Supporting Online Material for

Extending Earthquakes' Reach Through Cascading

David Marsan* and Olivier Lengliné

*To whom correspondence should be addressed. E-mail: david.marsan@univ-savoie.fr

Published 22 February 2008, *Science* **319**, 1076 (2008)
DOI: 10.1126/science.1148783

This PDF file includes:

Materials and Methods

SOM Text

Figs. S1 to S6

Table S1

References

We here test the method, using synthetic catalogs (section 1). We modify the method as described in the manuscript to account for the non-uniform distribution of background earthquakes over space (section 2). We run an analysis of the California earthquake dataset using 3D distances between the aftershock hypocenters and the mainshock rupture plane instead of the 2D epicentral distances as in the manuscript (section 3). Finally, we discuss the linearity hypothesis on which the method is based (section 4).

1 – A test of the method:

We demonstrate the ability of the method to decipher the multiple connections between earthquakes, by analyzing a synthetic catalogue using the Epidemic-Type Aftershock Sequence (ETAS) model (*S1, S2*). This model postulates that every earthquake triggers its own, local, aftershock sequence, therefore resulting in a cascade of triggering. Earthquakes are generated with a time and space-dependent rate density $\lambda(\underline{x},t)=\lambda_0+\sum_{t_i<t} K\, e^{\alpha m_i}(t-t_i+c)^{-p} f(r,m_i)$ computed by summing the influence (kernels) of all past earthquakes of indices i. The rate density $\lambda(\underline{x},t)$ integrated over a given space-time volume $V=(\underline{X},T)$ yields the mean number $\Lambda=\int_X d\,\underline{x}\int_T dt\ \lambda(\underline{x},t)$ of earthquakes occurring in V, so that the actual number is the realization of a Poisson law with mean Λ. Parameter λ_0 is the rate density of background earthquakes, i.e., earthquakes that occur randomly in space and time and that are

not triggered by previous earthquakes. We denote by $r=\left|\underline{x}-\underline{x_i}\right|$ the distance between earthquakes i and j, and by m_i and t_i the magnitude and occurrence time of earthquake i. Parameters K and α constrain the direct aftershock productivity of each mainshock. The rate of directly triggered aftershocks follows Omori-Utsu's law $(t+c)^{-p}$ with exponent p and time cut-off c. The spatial density is $f(r,m)=\dfrac{1}{2\pi r L \ln(1+\frac{1}{L})\times(1+\frac{r}{L})}$ for $0\le r\le 1$, and $f(r>1,m)=0$, so that $\int_0^1 dr\, 2\pi r\, f(r,m)=1$, with $L\sim 10^{0.5m}$ the rupture length at magnitude m. For this test, we used the parameters $\lambda_0=0.25$ earthquakes per unit time and unit surface, $\alpha=2$, $c=10^{-2}$, $p=1.2$, $K=0.0094$, giving a branching ratio (mean number of direct aftershock triggered by an earthquake) of 0.9. The magnitudes are distributed according to a Gutenberg-Richter law (*18*), hence a probability density $f_m(m)=b\ln(10)\times 10^{-bm}$, with b-value equal to 1, and minimum magnitude $m=0$. A total of $N=7010$ earthquakes were thus generated, over a duration $T=10^3$ (in arbitrary units; Fig. S1). The largest earthquake has magnitude $m=4.61$, and a rupture length set to $L=0.1$ (in arbitrary units). The earthquakes are constrained to occur on a $S=2\times 2$ surface, with periodic boundaries, i.e., a torus.

We apply the algorithm to this simulated catalogue. We further simplify the method by decoupling the

spatial and the temporal dependence of the kernel: $\lambda_i(\underline{x},t)=\lambda_s(|\underline{x}-\underline{x_i}|,m_i)\times\lambda_t(t-t_i,m_i)$. This allows for a more robust estimation of the weights, by significantly reducing the number of kernel values to determine. Since a direct dependence on magnitude can be ambiguously carried by either the spatial density λ_s (number of triggered earthquakes per unit area) or the temporal rate λ_t (number of triggered earthquakes per unit time), we constrain the spatial density to be normalized:

$\int_0^\infty\int_0^\infty dx\,dy\,\lambda_s(x,y,m)=\int_0^\infty dr\,2\pi r\,\lambda_s(r,m)=1$. The initial guess for the rates / densities is done by democratically setting equal weights to all preceding earthquakes (including the background term),

with no condition on their magnitudes: $w_{0,j}=w_{i,j}=\frac{1}{j},\forall i<j$, and then starting the first iteration of the algorithm directly at step 2 (i.e., updating of the rates). We postulate that convergence is reached when the logarithm of the densities $\lambda_s(|\Delta\underline{x}|,m)$ or rates $\lambda_t(\Delta t,m)$ do not change by more than 1% for any given separation $|\Delta\underline{x}|$ or time Δt , between two consecutive iterations. This criterion is found to be well adapted: using more constraining criteria did not lead to visually different rate estimates.

The solution obtained is in close agreement with the theoretical rates (Fig. S2). The method is therefore well able to separate the individual triggering of each mainshock in a complex time series which involves multiple, cascading, triggering. Again, this estimation was performed with no a priori constraints on the shape of the rates / densities. Discretization in space, time and magnitude for

computing the kernels is the only arbitrary input.

