Geophys. J. Int. (2003) **154,** 179–195

Seismicity response to stress perturbations, analysed for a world-wide catalogue

David Marsan* and Christopher J. Bean

Geology Department, University College Dublin, Ireland. E-mail: David.Marsan@univ-savoie.fr

Accepted 2003 February 17. Received 2003 February 17; in original form 2000 July 20

SUMMARY

A statistical analysis of the relative distance and time delay separating pairs of earthquakes is performed, to determine how a seismicity system evolves dynamically following a stress perturbation (earthquake). We analyse the space–time correlations of a worldwide seismicity data set (CNSS catalogue, 1963–1998, $M \geq 5$, depth $\leq$70 km), and show that seismic activity diffuses away from an earthquake as the time delay increases following its occurrence. Two regimes are observed: a slow diffusion at short timescales (up to 10 days) during which the mean distance $R(\Delta t)$ between the initial earthquake and the subsequent earthquakes grows as $R(\Delta t) \sim \Delta t^H$, with $H = 0.19$, and a second regime at longer timescales with $H = 0.4$. The growth exponent H in the first regime increases systematically with the size of the initial earthquake, but no such dependence is observed for the second regime. Associating the latter with the viscous diffusion of stress in the upper mantle, we obtain an estimate of about 10^{17} Pa s for the asthenospheric lateral viscosity. This diffusion indicates that the relaxation response of the crust to a stress step depends non-linearly, in respect of both its intensity and general form, on the perturbation. A positive correlation between the regional heat flow and diffusion exponent is found, suggesting a strong thermal control on the diffusion. The overall auto-decorrelation and this diffusion process are the two major mechanisms by which seismicity systems relax stress concentrations, by redistributing them in space and in time. Both processes exhibit typical power-law behaviours, supporting the notion of space–time scale invariance of stress exchanges between seismically active faults.

Key words: diffusion, fault interactions, scale invariance, seismicity.

1 INTRODUCTION

1.1 Seismicity changes in response to a stress perturbation

A remarkable feature of seismicity systems is the ubiquitous nature of the $1/t^p$ response, in the number of triggered earthquakes and at long enough timescales, to a stress perturbation. The modified Omori law (Utsu 1961) is found to characterize aftershock sequences in both laboratory fracturing experiments (Hirata 1987) and for tectonic systems, showing that the local stress step imposed by the occurrence of a seismic event induces non-elastic (that is, involving delays) readjustments of the system. Various models have been shown to be equally able to explain the power-law distribution of time delays (see Shaw 1993, and references therein; Dieterich 1994), extending typically over wide timescale ranges. While these observations and models have helped considerably in our understanding of this universal behaviour of seismically active systems, studies have generally focused on examining only the temporal response, rather than studying the complete dynamics, i.e. how the response changes in both space and time. One major difficulty in empirically estimating such a dynamic response stems from the fact that the spatio-temporal patterns of triggered aftershocks are controlled by (1) the temporal evolution of the stress perturbation, with both a dynamic (seismic waves) and a so-called 'static' regime, the latter being subject to visco and poro-elastic alterations and (2) the dynamic response of the crust to a stress step, which can involve numerous microscale mechanisms (for example, stress corrosion, or accelerating slip on a nucleating fault patch). Another hurdle is also encountered when trying to ensure the causality of any link between the mainshock (stress perturbation) and its aftershocks (seismic response). One is reduced either to studying only a few well-constrained aftershock sequences of strong, rather isolated, mainshocks for which the background seismicity level can be neglected, or to developing somewhat arbitrary clustering methods, implicitly assuming some form of causality. In this paper we present a method based on a space–time correlation analysis, for determining the response (averaged over global, planetary seismic activity) of the lithosphere to stress perturbations, without assuming a direct

*Now at: Laboratoire de Géophysique Interne et Tectonophysique, Université de Savoie, Le Bourget-du-Lac, France.

causal link between a stress step and the subsequent changes in the system.

1.2 Space–time correlations in seismicity systems

Our approach is to determine the degree of interaction between faults at various spatial and temporal scales by computing the space–time correlations of earthquake populations. The relatively frequent occurrence of pairs of earthquakes separated by a distance between r and $r + L$ and a time lag between Δt and $\Delta t + T$, where L and T are the resolution scales, indicates that stress exchanges between any two faults separated by r is on average enhanced at this particular timescale Δt. The overall temporal correlation decays algebraically with time lag Δt, indicating a temporal clustering of seismicity. Also, the spatial correlation pattern evolves significantly with Δt, as was shown in Marsan *et al.* (1999, 2000) for various systems: a relatively greater proportion of earthquakes at Δt is found to be triggered at large scales, corresponding to a diffusion mechanism by which stress is spatially relaxed in the post-seismic regime away from an initial earthquake, this effect being superimposed on the overall temporal dissipation.

In this paper, we study the space–time correlations between earthquakes for a worldwide catalogue (Council of the National Seismic System (CNSS) composite catalogue) spanning a 36 yr interval from 1963 to 1998. The analysis of such correlations offers some insights into the stochastic dynamics of earthquake populations over wide scale ranges both in space, from about 10 to 2000 km (see the Appendix for the upper limit scale) and in time, from minutes to years. More precisely, we define in Section 3 the Green function $G(r, \Delta t)$ as the probability that, given an initial earthquake (the 'main event') at $\mathbf{x}$ and t, and given that a subsequent, correlated earthquake (the 'afterevent') occurs at time $t + \Delta t$, the latter is located at distance r from the main event. We show that $G(r, \Delta t)$ spreads towards large r as Δt increases. G can be interpreted as the propagator of the perturbation of the stress field, as seen through the average seismic activity, triggered by the main event. Thus, by determining G, we probe the dynamic evolution of this initial perturbation. Since G is defined (see Section 3.1) by removing the spatial distribution of the earthquakes temporally uncorrelated with the initial earthquake, or 'main event', it accounts for the spatial structure of the *temporally correlated* subsequent 'afterevents' only. Also, G is merely a correlation measure, so does not carry *a priori* any information concerning the existence of a causal link between the main event and its afterevents: systematic spatio-temporal co-location does not infer causality. Specifically, two events A ('an earthquake occurs at $\mathbf{x}$ and t') and B ('an earthquake occurs at $\mathbf{x} + \Delta\mathbf{x}$ and $t + \Delta t$') may be correlated even though one did not 'trigger' the other, whenever a third event C is causally responsible for both A and B: the two cases $A \rightarrow B$ and $C \rightarrow (A, B)$ cannot be distinguished by the two-point correlation measures used in this paper.

We emphasize that we are not analysing aftershock sequences in the classical sense: our initial ('main event') and subsequent ('afterevent') earthquakes are not selected based on their relative sizes and/or interepicentral distances. While traditional aftershock studies make assumptions concerning what is and what is not an aftershock of a given mainshock, generally according to the relative magnitude of the events, our analysis is devoid of any such, arbitrary, selection rules: apart from Section 3.2.1, where we investigate the effect of the magnitude of the main event on the diffusion exponent, all the earthquakes present in the catalogue are considered as 'main events' and 'afterevents', in turn. This method is therefore based on the computation of a non-conditional space–time correlation product.

1.3 Space–time scale invariance

Scale invariance, either in space (fractal or multifractal clustering of hypocentres and of faults; see Turcotte 1992 and Main 1996 for reviews) or in time (Omori's law expressing the power-law decay of seismic rate following a large earthquake, Omori 1895; temporal fractal clustering of large earthquakes, see Kagan & Jackson 1991) has been reported in many instances of seismicity systems. Space–time scale invariance might therefore be expected to characterize worldwide seismicity and to lead to typical power-law behaviour of the Green function G (fully defined in Section 3). Defining an anisotropic space–time scale changing operator $T_\lambda : (\mathbf{x}, t) \rightarrow (\mathbf{x}', t') = (\lambda^H \mathbf{x}, \lambda t)$, G is expected to scale under the action of T_λ: $T_\lambda G \sim \lambda^a G$, where $T_\lambda G(r, t) = G(T_\lambda(r, t))$, and a is some scaling exponent characteristic of G. The parameter H measures the scaling anisotropy, that is, $H - 1$ is the degree of departure from self-similarity, in respect of space and time. Indeed, this type of behaviour is observed for the three systems studied in Marsan *et al.* (2000): an underground mine (Creighton Mine, Canada), a geothermal area (Long Valley Caldera and surroundings, eastern California) and a transform fault zone (southern California). In the monoscaling approximation, for which a unique T_λ is taken as valid for the whole space–time domain, H is estimated to be 0.1, 0.22 and 0.22, respectively. In two cases (Creighton Mine and southern California) H was also found to be sensitive to the magnitude difference between the main event and the afterevents. Multiscaling is found to be significant for the Long Valley Caldera and the southern California data, that is, nonlinearity of the structure function is unambiguous, implying the existence of a multiplicity of local scale changing operators. The parameter H can be determined by computing the mean distance $R(\Delta t) = \langle r(\Delta t) \rangle$, i.e. the mean of a randomly variable distance $r(\Delta t)$ following the probability distribution $G(r, \Delta t)$, so $R(\Delta t) = \int dr\, r G(r, \Delta t)$. It is straightforward to see, for example using a direct dimensional argument, that this mean distance scales as $R(\Delta t) \sim \Delta t^H$.

1.4 Outline and main results of the paper

We start in Section 2 with an example of the diffusion process along the Philippine Trench, a seismically active subduction zone. This regional process is studied using a method differing from the Green function analysis reported in Marsan *et al.* (2000) and in Section 3. This method, based on the diagonalization of cross-correlation matrices for different time lags, offers a clear illustration of the diffusion. However, it is much more computationally involved than the Green function analysis, described in Section 3 for the worldwide seismicity recorded in the CNSS catalogue starting on 1963 February 6 and finishing on 1998 December 30. Only events with magnitude $M \geq 5$ were considered, ensuring the completness of the total catalogue as no apparent break of scaling is observed at low magnitude in the magnitude–frequency distribution. Also we studied only shallow earthquakes (depths ≤ 70 km). The space–time scaling anisotropy of the system is determined via the scaling of $R(\Delta t)$. The system is found to possess two distinct space–time scaling regimes: one at timescales less than about 10 days, characterized by $H = 0.19$, and another at timescales longer than 10 days

(and up to at least 700 days), for which $H = 0.4$. The growth exponent of the first regime is found to increase with the size of the main event. A correlation is observed between this diffusion exponent and the regional heat flow, indicating the possibility of a significant thermal control on the process.

Tajima & Kanamori (1985), report a somewhat similar analysis for 44 large ($M \geq 7.5$) and five intermediate earthquakes, mostly in subduction zones. A fundamental difference between their work and ours is that they considered the diffusion from the mainshock, for a set of aftershock sequences, while we look at the diffusion of seismicity ('afterevents') away from all earthquakes ('main event') present in the catalogue, as explained in Section 1.2. They computed a coefficient of expansion $\eta(\Delta t)$ of the ruptured area, at four time lags $\Delta t = 1$, 10, 100 d and 1 yr after the main shock, by drawing a contour encompassing all the closest aftershocks for which the cumulated energy release was above a given energy threshold, which decreased with time according to Omori's law. An average η at 100 days of about $\eta(100\,\mathrm{d}) = \frac{S(100\,\mathrm{d})}{S(1\,\mathrm{d})} = 0.15$ can be obtained from their analysis, where $S(\Delta t)$ is the surface of the area ruptured at Δt. This is equivalent to $\frac{R^2(\Delta t = 100\,\mathrm{d})}{R^2(\Delta t = 1\,\mathrm{d})} = 0.15$, hence $R(\Delta t) \sim \Delta t^{0.075}$ for a power-law fit. Such a small H can partly be explained by the fact that the typical maximum mainshock–aftershock distance is of the order of a few hundreds of km (see the Appendix). By considering larger distributions, and, more fundamentally, by not making the traditional distinction between large mainshocks and smaller aftershocks, we retrieve larger growth exponents H, characteristic of an overall more vigorous diffusion regime. Also, such a small value of H implies that a logarithmic law for $\eta(\Delta t)$ fits the data equally well, as done by Tajima and Kanamori. While Tajima and Kanamori properly corrected for the temporal dissipation of correlation with increasing Δt by lowering the energy threshold, they made no attempt to correct for the temporally uncorrelated structure (that is, to account for the background seismicity). For large earthquakes, and considering values of Δt up to a maximum of 1 year, this simplification is probably not unreasonable.

2 A REGIONAL CASE STUDY: DIFFUSION OF SUBDUCTION ZONE SEISMICITY ALONG THE PHILIPPINE TRENCH

To illustrate the diffusion of seismic activity with time following an initial earthquake, we present an analysis performed on a regional subset of the CNSS catalogue. This analysis, based on the diagonalization of correlation matrices at different time lags and elliptical fitting of the eigenvectors, is computationally expensive and not well suited to the study of spatially sparse distributions, but is illustrative and permits the determination of a rotational component if present. A static analysis of the eigenmodes of California earthquake distributions can be found in Rundle *et al.* (1999). We extend this analysis, probing the space–time correlation patterns by considering the correlation for different time lag windows rather than computing the total correlation of the catalogue, that is, after removing all temporal information.

