An edited version of this paper was published by AGU. Copyright (2013) American Geophysical Union.
DOI: 10.1002/jgrb.50306

Triggering cascades and statistical properties of aftershocks

Chad Gu[1], Aicko Y. Schumann[1], Marco Baiesi[2,3] and Jörn Davidsen[1]

Abstract. Applying a simple general procedure for identifying aftershocks, we investigate their statistical properties for a high-resolution earthquake catalog covering Southern California. We compare our results with those obtained by using other methods in order to show which features truly characterize aftershock sequences and which depend on the definition of aftershocks. Features robust across methods include the p-value in the Omori-Utsu law for large mainshocks, Båth's law, and the productivity law with an exponent smaller than the b-value in the Gutenberg-Richter law. The identification of a typical aftershock distance with the rupture length is a feature we confirm as well as a power law decay in the spatial distribution of aftershocks with an exponent less than 2. Other results we obtain, but not common to all other works including *Marsan and Lengliné* [2008]; *Hainzl and Marsan* [2008]; *Zhuang et al.* [2008], are (a) p-values that do not increase with the mainshock magnitude, (b) the duration of bare aftershock sequences that scales with the mainshock magnitude, (c) an additional power-law in the temporal variation, at intermediate times, in the rate of aftershocks for mainshocks of small and intermediate magnitude and (d) a b-value for the Gutenberg-Richter law of background events that is sensibly larger than that of aftershocks. Tests on synthetic catalogues generated by the epidemic-type aftershock sequence model corroborate the validity of our approach.

1. Introduction

One of the hallmarks of seismicity is the clustering of earthquakes in space and time. This is particularly evident from the observation that the local rate of seismic activity increases significantly after large earthquakes. From a physical perspective, the spatiotemporal clustering of earthquakes typically indicates that the vast majority of them are triggered by the preceding ones due to static or dynamic stress changes, fluid flow, afterslip and/or other mechanisms [*Main*, 2006; *van der Elst and Brodsky*, 2010]. While it is known that earthquakes, independent of their sizes, can trigger other earthquakes, one commonly denotes the largest earthquake in such a cluster as the mainshock and all following (preceding) ones as aftershocks (foreshocks). As far as we know, there is no physical distinction in the *relaxation* mechanism between mainshocks, foreshocks and aftershocks [*Houghs and Jones*, 1997; *Helmstetter and Sornette*, 2003] such that the classification of earthquakes relies on correctly identifying *triggering* relationships between them. Moreover, some aftershocks classifications not only include those earthquakes that are *directly* triggered by the mainshock — the first generation of aftershocks — but also earthquakes that are *indirectly* triggered, e.g., earthquakes directly triggered by other triggered events. Thus, a sequence of aftershocks corresponds to a cascade of triggering that extends the reach of the initial mainshock. A long-standing challenge is to determine which earthquakes are connected, either directly or indirectly, and to identify aftershock sequences based on causation [*Parsons and Velasco*, 2009].

To separate earthquakes between mainshocks and aftershocks, several declustering algorithms have been proposed in the past [*Gardner and Knopoff*, 1974; *Keilis-Borok et al.*, 1980; *Reasenberg*, 1985; *Davis and Frohlich*, 1991; *Molchan and Dmitrieva*, 1992; *Zhuang et al.*, 2002, 2004; *Baiesi and Paczuski*, 2004, 2005; *Zaliapin et al.*, 2008; *Marsan and Lengliné*, 2008]. Some of these methods have arbitrary rules and are rich in parameters, calling their objectivity in question. For example, declustering schemes based on a sharp space-time aftershock window following a large earthquake may suffer from the loss of long-range triggering and/or the inappropriate selection of the window shape. To overcome such problems, a number of methods trying to estimate earthquake correlations were proposed [*Zhuang et al.*, 2002, 2004; *Baiesi and Paczuski*, 2004, 2005; *Zaliapin et al.*, 2008; *Marsan and Lengliné*, 2008]. It is this class of methods that we consider here. Specifically, we choose to study the algorithm proposed by *Baiesi and Paczuski* [2004] (BP), which is able to distinguish between populations of events triggered by other events and background activity as shown by *Zaliapin et al.* [2008]. We compare the statistical properties of aftershocks obtained in this way with those observed using other methods, especially the schemes proposed by Zhuang, Ogata and Vere-Jones (ZOV) [*Zhuang et al.*, 2002, 2004] and by *Marsan and Lengliné* [2008] (ML). While there are some formal similarities between these techniques, there are also significant differences: In the BP approach, one defines a space-time-magnitude nearest-neighbor distance between earthquakes, based on phenomenological laws of seismicity related to the Gutenberg-Richter law, while ZOV assume that Epidemic-Type-Aftershock-Sequence (ETAS) models with observed phenomenology at the level of earthquake activity can be used to describe seismicity; finally, ML implemented an iterative algorithm based on a linear superposition of triggering rates of events that does not rely on a specific model. A detailed comparison of our results with those obtained using the other methods allows us to identify statistical properties of aftershocks sequences that are apparently robust and well-defined across methods. However, we also show that several statistical features depend sensitively on how one

[1]Complexity Science Group, Department of Physics and Astronomy, University of Calgary, Canada
[2]Dipartimento di Fisica e Astronomia, Università degli Studi di Padova, Italy
[3]INFN, Sezione di Padova, Italy

0148-0227/13/$9.00

defines and identifies aftershocks. An additional motivation of our work is to test whether the results in the BP scheme are robust with respect to variations in the method itself.

The outline of the paper is as follows: in Section 2, we discuss in detail how we identify a trigger of a given earthquake. One particular focus is on the difference between directly and indirectly triggered events which allows one to define aftershocks in at least two different ways. In Section 3, we introduce the ETAS model which we use in the following sections to i) establish the robustness of the applied methodology and ii) for comparison with our analysis of a high-resolution catalog from Southern California. Section 1 presents our findings for the rate of aftershocks after a mainshock of a given magnitude. In particular, we discuss aftershock productivity. The temporal duration of aftershock sequences is investigated in Section 5. In Section 6, the frequency-magnitude distribution of aftershocks and background events are analyzed. Our findings regarding Båth's law are presented in Section 7. We discuss in detail the spatial distribution of aftershocks in Section 8 as well as its implication for the ongoing discussion on static and dynamic triggering of aftershocks. Finally, we present a discussion and our conclusions in Section 9.

2. Aftershock identification

To identify aftershocks, we follow the method first proposed by *Baiesi and Paczuski* [2004]. It is based on a metrics quantifying space-time-magnitude distances between seismic events in order to identify the nearest neighbor for any given earthquake. The distribution of all these nearest neighbor distances then allows one to identify potential triggering relationships [*Zaliapin et al.*, 2008]. Thus, this procedure is similar to other recent approaches used to identify functional networks in complex systems [*Eguíluz et al.*, 2005]. The general goal is to sparsify a weighted network — here described by the matrix of space-time-magnitude distances between all earthquakes — thereby keeping only a backbone composed by the set of "significant" correlations. The process of weight assignment must be based on a probabilistic argument, and thus should conform to the empirical laws appropriate to the systems of concern.

For our specific case, consider a time-ordered earthquake catalogue given as $(t_j, \vec{x}_j, m_j)$, where t_j is the time of occurrence of the j-th event, $\vec{x_j}$ its epicenter or hypocenter, m_j its magnitude and $1 \leq j \leq N$. For every earthquake j, we compute the expected number of events for every precedent index i [*Baiesi and Paczuski*, 2004, 2005; *Baiesi*, 2006]:

$$n_{ij} = c \cdot (r_{ij})^{D_f}\, t_{ij} 10^{-b\, m_i} \qquad (i < j)\,\&\,((t_{ij} - r_{ij}/\bar{v}) \geq 0), \tag{1}$$

where $\bar{v}$ is the propagation speed and the conditions ensure causality. This expression is based on the empirical Gutenberg-Richter law, which states that the expected number of earthquakes N bigger than a given magnitude m in a given (large) area over a given (large) time interval follows $N = a10^{-b\,m}$ with $b \approx 1$ [*Gutenberg and Richter*, 1949]. The prefactor a depends on the size of the considered area and on the length of the considered time interval [*Turcotte*, 1997]. On average, a increases linearly with the length of the time interval and approximately as a power law with fractal dimension D_f with the linear size of the considered area. These dependencies are directly reflected in Eq. (1), where $t_{ij} = t_j - t_i$ and $r_{ij} = d(\vec{x}_i, \vec{x}_j)$ are the time separation and spatial distance, respectively. It is important to realize, though, that the spatial distribution of seismicity is not a simple fractal but shows a) signs of multifractality if one considers epicenters [*Davidsen and Goltz*, 2004] and b) strong finite size effects and depth dependence if one considers hypocenters [*Kagan*, 2007]. Thus, there is huge uncertainty in the choice of D_f. To address this point, we consider different values of D_f centered around $D_f = 2.2$ in the following. The latter value has been estimated as the asymptotic correlation dimension of hypocenters for shallow-crust seismic activity in Southern California, which takes place at depths less than 30 km [*Kagan*, 2007]. Note that we explicitly focus on hypocenters and three-dimensional distances between them, allowing us to analyze values of $D_f > 2$ in contrast to [*Baiesi and Paczuski*, 2004; *Zaliapin et al.*, 2008].

In general one quantifies correlations by detecting deviations from randomness. In other words, one states a null hypothesis of uncorrelated events and measures departures from that [*Baiesi and Paczuski*, 2005]. In seismicity, based on Eq. (1), we can identify the most likely candidate i that triggered event j as the one that minimizes the expected number of earthquakes n_{ij}, since it is the least probable event pair to have actually occurred at random. The collection of all minimal pairs or nearest neighbors $\{k, j\}$, where k satisfies $n_{kj} = \min_{i<j}\{n_{ij}\} \equiv n_j^*$ for a given j, constitutes the pool of *possible* triggering relations. From this pool, one can identify those earthquake pairs for which the triggering relations are indeed statistically significant as shown by *Zaliapin et al.* [2008]. This allows one to separate earthquakes into two classes as discussed in detail below: triggered events and non-triggered or background events. This classification sets the BP method apart from other probabilistic methods used for aftershock identification including the schemes proposed by Zhuang, Ogata and Vere-Jones (ZOV) [*Zhuang et al.*, 2002, 2004] and by *Marsan and Lengliné* [2008] (ML). In the ZOV and the ML scheme, each event is assigned a set of non-zero probabilities. For a fixed event, each of these probabilities quantifies the likelihood of a specific preceding event being the trigger of the fixed event, with the exception of a single probability which quantifies the likelihood of the considered event being a background event. While the assignment of these probabilities in the ZOV and the ML scheme is formally similar to the inverse space-time-magnitude metric given in Eq. (1) — see, for example, [*Zhuang et al.*, 2004] — both schemes involve an iterative estimation of such a metric itself in the process. For the BP method, the estimation of the parameters quantifying the Gutenberg-Richter law in Eq. (1) is performed independently and discussed below.

2.1. Southern California

Here, we apply the BP analysis described above to seismic data obtained through the Southern California Seismic Network, where seismic activities from 1984 to 2005, within an area of approximately $356,528$ km^2 have been compiled [*Lin et al.*, 2007]. The events in the current catalogue have been relocated by computing waveform cross-correlation from seismographic data, yielding higher spatio-temporal resolution of low-magnitude events in active regions closer to the detection sources. We assumed completeness of measurements for all events above a lower magnitude cutoff $m_{th} = 2.5$ [*Wiemer and Wyss*, 2000; *Schorlemmer and Woessner*, 2008] such that we are left with $N = 27539$ events and $N - 1$ minimal pairs. In computing the $\{n_{ij}\}$'s according to Eq. (1), we also applied a lower cutoff for the inter-event time to avoid unphysical selection, keeping only inter-event times $t_{ij} > r_{ij}/\bar{v}$. Here, $\bar{v} = 6$ km/s is the propagation speed of P-waves in the Earth's crust, averaged from the seismic wave velocity model used by Lin *et al.* for the relocation procedure in the catalog [*Lin et al.*, 2007]. We also set $b = 1.09$, since an estimate for this catalog based on the maximum likelihood method discussed in [*Naylor et al.*, 2009] gave $b = 1.09 \pm 0.07$.

From the set $\{n_j^*\}$ one can identify two statistically distinct populations of earthquake pairs by defining the weighted relative distances in space and time, respectively:

$$\begin{aligned} \tau_{kj} &= t_{kj} 10^{-bm_k/2}, \\ l_{kj} &= r_{kj}^{D_f} 10^{-bm_k/2}. \end{aligned} \tag{2}$$

This follows from Fig. 1 which shows density plots of the set $\{n_j^*\}$ split into its τ_{kj} and l_{kj} contributions as first proposed by *Zaliapin et al.* [2008]. Note that $n_j^* = \tau_{kj} l_{kj}$ where we have set $c = 1$ without loss of generality. The denser regions in Fig. 1 correspond to the most probable occurrences of weighted relative time and distances of all extremal pairs, and a bimodal structure is clearly visible for all D_f. The existence of this bimodality indicates that the triggering relations of event pairs with $\log n_j^* < \log n^*$ (indicated by the straight solid lines in Fig. 1) are statistically significant while event pairs above the threshold n^* are not [*Zaliapin et al.*, 2008]. Thus, we may naturally identify all events j with $\log n_j^* < \log n^*$ as triggered events, while we can consider all other events — which have not been triggered by events in the catalog — as *background* events. All results presented in the following do not depend significantly on the exact choice of n^* as long as it correctly separates the bimodal structure in the density plots, see also [*Zaliapin and Ben-Zion*, 2013].

Due to the bimodality, we can reduce the number of potential triggering relationships to those that are significant. These remaining pairs $\{k, j\}$ of a triggering earthquake k and the event j it triggered allow us to represent seismicity in the form of triggering cascades. Mathematically speaking, seismicity can be represented by a sparse directed network, which consists of disconnected trees: each tree corresponds to a triggering sequence and the root of each tree corresponds to an event j, for which $n_j^* > n^*$. These events are background events, which are not triggered by any preceding earthquake in the catalog. For a given event, we define its *first generation aftershocks* as those directly triggered by it. In contrast, the full set of aftershocks of a given event consists of all directly or indirectly triggered events, where the latter is already intrinsic in the transitivity of the directed network structure (if C is triggered by B, and B is triggered by A, then C is indirectly triggered by A). Following the terminology in related literature, these are called the *bare* and *dressed* aftershock sequences, respectively [*Helmstetter and Sornette*, 2002; *Marsan and Lengliné*, 2008]. Note that our definition of aftershocks is *independent* of the magnitudes of the triggered events — similar to the approach by *Marsan and Lengliné* [2008] but in contrast to many other definitions that include the notion of foreshocks (see, for example, *Zaliapin and Ben-Zion* [2013]). In particular, *every* earthquake that triggers another one is considered a mainshock.