A realisation of the background seismicity, i.e., the earthquakes that initiate clusters of aftershocks, is computed by using the background probabilities $w_{0,i}$. Such a realisation is compared to the total set of earthquakes (Fig. S1). The background seismicity is clearly devoid of aftershock clusters as expected, and occurs at a nearly constant rate density of 0.242, which is close to the true value $\lambda_0 = 0.25$. Monte-Carlo simulations were performed to check whether the estimated λ_0 is always correctly determined. Forty runs of the ETAS model (keeping the same parameters) were analysed, yielding an estimate of 0.248 +/- 0.010 for the background rate density. No significant correlation (at the 90% confidence level) is observed between the estimate of λ_0 and the total number of earthquakes produced in each simulation (that ranged from 1747 to 9539).

The convergence is equivalent to minimizing a cost function J, classically defined as the opposite of the logarithm of a likelihood function:

$$J = \lambda_0 \times S \times T + \sum_{i=1}^{N} \int_0^{T-t_i} dt \; \lambda_t(t, m_i) - \sum_{i=1}^{N} \log\left(\lambda_0 + \sum_{j=1}^{i-1} \lambda_t(t_i - t_j, m_j) \times \lambda_s(r_{ij}, m_j)\right)$$

, where T is the total duration of the earthquake catalog (Fig. S3). Changing the spatial discretization does not significantly affect the obtained temporal rates: even in the case of a purely temporal method, i.e., not using any information on distances between earthquakes, the final rates are close to the theoretical rates (Fig. S4).

Exploiting the spatial information mostly speeds up the convergence, as can be seen by comparing Figure S4 to Figure S3.

Finally, the solution does not depend on the initial choice for the kernel. In particular, the same final solution is obtained starting from the democratic kernel as defined above, or starting from the discretized theoretical kernel (Figs. S3 and S4). The latter is not the best solution: it has a cost function value J of the order of +100 compared to the best solution found during the convergence of the algorithm.

2 – Non-uniform spatial distribution of background earthquakes:

We modify the method described in the main text to account for the fact that background earthquakes do not densely cover the region, although we still consider that they occur at a stationary rate (i.e., forming a Poisson process). The background term $\lambda_0(\underline{x})=\lambda_{0,t}\times\lambda_{0,s}(\underline{x})$ is then the product of the constant rate $\lambda_{0,t}$ with the local density $\lambda_{0,s}(\underline{x})$ conditioned on the position $\underline{x}$ relative to the background earthquakes: $\lambda_{0,s}(\underline{x})=\sum_{i=1}^{N} g(\left|\underline{x}_i-\underline{x}\right|)w_{0,i}$, with the normalisation $\int_0^\infty dr\,2\pi r\,g(r)=1$.

The goal is then to estimate $g(r)$. This is done by modifying the two steps of the algorithm: (step 1) attribute the background weights $w_{0,j}=\alpha_j\lambda_0(\underline{x}_j)$, and (step 2) compute $\lambda_{0,t}=\frac{1}{T}\sum_{i=1}^{N} w_{0,i}$ and

$g(r)=\frac{\sum_{i,j\in A} w_{0,i}w_{0,j}}{S(r,\delta r)\sum_{i,j} w_{0,i}w_{0,j}}$, where A is the set of earthquakes pairs such that $\left|\underline{x}_i-\underline{x}_j\right|=r\pm\delta r$. Additionally, a correction for changes in completness magnitude in the first few hours after large mainshocks ($m\geq 6$) is introduced: each aftershock counts for $10^{b(m_c(t)-3)}$ events, where the b-value of the Gutenberg-Richter law (*16*) is 1.05, and $m_c(t)\geq 3$ is the completness magnitude at the time t of the aftershock.

3 – Analysis of California earthquakes using distances to faults:

We here perform the same analysis of California earthquakes as in the manuscript, but, instead of using inter-epicentral distances, we here rather use distances between the causative fault and the target hypocenter. This allows to account for the spatial anisotropy of the triggering.

Computing the distances:

The distance $r_{i,j}$ between earthquake i and earthquake j is taken to be the shortest distance from the fault plane of earthquake i to the hypocenter of earthquake j, or equal to the inter-hypocentral distance if the latter is shorter than the former. For all $m\geq 6$ earthquakes (8 earthquakes), the fault plane geometry is taken from rupture models deduced from inversion of seismic and geodetic data, see Table S1 for references. All source models are provided by Martin Mai's finite source rupture model database (www.seismo.ethz.ch/srcmod).

We assumed the same fault plane for the Hector Mine mainshock and its biggest (m=6.7) recorded aftershock. For all the other ($m<6$) earthquakes, the fault plane is computed as followed: we first selected earthquakes occurring within half a day following the considered mainshock and with epicentral distances less than twice the length $d=10^{0.5\times(m-4)}$ in km. Note that this scaling is consistent with the scaling found with the method when using inter-epicentral distances, see main manuscript. The fault plane is then defined as the plane passing through the mainshock and minimizing the distances, in the least-square sense, to all the selected earthquakes. Because rupture on the fault plane can be unilateral or bilateral, the center of the fault plane is chosen as the mean position of all the selected earthquakes, instead of the hypocenter of the mainshock. The size of the (square) fault plane is defined by its half-length d as defined above.