We analyse the diffusion of seismicity in the Philippine Mobile Belt, where the seismic activity is mainly concentrated along the eastwards subduction of the Philippine Sea Plate at the Philippine Trench. Our choice of this region is motivated by the high level of activity experienced in this region, with 3274 $M5+$ shallow earthquakes (8.4 per cent of the total CNSS catalogue for shallow seismicity) occurring between 1963 and 1998 in the latitude N0°–N20°, longitude E118°–E130° zone. We define a 20×12 grid on this region, with 1×1 deg^2 sized cells, and compute the seismic activity time-series $s_i(t)$, $i \in [1; 240]$, for each of these cells. The elements $C_{ij}(\Delta t)$ of the correlation matrix $\mathbf{C}(\Delta t)$ are defined as the covariance of time-series s_i and s_j with time lag Δt:

$$C_{ij}(\Delta t) = \int dt [s_i(t) - \bar{s}_i][s_j(t + \Delta t) - \bar{s}_j], \tag{1}$$

where $\bar{s}_i$ is the mean of $s_i(t)$ for the temporal resolution chosen for the discretization of the activity in time-series (here taken equal to 1 h), *cf.* Fig. 1. Eigenvalues $\lambda_i(\Delta t)$ and associated eigen vectors $\mathbf{P}_i(\Delta t)$ of $\mathbf{C}(\Delta t)$ are computed. For each vector $\mathbf{P}_i(\Delta t)$, we compute its inertial centre

$$\mathbf{X}_i(\Delta t) = \frac{\sum_j |P_{ij}(\Delta t)| \mathbf{x}_j}{\sum_j |P_{ij}(\Delta t)|}, \tag{2}$$

where $\mathbf{x}_j$ is the position of the jth cell and $P_{ij}(\Delta t)$ is the jth component of the ith eigenvector; the elliptical fit of $\mathbf{P}_i(\Delta t)$ can also be determined, yielding two radii $R_i^{(1)}(\Delta t)$ and $R_i^{(2)}(\Delta t)$. A mean radius $R(\Delta t)$ can be retrieved, which measures the average radius of the eigenvectors at time lag Δt, by averaging the square root of the product of the two radii $R_i^{(1)}(\Delta t)$ and $R_i^{(2)}(\Delta t)$ of all the ellipses, weighted by the normalized absolute value of the corresponding eigenvalue:

$$R(\Delta t) = \frac{\sum_i |\lambda_i(\Delta t)| \sqrt{R_i^{(1)}(\Delta t) R_i^{(2)}(\Delta t)}}{\sum_i |\lambda_i(\Delta t)|}. \tag{3}$$

By taking the square root of $R_i^{(1)}(\Delta t) \times R_i^{(2)}(\Delta t)$, we compute the radius of the circle giving the same surface as the elliptical fit. The rationale for using an elliptical rather than a circular fit is to estimate, in addition to a 'mean' radius $R(\Delta t)$, any rotational component that might be present in the diffusion process, and any possible scaling anisotropy in space. The normalization of λ_i by $\sum_i |\lambda_i(\Delta t)|$ in eq. (3) allows us to define $R(\Delta t)$ as a measure of how much, on average, the correlation structure at time lag Δt is spatially spread, regardless of the dissipation of correlated seismic activity with time following the initial earthquake. We linearly decompose $\mathbf{C}(\Delta t)$ in an orthonormal basis because the eigenvectors are, by definition, independent at time lag Δt, so their average geometrical properties are representative of the global clustering properties of the system at Δt.

Fig. 2 shows the elliptical fits obtained for the five dominant eigenvectors (that is, associated with the five largest $|\lambda_i(\Delta t)|$) for three different time lag windows: 0–1 h, 12–16 h and 85 d 4 h–85 d 8 h. These five eigenvectors convey about 31, 50 and 37 per cent, respectively, of the trace of $\mathbf{C}(\Delta t)$ at those time lags. The ellipses are centred on the centre of the eigenvectors. The minimum radius is 55.5 km, which is half the spacing of the grid used to compute the time-series, a distance of 1° along a great circle corresponding to about 111 km. Two major phenomena can be observed in Fig. 2 as Δt increases: (1) the size of the ellipses grows, indicating spatial relaxation of the space–time correlation structure as time increases and (2) they migrate towards the centre of the earthquake distribution. The latter effect is purely a consequence of the first, since as the eigenvectors become spatially wider with increasing Δt, they tend to 'probe' more and more of the total spatial distribution, and their inertial centres therefore converge towards the inertial centre of the total seismicity distribution.

Fig. 3 displays the growth of the average of all ellipses, weighted by the normalized $|\lambda_i(\Delta t)| / \sum_i |\lambda_i(\Delta t)|$, at given Δt, as well as the

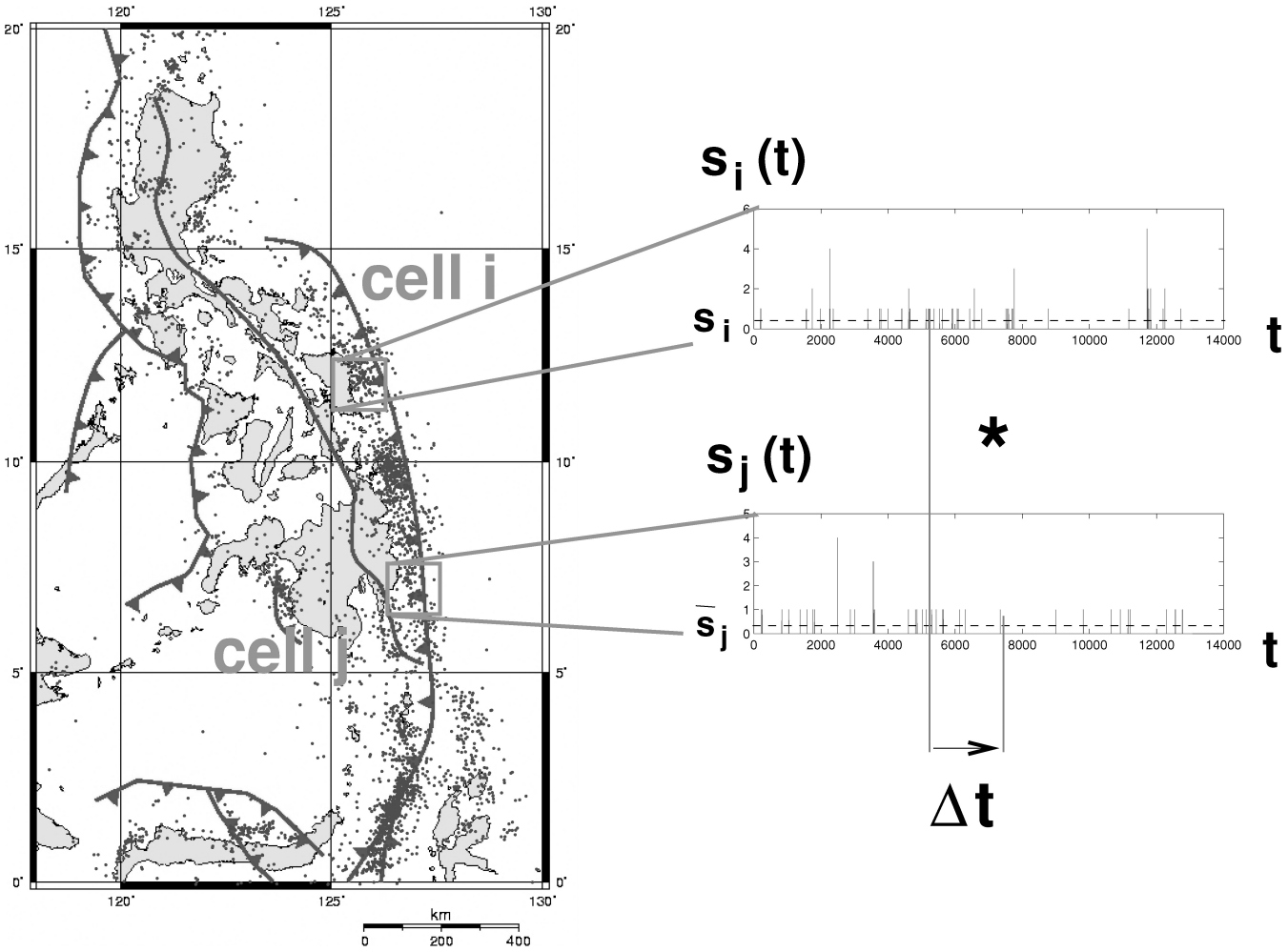

Figure 1. Computation of the cross-correlation matrix component $C_{ij}(\Delta t)$, defined as the correlation between the activity time-series $s_i(t) - \bar{s}_i$ and $s_j(t) - \bar{s}_j$ with time lag Δt, for two of the 240 cells of the grid (cell i and cell j).

associated variations of $R(\Delta t)$ with Δt ranging from 1 h to about 3 months (85 d). A clear growth of $R(\Delta t)$ is observed, following a power-law trend $R(\Delta t) \sim \Delta t^H$ with $H = 0.26$. We compare these results with those obtained using the analysis method based on empirical Green functions as detailed in the next section. A growth in $R(\Delta t) \sim \Delta t^{0.32}$ is obtained over the same range of Δt. There is no apparent break of scaling, at least for this time interval, and the mean ellipses show neither a rotational trend (with the exception of the innermost ellipse, for which the rotation is due to finite resolution effects) nor any significant changes in the aspect ratio, as would be expected in cases of scaling anisotropy in space: the simple self-affine case where $\Delta x \sim \Delta t^{H_x}$ and $\Delta y \sim \Delta t^{H_y}$, $H_x \neq H_y$, would here lead to $R(\Delta t) \sim \sqrt{\Delta x \Delta y} \sim \Delta t^{(H_x+H_y)/2}$, and to a systematic change in aspect ratio $\Delta x/\Delta y$ with Δt.

This example illustrates a case of diffusion of seismic activity away from any initial earthquake in a subduction zone. This diffusion is characterized by a growth exponent H, the value of which is estimated at between 0.26 and 0.32 depending on the method of analysis. These values of growth exponent indicate an anomalous diffusion, that is, slower than the normal (for instance, heat) diffusion for which $H = 0.5$: the initial stress concentration, which led to an earthquake, undergoes a slow, scale-invariant, spatial relaxation as it triggers other earthquakes on more and more distant faults. The mean distance between the focus of the initial earthquake and those of subsequent earthquakes grows from 55.5 km or less at $\Delta t \leq 1$ h to about 300 km at $\Delta t \simeq 3$ months. Along with this spatial relaxation, a temporal dissipation of the initial stress is also at work, leading to a decay of the overall correlation with Δt. These two scale-invariant phenomena are first-order effects controlling the space–time correlation patterns of the seismicity along the Philippine Trench. In the next section we extend the analysis to the whole CNSS catalogue, using a method based on the determination of empirical Green functions that is computationally better suited to the task.

3 GREEN FUNCTION METHOD

3.1 Definitions

To introduce the propagator, or Green function, $G(r, \Delta t)$, we first define the space–time correlation $N(r, \Delta t)$, which measures the average number of pairs of earthquakes at r and Δt away from each other:

$$N(r, \Delta t) = \frac{1}{T\mathcal{N}_{\Delta t}} \sum_{i=1}^{\mathcal{N}_{\Delta t}} \sum_{j=i+1}^{\mathcal{N}} \Theta(t_j - t_i \in [\Delta t; \Delta t + T]) \times \Theta(|\mathbf{x}_j - \mathbf{x}_i| \in [r; r + L]), \quad (4)$$

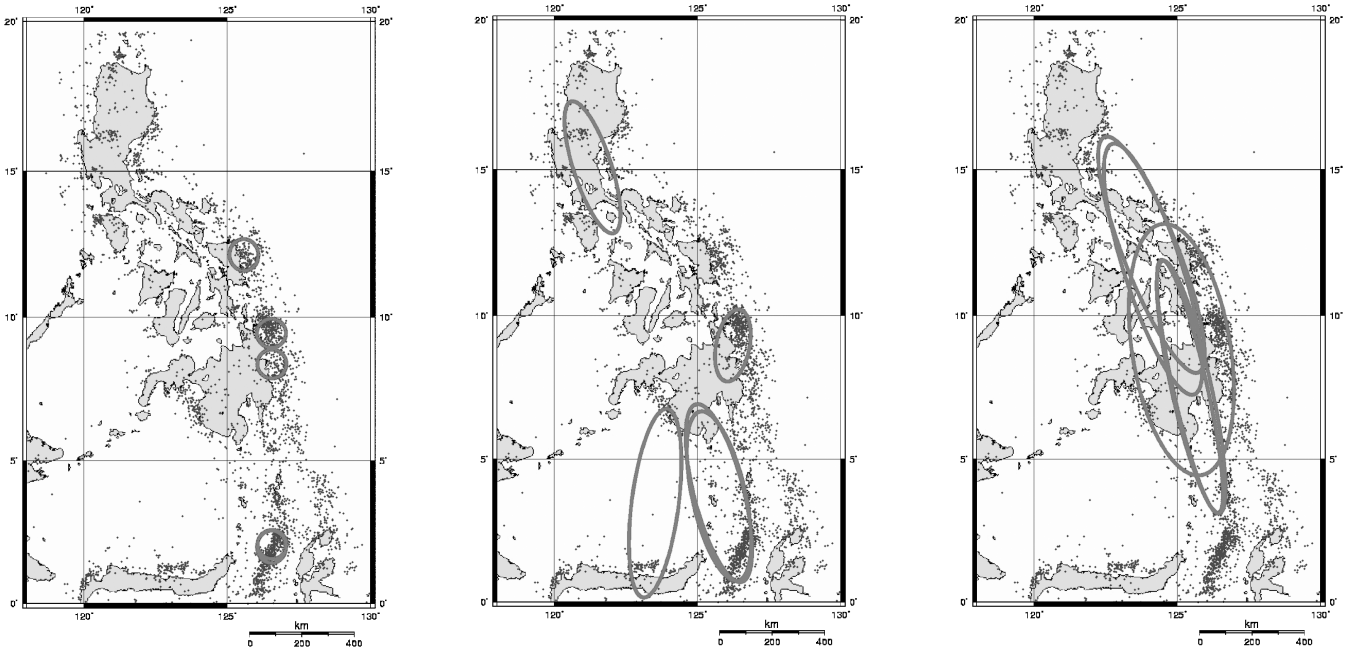

Figure 2. Elliptical fits of the five dominant eigenvectors, for three time windows: left, Δt between 0 and 1 h; centre, Δt between 12 and 16 h; and right, Δt between 85 days and 4 h and 85 days and 8 h. The overall earthquake distribution is shown by the dots. For Δt between 0 and 1 h, the ellipses are seen to be nearly circular with a radius of 55.5 km, indicating that the eigenvectors are spread over an area smaller than the resolution of the 20 × 12 grid used in this analysis.

where:

(1) the data set consists of $\mathcal{N}$ earthquakes, occurring at times $(t_1, t_2, \ldots, t_{\mathcal{N}})$, and locations $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{\mathcal{N}})$;

(2) T and L are the temporal and spatial sizes of the space–time window $\mathcal{W}$ defined for numerical purposes, so $N(r, \Delta t)$ is effectively the average number of events within $\mathcal{W}$ centred at a space–time distance $(r, \Delta t)$; N is a measure (in space) and a density (in time), so is a temporal rate with the dimension of the inverse of time;

(3) $\mathcal{N}_{\Delta t}$ is the largest index i such that $t_{\mathcal{N}} - t_i < \Delta t + T$;

(4) $\Theta(\mathcal{P}) = 1$ if $\mathcal{P}$ is true, 0 otherwise.