Similar to the findings in [*Zaliapin et al.*, 2008], the value of $\log n^*$ does not depend on m_{th} (not shown). Holding the fractal dimension fixed ($D_f = 2.3$), we found the same threshold values n^* in each case, by varying the lower magnitude threshold from $m_{th} = 2.5$. This is particularly true for $m_{th} = 3.0$ and $m_{th} = 3.5$, which we use for some analysis in the subsequent sections. We also would like to point out that, for $D_f \leq 2$, there are no qualitative differences between using hypocenters and three-dimensional Euclidean distances on one hand and using epicenters and geodesic distances on the surface of a sphere on the other hand.

3. Synthetic ETAS catalogs

To further establish the reliability of the above BP method to identify triggering cascades and recover their statistical properties as well as for comparison with our findings for Southern California, we also study surrogate catalogs generated by the Epidemic-Type Aftershock Sequence (ETAS) model [*Kagan and Knopoff*, 1987; *Helmstetter and Sornette*, 2002; *Ogata and Zhuang*, 2006; *Peixoto et al.*, 2010]. The ETAS model and related models (as, for example, [*Turcotte et al.*, 2007]) are based on the assumption that any earthquakes can trigger other earthquakes. This is quantified in terms of triggering rates which depend on space and time such that the ETAS model falls into the class of non-homogeneous spatio-temporal Poisson processes. The overall rate of activity is given by

$$\lambda(\vec{r}, t) = \lambda_b(\vec{r}) + \sum_{t_i \leq t} \phi_{m_i}(t - t_i, \vec{r} - \vec{r}_i), \tag{3}$$

where t_i, $\vec{r}_i$, and m_i refer to the time, location, and magnitude of earthquake i, respectively. $\phi_{m_i}(t-t_i, \vec{r}-\vec{r}_i)$ describes the occurrence rate of earthquakes triggered by event i while $\lambda_b(\vec{r})$ is the spatially varying background activity. For the ETAS model, the occurrence rate is defined as

$$\phi_{m_i}(t - t_i, \vec{r} - \vec{r}_i) = \rho(m_i)\Psi(t - t_i)\Phi_{m_i}(\vec{r} - \vec{r}_i), \tag{4}$$

where $\rho(m_i)$ is the number of directly triggered earthquakes (which is a random variable itself and its distribution depends on m_i), $\Psi(t - t_i)$ is the normalized temporal distribution of directly triggered events and $\Phi_{m_i}(\vec{r} - \vec{r}_i)$ is the normalized spatial distribution of directly triggered events in two dimensions. The functional form of these three distributions are derived from empirical observations.

To generate a surrogate catalog based on the ETAS model, we consider a time interval $[0, T]$ and a square spatial area of size $L \times L$ with periodic boundary conditions. The triggering activity is initialized by independent, spatially non-uniformly- and temporally Poisson-distributed background earthquakes with an average rate of $\langle \lambda_b \rangle$. The magnitudes, m_i, for each earthquake (background and triggered events) are drawn from a normalized Gutenberg-Richter probability density function

$$p(m_i) = b \ln(10) \cdot 10^{-b(m_i - m_0)} \tag{5}$$

with lower-threshold magnitude m_0 such that $m_i \geq m_0$. After initialization, triggered events are added time-progressively such that the productivity of each earthquake is Poisson-distributed with average

$$\langle \rho(m_i) \rangle = K 10^{\alpha(m_i - m_0)}, \tag{6}$$

where α and K are constants. The waiting times, $t - t_i$, between an event at time t_i and the earthquakes it triggers at times $t > t_i$ are distributed according to the normalized Omori-Utsu law [*Utsu et al.*, 1995]

$$\psi(t - t_i) = \frac{\theta c^\theta}{[(t - t_i) + c]^{1+\theta}}, \tag{7}$$

with $c > 0$ and $\theta > 0$. Following recent empirical studies [*Felzer and Brodsky*, 2006], triggered events are assumed to be distributed in space according to a probability density function $\Phi_{m_i}(\vec{r} - \vec{r}_i)$ that is isotropic and reduces to

$$\varphi_{m_i}(|\vec{r} - \vec{r}_i|) = \frac{\mu |\vec{r} - \vec{r}_i|}{l(m_i)^2 \left(\frac{|\vec{r} - \vec{r}_i|^2}{l(m_i)^2} + 1 \right)^{1+\mu/2}} \tag{8}$$

for the distribution of distances $|\vec{r} - \vec{r}_i|$. Here, $l(m_i)$ is the rupture length of the trigger given by $l(m_i) = l_0 10^{0.45 \cdot m_i}$ [*Peixoto et al.*, 2010] and μ is another constant. To minimize the effect of catalog incompleteness due to missing aftershocks triggered by events that occurred before the simulation period, we remove all data in the time interval $[0, t_c]$ from the catalog. t_c and all other parameter values we used here to generate ETAS catalogs are given in Table 1. The specific values were chosen to closely resemble those empirically observed for Southern California as was the spatially varying background rate. In the following, we mostly

Table 1. Parameters used for simulating the ETAS model (see Section 3 for details). They are chosen such that the surrogate catalogs resemble the observed activity in Southern California.

	Simulation time & space			Background earthquakes	Gutenberg-Richter law		Omori law		Aftershock productivity		Epicenter distribution	
Parameter	T	t_c	L	$\langle\lambda_b\rangle$	m_0	b	c	θ	K	α	μ	l_0
Unit	days	days	km	day^{-1}	–	–	days	–	–	–	–	m
Value	8000	365	600	1	2.5	1.09	0.024	0.2	0.155	0.9	0.6	15

focus on one specific ETAS catalog not only as an example but also for comparison with the catalog from Southern California. This ETAS catalog is also provided in the Supplementary Material.

As shown by *Zaliapin et al.* [2008], the BP method described in Section 2 allows one to identify two statistically distinct populations of events for the ETAS model corresponding to triggered and non-triggered or background earthquakes. This is confirmed by Fig. 2. Due to the statistical nature, there is typically some overlap between the two populations such that a perfect separation is not possible. Yet, we show below that this does not significantly affect the vast majority of statistical properties of the two classes for the relevant parameter values in the ETAS model.

4. Omori-Utsu law & aftershock productivity

In the language that clearly necessitates a definition of triggered earthquakes, the Omori-Utsu law is the empirical observation that the rate of aftershocks following a large earthquake scales as [*Utsu et al.*, 1995]

$$r(t) = \chi \cdot (t + c)^{-p}, \tag{9}$$

where t is the time after the mainshock. Though its validity has been documented in many cases, the exact nature of its parameters and their dependencies are not well-known. For example, the exact origin of the parameter c is still debated. The most recent results suggest that it is of physical rather than instrumental origin [*Kagan and Houston*, 2005; *Peng et al.*, 2006, 2007; *Peng and Zhao*, 2009] but a generally accepted physical interpretation of its meaning is still lacking despite some progress [*Narteau et al.*, 2009]. In particular, its quantification might sensitively depend on how one defines aftershocks [*Hainzl and Marsan*, 2008]. Moreover, its estimate is often affected by short-term aftershock incompleteness [*Kagan*, 2004]. In the limit $t \gg c$, one can instead study a rate equation of the form

$$r(t) \sim \chi \cdot t^{-p}, \tag{10}$$

which avoids the above complications.

4.1. Results for Southern California

Fig. 3 shows both the bare and dressed aftershock rates for the three largest earthquakes in the Southern California region over the time-span of the catalog, computed by applying the BP method for different fractal dimensions. In all cases, the rates can be well described by the Omori-Utsu law. Moreover, $p \approx 1.25$ in all cases as indicated by the insets. As expected, the value of χ is higher for the dressed aftershock rates compared to the bare rate.

Since the importance of small distances r_{ij} on identifying the trigger of a given earthquake increases with D_f as evident from Eq. (1) and Fig. 1, the value of D_f can have some impact on the statistical properties of triggering cascades as well. While p seems to be invariant with respect to the choice of D_f for these largest earthquakes, we find that this is not true for smaller mainshocks. As Fig. 4 shows, a power-law decay for sufficiently large t might not even be a good description for $D_f = 1.6$ if we consider the average rates of aftershocks directly triggered by mainshocks of magnitude $4.0 \leq m \leq 4.5$. Yet, Fig. 4 also shows that there are no significant differences in the bare aftershock rates for $2.0 \leq D_f \leq 3.0$. Hence, it makes sense to use the value $D_f \approx 2.2$ estimated for Southern California for calculating the metrics given by Eq. (1) and to remember that the relative results are not very sensitive to variations in this parameter. This also indicates that the influence of finite size effects (or even multifractality if present) as discussed in Section 2 is negligible.

Fig. 4 also allows a comparison between the bare and the dressed aftershock rate for $D_f = 2.3$. At $t < 10^{-2}$ day, only the directly triggered aftershocks contribute and thus we see an agreement between the bare and dressed rates. As expected, over the domain $t > 1$ day the dressed rates are much higher than the bare rates. Yet, as in the case of the largest earthquakes (see Fig. 3) we find in the same region that both the bare and dressed rates decay with a similar exponent $p \approx 1.2$ (see inset of Fig. 4). The agreement between the two rates in the above regimes is also valid for mainshocks over other magnitude ranges as follows from Fig. 5. On the other hand, in the intermediate regime $10^{-2} < t < 1$ days, the functional form of the dressed rates necessarily deviates from that of the bare rates. Even if short-term aftershock incompleteness of the catalog were to play a role in our specific observations, the differences between bare and dressed rates are necessarily generic. This implies that the exact functional form used to describe aftershock rates will depend on whether one includes higher-order aftershocks or not as confirmed by *Helmstetter and Sornette* [2002]; *Marsan and Lengliné* [2008]; *Hainzl and Marsan* [2008]. Our findings show, however, that the slope of the power-law decay in the rates for large t is independent of this choice and, thus, a robust feature.

Focusing on the robust behavior of the aftershock rates for $t > 1$ day, we can investigate the behavior of $p(m)$ and $\chi(m)$ as a function of mainshock magnitude m — similar to [*Hainzl and Marsan*, 2008]. From the data shown in Fig. 5, we obtained for 100 days $> t > t_0$ the best power law fits to Eq. (10) for aftershocks a) directly triggered (bare rates, $t_0 = 1$ day) and b) both directly *and* indirectly triggered (dressed rates, $t_0 = 4$ days) by earthquakes within a given magnitude range. The estimated parameters p and χ are shown in Fig. 6. For all three values of D_f, we observe a decreasing trend in $p(m)$ for the bare aftershock rates. The trend becomes less and less pronounced for larger D_f. In contrast, the dressed p-value appears to be independent of m for $D_f = 2.3$ in Fig. 6. The scaling of the productivity parameter χ, however, is unaffected by inclusion of indirectly triggered aftershocks: χ scales as $10^{\alpha m}$ for sufficiently large magnitudes where $\alpha \approx 0.9$ in both the bare and dressed cases.

Instead of evaluating the aftershock productivity by estimating χ, we can also analyze the total number of aftershocks, N_{as}, directly triggered by an event with a given

magnitude m. This number is expected to scale on average as $N_{as} \sim 10^{\alpha m}$ as well [*Baiesi and Paczuski*, 2005]. Indeed, Fig. 7 confirms this asymptotic scaling giving an estimate of $\alpha \approx 0.85$ for $D_f = 2.3$ for both bare and dressed aftershocks. Note that α does not depend on the lower magnitude cutoff, m_{th}, applied to the catalog. This follows from the inset of Fig. 7.

4.2. Results for ETAS catalogs

Fig. 8 shows that the most important parameters of the Omori-Utsu law — including those of the productivity law — can be recovered. In particular, panels A and B of Fig. 8 indicate that both the bare and dressed aftershock rates of the ETAS catalog can be fitted very well by Eq. (10) — independent of the specific magnitude ranges of the mainshock. More importantly, we find that $p = 1 + \theta$ within the uncertainties for large mainshocks indicating that the aftershock identification method allows us to successfully recover the parameter θ of the ETAS model. Panel C of Fig. 8 further indicates that there is a tendency to overestimate p for mainshocks of smaller magnitude, especially for directly triggered aftershocks. Yet, this feature is much less pronounced or even absent if we consider variations of the ETAS model for which the spatial distribution of triggered events, $\Phi_{m_i}(\vec{r} - \vec{r}_i)$, has a slightly different form or is anisotropic [*Moradpour et al.*, 2013], which is the more realistic scenario.

The expected exponential increase of χ with mainshock magnitude (see Eq. (6)) is documented by panel D of Fig. 8. Specifically within the uncertainties, we recover the ETAS parameter $\alpha = 0.9$ that characterizes the aftershock productivity. Instead of evaluating only the scaling of the aftershock productivity by estimating χ, we can directly analyze the total number of aftershocks, N_{as}, triggered by an event with a given magnitude m and compare it to Eq. (6). As Fig. 9 shows, both agree very well for the bare case. While the productivity is necessarily higher in the dressed case, one can observe a very similar behavior. These findings provide additional evidence that the BP method described in Section 2 gives reliable estimates of the statistical properties of triggering cascades or aftershocks.

4.3. Summary

The analysis of the ETAS catalogs shows that the BP method can reliably recover key parameters of the Omori-Utsu law and the productivity law such as p and α. The analysis of the catalog from Southern California further shows that the observed statistical properties of aftershocks are robust with respect to variations in the BP method itself as long as one considers hypocenters and a fractal dimension close to the estimated correlation dimension. While our analysis provides clear evidence for $\alpha \approx 0.9$ and, thus, $\alpha < b$ for Southern California, the presence of a variation in p with mainshock magnitude is not absolutely clear. For Southern California there is a monotonically decreasing trend in $p(m)$ (see Fig. 6) which, for $D_f > 2.0$, is similar to the bias observed in the ETAS catalog (see panel C in Fig. 8). This in combination with the true uncertainties of the fits and the possible role of spatial anisotropies discussed in Section 4.2 indicates that $p(m)$ is either slightly decreasing with m or constant around 1.2. Indeed, the dressed p-value appears to be independent of m for $D_f = 2.3$ in Fig. 6. Alternatively, this observation could indicate a potential statistical difference between the first generation of aftershocks and the full set of aftershocks.