Results of the analysis:

We use the same correction coefficients as with the analysis using inter-epicentral distances. Figures S5 and S6 are the equivalent, for the new analysis, of Figures 1 and 2 of the main manuscript. In comparison to the previous results, we note the following changes and similarities:

- the bare rates are larger for the $m\geq 6$ mainshocks (Fig. S5 a). The bare productivity of those earthquakes is significantly increased, yielding a scaling $\chi\sim 10^{0.86m}$ (Fig. S6 b), hence a larger a parameter (increasing from 0.60 to 0.86). The dressed rates are very similar to the previous ones, at least for $t>0.1$ days; for shorter time intervals, they are increased for $m\geq 6$ mainshocks and

decreased for $m<6$. The dressed productivity scales as $\chi \sim 10^{0.73m}$, which is close to the $a=0.66$ found previously. This does not change our earlier conclusion that small mainshocks collectively trigger more aftershocks than their larger counterparts, since the exponent *a* is still significantly less than the b-value of the Gutenberg-Richter law for southern California ($b=1.05$, Fig. 3b).

- The mean aftershock sequence duration is magnitude dependent for both the bare and the dressed kernels, but the latter is 6 to 10 times longer than the former (Fig. S6 c).

- The background rate of $m \geq 3$ earthquakes is lowered down to 0.11 day^{-1}, giving a 13% probability of being a background earthquake at this magnitude cut-off.

- The spatial densities cannot be directly compared between the two analyses, as one is a 2D while the other is a 3D kernel. The bare densities (at least at short distances $r<10$ km) and the dressed densities decay with distance as $\lambda_s(x, y, z, m) \sim r^{-3}$, at all magnitudes (cf. dashed lines in Figure S5 d). The median distance from the rupture plane to the directly triggered aftershocks is very small, ranging between 1.0 km and 3.4 km, with no clear dependence on magnitude, implying that most direct aftershocks occur very close to the rupture plane. The mean triggering distances are increased by an average factor of 7 when considering the full aftershock cascade rather than only the directly triggered aftershocks (Fig. S6 d). As with the previous analysis using epicentral

distances, aftershock diffusion away from the mainshock is only observed for the dressed kernel, although the process is significantly sub-diffusive, i.e., the mean distance grows as $r \sim t^{0.20}$ (Fig. S6 e).

- Interestingly, all significant differences between the two analyses affect the bare kernel. The dressed kernel characteristics are nearly identical: almost the same exponents are obtained for the duration scaling, for the aftershock zone expansion, and for the productivity. Since most changes to the bare kernel are found for large mainshocks while small mainshocks keep the same bare characteristics, this clearly confirms that the dressed aftershock sequences are dominated by small shocks.

In conclusion, considering the distance to the causative fault rather than the simpler epicentral distance affects the results of the analysis, in particular by giving more weight to the large mainshocks. This does not however modify the main conclusions of the previous analysis, i.e., (1) small earthquakes collectively trigger more direct and indirect aftershocks than large earthquakes; (2) the cascade of triggering significantly extend aftershock sequences (here by a factor of about 7 to 10) both in time and space, and causes a slow diffusion of the aftershock zone; (3) the background contribution is small, of the order of 13% to 19% at $m_c = 3$, and is expected to decay to even lower values at smaller magnitude cut-off.

4 – Linearity hypothesis:

The main assumption of this method is the linearity of the triggering process: the collective triggering of a set of triggers is simply the sum of their individual triggering, which stays the same whatever the history of the process. This assumption goes against the rate-and-state friction model (*S8*), which is non-linear. However, accounting for non-linearity would require an underlying model, hence making the method model-dependent. There is unfortunately no simple way to evaluate the degree of non-linearity of the triggering process directly from the data. The scaling of the productivity parameter χ however gives some hints: in case of linear triggering, this scaling (parameter *a*) must be the same for bare and for dressed aftershocks. We find that $a = 0.60 \pm 0.07$ (bare) and $a = 0.66 \pm 0.04$ (dressed) when looking at epicentral distances (see main manuscript), and $a = 0.86 \pm 0.09$ (bare) and $a = 0.73 \pm 0.04$ (dressed) when using distances to the fault (see above). We therefore cannot reject the hypothesis of linearity based on the value of this parameter.