Note that we assume spatial isotropy, so $N(r, \Delta t)$ depends spatially solely on the distance r between pairs of earthquakes. Only a significant spatial anisotropy in scaling would require this assumption to be dropped. The analysis of Section 2 showed, at least in the particular case of the seismicity at the Philippine Mobile Belt, the absence of a rotational trend and of a systematic change in aspect ratio. Such a rotation would be expected for a system possessing a significant spatial anistropy in scaling. The observations of Section 2 therefore support this hypothesis of isotropy.

For a temporally uncorrelated distribution of earthquakes, $N(r, \Delta t)$ does not depend on Δt, and therefore $N(r, \Delta t) = \bar{N}(r)$, where

$$\bar{N}(r) = \frac{1}{\mathcal{N}\mathcal{T}} \sum_i^{\mathcal{N}} \sum_j^{\mathcal{N}} \Theta(|\mathbf{x}_j - \mathbf{x}_i| \in [r; r + L]) \tag{5}$$

with $\mathcal{T} = t_{\mathcal{N}} - t_1$ being the time interval spanned by the catalogue. The cumulative sum of $\bar{N}(r)$ is similar to the well-known pair integral (Grassberger & Procaccia 1983). For distributions possessing some temporal correlation, it is expected that $N(r, \Delta t) > \bar{N}(r)$ and $N(r, \Delta t)$ relaxes towards $\bar{N}(r)$ as $\Delta t \to \infty$. We define $G(r, \Delta t)$ as the normalized difference $N(r, \Delta t) - \bar{N}(r)$:

$$G(r, \Delta t) = \frac{N(r, \Delta t) - \bar{N}(r)}{\int dr' [N(r', \Delta t) - \bar{N}(r')]}. \tag{6}$$

Fig. 4 illustrates the method. By taking the difference $N(r, \Delta t) - \bar{N}(r)$, we remove the contribution from the 'background', where 'background' seismicity here refers to a distribution, not a set of well-identified earthquakes. We therefore analyse the temporally correlated structure: an observed relaxation of $G(r, \Delta t)$ at increasing Δt towards large r is not due to a trivial relaxation from the initial clustering to the wider average background seismicity pattern, since we correct for the latter by subtracting $\bar{N}(r)$ from $N(r, \Delta t)$. The denominator $\bar{N}(\Delta t) = \int dr' [N(r', \Delta t) - \bar{N}(r')]$ in eq. (6) gives the temporal correlation with lag Δt. This normalization of $N(r, \Delta t) - \bar{N}(r)$ by $\bar{N}(\Delta t)$ ensures that $\int dr G(r, \Delta t) = 1$, $\forall \Delta t$, so temporal sections of $G(r, \Delta t)$ are probability densities. The numerical definition of $N(r, \Delta t)$ is such that $N(r, \Delta t)$, and therefore $G(r, \Delta t)$, are effectively measures rather than densities in the spatial domain, and $\int dr G(r, \Delta t) = 1$ should of course read $\sum_i G(r_i, \Delta t) = 1$ for r_i varying with an increment L. $G(r, \Delta t)$ is thus the probability that, knowing that an earthquake occurred at a given time and location, a temporally correlated subsequent earthquake occurring at a delay Δt after this initial earthquake will take place at a distance r away from it. Both stationarity and homogeneity of the system are implicitly assumed, so we only need to determine the propagator for relative distances r and time differences Δt. Relatively severe numerical problems can be encountered due to the limited number of data and are discussed in the Appendix.

3.2 Results

As mentioned in Section 1.3, the space–time scaling of the system manifests itself in the behaviour of $G(r, \Delta t)$ under the action of

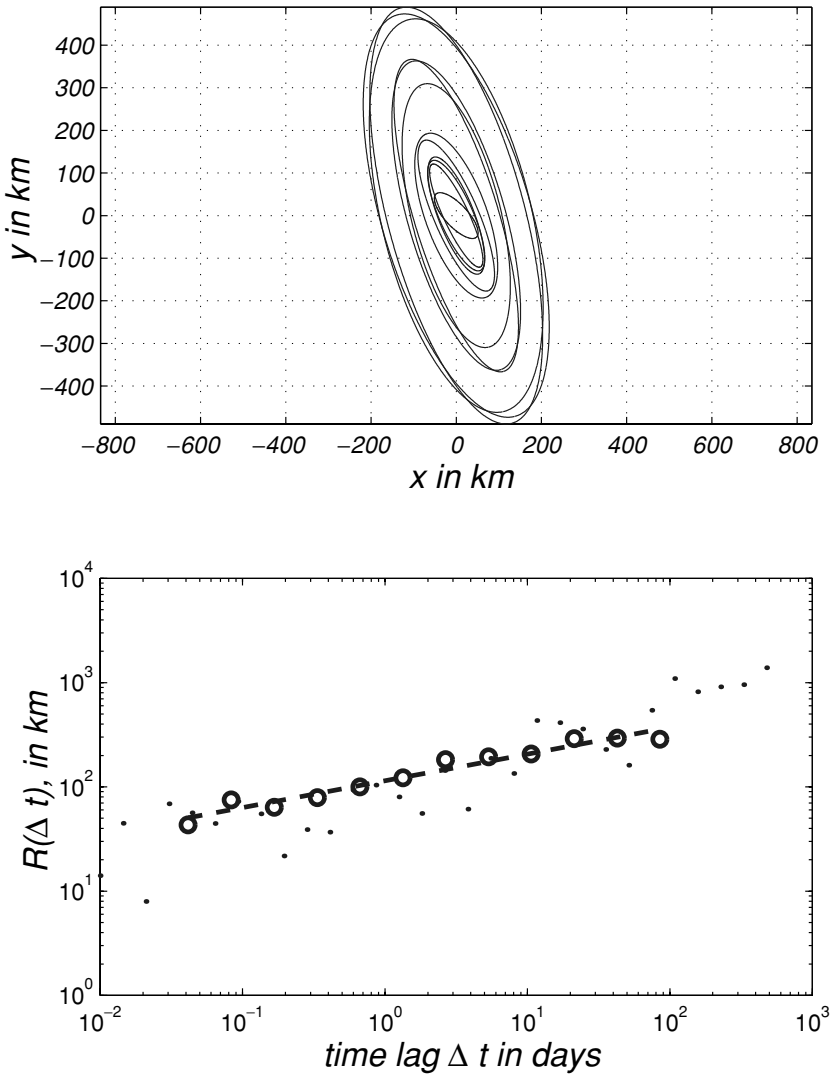

Figure 3. Top: mean ellipses (see the text for details) obtained at 12 different time lags Δt algebraically increasing. The first time lag window (innermost ellipse) is that of Fig. 1, left-hand graph, the fifth is that of Fig. 1, centre graph, and the 12th (outermost ellipse) is that of Fig. 1, right-hand graph. Bottom: corresponding mean radius $R(\Delta t)$, as black circles, and best power-law fit $R(\Delta t) \sim \Delta t^{0.26}$ as a thick dashed line. The mean radii obtained using the empirical Green function method are shown as small dots for a wider time range and gives a $R(\Delta t) \sim \Delta t^{0.32}$ fit for Δt in the same range as for the elliptical fit method.

the (local, in the case of multiscaling) space–time scale-changing operator T_λ. We determine this scaling operator T_λ by analysing the temporal behaviour of the mean distance $R(\Delta t) = \int dr\, r G(r, \Delta t)$, which scales as $R(\Delta t) \sim \Delta t^H$. H measures the space–time scaling anisotropy of T_λ. A more complete discussion on this subject can be found in Marsan *et al.* (2000). This definition of H is consistent with that described in Section 2 using the eigenvector decomposition of correlation matrices at time lags Δt, *cf.* eq. (3). In both cases $R(\Delta t)$ is insensitive to the overall temporal dissipation of correlation, and yields only an estimate of the typical length characteristic of the space–time correlation structure, hence a measure of the correlation length, at time lag Δt. We calculate H for the global catalogue, which mainly consists of subduction zone earthquakes, and analyse regional systems separately.

3.2.1 Global catalogue

Fig. 5 shows the functions $N(r, \Delta t)$, $\bar{N}(r)$ and $G(r, \Delta t)$ for Δt in three different windows: $0 < \Delta t < 0.1$ d, $9 < \Delta t < 10$ d and $90 < \Delta t < 100$ d. The relaxation of $N(r, \Delta t)$ towards $\bar{N}(r)$ as Δt increases is clearly seen, as is the widening of $G(r, \Delta t)$, spreading from a peak at small r (most of the afterevents being clustered around the main event at short timescales) to wider distributions as time increases. Also, $N(r, \Delta t)$ is found to be significantly greater than $\bar{N}(r)$, that is $N(r, \Delta t) - \bar{N}(r)$ is clearly above the noise level at distances r up to at least 300–400 km for the three time lag windows.

Fig. 6 shows the mean distance $R(\Delta t)$ for Δt varying between 10^{-3} and 700 d, along with the mean distances $R(\Delta t)$ found for the three systems studied in Marsan *et al.* (2000). Two scaling regimes are observed for the CNSS catalogue: at short timescales $\Delta t < 10$ d, $R(\Delta t) \sim \Delta t^{0.19}$, and at long timescales $\Delta t > 10$ d, and up to at least 700 d, $R(\Delta t) \sim \Delta t^{0.4}$. The first regime is similar to those found for both the Long Valley Caldera and the southern California systems, for which H was estimated to 0.22 for Δt ranging between 10^{-2} and 40 d, and between 10 min and 60 d, respectively. The second regime is close to the $R(\Delta t) \sim \Delta t^{1/2}$ law one would expect from a normal diffusion process driven by the viscous relaxation of either the

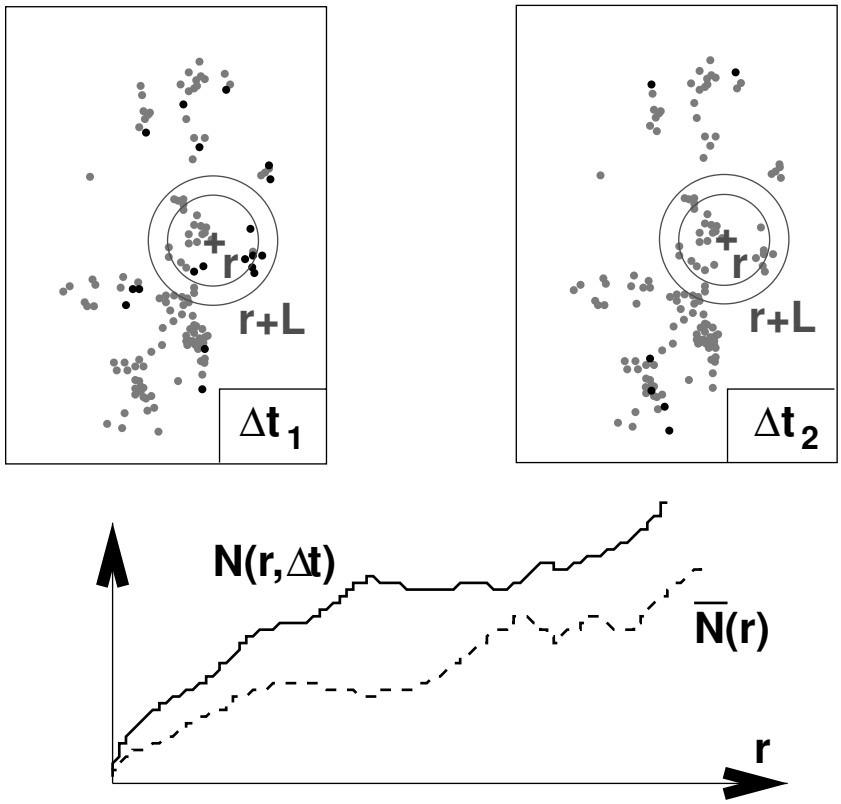

Figure 4. Construction of the Green function $G(r, \Delta t)$. For one main event (cross), we count the number of events (black dots) among all earthquakes (grey dots) occurring at a distance between r and $r + L$, and after a time gap between Δt_1 and $\Delta t_1 + T$ (left) and Δt_2 and $\Delta t_2 + T$ (right). Stacking these distributions obtained for all earthquakes as main events, and dividing by T, we retrieve the correlation $N(r, \Delta t)$ which lies above the distribution $\bar{N}(r)$ of temporally uncorrelated pairs of earthquakes, obtained by performing the same computation for all pairs without any constraints on the time delay Δt. $G(r, \Delta t)$ is defined as the normalized difference $N(r, \Delta t) - \bar{N}(r)$.

asthenosphere, as initially proposed by Elsasser (1969) for subduction zone earthquakes, or, at shallower depths, the plastosphere.