We also observe deviations from the Omori-Utsu law for intermediate and small mainshocks. For Southern California, the inset of Fig. 4 shows that the power-law decay as described Eq. (10) with $p \geq 1.2$ does not hold for $t < 1$ day. For the bare rates, there is evidence of another power-law decay with exponent ≈ 1.0 in the domain $10^{-2} < t < 1$ day. While this could be an artifact due to short-term aftershock incompleteness — we do not correct for this incompleteness in any way — a recent analysis of high-frequency waveforms in Japan has shown that only time scales less than 10^{-2} days after mainshocks of magnitude 3 to 5 are typically affected by this type of catalog incompleteness [*Peng et al.*, 2007]. Moreover, there is no indication of a comparable behavior for the ETAS catalogs we studied (see Fig. 8). This is even true if short-term aftershock incompleteness is taken into account (not shown), which can be mimicked by eliminating small events from the ETAS catalog that directly follow large earthquakes according to the time-dependent magnitude threshold of completeness described in [*Helmstetter et al.*, 2006].

Provided that all these observations are applicable to the catalog from Southern California we study here, it suggests that our findings are not an artifact and the Omori-Utsu law needs to be modified if one considers (directly) triggered earthquakes associated with mainshocks of intermediate magnitude as defined by the method used here. Further evidence comes from Fig. 5 where the dressed and bare aftershock rates are shown for different magnitude ranges of the mainshock. For the bare case, best fits over the domain $10^{-2} < t < 1$ day give exponents between 0.9 and 1.0 for all magnitude ranges with $m < 6.0$. Note that such a break in scaling implying systematic deviations from the Omori-Utsu law for small t has also been reported by *Peng et al.* [2007].

4.4. Comparison with other studies

Understanding whether and how the Omori exponent p varies with the mainshock magnitude has important physical implications. Within the framework of the rate-and-state model [*Dietrich*, 1994], one can show that the coseismic stress heterogeneity on the main fault as well as the afterslip determine the dependence of the parameters in the Omori-Utsu law on m [*Hainzl and Marsan*, 2008]. In particular, empirical stress distributions with high coefficient of variation in the distribution will give rise to $p(m) \approx constant$, in line with our results here. This is, however, different from other results.

The nonparametric aftershock identification scheme of ML [*Marsan and Lengliné*, 2008] — in some sense a generalization of the ETAS-based method introduced by ZOV [*Zhuang et al.*, 2002, 2004] — is model-independent and linear, since it identifies triggering relationships based on summing the influences of previous seismic events. Using the ML scheme and different space-time window methods, *Hainzl and Marsan* [2008] found an overall increasing trend in $p(m)$ as also reported in [*Ouillon and Sornette*, 2005] but the high dispersion in their estimates did not rule out the possibility that p is independent of the mainshock magnitude (especially if one takes systematic errors in the chosen fitting procedure into account). Since the magnitude ranges and the catalog studied in [*Hainzl and Marsan*, 2008] differ vastly from ours, a more appropriate comparison can be drawn with the p values reported in [*Marsan and Lengliné*, 2008], where an earlier version of the Southern California catalog was analyzed (for all events $m > 3.0$) using the ML scheme. For directly triggered aftershocks, their estimated p-values exhibit a non-monotonic dependence on m, varying between 0.8 to 1.4. Note that these values vary significantly if the distance to fault is used instead of the epicentral distance in the ML scheme and are, thus, not robust. On the other hand for both indirectly and directly triggered events, they found p to be generally lower by 0.2–0.4 with a less pronounced dependence on m ($0.6 < p < 0.9$). Most of these findings are in sharp contrast to our results presented here.

The analyzed schemes (ML [*Marsan and Lengliné*, 2008; *Hainzl and Marsan*, 2008], ZOV [*Zhuang et al.*, 2008], and

BP used here) agree on the inequality $\alpha < b$, which implies that frequent small mainshocks dominate the "production" of seismic events. While *Hainzl and Marsan* [2008] found that χ increases as $10^{\alpha m}$ with $\alpha = 0.66$, *Marsan and Lengliné* [2008] found $\alpha = 0.6$ for Southern California. Yet, this value varies significantly if the distance to fault is used instead of the epicentral distance in the ML scheme, giving $\alpha = 0.86$ in the former case. Our value of $\alpha \approx 0.9$ is instead closer to estimates given by most space-time window methods [*Hainzl and Marsan*, 2008]. These observations emphasize again that some statistical properties of triggering cascades and aftershock sequences sensitively depend on the exact selection method.

Our analysis of the ETAS catalog also shows that the BP method can recover the model parameter α. This is even true if the spatial distribution of triggered events in the ETAS model is anisotropic as for the catalog studied by *Hainzl et al.* [2008] and discussed in more detail in [*Moradpour et al.*, 2013]. This is a clear advantage over the ZOV method, which underestimates α dramatically in this case [*Hainzl et al.*, 2008].

5. Duration of aftershock sequences

To compare the relative temporal influence of bare and dressed aftershock sequences as well as its dependence on mainshock magnitude, it can be helpful to study the average time interval between a mainshock and its last aftershock. This is shown in Fig. 10 for Southern California. The averaging is performed over the number of corresponding mainshocks in each magnitude bin of width $\delta m = 0.5$. We observe an increasing trend in both the bare and dressed aftershock durations, with the dressed duration scaling as $10^{0.44m}$. The later is very similar to what has been observed using the ML scheme [*Marsan and Lengliné*, 2008]. For the bare duration, however, the ML method did not give robust results: Using epicentral distances, the bare duration was found to be independent of the mainshock magnitude and of the order of 10 – 15 days for $m > 3$, while using distance to fault instead gave results similar to the dressed case, albeit lower by a factor of 6 to 10. These inconsistencies together with our findings here provide clear evidence against the hypothesis put forward by *Marsan and Lengliné* [2008] that direct triggering mechanisms are only short-lasting. An increase in the period over which aftershocks typically occur with mainshock magnitude follows the intuition that larger earthquakes exert longer-lasting direct temporal influence.

In particular, such behavior is expected based on the Omori-Utsu law and the productivity law, which in the context of the ETAS model fully determine the average time interval between a mainshock and its last aftershock: From extreme value statistics it follows that the average duration of a directly triggered aftershock sequence scales asymptotically as $10^{(\alpha/p)m}$ [*Schumann et al.*, 2012]. Fig. 11 shows that our method to define aftershocks allows us to roughly recover such a scaling for the ETAS model. Using the estimated values of α and p, Fig. 10 shows that we can recover such a scaling for magnitudes larger than 4 for Southern California as well. Note that the average duration for the largest events ($m > 6$) is underestimated due to the finite length of the catalog. For smaller events, this effect is not absent but much less pronounced due to the shorter durations.

6. Aftershock magnitudes and background seismicity

One of the main features of seismicity is its energy-scale invariance as documented by the empirical Gutenberg-Richter (GR) law [*Gutenberg and Richter*, 1949]. It states that the frequency–magnitude distribution of earthquakes decays as 10^{-bm} with an exponent $b \approx 1$. Since the magnitude m is a logarithmic measure of the energy of an earthquake [*Turcotte*, 1997], this corresponds to a power law in terms of the energies which is a typical sign of scale-invariance. Moreover, the b value seems to be independent of the specific geographic area as long as one considers sufficiently large areas over sufficiently long time intervals [*Kanamori and Brodsky*, 2004; *Gulia and Wiemer*, 2010]. Recently, it has been shown that the GR law even holds down to magnitudes $m = -4.4$ [*Kwiatek et al.*, 2010]. An often made assumption — which is also one of the main assumptions in the ETAS model — is that both triggered and background events follow the Gutenberg-Richter law. Thus, it is important to test this assumption.

6.1. Results for Southern California

Using $D_f = 2.3$ in the BP method as before to compute the triggering cascades, Fig. 12 compares the frequency magnitude distributions of triggered and background events in the catalog from Southern California we study here. Using the maximum likelihood method discussed in [*Naylor et al.*, 2009], we find $b = 1.04 \pm 0.02$ for triggered events and $b = 1.19 \pm 0.02$ for background events. This result indicates a statistically significant difference between the two populations. Increasing the magnitude cutoff to $m_{th} = 3.0$ ($m_{th} = 3.5$), we obtain $b = 1.00 \pm 0.03$ ($b = 0.94 \pm 0.05$) for triggered and $b = 1.14 \pm 0.04$ ($b = 1.10 \pm 0.07$) for background events. Thus, the Gutenberg-Richter exponents for the two populations are not significantly affected by m_{th}, strongly suggesting that the difference between the two populations is a robust feature.

6.2. Results for ETAS catalogs

Fig. 13(a) shows the frequency-magnitude distribution for the two populations separately as identified for a single ETAS catalog. Clearly, there are no statistically significant differences between the b-values of the two earthquake populations. This is even true if short-term aftershock incompleteness is taken into account: Fig. 13(b) shows the frequency-magnitude distributions if short-term aftershock incompleteness is mimicked by eliminating small events from the ETAS catalog that directly follow large earthquakes according to the time-dependent magnitude threshold of completeness described in [*Helmstetter et al.*, 2006]. The b-value decreases only very slightly for the triggered events. This is true for all ETAS catalogs we tested — including variations in the model parameters given in Table 1 and other realistic scenarios like spatial anisotropy in $\Phi_{m_i}(\vec{r} - \vec{r_i})$ [*Moradpour et al.*, 2013] — and, thus, seems to be a robust result.

6.3. Summary & Discussion

The analysis of the synthetic ETAS catalogs shows that the BP method can reliably recover the correct and identical b-values for both triggered and background events, independent of the presence of short-term aftershock incompleteness. This clearly indicates that the difference in the b-values for triggered and background events in Southern California is not an artifact due to the BP method used to define aftershocks, which is also supported by findings in [*Shearer*, 2012].

One possible explanation for the higher b-value of the background events could be the fact that some of these events were actually triggered by events below the magnitude threshold, i.e., their trigger is not part of the catalog. As model studies show, the number of such "mislabeled" events can be significant and it is a general problem for identifying background events [*Wang et al.*, 2010a;

Zhuang et al., 2008; *Wang et al.*, 2010b]. Using a maximum likelihood estimator based on an ETAS model and/or the ZOV method, the same model studies found also that the proportion of background events depends sensitively on the applied magnitude threshold m_{th}. This is, however, not the case for the method used to identify background events here. Not only are our estimates of the b-values independent of m_{th} but also the percentage of background events seems to be independent of the magnitude threshold we apply before analyzing the catalog with the BP method: the background events consist of about 39.6% of the total number of $m > 2.5$ earthquakes and for $m_{th} = 3.0$, $m_{th} = 3.5$, $m_{th} = 4.0$ and $m_{th} = 4.5$, we obtain 39.3%, 39.9%, 38.8% and 41.8% background events, respectively. To better understand these percentages, we note that in our tests with ETAS catalogs, the background rate tends to be overestimated by the BP method. The overestimation is mostly due to wrongly identified singles (events that were not triggered and do not trigger other events).

Our findings are also very different from the results obtained by using the ML scheme [*Marsan and Lengliné*, 2008]. The authors of that method did not find any statistical differences in the b-value between triggered and background earthquakes. Additionally, the background events were found to consist of about 19.5% of the total number of $m \geq 3.0$ earthquakes and this percentage increased monotonically with m_{th} reaching 68% for $m_{th} = 5.0$. As these results show, the estimation of the seismic "background" is highly dependent on the exact method to define background events and such a classification might not even be well-defined. An additional complication is that such estimations — including those based on the ML and the ZOV scheme — are intrinsically unreliable even for the ETAS model [*Sornette and Utkin*, 2009].

7. Båth's law

Apart from the Omori-Utsu law and the Gutenberg-Richter law for aftershocks, a third empirical scaling law has been proposed for aftershock sequences. Båth's law states that the average magnitude difference between a mainshock and its largest aftershock is ≈ 1.2, independent of the mainshock magnitude [*Båth*, 1965]. Some progress has been made in understanding its origin based on the framework of the ETAS model — and branching models in general — where it arises as a natural consequence of the model assumptions and typical definitions of aftershocks [*Helmstetter and Sornette*, 2003; *Saichev and Sornette*, 2005; *Vere-Jones and Zhuang*, 2008]. In particular, it was shown that Båth's law only strictly holds for those parameter regimes in the ETAS model that are close to criticality and satisfy $\alpha > b/2$. In the subcritical case, the average magnitude difference between a mainshock and its largest aftershock actually increases with the main shock magnitude. Specifically, the average magnitude of the largest aftershock, m^*_{as}, depends on the mainshock magnitude, m, as $\langle m^*_{as} \rangle = \frac{\alpha}{b} m + d$, where d is some constant.

Fig. 14 shows that the BP method given in Section 2 allows us to recover this behavior for sufficiently large mainshock magnitudes from ETAS catalogs though Båth's law remains a reasonable approximation. This is independent of whether we consider only directly triggered aftershocks or all aftershocks. In both cases, the deviations from the two possible behaviors for small mainshocks are expected due to the lower magnitude cutoff of the catalog. This cutoff directly leads to a bias in our estimate of the magnitude of the largest aftershock since earthquakes triggering only events below the threshold of observation are not included in our estimate. Thus, the average magnitude of the largest aftershock will be overestimated.

To test the general validity of and the possibility of deviations from Båth's law for real seismicity, we use our definition of aftershocks based on the BP method for the catalog from Southern California. We find that Båth's law holds to a very good approximation for mainshocks with magnitude bigger than approximately $m_{th} + 2$, see Fig. 15. In particular, within the given uncertainties it is as good an approximation as the behavior expected based on the ETAS model, which predicts deviations from Båth's law in the subcritical regime. The corresponding behavior is shown as the dash-dotted line in Fig. 15 for comparison. For dressed aftershocks, the average magnitude of the largest aftershock is always bigger compared to the bare case. However, this difference is surprisingly small indicating that Båth's law is quite insensitive to the inclusion of indirectly triggered aftershocks.

8. Decay of aftershock density with distance

In addition to the temporal characteristics of aftershock sequences discussed in the sections above, we can enrich the description of the aftershock statistics by studying the spatial distribution of aftershocks relative to their mainshock. This distribution is a signature of the triggering process and, thus, plays an important role in discriminating between potential candidates for this mechanism [*Felzer and Brodsky*, 2006; *Gomberg and Felzer*, 2008; *Lippiello et al.*, 2009; *Marsan and Lengliné*, 2010; *Richards-Dinger et al.*, 2010; *Powers and Jordan*, 2010].

