Masking of earthquake triggering behavior by a high background rate and implications for epidemic-type aftershock sequence inversions

Masking of earthquake triggering behavior by a high background rate and implications for epidemic-type aftershock sequence inversions

S. Touati,[1] M. Naylor,[1] I. G. Main,[1] and M. Christie[2]

Received 15 March 2010; revised 1 December 2010; accepted 17 December 2010; published 8 March 2011.

[1] We examine the effects of the spontaneous background event rate and aftershock triggering characteristics on the temporal statistics of seismicity in the epidemic-type aftershock sequence model. Recent work has shown that the earthquake interevent time distribution is generally bimodal: a superposition of a gamma component from triggered aftershocks at short time intervals and an exponential component at longer intervals from spontaneous events and the overlapping of independent aftershock sequences. The relative size of these two components varies between catalogs, so there is no simple, universal scaling; at the extreme of high spontaneous rate, e.g., in large regions, the high probability of temporally overlapping aftershock sequences causes the exponential component to dominate. Here we further explore the effects of both the spontaneous rate and the aftershock triggering parameters. We show that the analytical theory of Saichev and Sornette (2007), although valid under their assumptions, gives the impression of a more "universal" behavior if used outside its stated range of applicability. We also show that within the high-overlap (high-spontaneous rate) regime, a maximum likelihood inversion of the model's temporal parameters is both less accurate and biased; specifically, the background rate is systematically overestimated. This has implications on the suitable range of region sizes for which parameter inversion may be reliable and must therefore be taken into account in any inversion for temporal variations in background rate in time-dependent hazard calculation.

Citation: Touati, S., M. Naylor, I. G. Main, and M. Christie (2011), Masking of earthquake triggering behavior by a high background rate and implications for epidemic-type aftershock sequence inversions, *J. Geophys. Res.*, *116*, B03304, doi:10.1029/2010JB007544.

1. Introduction

[2] Over the past decade, much literature has been published on the distribution of time intervals between successive earthquakes in a recorded catalog [*Bak et al.*, 2002; *Corral*, 2003, 2004; *Davidsen and Goltz*, 2004; *Hainzl et al.*, 2006; *Shcherbakov et al.*, 2005; *Molchan*, 2005]. It has become commonplace to model these interevent times using a single gamma distribution, which has led some authors to propose universality, extending from tectonic earthquakes [*Bak et al.*, 2002; *Corral*, 2003, 2004; *de Arcangelis et al.*, 2008; *Lennartz et al.*, 2008] to rock fracture [*Davidsen et al.*, 2007]. This idea follows from the observation that rescaling the data by region size and magnitude cutoff [*Bak et al.*, 2002] or the mean event rate [*Corral*, 2003] causes the distribution curves to collapse on top of one another. Analytical investigations by *Saichev and Sornette* [2006, 2007] confirmed an approximately universal form over the range of parameters considered typical. However, our recent work [*Touati et al.*, 2009] generalizes these observations beyond the data selection criteria used by Saichev and Sornette and shows that the long interevent times do not scale in a "universal" way from the short ones. If the distribution is to be described as universal, that term is only applicable in a different sense than originally proposed [*Bottiglieri et al.*, 2010].

[3] We start from the point of view that earthquakes are in two broad categories: First, a spontaneous category, which is usually an approximate Poisson process composed mainly of "background" tectonically driven events, but likely also includes contributions from other driving forces, e.g., fluids as well as events that are triggered by other events below the detection threshold. These events may be considered independent of each other. Secondly, a triggered category, consisting of events triggered by others through the time-dependent relaxation of the earth's crust, and whose rate follows the empirical Omori aftershock law. When considering interevent times, we are looking at temporal intervals between pairs of events, which, it follows, are also of two types: either the events are causally related (that is, part of a common sequence of triggered events) or else they are

[1]University of Edinburgh, School of GeoSciences, Edinburgh, UK.
[2]Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, UK.

0148-0227/11/2010JB007544

unrelated and independent of each other. We refer to the former as same-sequence pairs and the latter as intersequence pairs of events. *Touati et al.* [2009] showed that interevent time distributions can take on a range of shapes from an exponential form to a double-peaked curve, depending on the balance between aftershock production and the rate of spontaneous background events; the gamma distribution is thus an approximation, and universality does not strictly hold. We presented a new understanding of the earthquake interevent time distribution as an essentially bimodal mixture distribution, comprising the contributions from precisely the two categories of event pairs described at the start of this paragraph. Those same-sequence events that happen to occur consecutively in the catalog produce gamma-distributed interevent times, since the temporal decay of the aftershock rate follows a power law. (If each sequence were left uninterrupted by further spontaneous events, the interevent time distribution would also be a power law.) The occurrence of intersequence events consecutively, on the other hand, produces an exponential interevent time distribution, as long as the different spontaneous seeding events from which they arise are themselves a Poisson process, a reasonable assumption in the case of stationary tectonic loading conditions. Thus the observation of bimodal distributions in earthquake data can be explained in terms of the peaks of these two contributions being well resolved.