We further characterize the process by distinguishing subsets of main events by size (magnitude). We analysed four subsets: $5 \leq M \leq 5.5$, $5.5 \leq M \leq 6$, $6 \leq M \leq 6.5$ and $6.5 \leq M \leq 7$ earthquakes. A $M \geq 7$ group was also examined, but the small number of events (256) caused the associated propagator $G(r, \Delta t)$ to be extremely noisy, leading to most of the mean distances $R(\Delta t)$ being negative. The full set of 39 093 earthquakes was retained as the afterevent database. We computed the spatio-temporal correlation between the given subsets of main events (initial earthquake occurring at a given time t) and the total set of afterevents (subsequent earthquakes occurring at t plus time lag Δt). Fig. 7 shows the mean distances $R(\Delta t)$ for these four sets of main events. A systematic increase in the growth exponent H is observed: $H = 0.27$, 0.32, 0.32 and 0.4, respectively. While there is some suggestion of two scaling regimes, with transition at 10 days, for the first two subsets ($5 \leq M \leq 6$), no obvious break is observed for the two other subsets ($6 \leq M \leq 7$) for which a single regime extends from 10^{-3} to at least 700 days. Note that the mean distances $R(\Delta t)$ shown in Fig. 6 for the total CNSS catalogue mostly probe the large population of smaller earthquakes: 80 per cent of all the earthquakes in the catalogue are of magnitude $M \leq 5.5$. This can be seen directly by comparing the similar estimates of Fig. 6 for the total CNSS catalogue with those in Fig. 7 for the first subset of main events (with $5 \leq M \leq 5.5$): an $H = 0.29$ value is obtained when fitting the curve of Fig. 6 with a single scaling regime for the whole scaling range shown. A possible interpretation for the apparent disappearance of the transition at 10 days in the two subsets for $M > 6$ is that only the first regime is magnitude dependent, while the second regime is independent of magnitude with $R(\Delta t) \sim \Delta t^{0.4}$. As the H exponent of the first regime increases with M, and gets closer to 0.4, the break of scaling becomes less and less evident. Also, our analysis does not rule out the possibility that large $M7+$ earthquakes might be characterized by an initial regime with H larger than 0.4. For a mean asthenospheric shear modulus of 100 GPa, a viscous regime developing from a Maxwell time of 10 days to larger timescales indicates a lateral viscosity of the order of 10^{17} Pa s, which is a rather low estimate compared with the already low 5×10^{17} Pa s value of Pollitz *et al.* (1998).

The observed diffusion of afterevents, albeit with a slight increase of H, for increasing magnitude cut-off for the main events, implies that the existence of seismicity diffusion is little dependent on the magnitude cut-off (chosen to be equal to 5, *cf.* Section 1.4) for the whole catalogue. While the catalogue may not be complete to $M = 5$ for all regions and time periods, this can play at most a limited role in this observation.

3.2.2 Ocean ridges

An analysis of the diffusion along ocean ridges, including spreading centres and associated transform and transcurrent fault systems, was also performed, to test the ubiquity of this phenomenon. We examined a set of 4451 earthquakes in the CNSS catalogue occurring along the major spreading centres: the Atlantic ridge, the southwest,

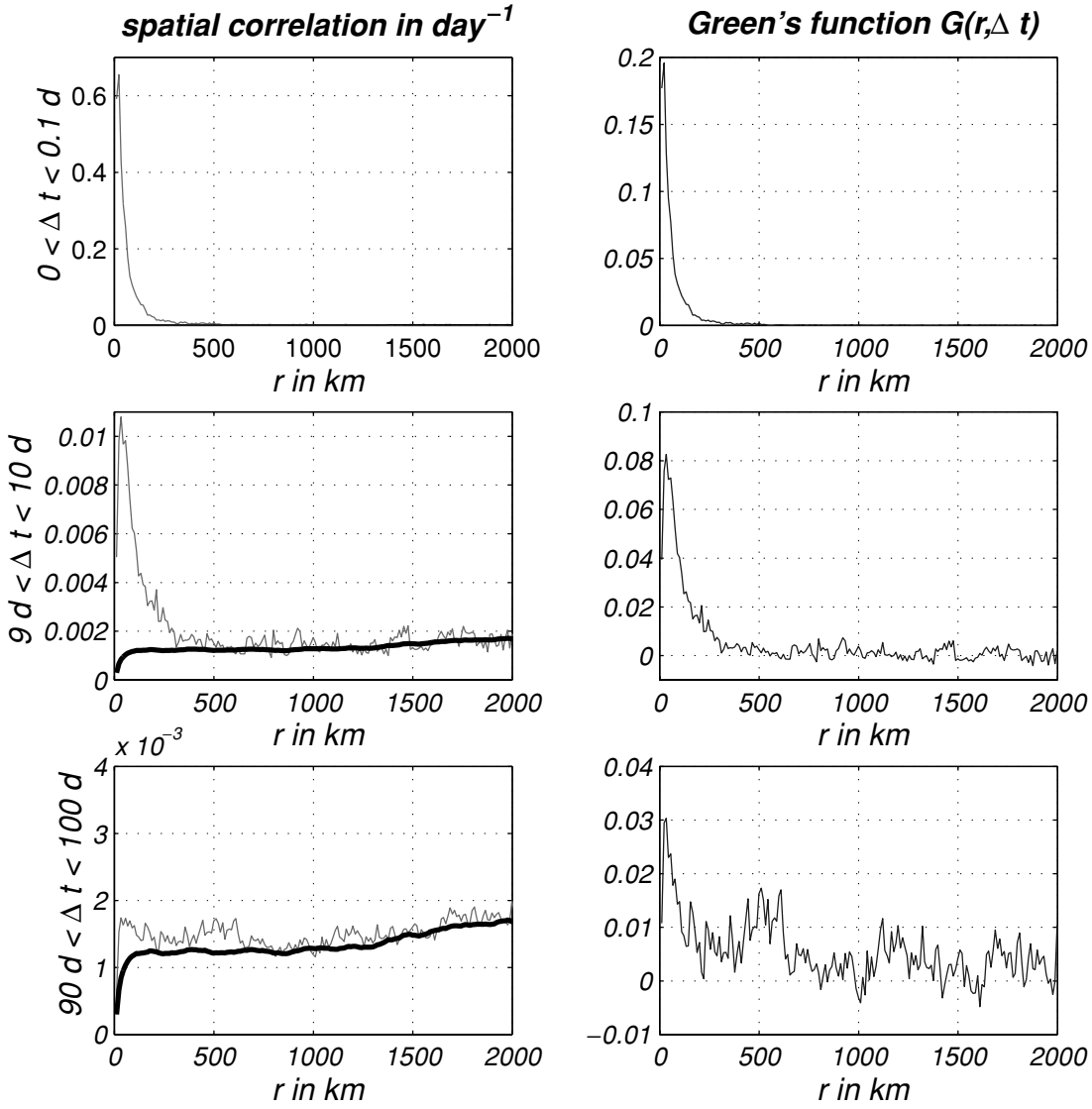

Figure 5. Functions $N(r, \Delta t)$ (grey) and $\bar{N}(r)$ (thick black) (left) and corresponding $G(r, \Delta t)$ (right), for three time lag windows: top, $0 < \Delta t < 0.1$ d; centre, $9 < \Delta t < 10$ d; bottom, $90 < \Delta t < 100$ d.

central and southeast Indian ridges, and along the East Pacific Rise. A single scaling regime with $H = 0.37$ is observed, as shown in Fig. 8. It is interesting that the seismic activity along ocean ridges is characterized by a larger growth exponent than subduction zone earthquakes, the latter forming most of the CNSS catalogue and therefore being those mostly probed in the previous analyses. This suggests that the diffusion might be controlled by temperature (see the next section). Again, in the case of ocean ridge seismicity, it is not clear whether the absence of a break at 10 days is due to the close similarity of the two regimes, for this data set. An analysis involving subsets of ocean ridge earthquakes within different magnitude bands unfortunately runs into the problem of limited data. It can also be seen in this case that the second regime $R(\Delta t) \sim \Delta t^{0.4}$ at $\Delta t \geq 10$ d is recovered.

3.3 Temperature control on diffusion

The diffusion processes discussed above extend from minutes to years in the post-seismic regime, and one could expect a rather significant control by the rheology on the diffusion exponent H. For example, where regional temperatures are high, high H values might be expected. Lower temperatures might then be associated with lower values of H. The relatively high value of $H = 0.37$ found for ocean ridge seismicity is consistent with such an interpretation, as large temperatures and heat flow are characteristic of such zones.

We searched for a possible correlation between regional surface heat flow Q, taken as an indicator of the average regional temperature at seismogenic depths, and the diffusion exponent H, by calculating the latter for nine large zones, mostly corresponding to subduction trenches: the Aleutian trench (denoted by ALE), the Kuril-Kamtchatka and Japan trenches (KUR), the Tonga-Kermadec trench (TON), the New Guinea, New Britain, South Solomon and New Hebrides trenches (NEW), the Mariana trench (MAR), the Ryukyu trench (RYU), the Java–Sumatra trench (JAV), a zone extending from the subducting Cocos plate to the Juan de Fuca plate, and containing the San Andreas Fault system, thus covering most of the central and north American seismicity (NAM) apart from the activity along the Aleutian trench, and the Peru and Chile trenches

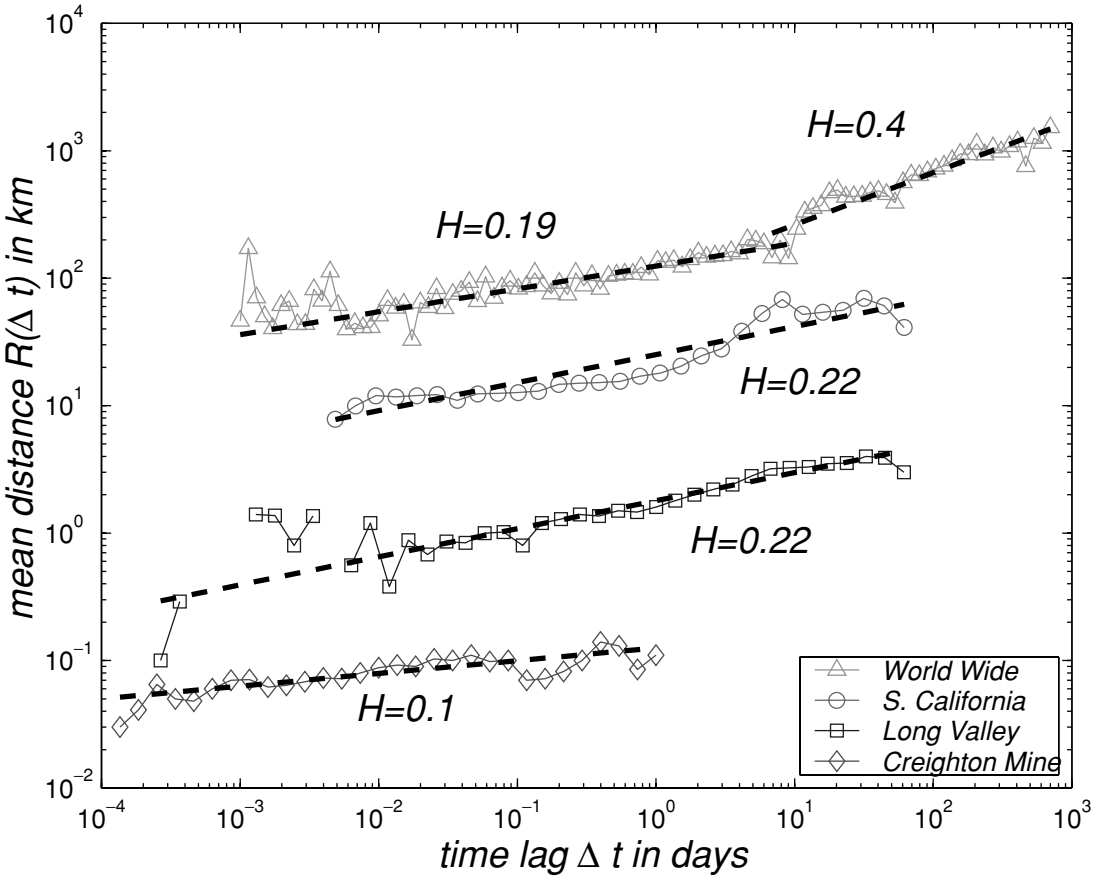

Figure 6. Mean distance $R(\Delta t)$ for worldwide seismicity along with the results obtained in Marsan *et al.* (2000), for three other systems (see the text for details). Dashed lines: best power-law fits, with the resulting estimates of the H exponent.