8.1. Results for Southern California

As before, we consider the three-dimensional distance between hypocenters, r_{ij}, separating an aftershock j from its mainshock i. For directly triggered aftershocks, the left panel of Fig. 16 shows the spatial probability density function of aftershocks, $P(r)$, for different ranges of the mainshock magnitudes. It is evident that these distributions have a shape almost independent of the magnitude of the mainshock. Of course, increasing this magnitude m_i, their maximum shifts toward larger distances, but upon rescaling r by a factor $10^{-\sigma m_i}$ with $\sigma = 0.42(3)$, one obtains a distribution of the rescaled distances $\lambda = r\, 10^{-\sigma m_i}$ with little trace of m_i, see right panel of Fig. 16. The only non-scaling feature is the cutoff at large distances, which seems to arise for $r \gtrsim$ 10-20 km (left panel). A possible explanation is that the finite width of the earth crust starts to be relevant for this statistics for distances longer that tens of km, similar to the known depth-dependencies and finite size effects reported by *Kagan* [2007]. This is supported by the absence of this feature for the ETAS catalogs — which use epicenters — as shown below.

The generic shape of the distance distributions, however, is characterized by two regimes. For short distances, it increases as a power $\sim r^{1.5}$. This indicates a fractal organization of aftershocks on short spatial scales since the increase corresponds to a correlation dimension $D_C \approx 1.5 + 1 = 2.5$. This is clearly incompatible with a homogeneous distribution of events in three dimensions ($D_C = 3.0$) as well as with aftershocks concentrated on a single two-dimensional fault plane ($D_C = 2.0$).

As follows from Fig. 16, the transition from the regime at short distances to the second regime at larger distances occurs at $L(m) \approx 8 \times 10^{0.42m}$m. This length scale and its scaling with magnitude is close to the estimated behavior of the rupture length, $L_R(m)$, of an earthquake of magnitude m. *Kagan* [2002] gives $L_R(m) \approx 20 \times 10^{m/2}$m, while *Wells and Coppersmith* [1994] found $L_R(m) = \sqrt{A_R} \approx 18 \times 10^{0.46\ m}$m, where A_R is the rupture area of the earthquake. Another estimate was obtained by *Davidsen et al.* [2008] giving $L_R(m) \approx 12 \times 10^{0.45m}$m. Thus, we have $L(m) \approx \frac{1}{2} L_R(m)$

indicating that the borders of the rupture area of the mainshock are the boundary between the spatial region giving rise to the first regime in $P(r)$ and the region where one detects the subsequent decay with r.

For distances longer than the typical distance $L(m)$, the right panel of Fig. 16 shows that the aftershock densities decay as a power-law $r^{-\mu}$, at least for $r \lesssim$ 10-20 km. The exponent we estimate is $\mu \simeq 1.6$. A similar behavior has been found in [*Marsan and Lengliné*, 2010] for the aftershock sequence of Landers using the ML scheme for aftershock identification. Our results do not change significantly if we consider the dressed aftershock density with distance, see Fig. 17, where we estimate $\sigma = 0.40(3)$ for this case.

To understand the origin of the observed exponent μ better, we consider the spatial probability density function of aftershocks conditioned on their time of occurrence, $P_m(r|T)$. Thus, $P_m(r|T)$ represents the distribution of spatial distances between a mainshock of magnitude m and only those of its aftershocks that occur within a time interval T after the mainshock. As an example, $P_m(r|T)$ averaged over all mainshocks within the magnitude range $[3, 4[$ is shown in Fig. 18 for different time intervals. For other magnitude ranges, we observe qualitatively similar results. Most strikingly, $P_m(r|T)$ does not decay with an exponent $\mu \simeq 1.6$ and the distribution changes significantly with T. While the location of the maximum is invariant with respect to T, the decay for larger values of r is not. Considering only aftershocks occurring within an hour of the mainshock, there is a power-law decay with an exponent $\tilde{\mu} \approx 1.3$ up to approximately 10km, followed by a much steeper decay. This transition point does not remain at 10km for aftershocks occurring at later times, but instead it moves to smaller values, reaching a few hundred meters for $T = [1 \text{ month}, \infty]$. This clearly indicates that $\mu \simeq 1.6$ is an *effective* exponent that arises due to a weighted superposition of different distributions which obey a different power-law with exponent $\tilde{\mu}$ over varying and rather limited ranges.

8.2. Results for ETAS catalogs

For the ETAS model, the spatial probability density function of directly triggered aftershocks is given by Eq. (8). Fig. 19 shows that the BP method can reconstruct fairly well its form, with some systematic deviations. While the scaling of the most likely distance with the rupture length as well as the full scaling collapse of the distributions for different mainshock magnitudes are recovered, the power-law increase at shorter distances has a smaller exponent and the power-law decay at larger distances has a bigger exponent compared to Eq. (8). The deviations in the exponents are about 0.35 in both cases for the ETAS catalogs considered. It is important to realize, though, that these deviations *can* change significantly if different spatial distributions of triggered events, $\Phi_{m_i}(\vec{r} - \vec{r_i})$, are used in the ETAS model [*Moradpour et al.*, 2013].

Yet, these deviations are absent if one considers the spatial probability density function of aftershocks conditioned on their time of occurrence, $P_m(r|T)$, with $T < 1$ hour, see Fig. 20. It is clear that any aftershock detection method will be more effective for shorter times, because a long delay after the mainshock implies that many other events have occurred in the meantime which may potentially "shadow" the correct assignment of an aftershock. In our case, according to the space-time-magnitude metrics given by Eq. (1), an earthquake occurring long after an event can only be recognized as an aftershock of that event (i.e. not being an aftershock of some other event or distinguished from background seismicity) if the two are close enough in space. This explains the deviations from the true behavior for $P(r)$, which considers aftershocks at all times.

8.3. Summary & Discussion

The analysis of the ETAS catalogs shows that the BP method can reliably recover the spatial probability density function of directly triggered aftershocks if one considers only those aftershocks that occur close in time to their respective mainshock. Including all aftershocks can lead to a bias, which gives rise to an overestimation in the absolute value of the power-law exponent characterizing the behavior at large distances, while the most likely distance being half the rupture length is not affected.

Indeed, we observe a similar behavior for the catalog from Southern California. In particular, the association of the rupture length with the most likely distance of aftershocks is a robust finding with respect to variations in our method to define aftershocks [*Baiesi and Paczuski*, 2004, 2005; *Baiesi*, 2006; *Lippiello et al.*, 2009] and provides a solid justification of the often-made assumption that the aftershock zone scales with the rupture zone [*Kagan*, 2002]. Moreover, a similar scaling exponent $\sigma = 0.43$ was found independently with the ML scheme to define aftershocks [*Marsan and Lengliné*, 2008]. This is clear evidence that these are characteristic features of the spatial distribution of aftershocks.

The results for Southern California also show that the exponent of the power-law increase for short distances is indistinguishable from the correlation dimension $D_f \approx 2.2$ established for seismicity in Southern California [*Kagan*, 2007] if the bias is taken into account (see Fig. 18). This indicates a clear self-consistency between the spatial distribution of aftershocks and of seismicity in general on short spatial scales.

Moreover, we find that the power-law decay for distances larger than the rupture length is characterized by an exponent $\mu \approx 1.3$, taking the bias into account. While initially it was thought that an exponent of $\mu < 2$ indicates that triggering of aftershocks by dynamic stress plays an important role [*Felzer and Brodsky*, 2006], it was argued more recently that triggering of aftershocks by static stress changes alone can also lead to values of μ as low as one [*Richards-Dinger et al.*, 2010], depending on the b-value and the productivity exponent α. Specifically, it was argued that $\mu = 1.49(-b + \alpha) - 1$ if only static stress changes are taken into account. For the catalog studied here, this would imply that $\mu \approx 1.49(-1.09 + 0.87) - 1 = 1.33$. This is identical to the directly measured value $\mu \approx 1.3$ and seems to indicate that static stress changes are sufficient to explain the triggering of aftershocks. However, the approach by *Richards-Dinger et al.* [2010] assumed that aftershocks are uniformly distributed along the rupture surface – this assumption is clearly violated for the catalog studied here since $D_c > 2$. Thus, it is still conceivable that dynamic stress changes significantly influence or even dominate the triggering of aftershocks.

Our observed values of μ and $\tilde{\mu}$ for Southern California are very different from the value 1.94 ± 0.04 reported in [*Lippiello et al.*, 2009] for aftershocks occurring within 5 hours of the mainshock. While the method used to identify aftershocks and catalog studied in their work are very similar to ours, it is likely that the differences arise due to the sensitive dependence on the fractal dimension D_f discussed in Section 1. In [*Lippiello et al.*, 2009], only $D_f \leq 1.6$ and epicenters were considered. $P_m(r|T)$ was also estimated using the ML method to identify aftershocks and for $T = [0; 15 \text{ min}]$ an exponent of 1.76 ± 0.35 was estimated for mainshocks within the magnitude range $[3, 4[$, while for $T = [12 \text{ h}; 24 \text{ h}]$ an exponent 1.97 ± 1.11 was obtained [*Marsan and Lengliné*, 2010]. Within these huge statistical uncertainties, the exponents are, hence, comparable to ours, though the functional form of $P_m(r|T)$ for small r could not be resolved. These different observations for the approximate power law decay of $P(r)$ and $P_m(r|T)$ for values of r larger than the rupture

length indicate that the behavior depends crucially on how one defines aftershocks. Moreover, these probability density functions might not even follow a well-defined power law over extended ranges and an analysis to distinguish between different triggering mechanisms purely based on the value of the corresponding exponent might be questionable from the start. Further support for this comes from our current understanding of the overall spatial distribution of seismicity which is not a simple fractal but shows signs of multifractality if one considers epicenters [*Davidsen and Goltz*, 2004] and strong finite size effects and depth dependence if one considers hypocenters [*Kagan*, 2007].

In addition, an approximate power-law decay of $P(r)$ with an exponent less than or equal to 2 also indicates that a quantity such as the mean distance of aftershocks $\langle r \rangle$ is somewhat ill-defined, as this average is mostly determined by the upper cutoff in the power-law. This cutoff can depend, for example, on the time-window used to collect aftershocks (see Fig. 18) as well as on the finite spatial range of the catalog. The very limited content in $\langle r \rangle$ is also reflected in the fact that the corresponding standard deviation is at least as large as $\langle r \rangle$ — of course this is again due to the power-law tail in $P(r)$. Since almost all published methods have obtained probability density functions for aftershock distances decaying with power-laws comparable to r^{-2} or weaker [*Felzer and Brodsky*, 2006; *Gomberg and Felzer*, 2008; *Lippiello et al.*, 2009; *Marsan and Lengliné*, 2010; *Richards-Dinger et al.*, 2010; *Powers and Jordan*, 2010], this seems to be a generic feature. We conclude that the mean aftershock distance is a quantity much influenced by uncontrolled parameters, mathematically ill-defined for $P(r) \sim r^{-\mu}$ with $\mu \leq 2$, and thus with little information on the seismic process. This casts doubts on the analysis of aftershock "diffusion" as, for example, presented in [*Marsan and Lengliné*, 2008].

9. Discussion and Conclusions

We have tested the robustness of the statistical properties of aftershocks as defined by the BP method first proposed in [*Baiesi and Paczuski*, 2004, 2005] and extended by *Zaliapin et al.* [2008]. The method is very simple to apply and does not require any assumptions on the underlying physics of earthquakes, being only based on known empirical laws and on general statistical arguments: correlations are measured as deviations from randomness, using the space-time-magnitude metrics given by Eq. (1). We have tested the performance of the BP scheme by considering synthetic catalogs generated by the ETAS model, showing that the laws assumed in the aftershock generation are recovered correctly in the statistics, and that other empirical laws of seismicity are also found. Moreover, the scheme based on Eq. (1) also gives meaningful results for real earthquake catalogs and the statistical properties of aftershocks are even robust with respect to variations in the method itself as long as one considers hypocenters and a fractal dimension close to the estimated correlation dimension.

However, one has to be careful in interpreting findings based on a single identification scheme [*Wang et al.*, 2010b], because aftershocks may be defined in different ways. Currently we are not aware of precise general physical arguments distinguishing correct identifications of aftershocks from incorrect ones, especially outside specific modeling frameworks. Thus, one is left with the option to compare the results in the literature to identify commonalities and features that are robust across several definitions of aftershocks. In order to be concise, we choose to compare our findings mostly with those obtained by the ZOV declustering technique [*Zhuang et al.*, 2004, 2008] and especially the recently proposed ML method [*Marsan and Lengliné*, 2008; *Hainzl and Marsan*, 2008] in which long range effects in space and time are also contemplated, similar analysis are carried out and the same seismic catalogue is studied.

Concerning the seismicity observed in Southern California, the aftershocks as defined by the BP method follow asymptotically an Omori-Utsu law with a p-value that is either slightly decreasing with m or constant around 1.2, while ML find a more variable p-value with $p < 0.9$ if directly and indirectly aftershocks are considered. We also observe a productivity parameter $\alpha \approx 0.9$, in line with other studies and with the conclusion that, since $\alpha < b$ (of the Gutenberg-Richer law), frequent small events are responsible for triggering most of the observed seismic activity, thus, dominating the production of aftershocks. The same conclusion is reached by ZOV and ML. For ML, this is despite the fact that their estimates of α and p for directly triggered aftershocks as well as the temporal duration of bare aftershock sequences are not robust with respect to variations in their method. We further provide evidence that the temporal duration of aftershock sequences scales with the mainshock magnitude similar to what is expected for the bare case based on the Omori-Utsu law and the productivity law and also similar to what has been observed by ML for the dressed case. Our findings also show that in contrast to what is often assumed, some of the properties of aftershock sequences of small mainshocks might actually be different from those of large mainshocks as systematic deviations from the Omori-Utsu law at small to intermediate time scales indicate — similar to findings by *Peng et al.* [2007].

We also find a percentage of background activity that is not changing with the lower magnitudes cutoff. This supports the picture that seismicity has a hierarchical structure, and it is again at variance with results by ML. Another deviation is the sensible difference that we note between the b-value of the Gutenberg-Richer law for aftershocks and for background events. The latter are found with a larger b-value, independently on the lower cutoff. Since this phenomenon is absent in the analyzed ETAS catalogs (where we know by construction that there is a unique b) even if short-term aftershock incompleteness is taken into account, we conclude that the observed b-value of background seismicity being larger than that of aftershocks is likely to be a real feature of seismicity.

One of the most solid results, found in the literature and confirmed here, is the appearance of a *typical* length scale of aftershocks, scaling with the mainshock magnitude m as $10^{0.42m}$ (the index 0.42 being compatible with values from other methods). This length scale can be identified with the rupture length of the mainshock. On the other hand, it is important to note that the *average* aftershock distance is an ill-defined quantity and hence that it is not meaningful to use it for studying diffusion of aftershocks. The underlying reason for this problem is a power-law tail in the aftershock distance distribution $\sim r^{-\mu}$ with $\mu < 2$ (all studies in the literature find this) indicating the absence of a finite first moment $\langle r \rangle$. We find $\mu \approx 1.3$ if measurement biases are eliminated. As a direct comparison shows, the exact value of μ depends sensitively on the chosen method. Thus, for the moment, results concerning the form of the tail in the aftershock distance distribution cannot help much in the debate on whether dynamic or static stress changes are responsible for triggering far events. From our findings we are inclined to assign a significant role to static stress changes, though clearly this is still an open point.