[4] In this paper we extend this analysis of interevent time distributions. We start by describing the epidemic-type aftershock sequence (ETAS) model (section 2), a stochastic model utilized throughout the paper in which aftershock triggering is viewed as a multigenerational branching process. We then show the effect of the spontaneous event rate on the temporal statistics in section 3. We test our results against the analytical model for earthquake interevent times of *Saichev and Sornette* [2007] and confirm that the latter is applicable within its stated assumptions, i.e., only at higher interevent times; at low interevent times inappropriate extrapolation of their model would produce an overestimate of occurrence, and give a false impression of broader-band universal scaling. In section 4 we demonstrate the relevance of our paradigm to real data and show the relationship between seeding rate in the ETAS model and region size in real data. We then explore the interaction between the effects of the triggering parameters and the seeding rate in section 5. We show that the effect of any changes in the branching parameters is diminished as the seeding rate is increased; in other words, seismicity that is characterized by an exponential interevent time distribution has a high redundancy in terms of the effective triggering parameters. This simple observation sparked by our understanding of the interevent time distribution turns out to have deep consequences for the inversion of temporal ETAS parameters, in that significantly less accurate results are obtained for higher seeding rates (as in larger or more tectonically active regions). We point out the implications for the possibility of detecting a varying background rate.

2. ETAS Model

[5] The ETAS model is a stochastic point process model based on well-known empirical laws of earthquake occurrence. The process is seeded by spontaneous, commonly called "background," events, occurring as a Poisson process in time with constant rate μ, to represent steady state tectonic driving. Aftershocks may then be triggered by all events; the aftershock rate $n(t)$ decays as a power law in time after a sizable event according to the Omori law

$$n(t) = \frac{K}{(c+t)^p}, \tag{1}$$

where K, c and p are constants and t refers to time. The number of aftershocks produced depends exponentially on the magnitude of the parent event. Each event's magnitude is selected independently from the Gutenberg-Richter distribution,

$$\log N = a - bm, \tag{2}$$

where N is the number of events in a given time period with magnitude $\geq m$, and a and b are constants. There is a lower threshold magnitude m_0; there may or may not be an upper magnitude cutoff. The ETAS conditional intensity function λ is a function of time and of the history of the process H_t at time t:

$$\lambda(t|H_t) = \mu + K \sum_{i:t<t_i} \exp(\alpha(m_i - m_0))(t - t_i + c)^{-p}, \tag{3}$$

where t_i are the times of the past events and m_i are their magnitudes. Thus, the five ETAS parameters are μ, the spontaneous event rate; c, p and K, the Omori law parameters; and α, the productivity parameter. Because we cannot distinguish a priori between foreshocks, main shocks and aftershocks, the model regards all events as capable of triggering further events. Thus, each spontaneous event may result in a cascade of nested aftershock sequences known as a "global" aftershock sequence [*Helmstetter and Sornette*, 2002].

[6] An alternative parameterization which we will use is obtained by making the substitution $A = K/c^p$:

$$\lambda(t|H_t) = \mu + A \sum_{i:t<t_i} \exp(\alpha(m_i - m_0))\left(1 + \frac{t - t_i}{c}\right)^{-p}. \tag{4}$$

For the spatiotemporal ETAS model, the conditional intensity varies with spatial coordinates x and y in such a way that the total conditional intensity for the entire region is equal to the temporal-only ETAS conditional intensity (equation (4)). Spatial clustering of aftershocks is accounted for by a spatial kernel $f_i(x - x_i, y - y_i)$, one implementation of which is

$$f_i(x - x_i, y - y_i) = \frac{(q-1)d^{2(q-1)}}{\pi[x^2 + y^2 + d^2]^q}, \tag{5}$$

where d and q are parameters [*Hainzl et al.*, 2008].

[7] A useful quantity which we will make use of in this paper is the branching ratio, defined as the average number of aftershocks generated by each event, which is obtained by integrating $A \exp(\alpha m)(1 + \frac{t}{c})^{-p}$ over both times and magnitudes from 0 to ∞, and is given by

$$n = \frac{Ac}{p-1}\frac{\beta}{\beta - \alpha}. \tag{6}$$

The parameter μ, being the average frequency of spontaneous events, has no effect on aftershock generation; it instead represents the average rate at which aftershock sequences are initiated and so determines the temporal overlap extent of separate sequences. It represents the roughly stationary component due mainly to tectonic loading, and we show in this paper that it may be identified as a proxy for region size. We first show the effect of μ on the interevent time distribution and then consider the effects of the other parameters.

3. Bimodality and the Seeding Rate

[8] To demonstrate the effect of the seeding rate parameter μ on the interevent time distribution for the ETAS model, we simulate synthetic earthquake catalogs of 100,000 events each, for four different values of μ, and plot histograms of the interevent times τ (Figure 1). We plot both the logarithm of the count in logarithmic bins, which optimally shows the bimodality, and the more usual probability density function (PDF) obtained by normalizing the counts by the bin widths, which tends visually to mask the bimodal form of the data [*Touati et al.*, 2009]. The effect of plotting a simple histogram is merely to increase the slope by 1 with respect to the equivalent PDF plot [*Bonnet et al.*, 2001, Figure 2b].