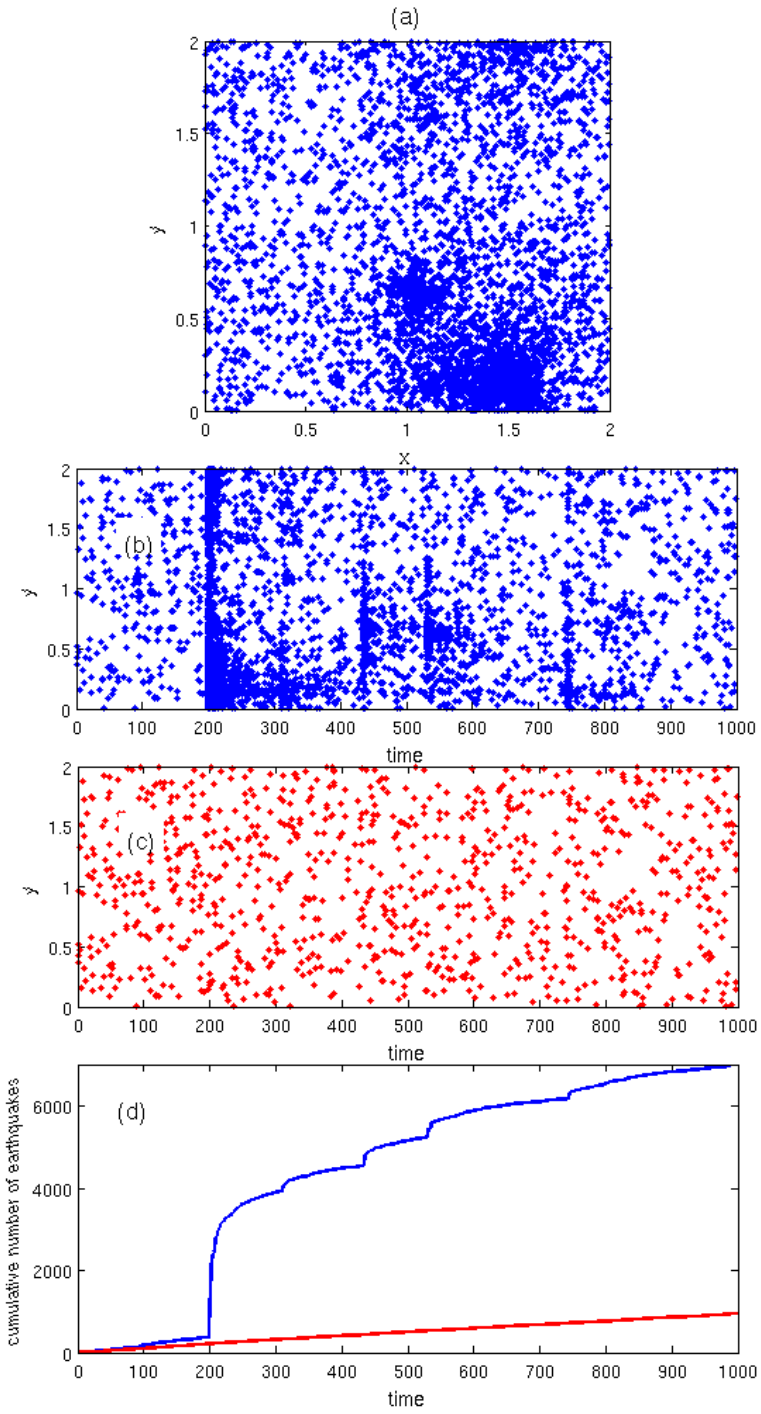

Figure S1: **synthetic earthquake catalogue.** (Blue) simulated catalogue using an ETAS model, compared to the catalogue declustered using the present method (in red). (a) Location of the epicentres; (b) and (c) space-time plots (along the y-direction); (d) cumulative time series. Aftershocks are efficiently removed by the declustering. The remaining background seismicity occurs at an estimated rate of 0.968 events / unit time, very close to the true 1 event / unit time imposed in the model.