The spatial organization of hypocenters in Southern California is close to be a monofractal with fractal dimension $D_f \approx 2.2$, but other options are possible (multifractality, for example). For our purposes this is not much relevant, as we show that the results based on Eq. (1) do not depend sensibly on this value of D_f as long as $D_f > 2.0$. Moreover, a posteriori, we recover a correlation dimension of aftershocks

≈ 2.2 in the aftershock region within the typical distance from the mainshock.

To summarize, the comparison of our findings with those obtained using different schemes to identify aftershocks clearly shows that many features of aftershock sequences depend sensitively on their identification. Fortunately, a significant number of statistical features of aftershock sequences seems to be robust. Most importantly, these include a productivity law with an exponent less than the b-value in the frequency-magnitude distribution — indicating that smaller earthquakes dominate the seismic energy budget due to their higher frequency — and a clustering of triggered events within a distance comparable to the rupture length of a mainshock.

Acknowledgments. We thank J. Moradpour for testing the numerical code. JD thanks S. Hainzl, T. Peixoto, and I. Zaliapin for helpful discussions. We also thank two reviewers and the Associated Editor for helpful comments. This project was financially supported by Alberta Innovates – Technology Futures (formerly Alberta Ingenuity).

References

Baiesi, M. (2006), Scaling and precursor motifs in earthquake networks, *Physica A*, *360*, 534.

Baiesi, M., and M. Paczuski (2004), Scale-free networks of earthquakes and aftershocks, *Physical Review E*, *69*, 066106.

Baiesi, M., and M. Paczuski (2005), Complex networks of earthquakes and aftershocks, *Nonlin. Proc. Geophys.*, *12*, 1.

Båth, M. (1965), Lateral inhomogeneities in the upper mantle, *Tectonophysics*, *2*, 483.

Clauset, A., C. R. Shalizi, and M. E. J. Newman (2009), Power-law distributions in empirical data, *SIAM Review*, *51*, 661.

Davidsen, J., and C. Goltz (2004), Are seismic waiting time distributions universal?, *Geophysical Research Letters*, *31*, L21612.

Davidsen, J., P. Grassberger, and M. Paczuski (2008), Networks of recurrent events, a theory of records, and application to finding causal signatures in seismicity, *Physical Review E*, *77*, 066104.

Davis, S. D., and C. Frohlich (1991), Single-link cluster analysis, synthetic earthquake catalogues, and aftershock identification, *Geophysical Journal International*, *104*, 289.

Dietrich, J. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, *Journal of Geophysical Research*, *99*, 2601.

Eguíluz, V. M., D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian (2005), Scale-free brain functional networks, *Physical Review Letters*, *94*, 018102.

Felzer, K. R., and E. E. Brodsky (2006), Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature (London)*, *441*, 735.

Gardner, J. K., and L. Knopoff (1974), Is the sequence of earthquakes in Southern-California, with aftershocks removed, Poissonian?, *Bulletin of the Seismological Society of America*, *64*, 1363.

Gomberg, J., and K. Felzer (2008), A model of earthquake triggering probabilities and application to dynamic deformations constrained by ground motion observations, *Journal of Geophysical Research*, *113*, B10317.

Gulia, L., and S. Wiemer (2010), The influence of tectonic regimes on the earthquake size distribution: a case study for Italy, *Geophysical Research Letters*, *37*, L10305.

Gutenberg, B., and C. Richter (1949), *Seismicity of the Earth*, Princeton University Press, Princeton, NJ.

Hainzl, S., and D. Marsan (2008), Dependence of the Omori-Utsu law parameters on mainshock magnitude: Observation and modeling, *Journal of Geophysical Research*, *113*, B10309.

Hainzl, S., A. Christophersen and B. Enescu (2008), Impact of earthquake rupture extensions on parameter estimation of point-process models, *Bulletin of the Seismological Society of America*, *98*, 2066.

Helmstetter, A., and D. Sornette (2002), Diffusion of epicenters of earthquake aftershocks, Omori's law, and generalized continuous-time random walk models, *Physical Review E*, *66*, 061104.

Helmstetter, A., and D. Sornette (2003), Båth's law derived from the Gutenberg-Richter law and from aftershocks properties, *Geophysical Research Letters*, *30*, 2069.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2006), Comparison of short-term and time-independent earthquake forecast models for southern California, *Bulletin of the Seismological Society of America*, *96*, 90.

Houghs, S. E., and L. M. Jones (1997), Aftershocks; are they earthquakes or afterthoughts?, *EOS, Transactions American Geohysical Union*, *78*, 505.

Kagan, Y. Y. (2002), Aftershock zone scaling, *Bulletin of the Seismological Society of America*, *92*, 641.

Kagan, Y. Y. (2004), Short-term properties of earthquake catalogs and models of earthquake source, *Bulletin of the Seismological Society of America*, *94*, 1207.

Kagan, Y. Y. (2007), Earthquake spatial distribution: the correlation dimension, *Geophysical Journal International*, *168*, 1175.

Kagan, Y. Y., and H. Houston (2005), Relation between mainshock rupture process and Omori's law for aftershock moment release rate, *Geophysical Journal International*, *163*, 1039.

Kagan, Y. Y., and L. Knopoff (1987), Statistical short-term earthquake prediction, *Science*, *236*(4808), 1563–1567, doi: 10.1126/science.236.4808.1563.

Kanamori, H., and E. E. Brodsky (2004), The physics of earthquakes, *Reports on Progress in Physics*, *67*, 1429.

Keilis-Borok, V. I., L. Knopoff, and I. M. Rotwain (1980), Bursts of aftershocks, long-term precursors of strong earthquakes, *Nature (London)*, *283*, 259.

Kwiatek, G., K. Plenkers, M. Nakatani, Y. Yabe, G. Dresen, and JAGUARS-Group (2010), Frequency-magnitude characteristics down to magnitude –4.4 for induced seismicity recorded at Mponeng gold mine, South Africa, *Bulletin of the Seismological Society of America*, *100*, 1165.

Lin, G., P. M. Shearer, and E. Hauksson (2007), Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *Journal of Geophysical Research*, *112*, B12309.

Lippiello, E., L. de Arcangelis, and C. Godano (2009), The role of static stress diffusion in the spatio-temporal organization of aftershocks, *Physical Review Letters*, *103*, 038501.

Main, I. (2006), A hand on the aftershock trigger, *Nature (London)*, *441*, 704.

Marsan, D., and O. Lengliné (2008), Extending earthquake' reach through cascading, *Science*, *319*, 1076.

Marsan, D., and O. Lengliné (2010), A new estimate of the decay of aftershock density with distance to the mainshock, *Journal of Geophysical Research*, *115*, B09302.

Molchan, G. M., and O. E. Dmitrieva (1992), Aftershock identification - methods and new approaches, *Geophysical Journal International*, *109*, 501.

Moradpour, J., S. Hainzl, and J. Davidsen (2013), in preparation.

Narteau, C., S. Byrdina, P. Shebalin, and D. Schorlemmer (2009), Common dependence on stress for the two fundamental laws of statistical seismology, *Nature (London)*, *462*, 642.

Naylor, M., J. Greenhough, J. McCloskey, A. F. Bell, and I. G. Main (2009), Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions, *Geophysical Research Letters*, *36*, L20303.

Ogata, Y., and J. Zhuang (2006), Space-time ETAS models and an improved extension, *Tectonophysics*, *413*, 13.

Ouillon, G., and D. Sornette (2005), Magnitude-dependent Omori law: theory and empirical study, *Journal of Geophysical Research*, *110*, B04306.

Parsons, T., and A. A. Velasco (2009), On near-source earthquake triggering, *Journal of Geophysical Research*, *114*, B10307.

Peixoto, T. P., K. Doblhoff-Dier, and J. Davidsen (2010), Spatiotemporal correlations of aftershock sequences, *Journal of Geophysical Research*, *115*, B10309.

Peng, Z., J. E. Vidale, and H. Houston (2006), Anomalous early aftershock decay rate of the 2004 Mw6.0 Parkfield, California, earthquake, *Geophysical Research Letters*, *33*, L17307.

Peng, Z., J. E. Vidale, M. Ishii, and A. Helmstetter (2007), Seismicity rate immediately before and after mainshock rupture from high-frequency waveforms in Japan, *Journal of Geophysical Research*, *112*, B03306.

Peng, Z. G., and P. Zhao (2009), Migration of early aftershocks following the 2004 Parkfield earthquake, *Nature Geosciences*, *2*, 877.

Powers, P. M., and T. H. Jordan (2010), Distribution of seismicity across strike-slip faults in California, *Journal of Geophysical Research*, *115*, B05305.

Reasenberg, P. (1985), Second-order moment of central California seismicity, 1969 – 1982, *Journal of Geophysical Research*, *90*, 5479.

Richards-Dinger, K., R. S. Stein, and S. Toda (2010), Decay of aftershock density with distance does not indicate triggering by dynamic stress, *Nature (London)*, *467*, 583.

Saichev, A., and D. Sornette (2005), Distribution of the largest aftershocks in branching models of triggered seismicity: Theory of the universal Båth law, *Physical Review E*, *71*, 056127.

Schorlemmer, D., and J. Woessner (2008), Probability of detecting an earthquake, *Bulletin of the Seismological Society of America*, *98*, 2103.

Schumann, A. Y., N. R. Moloney and J. Davidsen (2012), Extreme value and record statistics in heavy-tailed processes with long-range memory, in *Extreme events and natural hazards: The complexity perspective*, *Geophysical Monograph Series*, vol. 196, edited by A. Surjala Sharma et al., pp. 315–334, AGU, Washington, D.C.

Shearer, P. M. (1985), Self-similar earthquake triggering, Bath's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for southern California, *Journal of Geophysical Research*, *117*, B06310.

Sornette, D., and S. Utkin (2009), Limits of declustering methods for disentangling exogeneous from endogeneous events in time series with foreshocks, main shocks, and aftershocks, *Physical Review E*, *79*, 061110.

Turcotte, D. L. (1997), *Fractals and chaos in geology and geophysics*, 2nd ed., Cambridge University Press, Cambridge, UK.

Turcotte, D. L., J. R. Holliday, and J. B. Rundle (2007), BASS, an alternative to ETAS, *Geophysical Research Letters*, *34*, L12303.

Utsu, T., Y. Ogata, and R. S. Matsu'ura (1995), The centenary of the Omori formula for a decay law of aftershock activity, *J. Phys. Earth*, *43*, 1.

van der Elst, N. J., and E. E. Brodsky (2010), Connecting near-field and far-field earthquake triggering to dynamic strain, *Journal of Geophysical Research*, *115*, B07311.

Vere-Jones, D., and J. Zhuang (2008), Distribution of the largest event in the critical epidemic-type aftershock-sequence model, *Physical Review E*, *78*, 047102.

Wang, Q., D. D. Jackson, and J. Zhuang (2010a), Missing links in earthquake clustering models, *Geophysical Research Letters*, *37*, L21307.

Wang, Q., D. D. Jackson, and J. Zhuang (2010b), Are spontaneous earthquakes stationary in California?, *Journal of Geophysical Research*, *115*, B08310.

Wells, D. L., and K. J. Coppersmith (1994), New empirical relationships between magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bulletin of the Seismological Society of America*, *84*, 974.

Wiemer, S., and M. Wyss (2000), Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan, *Bulletin of the Seismological Society of America*, *90*, 859.

Zaliapin, I., A. Gabrielov, V. Keilis-Borok, and H. Wong (2008), Clustering analysis of seismicity and aftershock identification, *Physical Review Letters*, *101*, 018501.

Zaliapin, I., and Y. Ben-Zion (2013), Earthquake clusters in southern California, I: Identification, *submitted to Journal of Geophysical Research*.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.*, *97*, 369.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2004), Analyzing earthquake clustering features by using stochastic reconstruction, *Journal of Geophysical Research*, *109*, B05301.

Zhuang, J., A. Christophersen, M. K. Savage, D. Vere-Jones, Y. Ogata, and D. D. Jackson (2008), Differences between spontaneous and triggered earthquakes: their influences on foreshock probabilities, *Journal of Geophysical Research*, *113*, B11302.

C. Gu, Department of Physics and Astronomy, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada.

A. Y. Schumann, Department of Physics and Astronomy, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada.

M. Baiesi, Dipartimento di Fisica e Astronomia, Università degli Studi di Padova, Via Marzolo 8, I-35131 Padova, Italy (baiesi@pd.infn.it)

J. Davidsen, Department of Physics and Astronomy, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada. (davidsen@phas.ucalgary.ca)

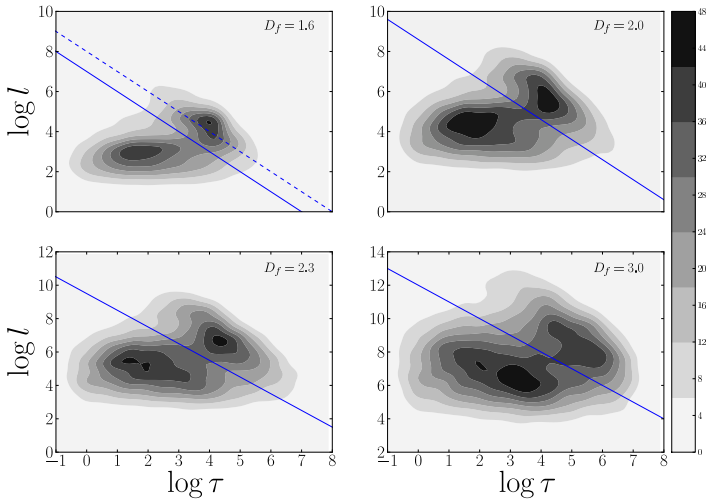

Figure 1. Southern California: Density plots of the set $\{n_j^*\}$ represented in $\log\tau$–$\log l$ space as defined in Eq. (1) for $D_f = 1.6, 2.0, 2.3, 3.0$. Time is measured in seconds and distances are measured in meters. Darker shading represent higher densities as quantified in the legend. In all cases, two different populations are visible, which can be separated as indicated by the straight lines. The straight lines correspond to constant thresholds n^* such that $\log\tau + \log l = \log n^*$. For increasing D_f, the blue lines are given by $\log n^* = 7.0$, 8.6, 9.5 and 12.0 respectively. Between 60 and 62% of each set $\{n_j^*\}$ are below the respective threshold and correspond to significant values of n_j^*. The complement of $\{n_j^*\}$ are then considered to be background events according to our definition. The dashed line in the top left panel indicates the threshold $\log n^* = 8.0$ used previously in [*Baiesi and Paczuski*, 2004]. This choice results in an additional inclusion of 15% from the set $\{n_j^*\}$ to the significant set.