[9] The origin of the histogram shape and its variation with μ becomes clear when we superpose the two subsets of data arising from the two categories of event pairs (Figure 1, left). We henceforth refer to same-sequence and intersequence intervals, following our stated convention for the event pair categories giving rise to these intervals. In increasing the seeding rate, more earthquake sequences overlap within the data set; the interevent times thus become dominated by those exponentially distributed ones arising from intersequence event pairs (see Figure 2), and the overall distribution becomes unimodal and exponential-like. The same-sequence interevent time distribution becomes truncated at shorter interevent times due to the typical sequence being interrupted earlier in its progression (note the change in scale of the τ axis between the different plots).

[10] Figure 1 (right) shows comb diagrams of the events in a 1000-event subset of each catalog, along with a cumulative count of the events, giving further visual illustration of the effect. As μ is increased there is clearly less variability in the temporal occurrence of events, and the cumulative count becomes a smooth increasing function of time. This is in line with an exponential-like distribution of interevent times. The stepped cumulative distribution at low μ, on the other hand, is characteristic of well-resolved aftershock sequences.

[11] To put this in the context of previous reported results, first we note that several studies motivated by the work of *Corral* [2003] have selected, from real earthquake data sets and experimental data sets, periods showing short-term stationarity in the event rate. (Short-term "stationary" periods emerge by chance in any data set, when there is a local absence of large events which would cause the event rate to change rapidly. In the longer term, of course, all tectonically driven regimes can be considered stationary.) This requirement allows a straightforward rescaling of the data by the mean event rate, but it appears that as a side effect, the data selection procedure generally results in a histogram similar to the $\mu = 1$ case in Figure 1, which approximates a "universal" gamma distribution. Thus, an apparently benign and otherwise quite reasonable filter for apparently "stationary" periods tunes the search to regimes where there is sufficient overlap of aftershock sequences to blur the underlying bimodality.

[12] Second, we consider the analytical PDF of *Saichev and Sornette* [2007], who report that the previously discussed universality is merely approximate and relies on a value of p close to (but greater than) 1. From our point of view, a smaller p value would certainly reduce the appearance of bimodality by reducing the power law decay exponent of the same-sequence distribution (see section 5), although we would suggest that there are additional alternative ways of observing a single power law, such as by controlling the effective overlap level through data selection. We compare our simulations with the analytical formulation of *Saichev and Sornette* [2007, equation (36)]) in Figure 1 (middle), which shows the data in the more commonly used format of a normalized PDF rescaled by the mean event rate. We note that there is a divergence between the theoretical and simulated curves at interevent times comparable to c and smaller, due to the clearly stated key assumption $\tau \gg c$ in their derivation. At high μ this assumption becomes more limiting due to the average interevent time shortening toward c, reducing the bandwidth over which the distribution function is valid, and giving the false impression of a more universal form if the range of applicability is mistakenly overlooked. Thus we present here a generalization of the work by these authors and Corral, extending the analysis beyond data sets that appear stationary in the short term, or have p close to 1, and extending the bandwidth of interevent times below c.

[13] We can further elucidate the relation between *Saichev and Sornette*'s [2007] theory and our understanding by plotting its two components separately: first, $\phi(x, m)$, which is the exponential probability of observing no events in the rescaled time interval $x = \lambda(m)\tau$; and second, the other component, which comes from the twice differentiation of $\phi(x, m)$ in obtaining the interevent time PDF. Figure 3 shows that $\phi(x, m)$ corresponds to the exponential intersequence component of the distribution. This follows from the fact that $\phi(x, m)$ is concerned with the longest observable interevent times, contingent on both the spontaneous rate and aftershock production pattern. In differentiating twice with respect to time, the constant background rate component of the exponent is eliminated and the effect of the aftershocks is revealed in the other component: the same-sequence power law decay is captured along with a shallower power law leading toward the intersequence peak.

[14] Finally, we note that the universality of *Bottiglieri et al.* [2010] relies on being able to compensate for the variation of μ through a corresponding variation in c, thereby, as we understand it, maintaining a roughly constant overlap extent between the separate aftershock sequences in the catalog. This is generally not the case in selecting real data, particularly in catalogs of different spatial extent, as we show next. To summarize our contribution so far, we replace the universality paradigm, which has been shown both by *Saichev and Sornette* [2007] and by Bottiglieri et al. to be of limited scope, with an intuitive, physically motivated