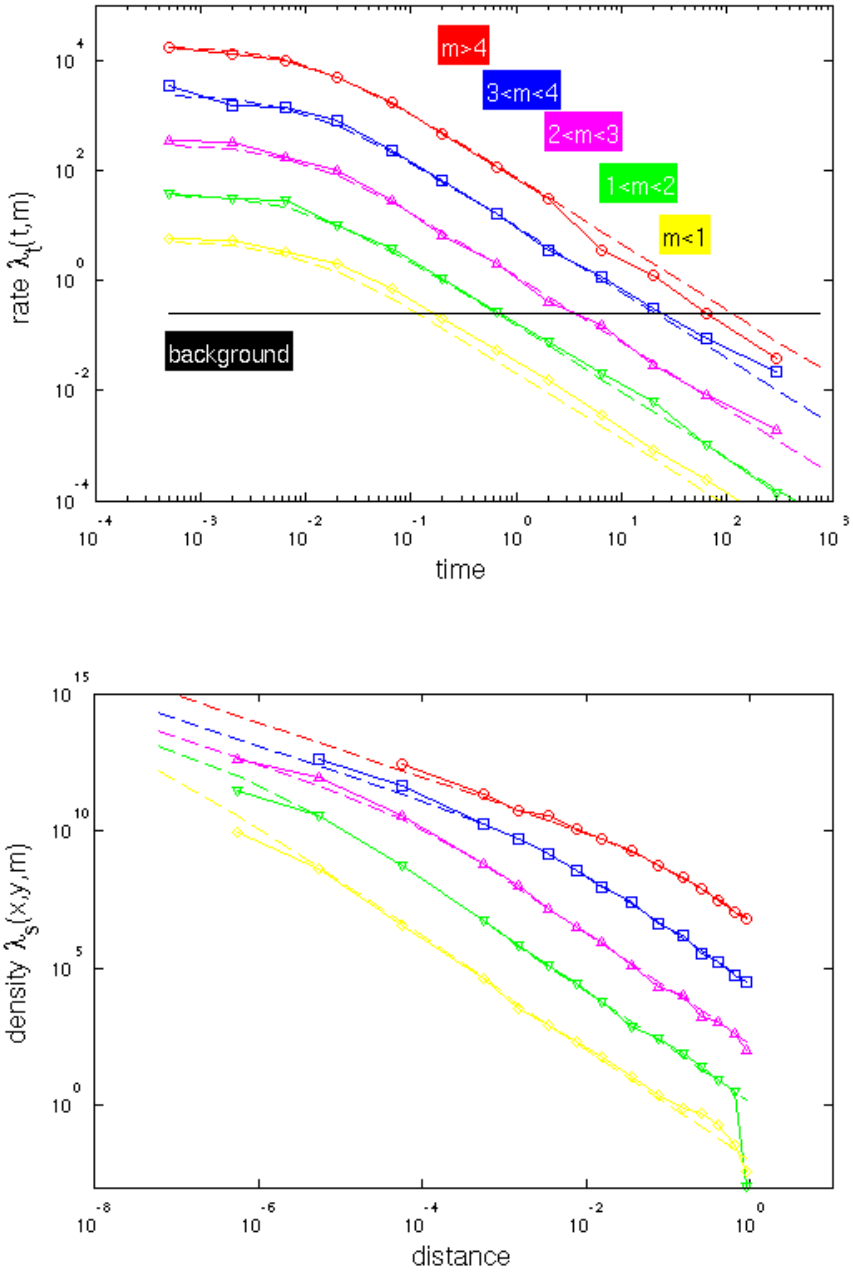

Figure S2: **estimated rates and densities for the synthetics.** Temporal rates $\lambda_t(t,m)$ (top graph, in earthquakes / unit time) and spatial densities $\lambda_s(x,y,m)$ (bottom graph, in earthquakes / unit area) estimated using the present method, compared to the theoretical $\lambda_t(t,m)=K\,e^{\alpha m}(t+c)^{-p}$ and $\lambda_s(x,y,m)=f(r,m)$ (dashed lines), for the synthetic catalogue of Figure S1. The spatial densities are shifted for clarity.