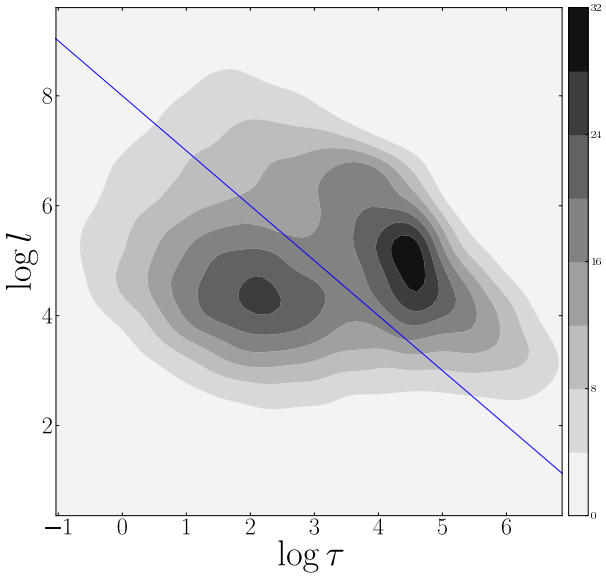

Figure 2. As Fig. 1, but for a surrogate catalog generated by the ETAS model with the parameters given in Table 1 ($N = 26986$). The blue line represents $\log n^* = 8.0$. About 50% of the set $\{n_j^*\}$ constructed from the catalog are below this threshold. More than 96% of all background events in the simulated ETAS catalog are correctly identified as such.

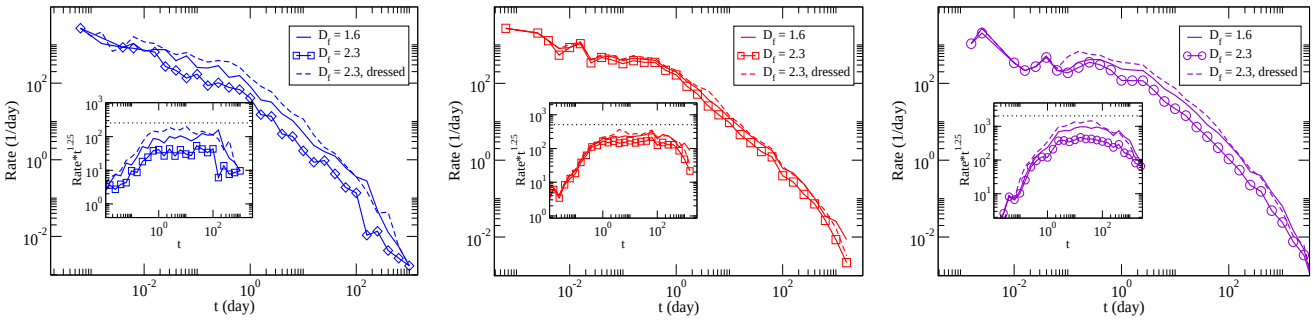

Figure 3. Southern California: Aftershock rates of three major events in the catalogue: Northridge (m=6.7), Hector Mine (m=7.1) and Landers (m=7.3) (from left to right). Bare rates for $D_f = 1.6$ are represented by the solid curves, while bare and dressed rates for $D_f = 2.3$ are represented by coloured symboled and dashed curves, respectively. Hypocentre distances are considered in both cases. The insets show the rates rescaled with the respective p value of the Omori-Utsu law.

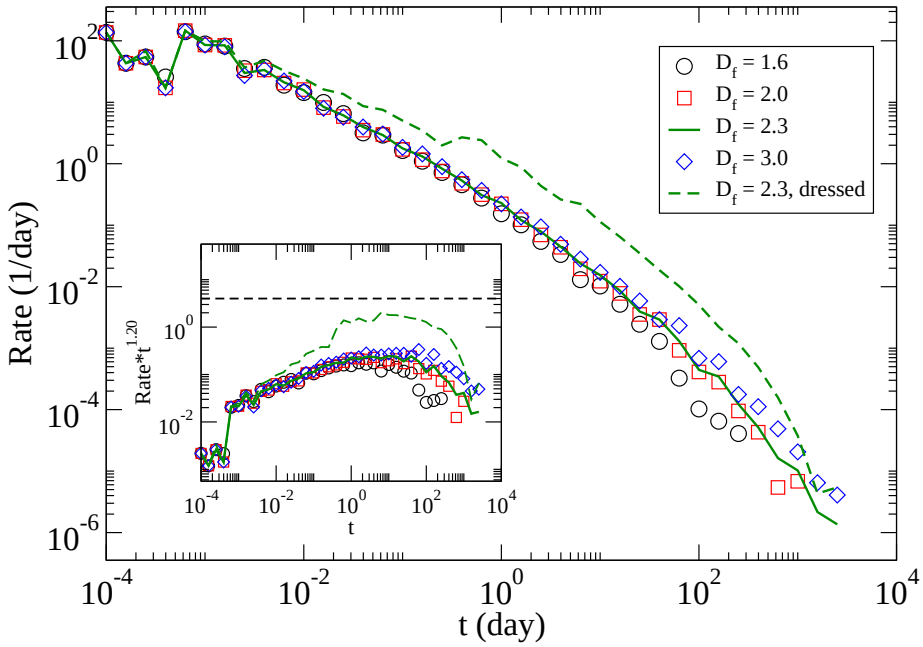

Figure 4. Southern California: Average aftershock rates for mainshock magnitudes $4.0 \leq m \leq 4.5$, computed with different fractal dimensions. The dressed (dark green dashed) curve for $D_f = 2.3$ overlaps with its bare counterpart (dark green) up to roughly 10^{-2} days. The dressed curve decays softly at first in the intermediate time range but for $t > 1$ day all curves for which $D_f \geq 2.0$ share a similar decay exponent. The inset shows the rescaled rates.

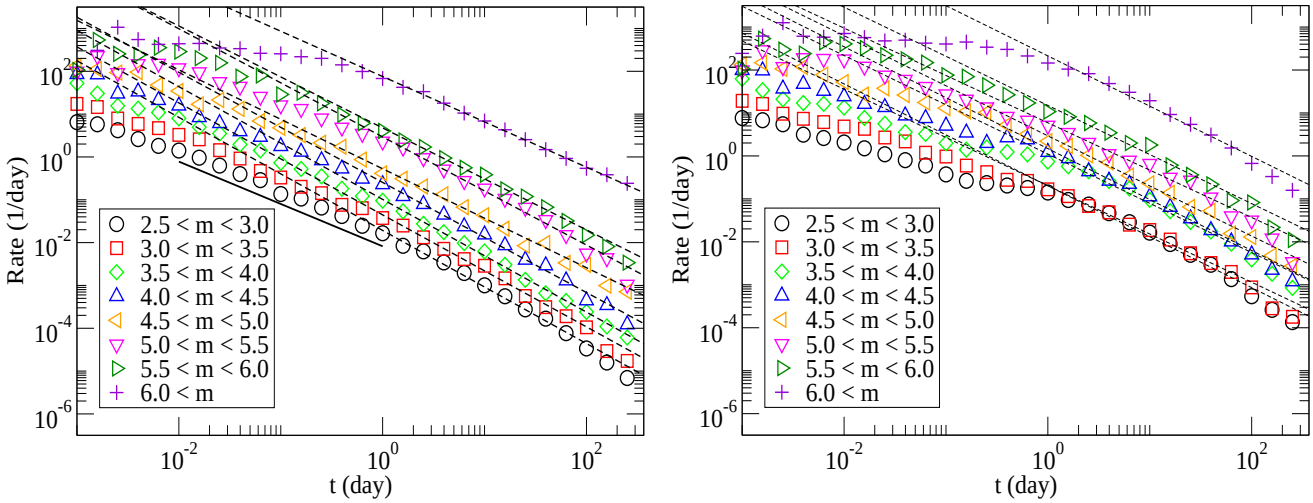

Figure 5. Southern California: Average aftershock rates of mainshocks within the given magnitudes increments, computed using $D_f = 2.3$, considering directly triggered aftershocks only (left panel) and all aftershocks (right panel). Best power law fits to $\chi \cdot t^{-p}$ in the regime of $t_0 < t < 100$ days are shown as dashed lines for the bare and the dressed case, respectively. For magnitude ranges with $m < 6.0$, clear deviations from the Omori-Utsu law (Eq. (9)) can be observed for shorter time scales. For the bare case, the deviations in the regime of $10^-2 < t < 1$ days follow another power law with an exponent about 1.0 as indicated by the solid line.

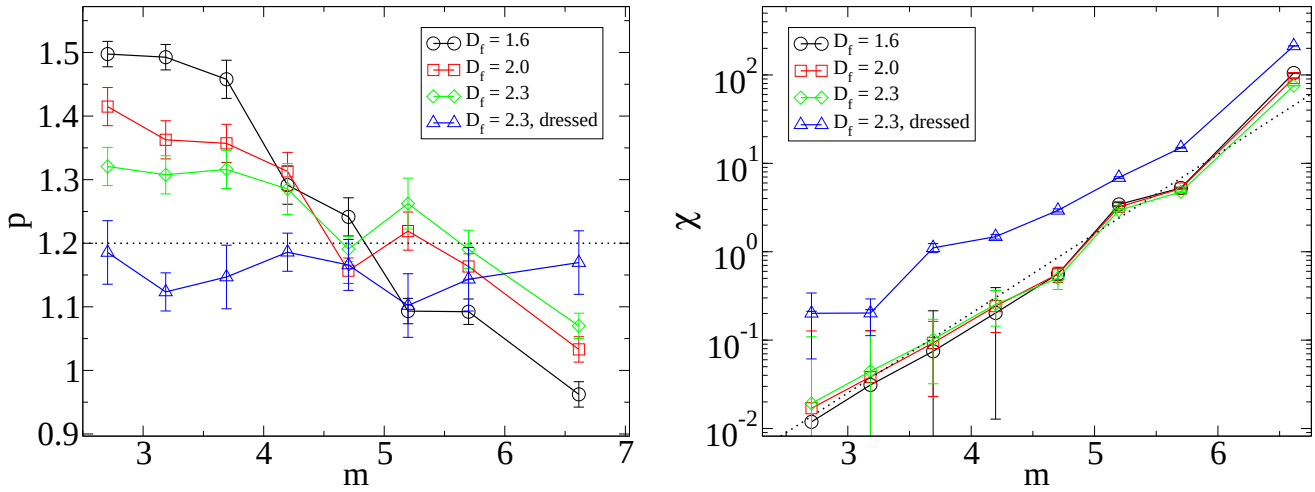

Figure 6. Southern California: i) Left panel: Variation of the exponent p in the Omori-Utsu law with mainshock magnitude m — see Fig. 5 for a specific example of how p was estimated. For $D_f < 2.0$, the statistical properties of the detected aftershock sequences are not robust for small m — see text for details. For $D_f \geq 2.0$, $\langle p \rangle \approx 1.2$ (indicated as the horizontal dotted black line in the panel) for both bare and dressed rates. ii) Right panel: Increasing behaviour of χ with mainshock magnitudes for different values of the fractal dimension. The black dotted line shows a scaling $\sim 10^{0.9m}$ as a guide to the eye. Note that the given error bars in (a) and (b) correspond to the statistical errors in the best fits and, thus, underestimate the true uncertainties (see, for example, [*Clauset et al.*, 2009]).

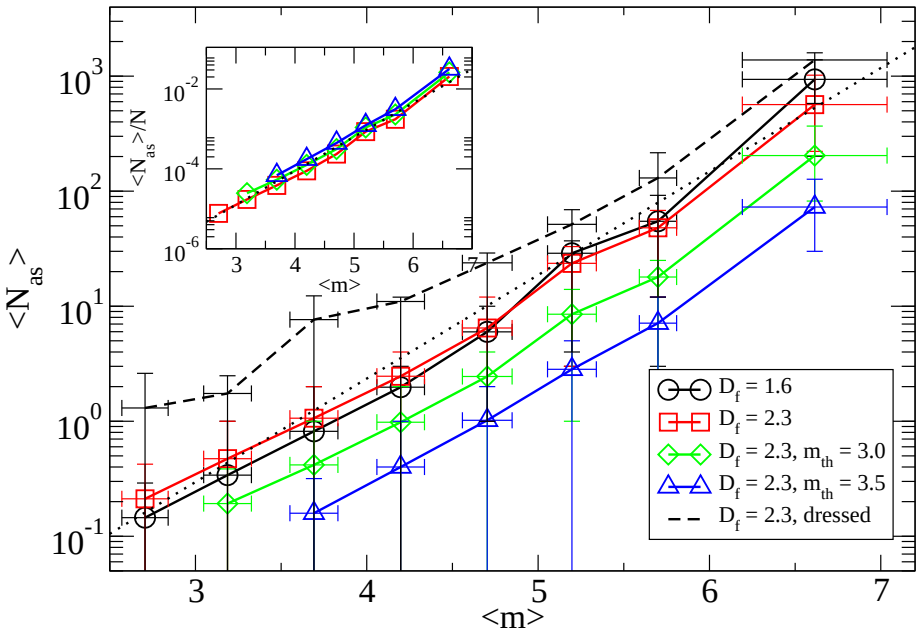

Figure 7. Southern California: Average number of earthquakes, $\langle N_{as} \rangle$, triggered by earthquakes within a given magnitude range, incremented by $\delta m = 0.5$ from $m = 2.5$ which is the lower magnitude threshold applied to the catalog. Error bars in $\langle N_{as} \rangle$ are taken at the 20/80 quantiles. Dotted curve represents a $10^{0.85m}$ scaling. The inset shows the rescaled average number of directly triggered earthquakes for different magnitude thresholds and fixed $D_f = 2.3$. N is the number of events in the respective catalog.