(CHI). Continental seismicity in central Asia and the Middle East, as well as European seismicity, were also considered, but the small numbers of $M5+$ events and the rather diffuse character of seismicity beneath central Asia led to poorly constrained estimates of H, and those regions were therefore discarded in this analysis. The Philippine Trench was examined in Section 2. In Fig. 9 we plot the mean distance $R(\Delta t)$ versus Δt for these nine zones, for Δt ranging from 10^{-2} d (about 15 min) to 450 days. Power-law fits yield H exponents varying from 0.13 (ALE) to 0.38 (RYU). Table 1 gives the estimates of H, along with the latitudes and longitudes defining the different zones, and the regional heat flow as estimated from Bott (1982). Fig. 10 demonstrates a positive correlation between surface heat flow and the diffusion exponent, for the nine zones plus the Philippine trench (PHI) as well as the seismicity along ocean ridges (OCE). Denoting by Q the heat flow in mW m^{-2} we found that $H = (2.8 \times 10^{-3} \pm 0.9 \times 10^{-3})\ Q + (0.079 \pm 0.06)$, and the null hypothesis of no correlation between H and Q is rejected at a 98.7 per cent significance level. This dependence of H on the regional heat flow suggests the existence of a strong thermal control on the diffusion process. However, the analysis of Marsan *et al.* (2000) showed a rather slow diffusion with $H = 0.22$ for the Long Valley Caldera, which is characterized by high heat flows.

3.4 Effective diffusion exponent

The number of temporally correlated afterevents occurring within a time interval Δt following the main event grows on average as $n(\Delta t) = \int_0^{\Delta t} d\delta t \int dr[N(r, \delta t) - \bar{N}(r)] \sim \Delta t^{1-p}$, where p is a characteristic scaling exponent reminiscent of the scaling exponent in the generalized Omori's law (Utsu 1961). The diffusion in $R(\Delta t) \sim \Delta t^H$ is therefore equivalent to $R(n) \sim n^{h=H/(1-p)}$, where h is the effective diffusion exponent for the random walks originating from the main event focus with a frequency of jumps decaying as Δt^{-p}. Here we determine the value of the effective exponent h for each of the regional seismicity systems selected in the previous section, by calculating the exponent p. We compute the dual, in the Fourier domain, of the autocorrelation: decay of the power spectrum $E(\omega)$ as $E(\omega) \sim \omega^{-\beta}$ yields a p exponent equal to $p = 1 - \beta$. Fig. 11 shows the power spectra used for estimating β and p in the range 12 h $< \Delta t <$ 11 yr. Table 1 gives the estimates of β and the corresponding effective diffusion exponents h. Values of h are found to range from 0.29 (Aleutian trench) to 4.11 (ocean ridge seismicity), and are generally greater than 0.5, indicating strongly super-diffusive processes when correcting for the operational time of the system. A plot of p versus Q (Fig. 10) also suggests a strong correlation between these two quantities, as previously reported by Mogi (1967) and Kisslinger & Jones (1991) at smaller spatial scales, and for aftershock sequence analyses.

4 SEISMIC ACTIVITY CORRELATED WITH A STRESS CHANGE

The analysis conducted in this paper allows us to determine empirically the response of the lithosphere to the stress change caused

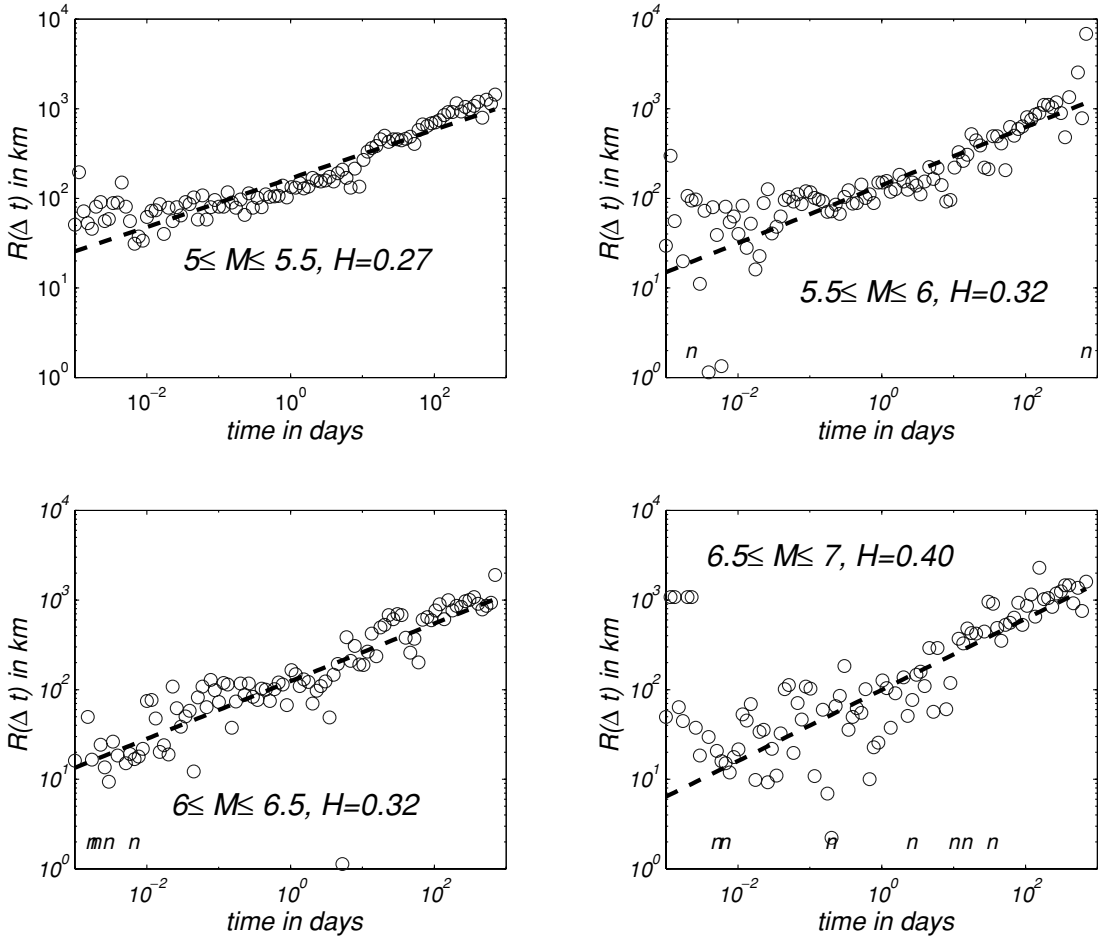

Figure 7. Mean distances $R(\Delta t)$ for the four subsets of main events, as indicated in the graphs. The best power-law fits are shown as dashed lines, with the corresponding H exponents. For the fourth subset $6.5 \leq M \leq 7$, the four outlier points at short timescales $\Delta t \leq 10^{-2}$ d and giving $R \simeq 1000$ km were discarded when computing the fit. Negative values of $R(\Delta t)$ obtained due to noise in $G(r, \Delta t)$ are indicated by an 'n' at the corresponding time lags.

by an earthquake, in terms of seismic activity. As the term 'response' normally implies a causal link between the initial perturbation and the subsequent changes of the system, and since, as already pointed out in the introduction, space–time correlation analyses cannot *a priori* recognize the existence of such causality, here we study the seismic activity *correlated* with (rather than triggered by) a stress change. The relative change in the temporally correlated post-seismic activity $K(r, \Delta t) = \frac{N(r,\Delta t) - \bar{N}(r)}{\bar{N}(r)}$ is the sum of two contributions:

$$K(r, \Delta t) = \mathcal{F}(\sigma(r) \mid \Delta t) + \int_0^{\Delta t} dt \mathcal{F}(\partial_t \sigma(r) \mid \Delta t - t), \qquad (7)$$

where $\mathcal{F}(\sigma(r) \mid \Delta t)$ gives the temporally correlated activity at position r following a static stress step $\sigma(r)$, and at Δt after the occurrence of the stress step. The second term is the activity correlated with the temporal evolution of the initial stress change, this evolution encompassing visco-elastic relaxation and fluid migration mechanisms, as well as the addition of subsequent stress changes due to temporally correlated afterevents. The latter contribution is thought to be rather weak, as typically no more than about 5–10 per cent of the main event energy is released in subsequent, temporally correlated afterevents. Also, viscous relaxation effects can be assumed to be negligible for the first diffusion regime in $R(\Delta t) \sim \Delta t^{0.19}$ for $\Delta t < 10$ d. Fig. 12 shows $K(r, \Delta t)$ for 10^{-2} d $\leq \Delta t \leq$ 3.5 yr, for four intervals of r, and for $M5+$ main and afterevents. $K(r, \Delta t)$ is found to decay from high values (up to 10^4 times the background seismicity level for $r \leq 10$ km at $\Delta t = 10^{-2}$ d) to sub-background seismicity levels $K(r, \Delta t) < 1$ at Δt around 10–100 d, depending on the distance r. The diffusion of post-seismic activity is clearly indicated here by the relatively slower decay at large r: as expected, the relaxation with Δt of $\mathcal{F}(\sigma(r) \mid \Delta t)$ depends non-linearly on σ, e.g. the system relaxes in $\sim \Delta t^{-0.7}$ for strong stress steps, at $r = 10$ km and in $\sim \Delta t^{-0.4}$ for smaller perturbations, at $r = 300$ km, for $\Delta t < 10$ km. As mentioned earlier, no causal deduction *stress step* (due to main event) $\rightarrow$ *post-seismic activity* (afterevents) can be made for this type of analysis. It can, however, be expected that, as the relative size of the main event compared with the afterevent increases, a more causal type of mechanism can be invoked; in which case $\mathcal{F}(\sigma(r) \mid \Delta t)$ can be roughly associated with an effective causal response of the system to a stress step. Fig. 13 displays $K(r, \Delta t)$ for $M6.5+$ main events and $M5+$ afterevents. Here also, a relatively slower relaxation is found at larger r, though a global trend in $\sim \Delta t^{-1}$ is observed at all r, at least for $\Delta t \geq 0.1$–1 d.

5 CONCLUSIONS

Interactions between earthquakes are neither only static nor local: they vary systematically with time through both dissipation and

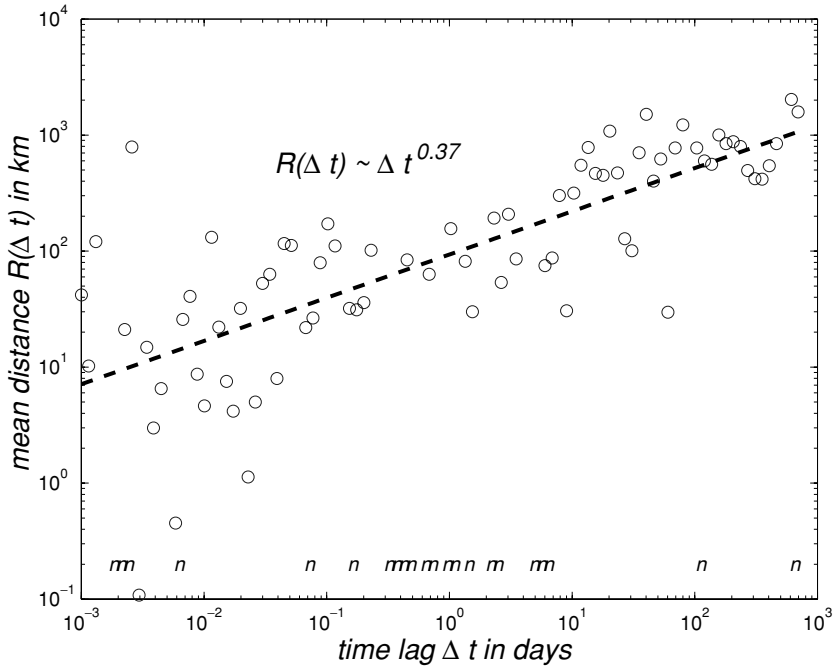

Figure 8. Mean distance $R(\Delta t)$ for the 4451 earthquakes along ocean ridges (see the text for details), together with the best power-law fit $R(\Delta t) \sim \Delta t^{0.37}$ as a dashed line. Negative values of $R(\Delta t)$ are indicated by an '*n*' at the corresponding time lags.