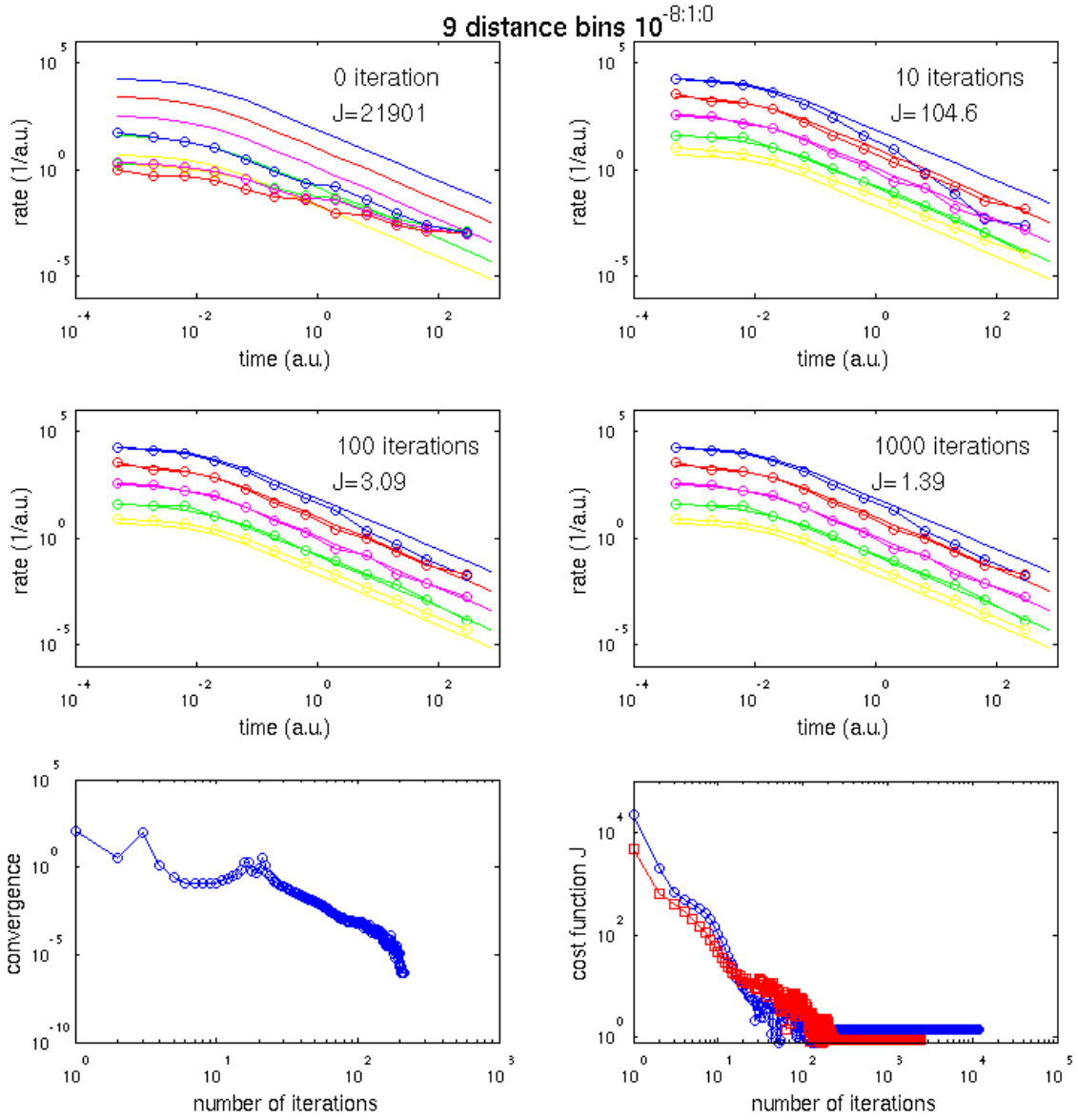

Figure S3: **snapshots of the temporal rates during convergence, for a space-time kernel discretized using 9 distance intervals.** Top four graphs: rate $\lambda_t(t,m)$ at starting point, and after 10, 100 and 1000 iterations, compared to the theoretical rates. Bottom left graph: convergence criterion vs. number of iterations. Bottom right graph: cost function J vs. number of iterations, for two different initial conditions (democratic kernel in blue, theoretical kernel in red). The cost J is shifted so that $J=0$ corresponds to the minimum value found during the convergence (the same for the two starting points).

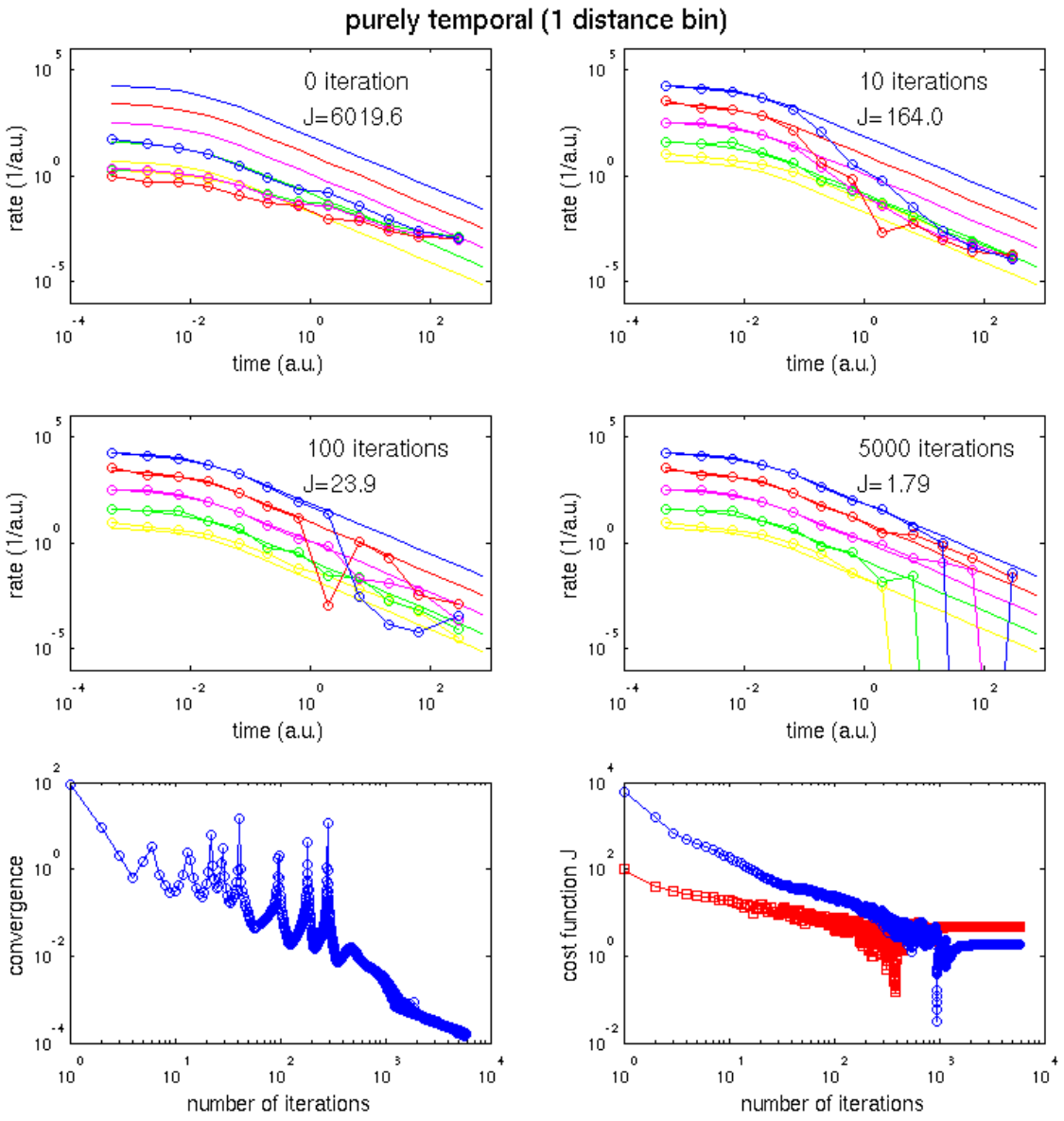

Figure S4: **snapshots of the temporal rates during convergence, for a purely temporal kernel.** Same as Figure S3, but for a kernel not using any information on distances between earthquakes. The snapshots are taken at starting point, and after 10, 100 and 5000 iterations.