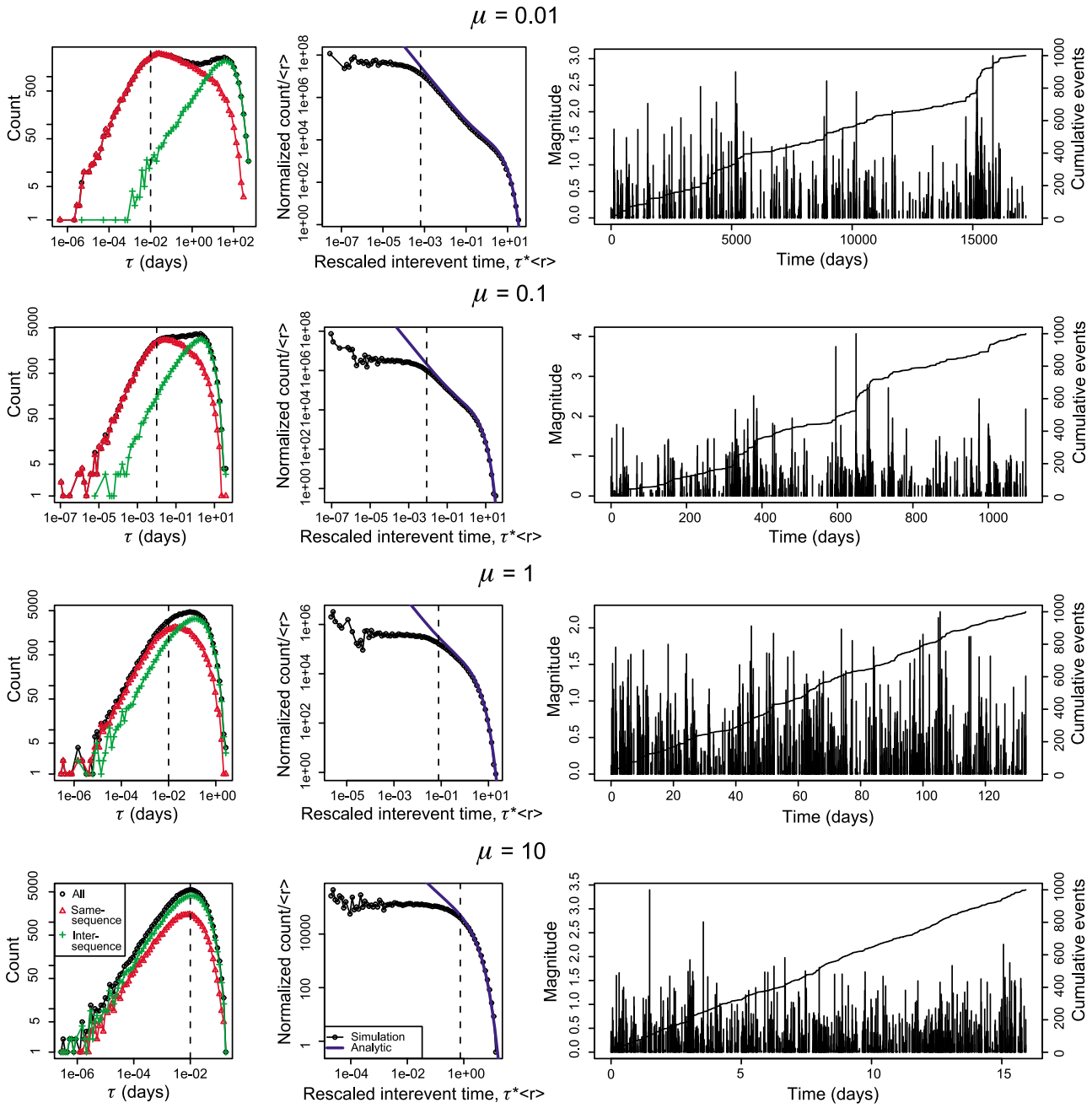

Figure 1. The effect of varying the ETAS parameter μ on (left) the interevent time histogram, with sub-distributions for correlated (red triangle) and uncorrelated (green cross) event pairs superposed; (middle) the normalized PDF rescaled by the mean event rate, plotted along with the analytic solution of *Saichev and Sornette* [2007]; and (right) the time series and cumulative event count for a sample of 1000 events. The value of c ($c\langle r\rangle$) is indicated by a dotted vertical line in the histogram (PDF) plots, which represents a lower limit of Saichev and Sornette's solution's range of applicability. The higher μ is, the more exponential the interevent time distribution and the smoother the cumulative event count increase. Values of μ used were as shown; other ETAS parameters were held fixed at $A = 10$, $\alpha = 1$, $c = 0.01$, $p = 1.2$; 100,000 events were simulated in each case.