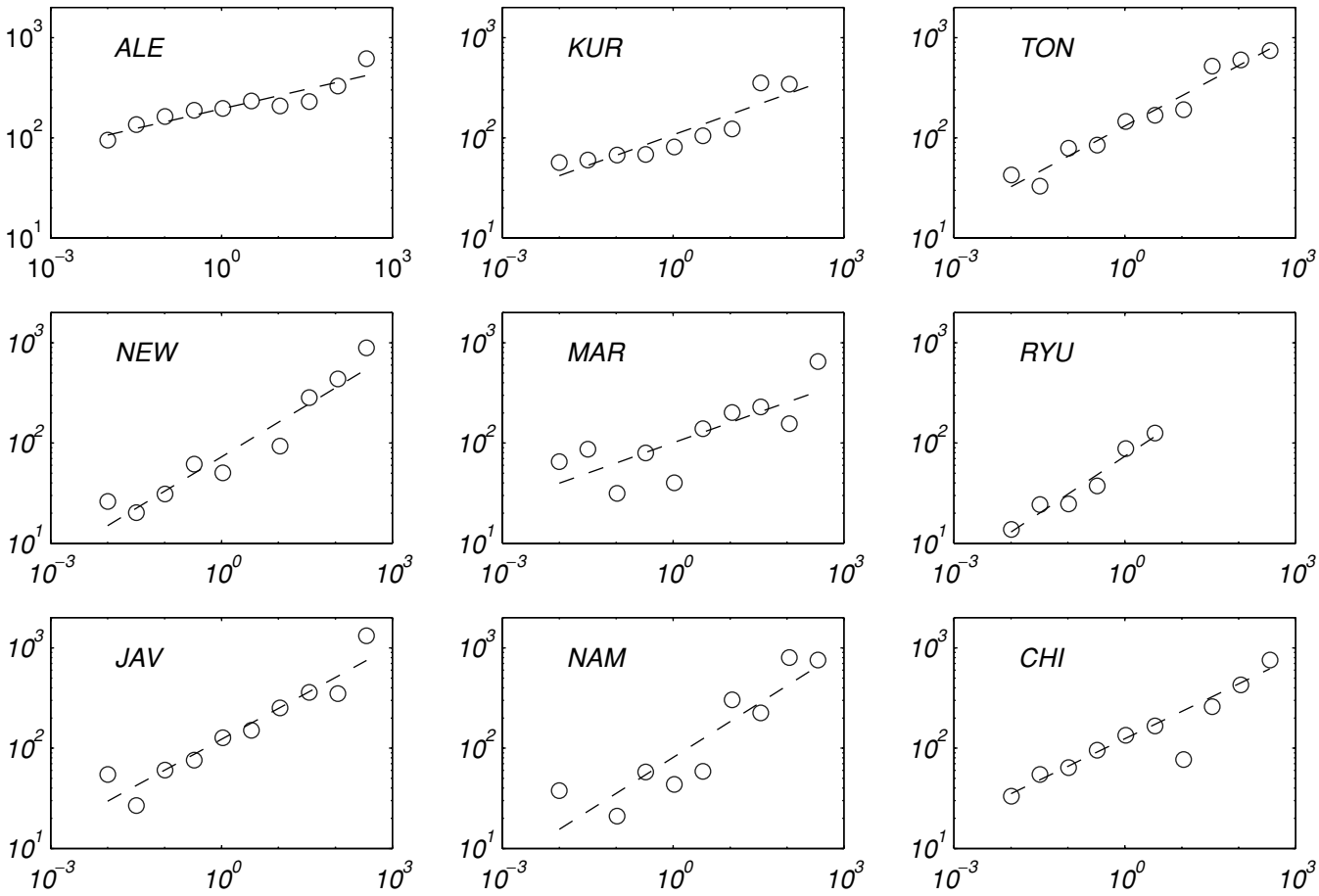

Figure 9. Mean distance $R(\Delta t)$, with Δt varying from 10^{-2} to 450 d, for the nine zones defined in the text. The best power-law fits are shown as dashed lines. The retrieved exponents are given in Table 1. Δt is in days; $R(\Delta t)$ is in km.

Table 1. Values of the regional diffusion exponents H, heat flow Q as estimated from Bott (1982), β exponents and corresponding effective diffusion exponents h, for the zones analysed.

Zone	Latitude, longitude	H	Q (mW m^{-2})	β	$h = H/\beta$
ALE	47° to 70°, 170° to −140°	0.13 ± 0.02	45 ± 10	0.45 ± 0.022	0.29 ± 0.06
KUR	37° to 60°, 135° to 170°	0.2 ± 0.04	60 ± 5	0.37 ± 0.011	0.54 ± 0.12
TON	−50° to −10°, 175° to −160°	0.3 ± 0.01	80 ± 5	0.22 ± 0.011	1.34 ± 0.11
NEW	−22° to 2°, 135° to 180°	0.34 ± 0.065	65 ± 10	0.31 ± 0.01	1.11 ± 0.26
MAR	10° to 37°, 135° to 150°	0.2 ± 0.065	60 ± 5	0.24 ± 0.005	0.84 ± 0.3
RYU	20° to 37°, 120° to 135°	0.38 ± 0.02	70 ± 5	0.15 ± 0.017	2.57 ± 0.48
JAV	(a) 0° to 20°, 90° to 110°	0.31 ± 0.03	70 ± 10	0.35 ± 0.013	0.89 ± 0.13
	(b) −15° to 0°, 90° to 130°				
NAM	10° to 50°, −130° to −80°	0.36 ± 0.05	100 ± 15	0.15 ± 0.01	2.42 ± 0.53
CHI	−45° to 10°, −85° to −60°	0.27 ± 0.02	65 ± 20	0.23 ± 0.007	1.17 ± 0.12
PHI	0° to 20°, 118° to 130°	0.32 ± 0.03	80 ± 5	0.33 ± 0.007	0.97 ± 0.11
OCE	See text	0.37 ± 0.03	120 ± 40	0.09 ± 0.009	4.11 ± 0.83

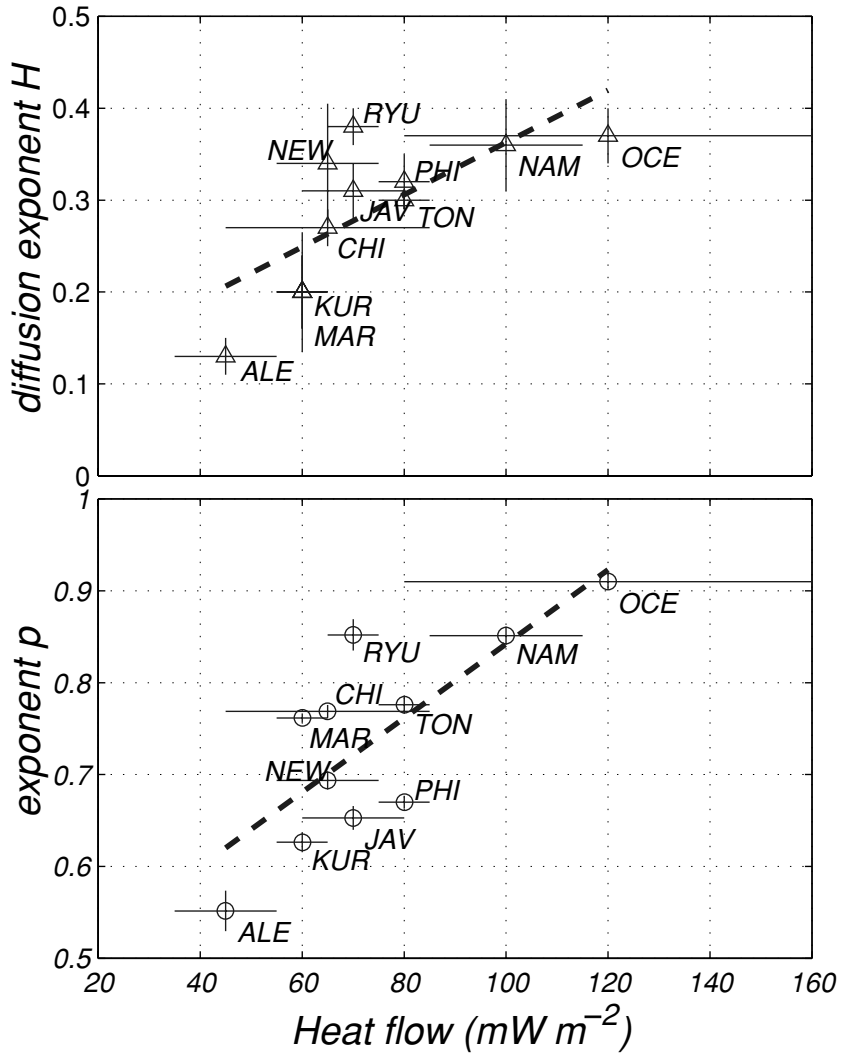

Figure 10. Plot of (top) diffusion exponent H and (bottom) exponent p versus the regional surface heat flow Q (in mW m^{-2}), for the nine zones of Fig. 7 plus the Philippine trench (PHI) and ocean ridge seismicity (OCE). The best linear fits give $H = 2.8 \times 10^{-3} Q + 0.079$ and $p = 4 \times 10^{-3} Q + 0.44$, as shown by the dashed lines.

relaxation, and they are found to be non-negligible up to distances of several rupture lengths. By observing dynamic changes in the spatial patterns of correlated subsequent seismic activity, we measure how the relative degree of interaction between earthquakes is systematically enhanced towards larger distances as the time difference Δt increases.

The question of whether the typical maximum triggering distance also grows with Δt is not easy to address empirically: one would

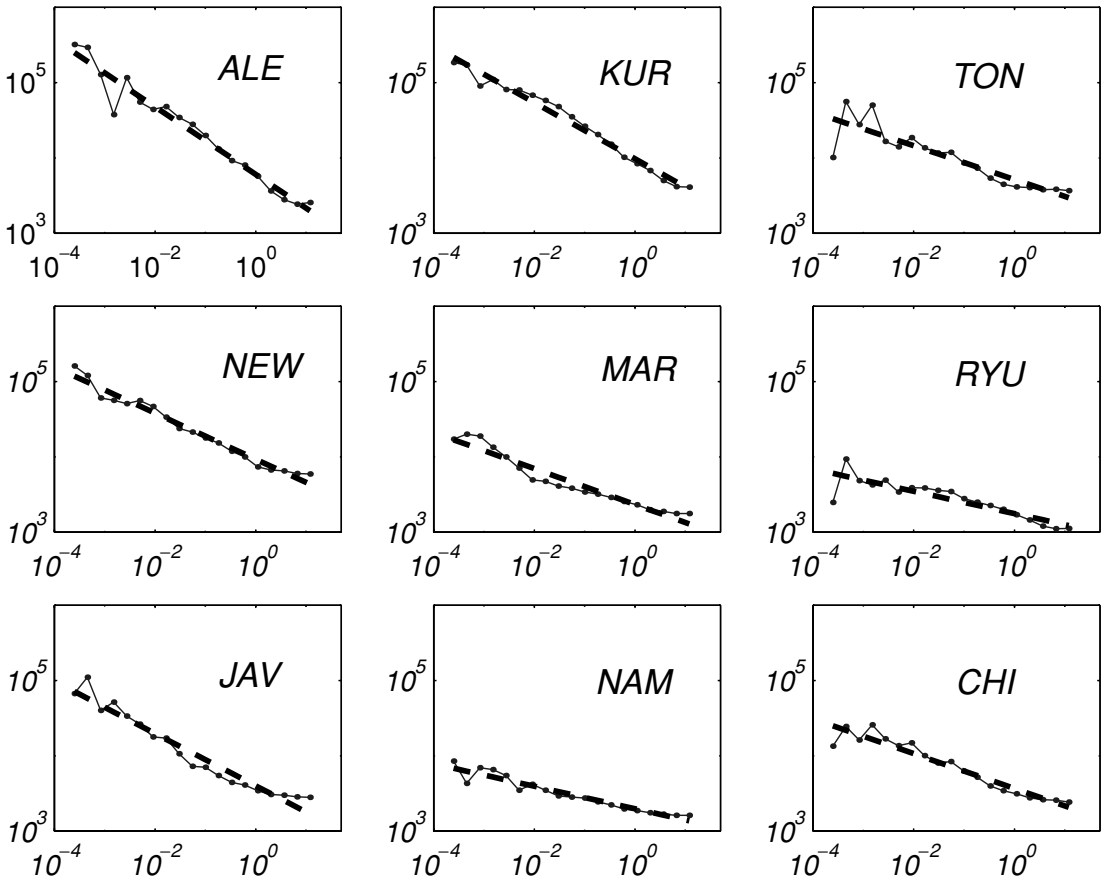

Figure 11. Power spectra of the seismicity time-series for the nine selected zones (see the text). x-axis: frequency ω in d^{-1}. The dashed lines give the power-law fits used in order to estimate the exponents β, for a scaling range from 12 h to 11 yr.

need to determine the minimum r at varying Δt for which $N(r, \Delta t)$ is expected to be within the noise band around $\bar{N}(r)$. Unfortunately such a measure is not robust. On the other hand, the mean distance $R(\Delta t)$ given by the mean of a normalized $N(r, \Delta t) - \bar{N}(r)$ yields more robust estimates, even though numerical problems associated with noise are also encountered. One might also argue that the maximum triggering distance could also follow a power-law growth with Δt, as a consequence of the space–time scaling invariance of the system. More sophisticated analyses would be required to investigate this issue.

Two types of interaction mechanisms can be distinguished, depending on the interaction length. Short-range interactions are those occurring between earthquakes within the rupture zone or its vicinity. We might reasonably expect these to be the result of direct local mechanisms. For example, residual high shear stresses following an initial earthquake will generate further earthquakes, and might lead to propagation of seismic activity by cascading triggering. The diffusion characteristics would then depend on the local distribution of asperities, as well as on the effective delaying mechanism. Also, crustal fluid movements triggered by the sudden change in normal stress can lead to fluid-induced seismicity. Such a mechanism is proposed by Nur & Booker (1972) to explain the temporal pattern of aftershock occurrence, and by Hudnut *et al.* (1987) and Noir *et al.* (1997) to explain delayed triggering in historical earthquake sequences. Two remarks can be made about such a mechanism: (1) a scale-invariant permeability structure is required to account for the observed low diffusion exponents, the fluid diffusion then being seen as taking place in a heterogeneous medium (e.g. percolation models are typical examples of how such scale-invariant media can alter the diffusion exponent) and (2) a significant rotational component is expected, corresponding to fluid transport from compressional to dilatational quadrants. Let us recall that, in the case of the Philippine Mobile Belt, no such component is observed.