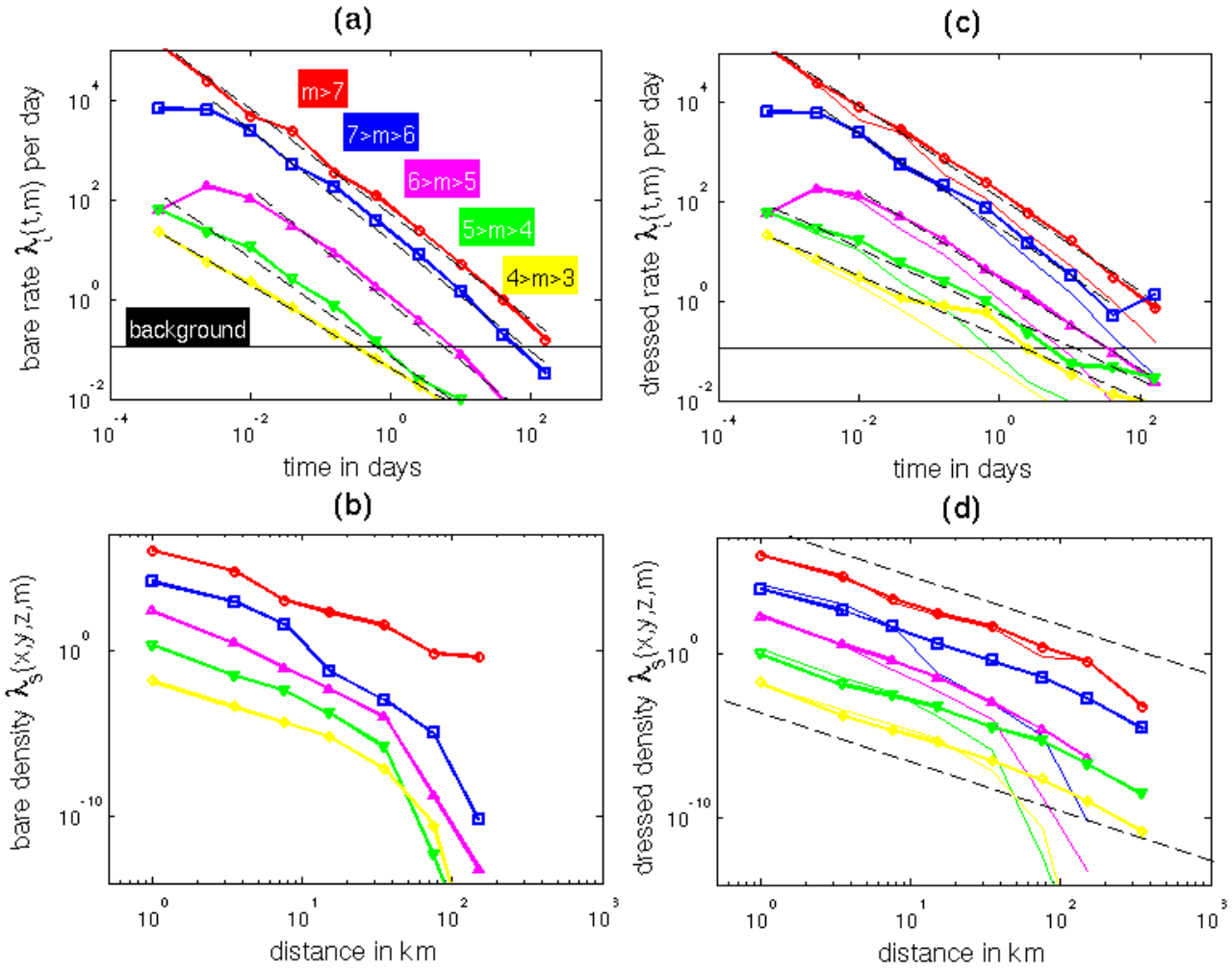

Figure S5: **rates and densities for California earthquakes using distances to faults.** As in Figure 1, for distance between earthquakes computed as the shortest distance from the causative fault to the target hypocenter (rather than the epicentral distance). The dashed lines in graph (d) indicate a $1/r^3$ scaling.