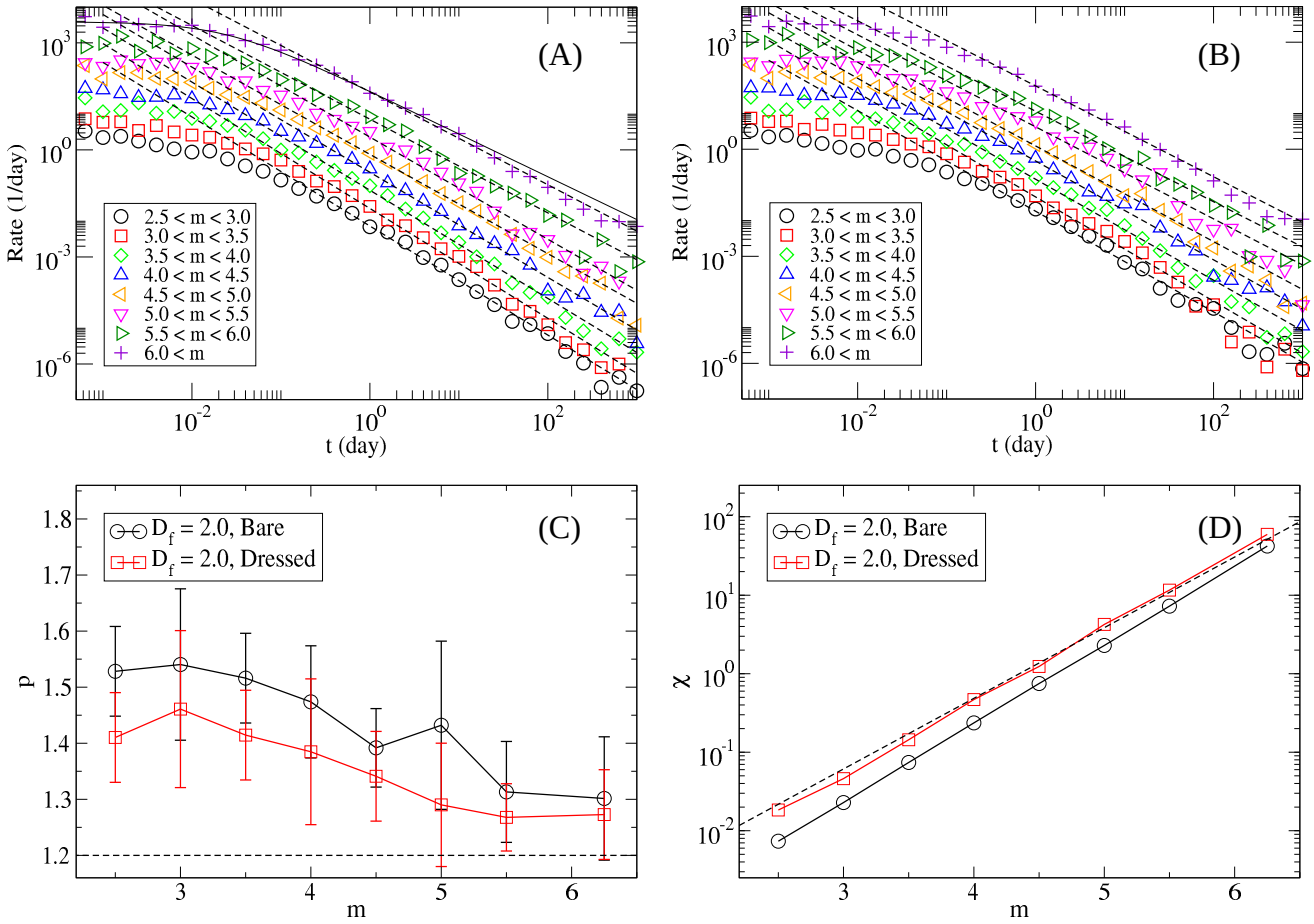

Figure 8. Average aftershock rates of mainshocks within the given magnitudes increments for the same ETAS catalog as in Fig. 2, computed using $D_f = 2.0$, considering directly triggered aftershocks only (panel A) and all aftershocks (panel B). Best power law fits to $\chi \cdot t^{-p}$ in the regime of $0.1 < t < 365$ days are shown as dashed lines for the bare and the dressed case, respectively. The solid line in panel A corresponds to the expected behavior for the given ETAS parameters. The estimated parameters are summarized in panels C and D, corresponding to an average over an ensemble of 10 ETAS catalogs generated with the same parameters (see Table 1). Note that the given error bars for p are estimated based on this ensemble and are, hence, much closer to the true uncertainties than the error bars for Southern California given in Fig. 6. In panel D, the dashed curve shows a scaling $\chi \sim 10^{0.9m}$ as a guide to the eye.

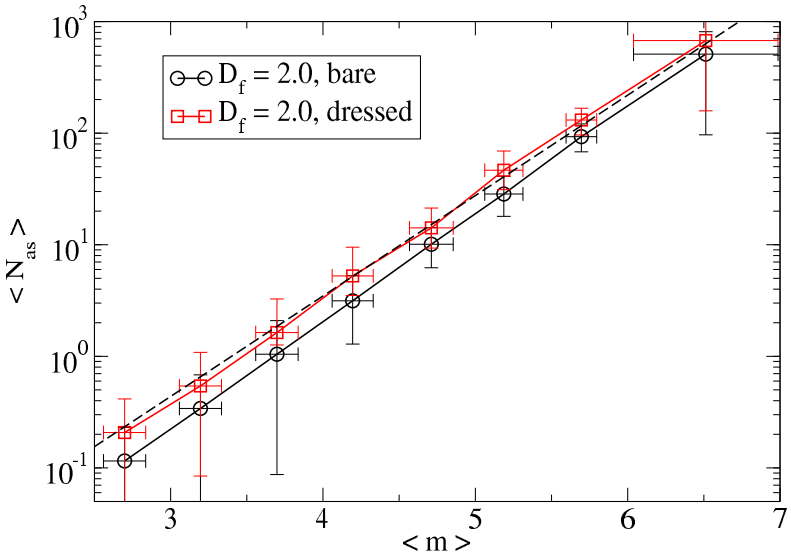

Figure 9. Average number of earthquakes, $\langle N_{as} \rangle$, triggered by earthquakes within a given magnitude range for the same ETAS catalog as in Fig. 2. The magnitude ranges are incremented by $\delta m = 0.5$ from $m_0 = 2.5$. Error bars in $\langle N_{as} \rangle$ are taken at the 20/80 quantiles. Black dashed curve represents Eq. (6) for the ETAS parameters given in Table 1.

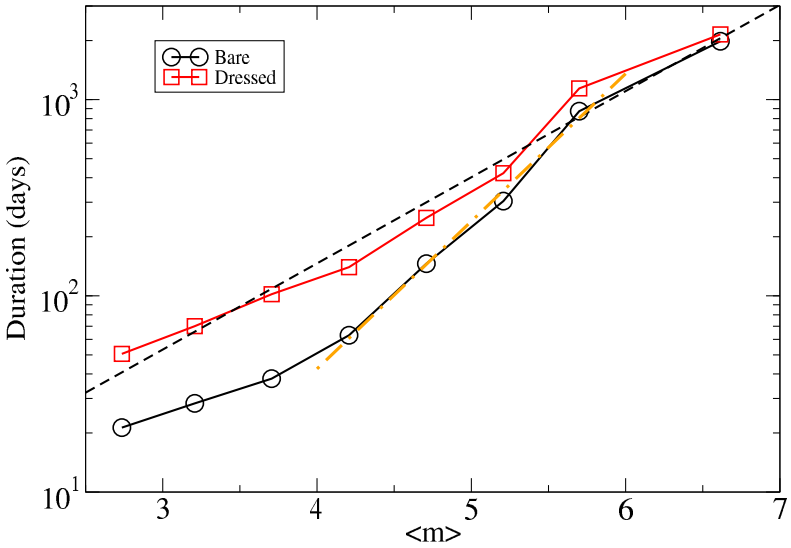

Figure 10. Southern California: Average duration of bare and dressed aftershock sequences, as a function of mainshock magnitudes. The durations of aftershock sequences were computed by calculating for all mainshocks the delays after which the last direct and last indirect aftershock occurred. $D_f = 2.3$ was used in computing the triggering cascades. These delays were then averaged conditioned on the magnitude of the mainshock: each mainshock is put into bins of width $\delta m = 0.5$, incremented from $m_{th} = 2.5$ according to their magnitudes. The dressed duration scales with mainshock magnitude approximately as $10^{0.44m}$, as indicated by the dashed curve. The bare duration scales as $10^{0.74m}$ as indicated by the dash-dotted curve, which agrees with the expected asymptotic scaling of $10^{(\alpha/p)m}$ (see text for details).

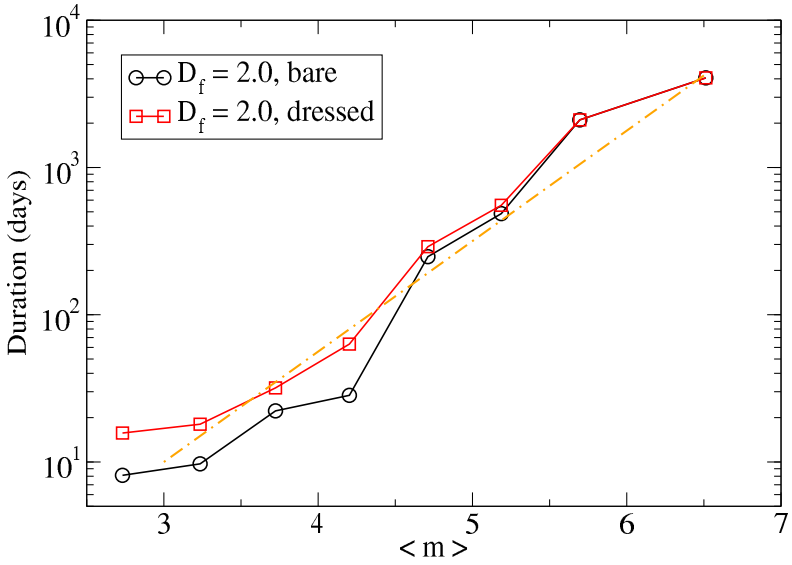

Figure 11. Average duration of bare and dressed aftershock sequences, as a function of mainshock magnitudes for the same ETAS catalog as in Fig. 2. The durations of aftershock sequences were computed by calculating for all mainshocks the delays after which the last direct and last indirect aftershock occurred. These delays were then averaged conditioned on the magnitude of the mainshock: each mainshock is put into bins of width $\delta m = 0.5$, incremented from $m_{th} = 2.5$ according to their magnitudes. As indicated by the dash-dotted curve, the bare and dressed durations scale roughly as $10^{(\alpha/p)m}$ as expected for the bare case (see text for details).

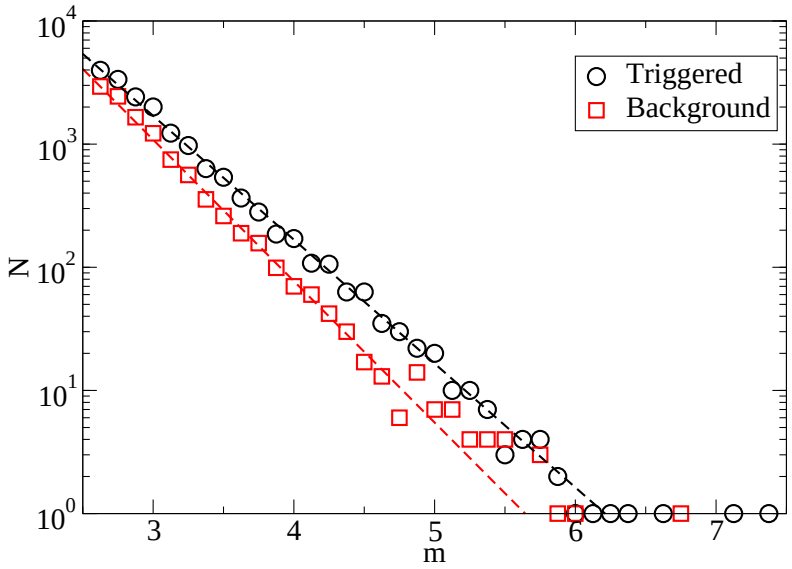

Figure 12. Southern California: Histogram of earthquake magnitudes for triggered and background events, from the triggering cascades computed using $D_f = 2.3$. The dashed curves indicate a Gutenberg-Richter scaling of the form 10^{-bm} with $b = 1.04 \pm 0.02$ for the triggered events and $b = 1.19 \pm 0.02$ for the background events. The b-values and their errors are determined using a maximum likelihood method [*Naylor et al.*, 2009]. Clearly, there are statistically significant differences between the b-values of the two earthquake populations.

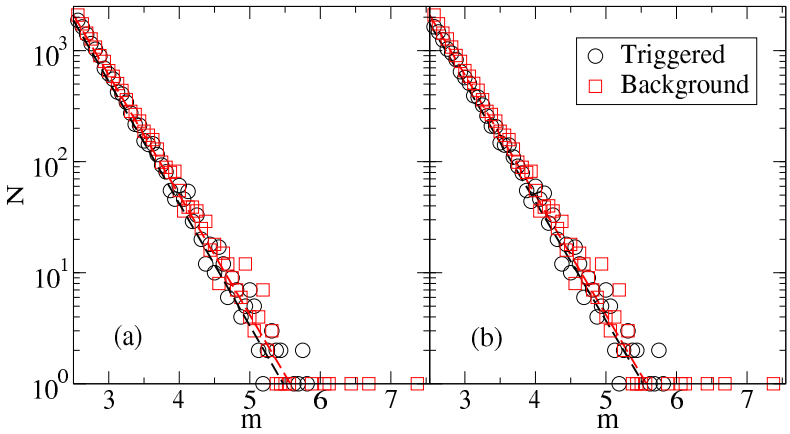

Figure 13. (a) Histograms of earthquake magnitudes for triggered and background events, from the triggering cascades for the same ETAS catalog as in Fig. 2. The dashed curves indicate a Gutenberg-Richter scaling of the form 10^{-bm} with $b = 1.08 \pm 0.02$ for background events and $b = 1.10 \pm 0.02$ for triggered events. The b-values and their errors are determined using a maximum likelihood method [*Naylor et al.*, 2009]. Note that background events constitute roughly 50% of the total number of events in the catalog. (b) The same as in (a) but with short-term aftershock incompleteness in the ETAS catalog (see text for details). In this case, $b = 1.08 \pm 0.02$ for background events and $b = 1.07 \pm 0.02$ for triggered events.