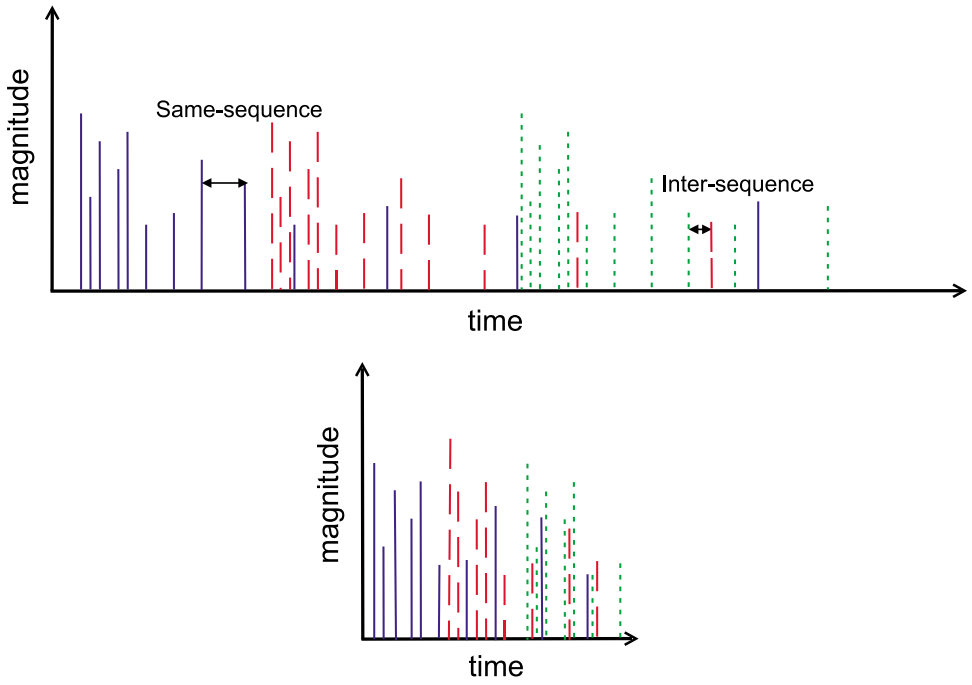

Figure 2. Illustration of the effect of seeding rate: when changing the seeding rate from (top) low to (bottom) high, distinct aftershock sequences, shown in different colors and line patterns, are initiated more frequently; this is effectively a translation of the sequences along the time axis, without compression. (Figure 2 (bottom) spans a shorter time to avoid the need for further sequences to be introduced.) Thus, a higher overlap between the aftershock sequences causes more of the interevent times to result from intersequence event pairs.

paradigm based on the classification of interevent times as either same-sequence or intersequence intervals.

4. Seeding Rate and Region Size

[15] We now turn our attention to real data from the worldwide Preliminary Determination of Epicenters (PDE) catalog, and demonstrate that the ETAS model is, in principle, a valid description of seismicity even on large spatial scales, with the parameter μ acting as a proxy for the size of the region.

[16] We choose data between 1 January 1969 and 1 January 2005 and apply a minimum cutoff magnitude of 5.0. The catalog is complete over this time period and magnitude range, as indicated by the absence of any roll-off in the magnitude-frequency relation toward smaller magnitudes. We first select events belonging to a circular region centered on 130° longitude and 0° latitude. This position is arbitrary but is in the most active seismic region, giving us a maximal amount of data. By varying the radius of this circle, we can tune the number of events included to correspond to the variation in μ in Figure 1; that is, approximately a tenfold increase in the event rate at each radius increase. (We cannot simply increase the area tenfold to create an analogy with Figure 1, since earthquake occurrence is highly inhomogeneous.)

[17] We must remove all events below magnitude 5.0 due to incompleteness of the catalog. Figure 4 shows the result; the error bars shown are based on 95% confidence limits of a binomial error distribution. Since the apparent branching ratio decreases as the magnitude threshold is raised [*Sornette and Werner*, 2005], the relatively high magnitude threshold would be expected to significantly reduce the same-sequence component and the appearance of bimodality in the interevent time distribution. However, we still find clear indication

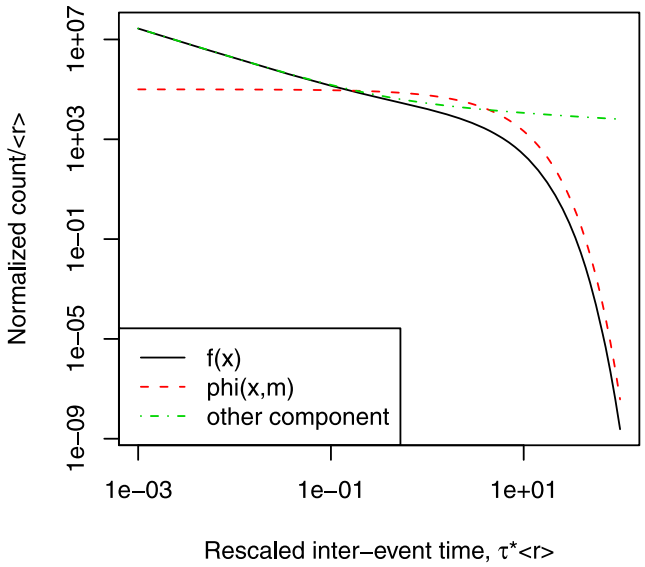

Figure 3. *Saichev and Sornette*'s [2007, equation (36)] analytical PDF, along with its two components: the exponential $\phi(x, m)$ and the other term which results from twice differentiation of $\phi(x, m)$. ETAS parameters were $\mu = 0.1$, $A = 10$, $\alpha = 1$, $c = 0.01$, $p = 1.2$.

of bimodality at smaller region sizes, with a transition to a unimodal form as the region size grows, just like in an ETAS model with increasing μ.

[18] It would be an oversimplification for most modeling purposes to regard global seismicity as a single ETAS parameterization, and we are not promoting such an idea. However, the use of the whole world presents a clear end-member case that catalogs of progressively smaller regions generally tend away from, and Figure 4 shows that this end-member case is at least qualitatively comparable to simply using a high μ in the ETAS model. The effective parameters of seismicity are not constant with respect to space, yet we can say that the global averaging of the effective parameters through this interevent times analysis produces identifiable bimodality as expected by the simple ETAS model. Stacking together seismicity of different Poissonian seeding rates through the concatenation of smaller regions should not worry us because the sum of several Poisson processes will be yet another Poisson process. The point here is to illustrate our paradigm for examining interevent times, and to demonstrate that the transition from bimodality to unimodality with an increasing μ has some genuine physical meaning in terms of the region size, and hence that the often subjective choice of what constitutes a "region" can strongly condition the model parameters.