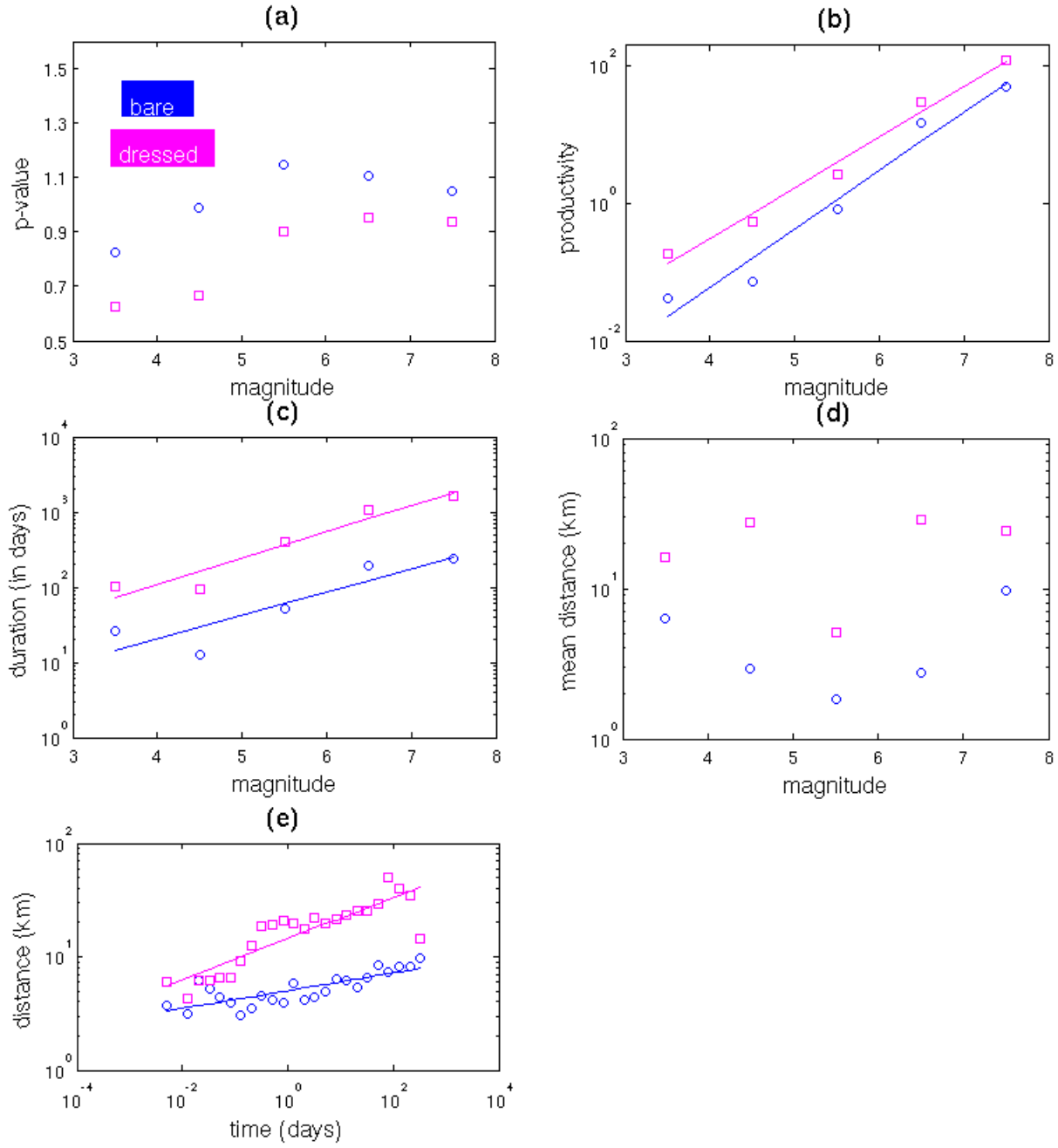

Figure S6: **aftershock sequence characteristics for California earthquakes usin distances to faults.** As in Figure 2, for distance between earthquakes computed as the shortest distance from the causative fault to the target hypocenter (rather than the epicentral distance). The fits are: (b) productivity $\chi \sim 10^{0.86m}$ (bare rate) and $\chi \sim 10^{0.73m}$ (dressed rate), (c) mean duration $t \sim 10^{0.31m}$ (bare) and $t \sim 10^{0.35m}$ (dressed), (e) mean mainshock – aftershock distance $r \sim t^{0.07}$ (bare) and $r \sim t^{0.20}$ (dressed).

Date	*Location*	*Magnitude*	*Reference*
November 23, 1987	Elmore Ranch	6.2	S6
November 24, 1987	Superstition Hill	6.6	S6
April 22, 1992	Joshua Tree	6.1	S3
June 28, 1992	Landers	7.3	S7
June 28, 1992	Big Bear	6.3	S4
January 17, 1994	Northridge	6.7	S7
October 16, 1999	Hector Mine	7.1	S5
October 16, 1999	Hector Mine aftershock	6.7	S5

Table S1: references to rupture models for all m ≥ 6 earthquakes.

References:

S1. Y. Y. Kagan, L. Knopoff, *J. Geophys. Res.* **86**, 2853 (1981).

S2. Y. Ogata, *J. Am. Stat. Assoc.* **83**, 9 (1988).

S3. S. E. Hough, D. S. Dreger, *Bull. Seismol. Soc. Am.* **85**, 1576 (1995).

S4. L. E. Jones, S. E. Hough, *Bull. Seismol. Soc. Am.* **85**, 668 (1995).

S5. A. Kaverina, D. Dreger, E. Price, *Bull. Seismol. Soc. Am.* **92**, 1266 (2002).

S6. S. Larsen, R. Reilinger, H. Neugebauer, W. Strange, *J. Geophys. Res.* **97**, 4885 (1992).

S7. Y. Zeng, J. Anderson, Evaluation of numerical procedures for simulating near-fault long-period ground motions using Zeng method, *Report 2000/01 to the PEER Utilities Program, available at http://peer.berkeley.edu* (2000).

S8. J. Dieterich, *J. Geophys. Res.* **99**, 2601 (1994).