthquakes and Faulting*, Cambridge Univ. Press, New York, 1990.

Shaw, B. E., Generalized Omori law for aftershocks and foreshocks from a simple dynamics, *Geophys. Res. Lett.*, *20*, 907, 1993.

Tajima, F., and H. Kanamori, Global survey of aftershock area expansion patterns, *Phys. Earth Plan. Inter.*, *40*, 77, 1985.

Tryggvason, A., Seismic tomography; inversion for P and S wave velocities, Ph.D. thesis, Uppsala Univ., Uppsala, Sweden, 1998.

Wang, J.-H., and C.-W. Lee, Multifractal measures of earthquakes in west Taiwan, *Pure Appl. Geophys.*, *146*, 131, 1996.

C. J. Bean and D. Marsan, Department of Geology, University College Dublin, Belfield, Dublin 4, Ireland. (chris.bean@ucd.ie; david.marsan@ucd.ie)

J. McCloskey and S. Steacy, School of Environmental Studies, University of Ulster, Coleraine, Co. Derry, Northern Ireland, BT52 1SA, UK. (j.mccloskey@ulst.ac.uk; s.steacy@ulst.ac.uk)

(Received October 26, 1999; revised June 7, 2000; accepted June 26, 2000.)