[19] We can use curve fitting to demonstrate that the decay in the interevent time distribution at large τ is exponential at all region sizes, which further supports the comparison with ETAS. It is convenient to note two points: first, when plotting exponentially distributed data as a logarithmic histogram without normalization, there is a peak; analytically, the x value at which this peak occurs is $1/\lambda$, where λ is the exponential parameter. Secondly, in the case of ETAS interevent times, the peak in the underlying intersequence component coincides with a peak in the overall distribution (Figure 1). We therefore assume that we can fit an exponential distribution to the latter portion of the histogram simply by reading off the peak value of interevent time. The red curves in Figure 4 result from this fitting; in each case, the fit is visually convincing, and so we consider these fitted exponentials to represent the inferred contribution from intersequence events, with a possible amplitude adjustment in the larger region cases. (Note, however, that we cannot infer the spontaneous event rate from this component of the distribution because the intersequence component arises from both aftershocks and spontaneous events; see section 5.) It is also clear from Figure 4 that the additional contribution from same-sequence event pairs exists, and grows more prominent as the region size is decreased.

[20] We may question whether the unimodal, exponential-like distribution observed for the largest region can be explained by the fact that in such a large spatial area, we are combining data from different seismic regions, with different effective triggering parameters and spontaneous rates, resulting in a Poisson process through this mixing. A second approach to check this is to repeat the analysis with mixed

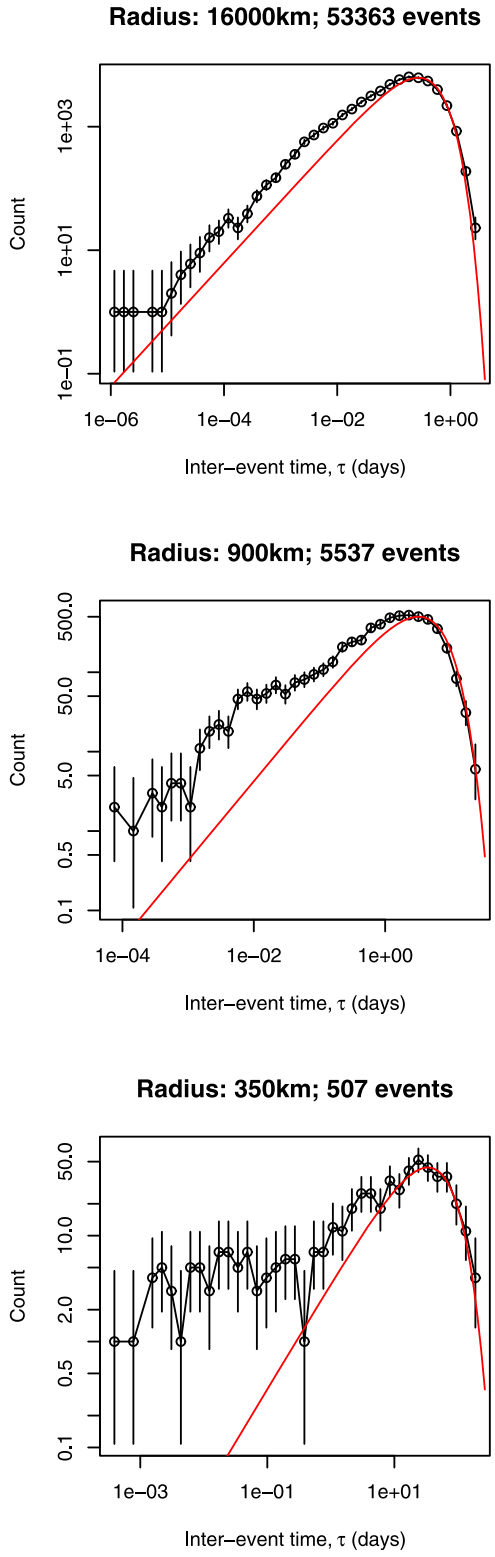

Figure 4. The effect of varying the spatial window size on the interevent time histogram, for the PDE catalog from 1 January 1964 to 1 January 2005. Red line shows the exponential fit for the background event pairs, based on the peak value, as described in the text. The smaller the region size is, the greater the deviations from this exponential distribution at short interevent times, indicating a greater number of dependent event pairs occurring sequentially. Circular regions centered on 180° longitude and 0° latitude were used. A minimum magnitude cutoff of 5.0 was applied to ensure catalog completeness.