Long-range interactions at long timescales can be due to the viscous relaxation of the asthenosphere/plastosphere, and have also been observed at short timescales following large earthquakes, e.g. Hill *et al.* (1993) and Linde & Sacks (1998). In both cases, fluid-induced seismicity is remotely triggered in geothermal areas. Again, it is not obvious to determine whether such remotely triggered seismicity patterns dynamically evolve as the time lag Δt increases. A cascading mechanism by which triggered activity in turn remotely triggers other areas could be proposed. However, the difference in magnitude between the first generation of triggered events and the triggering earthquake is quite significant in the case of the Landers earthquake (Hill *et al.* 1993), implying a strong dissipation in the level of triggered activity as we look at further generations.

At short timescales $\Delta t < 10$ d, for which asthenospheric viscoelastic effects are likely to be negligible, the main mechanism driving both the temporal dissipation and the spatial relaxation could be a non-linear response, seen as a change in the production rate of earthquakes, of the lithosphere subject to a stress step generated by the occurrence of a seismic event. Nucleation phase models, for example subcritical crack growth or rate and state friction models, have been proposed to account for the typical $1/\Delta t$ decay in global production rate following a stress perturbation, but have somehow failed, in their present form, to reproduce the typical Δt^H

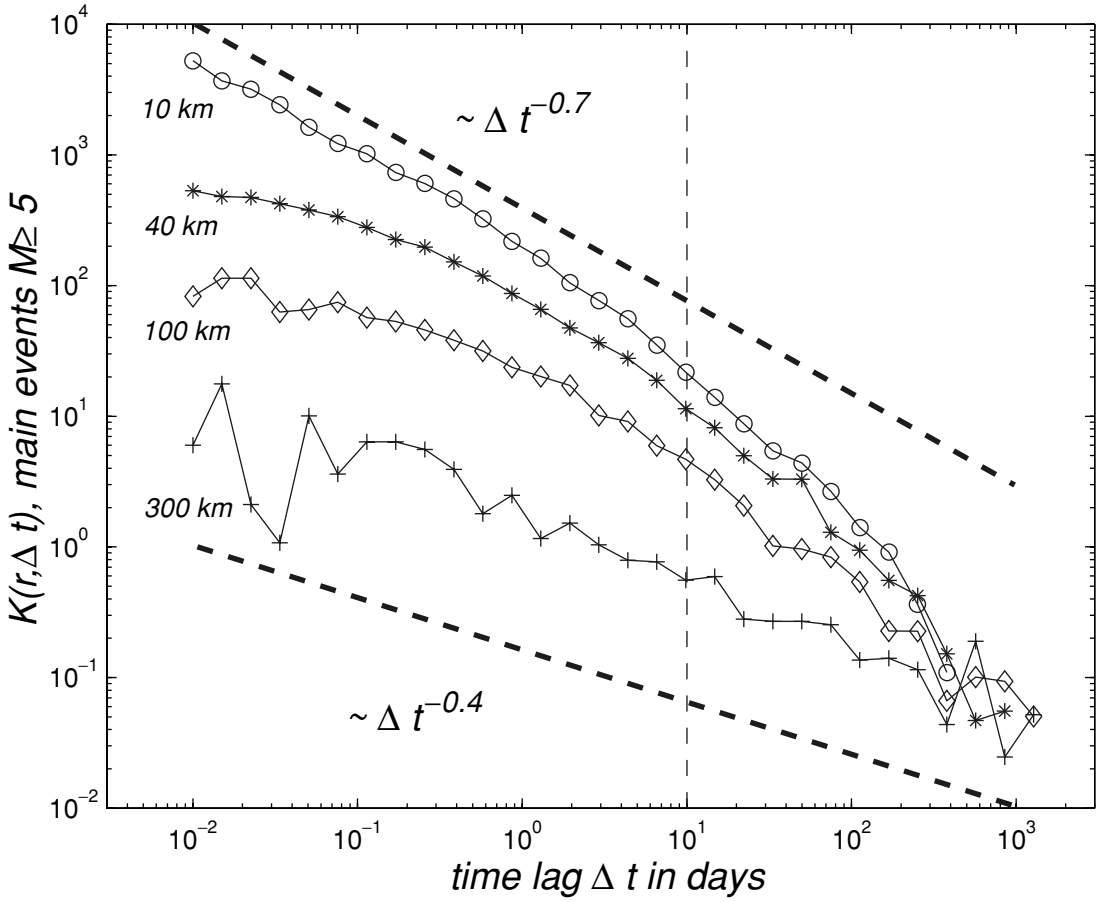

Figure 12. $K(r, \Delta t)$ for four intervals of r: (○) $0 \leq r \leq 10$ km, (∗) $30 \leq r \leq 40$ km, (◇) $90 \leq r \leq 100$ km and (+) $290 \leq r \leq 300$ km, for main and afterevents of magnitude 5 and above.

post-seismic activity diffusion (Marsan *et al.* 2000). As shown in Section 4, the relaxation at large r (small stress perturbation) is found to be relatively slower than at small r (strong perturbation). This implies that the time advance for a nucleating slip event due to a stress perturbation should increase more slowly with increasing time-to-failure as the stress perturbation is stronger: denoting by $n(t) = \bar{n}$ the constant rate of nucleating earthquakes within a given volume, with a time-to-failure t, and by $\delta t_\sigma(t)$ the time advance caused by a (e.g. Coulomb) stress perturbation σ (assuming here that the perturbation effectively causes an advance rather than a delay), the new time-to-failure t' is

$$t \rightarrow t' = t - \delta t_\sigma(t) \tag{8}$$

$$\Rightarrow \frac{dt}{dt'} = 1 + \frac{d\delta t_\sigma(t)}{dt'} \tag{9}$$

and the new rate of earthquake nucleations $n'_\sigma(t')$ is therefore $n'_\sigma(t')\,dt' = \bar{n}\,dt$, and is related to the function $\mathcal{F}$ of Section 4 as

$$\mathcal{F}(\sigma \mid t') = \frac{n'_\sigma(t') - \bar{n}}{\bar{n}} = \frac{d\delta t_\sigma(t)}{dt'}. \tag{10}$$

Hence the diffusion of post-seismic activity implies that, $\forall \lambda > 1$, $d\frac{\mathcal{F}(\lambda\sigma \mid t')}{\mathcal{F}(\sigma \mid t')}/dt' \leq 0$, hence

$$\frac{d}{dt'}\left[\frac{d\delta t_{\lambda\sigma}(t')}{d\delta t_\sigma(t')}\right] \leq 0. \tag{11}$$

While the model of Dieterich (1994) does indeed meet this condition, and therefore reproduces a growth of the aftershock/afterevent zone with time, it does not properly account for the absence of any typical space or timescale, within wide enough scale ranges, responsible for the global space–time scale invariance of the system, as is observed in our analyses.

Finally, let us point out the equivalence between the mean distance $R(\Delta t)$ and the correlation length characteristic of seismicity catalogues. Such a correlation length is defined in statistical physics (for example, percolation or Ising models) as a measure of the typical size of clusters present in the model, growing with the maximum size of the clusters, as the control parameter gets closer to the critical value. For seismicity systems, the clusters represent temporally correlated groups of earthquakes: the occurrence of an earthquake influences the future occurrence of other earthquakes within its cluster. The methods detailed in this paper hence correspond to an objective analysis of the temporal evolution of the correlation length in earthquake populations, and more particularly of the growth of this influence region (or cluster) with time after the occurrence of an earthquake. Extensive use of these methods for the understanding of the evolution of the typical size of the influence region at various stages during the seismic cycle appears particularly promising.

ACKNOWLEDGMENTS

The authors thank the Northern California Earthquake Data Center (NCEDC) and the Member networks of the Council of the National Seismic System (CNSS) for the use of the seismicity catalogue; Ian Main and Franz Lahaie for discussion; Gareth O'Brien for help

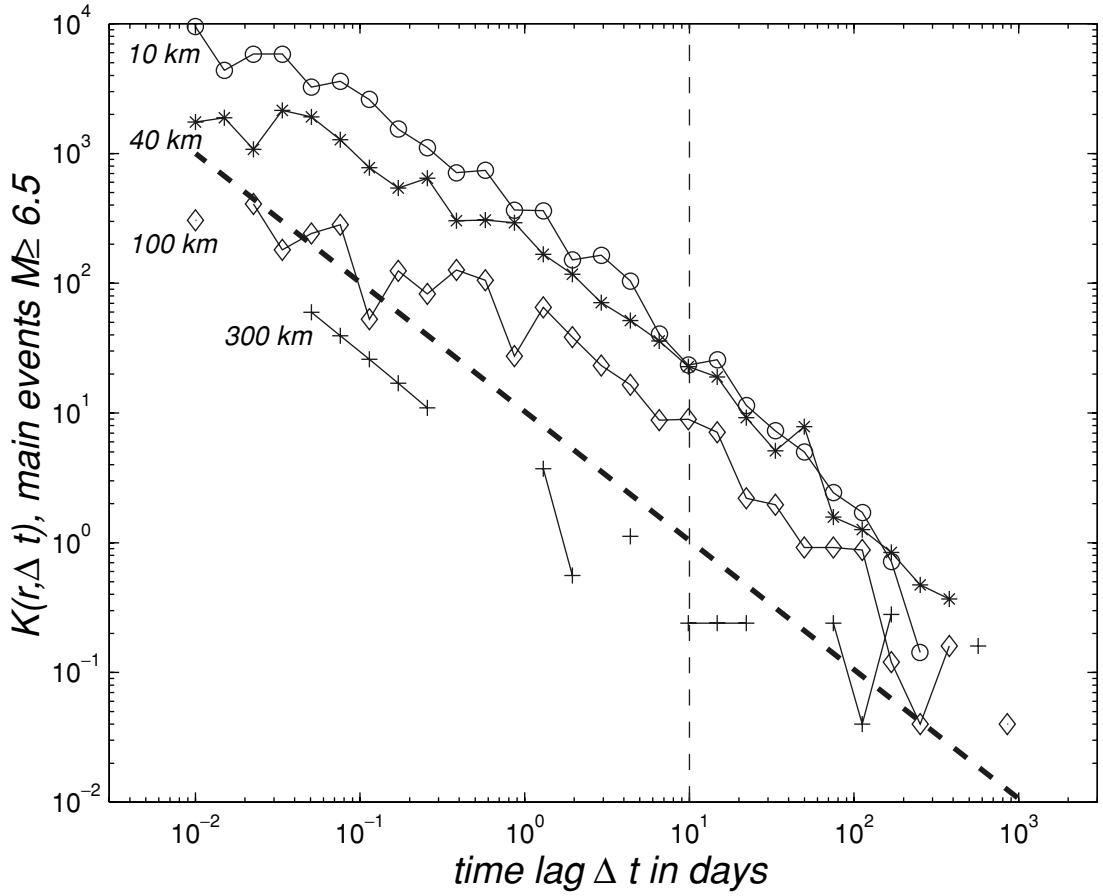

Figure 13. $K(r, \Delta t)$ for four intervals of r: (○) $0 \leq r \leq 10$ km, (∗) $30 \leq r \leq 40$ km, (⋄) $90 \leq r \leq 100$ km and (+) $290 \leq r \leq 300$ km, for main events of magnitude 6.5 and above and afterevents of magnitude 5 and above. The dashed line gives a $\sim \Delta t^{-1}$ decay.

in parallelizing the code used in Section 2; and the editor and two anonymous reviewers for many suggestions and corrections that helped clarifying the text. Funding of DM by Forbairt is acknowledged.

REFERENCES

Anderson, D.L., 1975. Accelerated plate tectonics, *Science,* **187,** 1077.

Bott, M.H.P. & Dean, D.S., 1973. Stress diffusion from plate boundaries, *Nature,* **243,** 339.

Bott, M.H.P., 1982. *The Interior of the Earth, its Structure,* Constitution and Evolution, Arnold, London.

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

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

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

Hill D.P., *et al.*, 1993. Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, *Science,* **260,** 1617.

Hirata, T., 1987. Omori's power law aftershock sequences of microfracturing in rock fracture experiments, *J. geophys. Res.,* **92,** 6215.

Huc, M. & Main, I.G., 2000. Anomalous stress diffusion in earthquake triggering, *EGS XXV General Assembly*, abstract.

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

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

Kisslinger, C. & Jones, L.M., 1991. Properties of aftershock sequences in southern California, *J. geophys. Res.,* **96,** 11 947.

Linde, A.T. & Sacks, I.S., 1998. Triggering of volcanic eruptions, *Nature* **395,** 888.

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

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

Marsan, D., Bean, C.J., Steacy, S. & McCloskey, J., 2000. Observation of diffusion processes in earthquake populations, and implications for the predictability of seismicity systems, *J. geophys. Res.,* **105,** 28 081.

Mogi, K., 1967. Earthquakes and fractures, *Tectonoph.,* **5,** 35.

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

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

Omori, F., 1895. On the aftershocks of earthquakes, *J. Coll. Sci.,* **7,** 111.

Pollitz F.F., Bürgmann, R. & Romanowicz, B., 1998. Viscosity of oceanic asthenosphere inferred from remote triggering of earthquakes, *Science,* **280,** 1245.

Romanowicz, B., 1993. Spatiotemporal patterns in the energy release of great earthquakes, *Science,* **260,** 1923.

Rundle, J.B., Tiampo, K.F. & Klein, W., 1999. Non-local space–time patterns in seismicity, with application to Bufe-Varnes scaling, *EOS, Trans. Am. geophys. Un.,* **80,** 924.

Sanders, C.O., 1993. Interaction of the San Jacinto and San Andreas fault zones, southern California: triggered earthquake migration and coupled recurrence intervals, *Science,* **260,** 973.