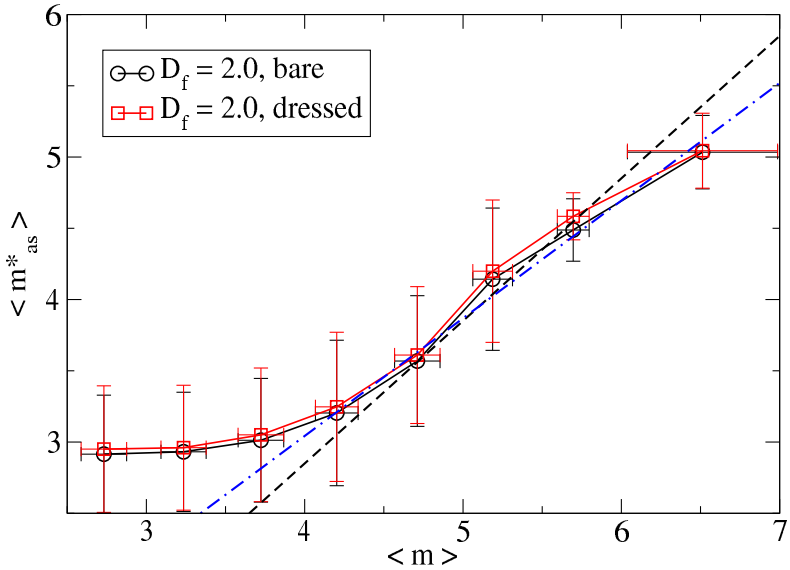

Figure 14. The magnitude of the largest aftershock, m^*_{as}, averaged over all mainshocks in magnitude bins of width $\delta m = 0.5$ for the same ETAS catalog as in Fig. 2. Dashed line corresponds to $\langle m^*_{as} \rangle = \langle m \rangle - 1.15$ and the dash-dotted line corresponds to $\langle m^*_{as} \rangle = \frac{\alpha}{b} \langle m \rangle + d$ where $d = -0.26$ is a fitted parameter and α and b are the corresponding parameters of the ETAS model.

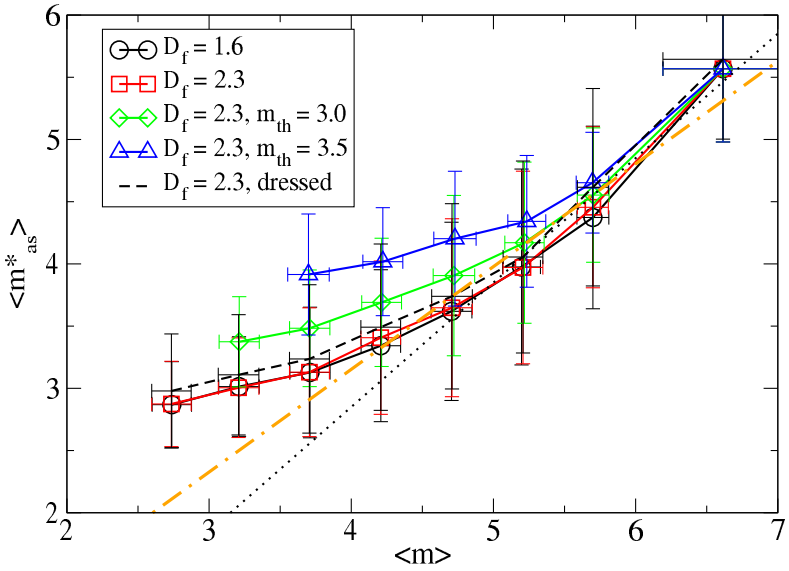

Figure 15. Southern California: The magnitude of the largest aftershock, m^*_{as}, averaged over all mainshock in magnitude bins of width $\delta m = 0.5$. The dotted straight line corresponds to $\langle m^*_{as}\rangle = \langle m\rangle - 1.15$, which represents Båth's law. The dash-dotted line corresponds to $\langle m^*_{as}\rangle \propto \frac{\alpha}{b}\langle m\rangle$, which is the behavior expected based on the ETAS model — see Fig. 14 for comparison. Here, $\alpha = 0.9$ and $b = 1.09$ as estimated above.

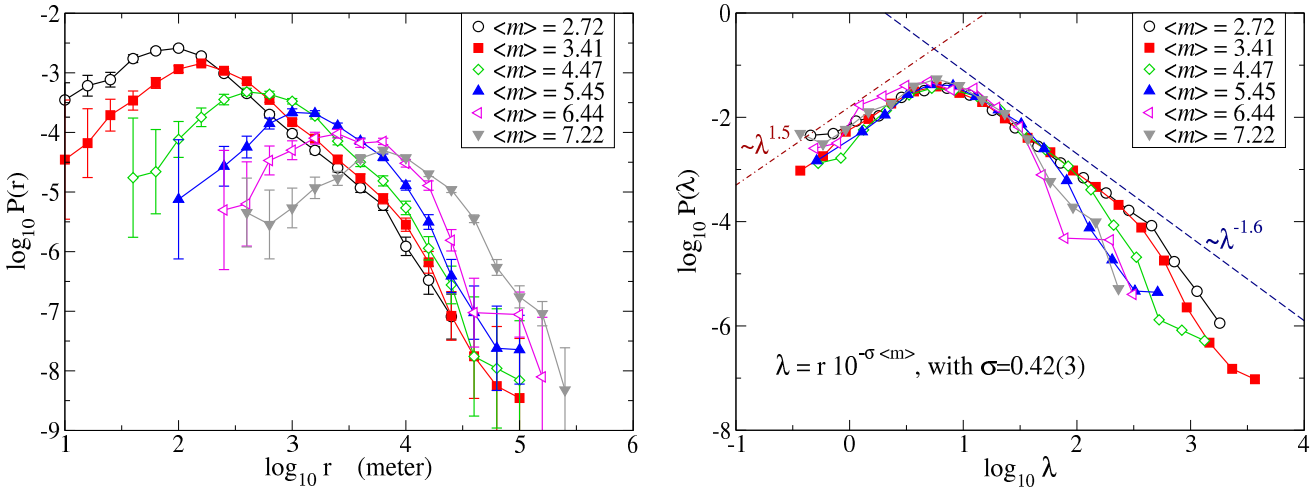

Figure 16. Southern California: i) Left: Probability density function of spatial distances between directly triggered aftershocks and their mainshocks, for different ranges of mainshock magnitudes. Bins of these magnitudes are $[2.5, 3[$, $[3, 4[$, $[4, 5[$, etc. and the average magnitudes within these bins are indicated in the legend. ii) Right: Probability density function of spatial distances between directly triggered aftershocks and their mainshocks, where the distances (in meters) are rescaled by a factor $10^{-\sigma\langle m\rangle}$ with $\sigma = 0.42$ and $\langle m\rangle$ is the average mainshock magnitude for the considered magnitude range.

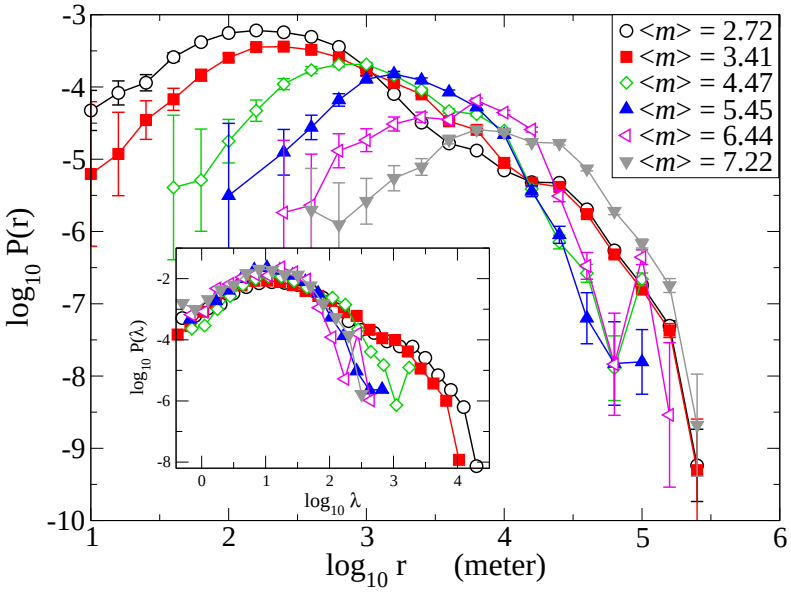

Figure 17. Southern California: Probability density function of *dressed* aftershock distances, for different mainshock magnitudes. The inset shows the probability density function of $\lambda = r10^{-\sigma\langle m\rangle}$ for $\sigma = 0.40$, where r is the dressed aftershock distance and $\langle m\rangle$ is the average mainshock magnitude for the considered magnitude range. For small λ, the function scales approximately as $\lambda^{1.5}$ and the decay after the maximum is proportional to $\lambda^{-1.6}$ for not too large λ.

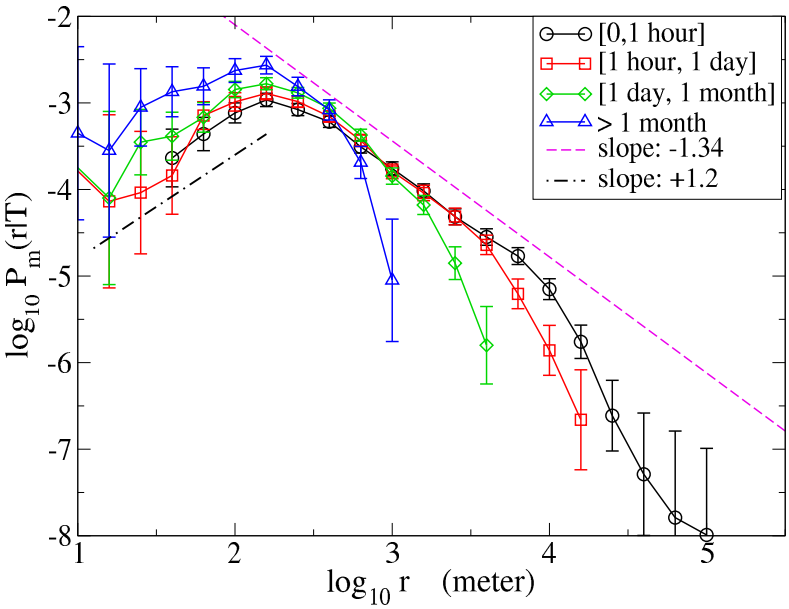

Figure 18. Southern California: Conditional probability density function of the spatial distance between mainshocks and their directly triggered aftershocks for mainshock magnitudes between 3 and 4. Different conditions with respect to the time of occurrence of the aftershocks are shown.

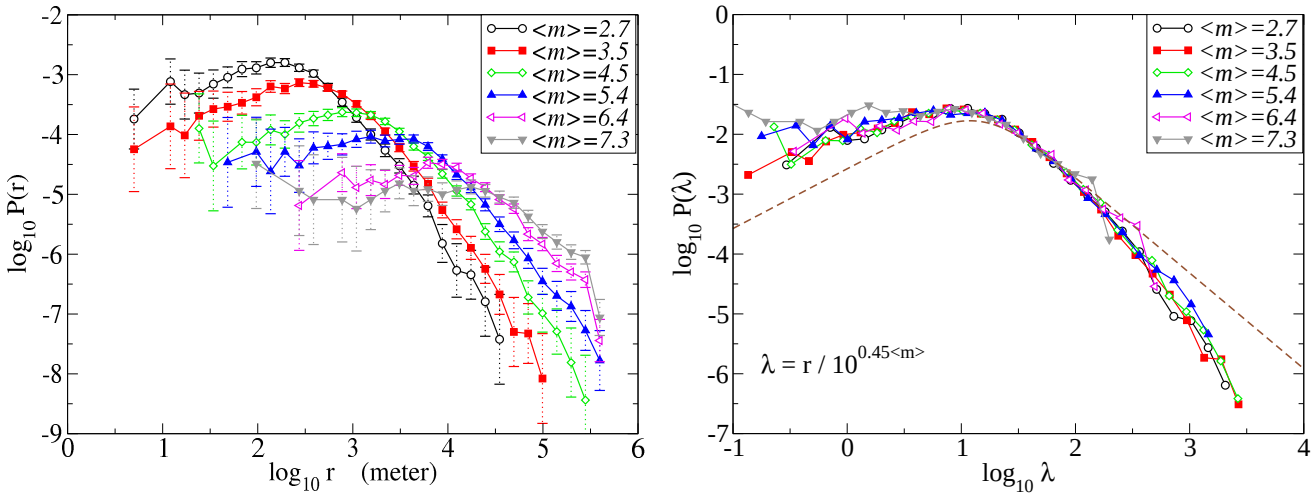

Figure 19. ETAS: i) Left: Probability density function of spatial distances between directly triggered aftershocks and their mainshocks, for different ranges of mainshock magnitudes. Bins of these magnitudes are $[2.5, 3[$, $[3, 4[$, $[4, 5[$, etc. and the average magnitudes within these bins are indicated in the legend. ii) Right: Probability density function of spatial distances between directly triggered aftershocks and their mainshocks, where the distances (in meters) are rescaled by a factor $10^{-\sigma\langle m\rangle}$ with $\sigma = 0.45$ and $\langle m\rangle$ is the average mainshock magnitude for the considered magnitude range. The dashed line corresponds to the expected behavior given by Eq. (8).

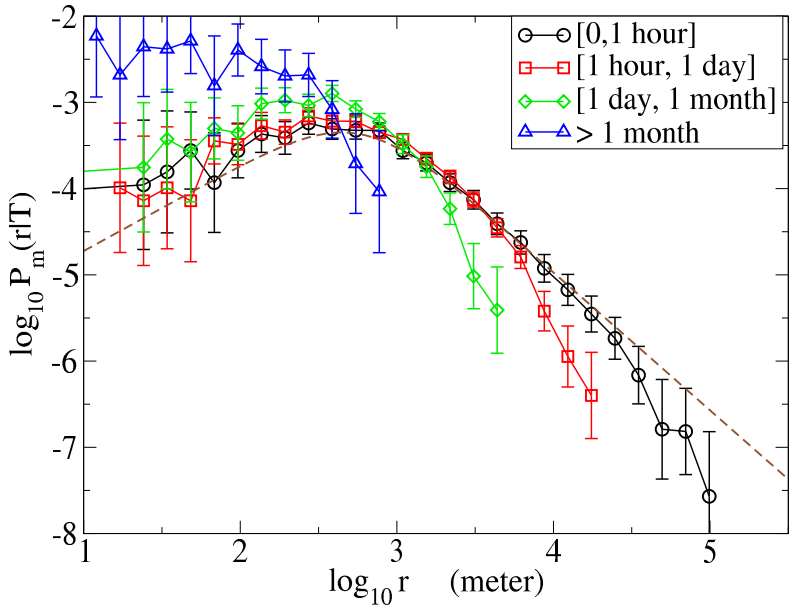

Figure 20. ETAS: Conditional probability density function of the spatial distance between mainshocks and their directly triggered aftershocks for mainshock magnitudes between 3 and 4. Different conditions with respect to the time of occurrence of the aftershocks are shown. The dashed line corresponds to the expected behavior given by Eq. (8).