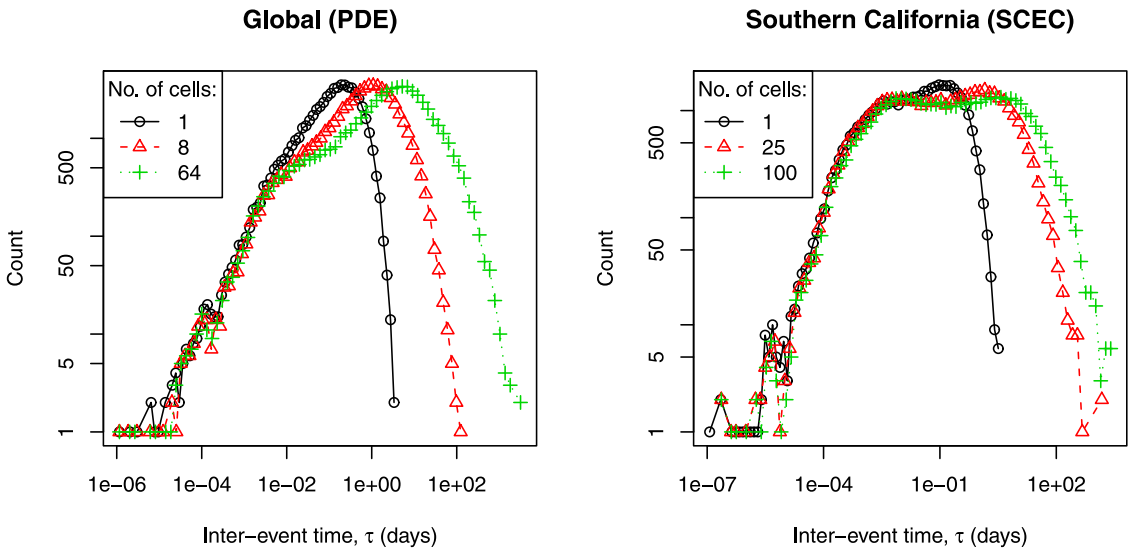

Figure 5. The effect of varying the spatial window size on the interevent time histogram (left) for the PDE catalog from 1 January 1964 to 1 January 2005 and (right) for the SCEC catalog from 1 January 1984 to 1 December 2000. In the case of more than one spatial cell, the plots were generated by mixing interevent times from the cells as described in the text. A minimum magnitude cutoff of 5.0 (2.2) was applied to ensure catalog completeness for the PDE (SCEC) data.

data at all region sizes. This additionally allows us to use the events from the whole globe and avoid such large counting errors in the histograms.

[21] We split the earth into two-dimensional spatial cells of equal size. It is most convenient to use, as cell boundaries, only those great circles that are lines of constant longitude or latitude; that is, the equator, and various lines of longitude. The interevent times are generated for each cell separately, and then a histogram is produced after concatenating data from all cells, with the procedure repeated for different cell sizes. The results of this are shown in Figure 5 (left). There is remarkably clear bimodality at smaller region sizes, despite the inevitable smoothing due to the mixing of data from different cells. This confirms that mixing of data alone cannot explain the exponential-like distribution of interevent times coming from large regions. Incidentally, Figure 5 (right) shows a similar plot for a regional catalog (Southern California Earthquake Center (SCEC), using events between 1 January 1984 and 1 December 2000, up to longitude 245°, and magnitudes above 2.2) broken into rectangular cells of (approximately, on the curved Earth surface) equal area; for the whole region (i.e., 1 cell) the form is similar to that for the PDE catalog split into 64 cells, and so Figure 5 (right) shows a kind of continuation of the spatial shrinking with the result that the same-sequence peak grows more dominant.

[22] We can conclude, then, that in a large region, the vast majority of interevent times come from unrelated, intersequence event pairs, and infer that the effective μ increases with region size. This means μ is in effect not a material parameter at the regional scale, since an arbitrary choice of region size can reduce or increase its value. In principle its spatial density function may be more representative as a material parameter that varies with tectonic region, but less so with region size.

[23] Although the interevent time distribution becomes more exponential-like as region size is increased, this is not to say that large regions have a lower effective branching ratio. It would seem to be obvious that the branching ratio in a large region cannot be lower than that of the smaller spaces of which it is composed; certainly in our ETAS analogue, we do not vary the branching ratio in altering μ; we know that the clustering in the catalog has merely become "hidden" due to an increased proportion of intersequence event pairs occurring sequentially. However, as we show in section 5, inversion of ETAS parameters can give the false impression that the branching ratio decreases with increasing region size.

[24] Spatial information can be used to overcome the temporal overlapping of aftershock sequences in declustering techniques, since triggering probability decreases rapidly with distance [*Huc and Main*, 2003]. These techniques provide a way of categorizing intervals as same-sequence or intersequence for real data, to further investigate the relevance of our ETAS-based paradigm. *Zaliapin et al.* [2008] presented a method for analyzing the spatiotemporal clustering within a catalog and using this information to decluster. They define a relationship ("distance" n_{ij}) between one event and a later event in terms of the time and space interval between the events; an event's nearest neighbor is then defined as the event for which this n_{ij} value is minimal. They find that for ETAS catalogs, the values fall naturally into two clusters; one for same-sequence aftershocks, which occur close in space and time to each other, and another for intersequence events which are more separated in space and time. This observation of bimodality, in the clusters and their marginal distributions, complements very well our interevent times analysis above.

[25] Figure 6 (top left) shows space and time components of the simple nearest-neighbor distance defined as $n_{ij} = r\tau$, for a spatial ETAS simulation (using the kernel (5)). In Figure 6 (right), we show a histogram of the time components, that is, the time interval to the nearest neighbor for each event. We can label each interval as same-sequence or