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

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

Turcotte, D.L., 1992. *Fractals and Chaos in Geology and Geophysics*, Cambridge University Press, Cambridge.

Utsu, T., 1961. A statistical study on the occurence of aftershocks, *Geophys. Mag.,* **30,** 521.

APPENDIX : NOTE ON THE DETERMINATION OF THE MEAN DISTANCE $R(\Delta t)$

As a result of statistical noise in the function $N(r, \Delta t)$, especially at large Δt, and therefore also in $G(r, \Delta t)$, the estimate of $R(\Delta t)$ is rather sensitive to the maximum scalelength $\mathcal{L}$ defining the interval on which $R(\Delta t)$ is computed, that is

$$R(\Delta t) = \int_0^{\mathcal{L}} dr\, r\, G(r, \Delta t). \tag{A1}$$

While ideally we would like to estimate $R(\Delta t)$ with the largest $\mathcal{L}$ available $\simeq$20 000 km, this is not an optimal empirical estimate, as illustrated in Fig. 14: in the upper graph we show $N(r, \Delta t)$ and $\bar{N}(r)$ for one of the time lag windows ($0.89 < \Delta t < 1.02$ d) effectively used in producing the $R(\Delta t)$ curve of Fig. 6. Noise is clearly present; while the peak at small r is (especially at short Δt) much stronger than the noisy fluctuations of $N(r, \Delta t)$ around $\bar{N}(r)$ at large r, the fact that $R(\Delta t)$ corresponds to the integral of $G(r, \Delta t)$ times r implies an enhancement of the noise at large r. The bottom graph of Fig. 14 shows the dependence of $R(\Delta t)$ on the scalelength $\mathcal{L}$ used for its computation, *cf.* eq. (A1), for the same time window as in the upper graph. As expected, $R(\Delta t)$ initially increases with $\mathcal{L}$, then

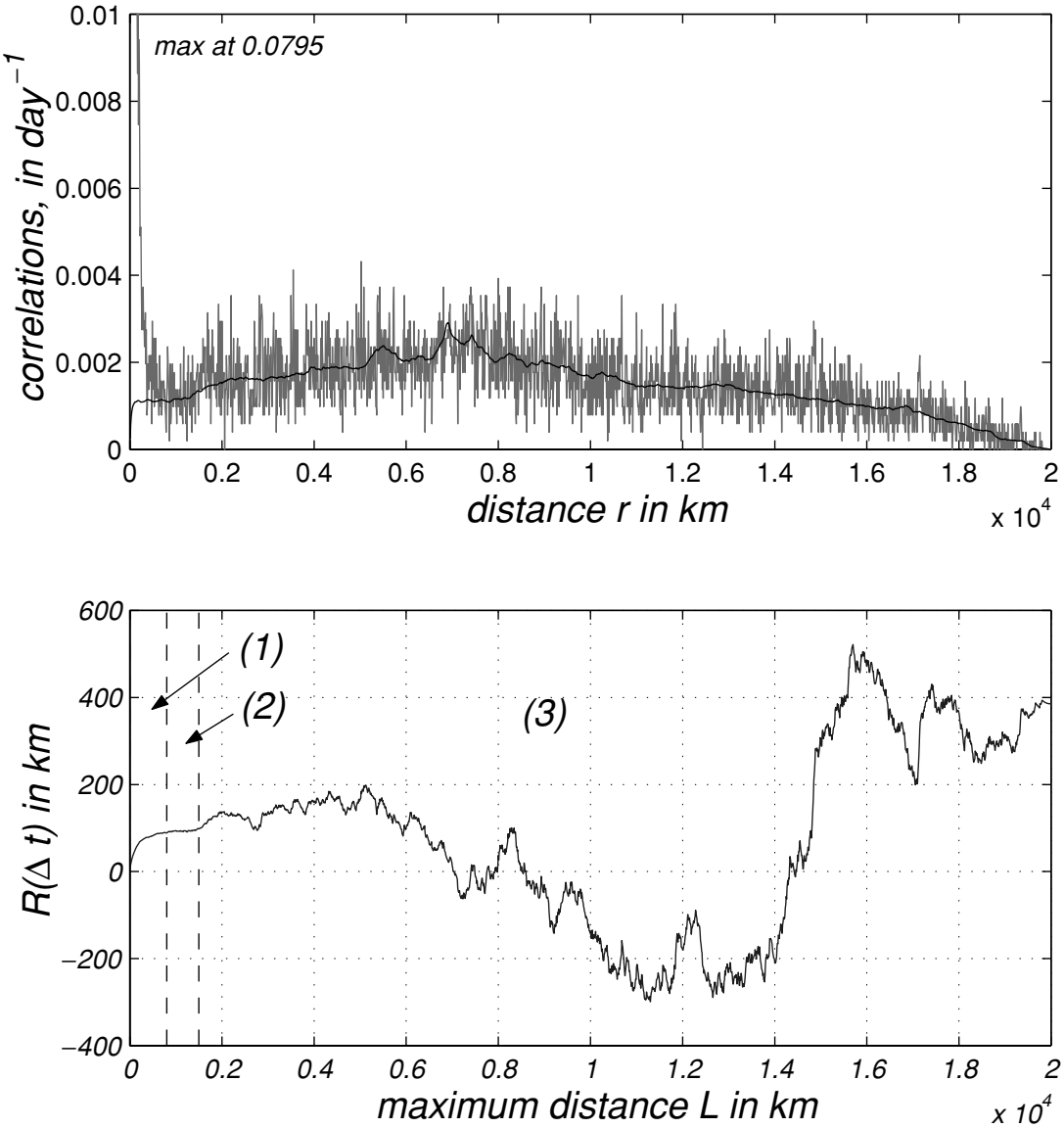

Figure 14. Divergence of the estimate of $R(\Delta t)$, for $0.89 < \Delta t < 1.02$ d. Upper graph: $N(r, \Delta t)$ (grey) and $\bar{N}(r)$ (black line), in d^{-1}. Note that the graph is truncated at 0.01 d^{-1}, but that the initial (at small r) peak of $N(r, \Delta t)$ has a maximum value of 0.0795 d^{-1}. Bottom graph: variations of $R(\Delta t)$ with maximum length $\mathcal{L}$. Three successive regimes can be distinguished: (1) $0 \leq \mathcal{L} \leq 800$ km, growth of $R(\Delta t)$ with increasing $\mathcal{L}$; (2) $800 \leq \mathcal{L} \leq 1500$ km, stabilization; and (3) $\mathcal{L} \geq 1500$ km, random fluctuations due to noise in $G(r, \Delta t)$.

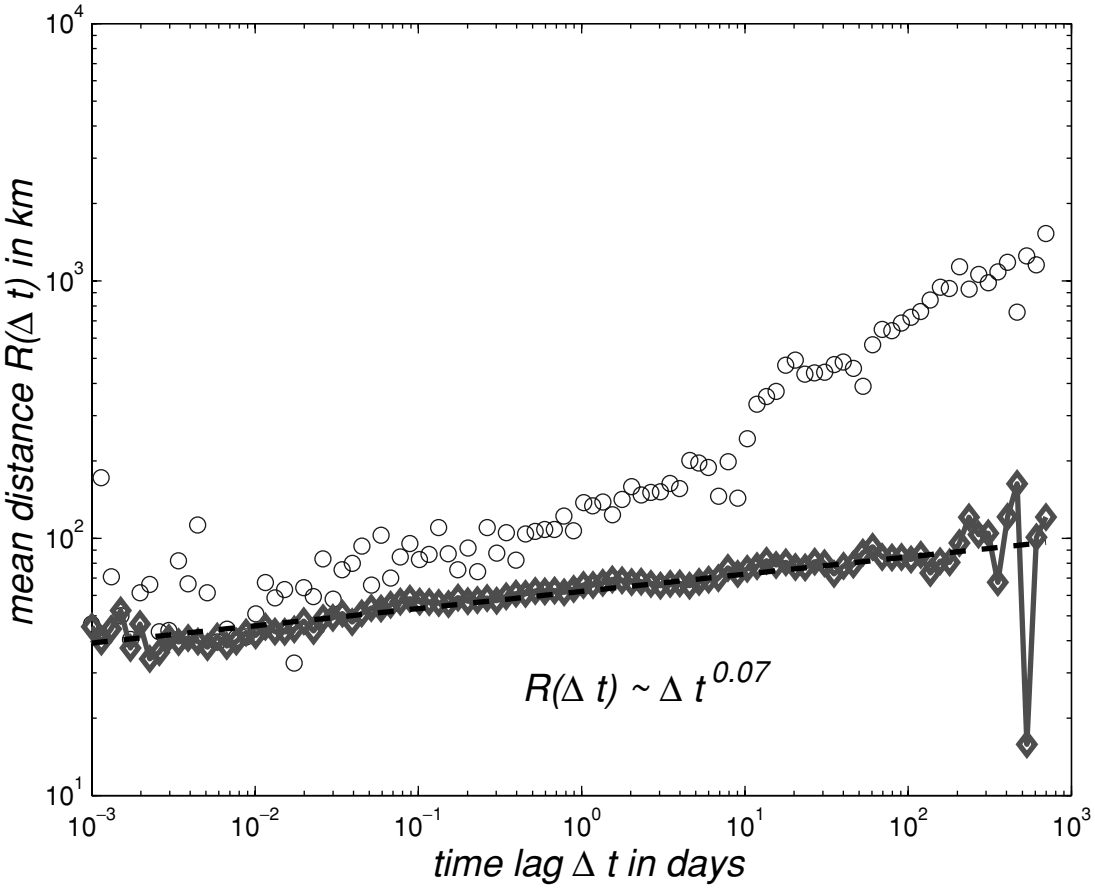

Figure 15. Mean distance $R(\Delta t)$ obtained for a maximum distance $\mathcal{L} = 200$ km (◇), together with (○) the curve obtained using $\mathcal{L} = 2000$ km, *cf.* Fig. 4. The best power law fit in dashed line gives $H = 0.07$. At small Δt (10^{-3} to 10^{-2} d), $R_{\mathcal{L}=2000\,\mathrm{km}}(\Delta t)$ converges towards $R_{\mathcal{L}=200\,\mathrm{km}}(\Delta t)$, with some scatter due to random fluctuations in $G(r, \Delta t)$ for $\mathcal{L} = 2000$ km at such time lags: we are in the stabilization regime for $\mathcal{L} = 200$ km and in the random fluctuations regime for $\mathcal{L} = 2000$ km. At larger time lags, we are in the stabilization regime for $\mathcal{L} = 2000$ km and in the growth regime for $\mathcal{L} = 200$ km, hence a too slow increase of $R(\Delta t)$ at $\mathcal{L} = 200$ km.

stabilizes, but eventually undergoes a strongly fluctuating random walk. The latter is responsible for the negative values of $R(\Delta t)$, as is also seen in Fig. 7 for $\mathcal{L} = 2000$ km. Negative values of $R(\Delta t)$ can also occur at large Δt due to insufficient sampling, see, for example, Figs 7 and 8.

These strong random fluctuations of $R(\Delta t)$ start when the level of noise in G becomes comparable to the actual value of G that would be obtained in the statistical limit of an infinite number of samples. To avoid this problem, we choose a value of $\mathcal{L}$ such that $R(\Delta t)$ is apparently stabilized, but small enough that the random fluctuations have not become significant. $\mathcal{L} = 2000$ km is a good candidate for the time range $10^{-3} < \Delta t < 700$ d. Previous studies and models giving estimates of the maximum distance reached by reported instances of viscous relaxations/diffusions: 1500 km after 10^4 yr (Elsasser 1969); 100 km after 100 yr (Bott & Dean 1973); 390 km after 5 yr (Anderson 1975); several thousands of km in 10–20 yr (Romanowicz 1993); about 1500 km in 15 yr and 2500 km in 25 yr as estimated from Pollitz *et al.* (1998); 'stress pulses' propagating along fault zones from 1.7 km yr^{-1} to about 150 km yr^{-1}, see Sanders (1993), suggest that this value of 2000 km is not too small for Δt up to 700 d, and is probably mostly in the stabilization regime. Finally, $\mathcal{L} = 2000$ km explains the fluctuations of $R(\Delta t)$ at small Δt, particularly for $\Delta t < 10^{-2}$ d in Fig. 6. A smaller value of $\mathcal{L}$ would lead to less scatter around the mean power-law dependence, but can give false information regarding the actual diffusion characteristics if probing before the 'stabilization' of $R(\Delta t)$ with $\mathcal{L}$. An example is given in Fig. 15, which displays $R(\Delta t)$ for $\mathcal{L} = 200$ km. A unique scaling regime with $H = 0.07$ is obtained, indicating a much slower diffusion process than it actually is. Interestingly, this low growth exponent is very close to that given by both Tajima & Kanamori (1985) with $H \simeq 0.075$, and Huc & Main (2000), $H = 0.06$, with maximum distances $\mathcal{L}$ equal to a few hundred km and 150 km, respectively. Huc & Main (2000) also correct for geometrical spreading, and their method is therefore not directly suited for an accurate determination of the diffusion exponent. The fact that significant triggering is readily evident up to at least 300 km (*cf.* Fig. 5) and that $R(\Delta t)$ appears to be inside the stabilization regime for most of the 10^{-3} to 700 d range at $\mathcal{L} = 2000$ km indicates that earthquake interactions are non-localized in the immediate coseismic rupture zone, but extend well beyond that (see also Hill *et al.* 1993 for an account of seismicity triggered up to distances of the order of 1250 km following the $M = 7.3$ Landers earthquake; see also Rundle *et al.* 1999 for a study of non-local interactions between earthquakes).