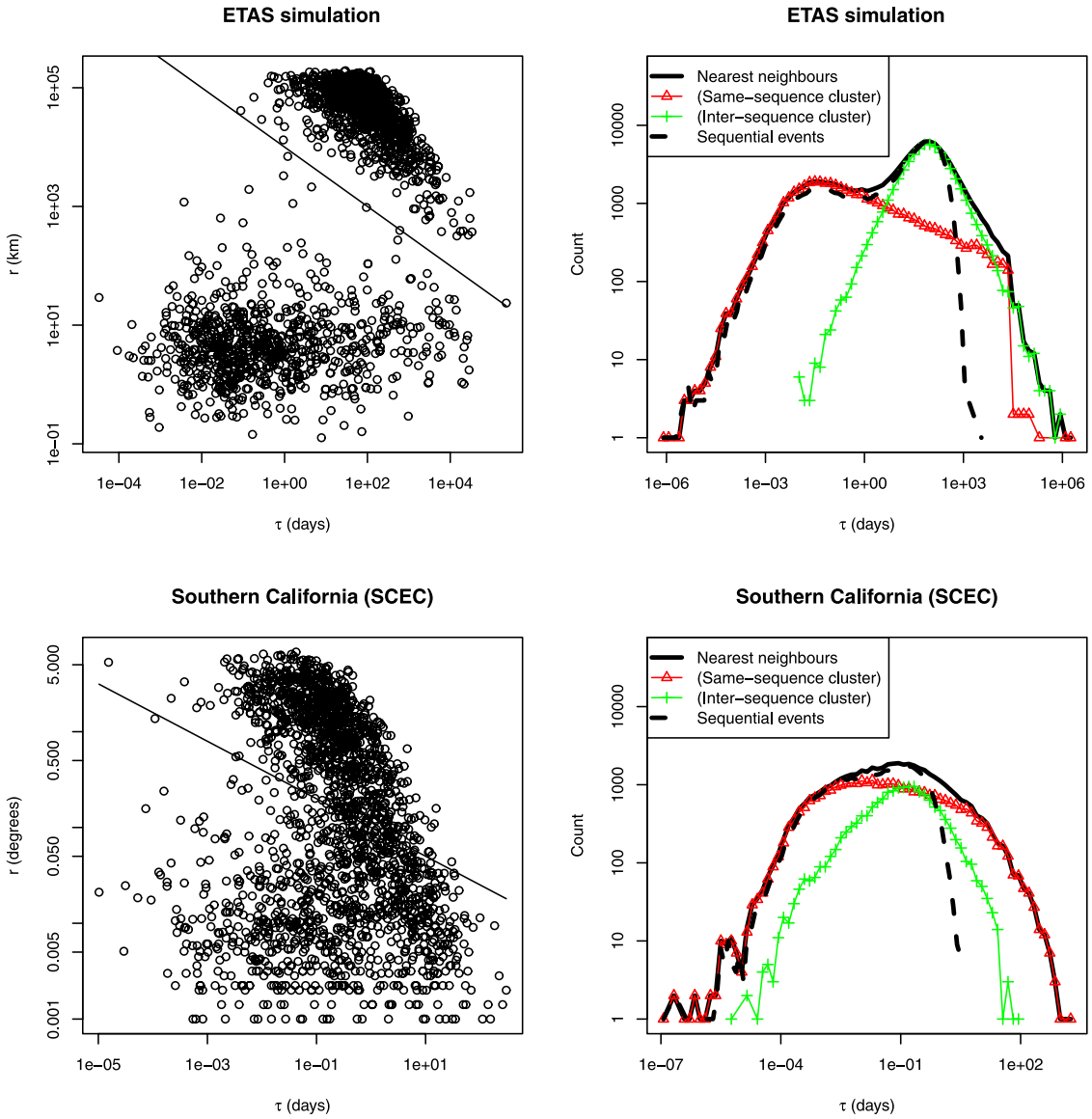

Figure 6. (left) Space (r) and time (τ) distances to the nearest spatiotemporal neighbor for each event of (top) a spatial ETAS simulation and (bottom) a Southern California catalog. The first 2000 events of each catalog only are shown. (right) Histogram of time intervals to the nearest neighbor (solid black line). Same-sequence and intersequence event pairs are superposed in red triangles and green crosses, respectively, classified according to which cluster they fall into as defined by the dividing lines shown in Figure 6 (left). The straightforward interevent time histogram is also shown by the dashed black line. Note that the intervals are not identical, as the spatial information allows more same-sequence events to be selected as nearest neighbors despite not being sequential in time alone. ETAS parameter values used are $\mu = 0.01$, $A = 5$, $\alpha = 1$, $c = 0.01$, $p = 1.2$, $d = 2$, $q = 1.5$; SCEC events between 1984 and 2000 of magnitude 2.2 and above were selected.

intersequence based on which cluster it falls into, as determined by setting a dividing line as shown in Figure 6 (left) and superpose histograms for these categories. The interevent time histogram is shown as a dashed line for comparison. Note that the intervals to the nearest neighbors are in general different from the interevent times, as the nearest neighbor may not be the next event in time if there is a later event that is closer in space. The intervals to the nearest neighbors therefore contain a higher proportion of same-sequence relationships, which is evident from the extended power law segment, particularly at longer intervals for which there are no interevent times. The division between the clusters is somewhat arbitrary, and inevitably overlap of the clusters will lead to mislabeling of some intervals, but a clear power law is recovered nonetheless.

[26] Figure 6 (bottom) repeats this analysis for the Southern California Earthquake Center (SCEC) catalog. The two clusters overlap to a greater extent, likely due to variations in the seismicity parameters throughout the region and also the inappropriateness of the hypocentral distance for aftershocks of large events, and so mislabeling of intervals is a bigger problem, but a clear power law can nevertheless be seen in the

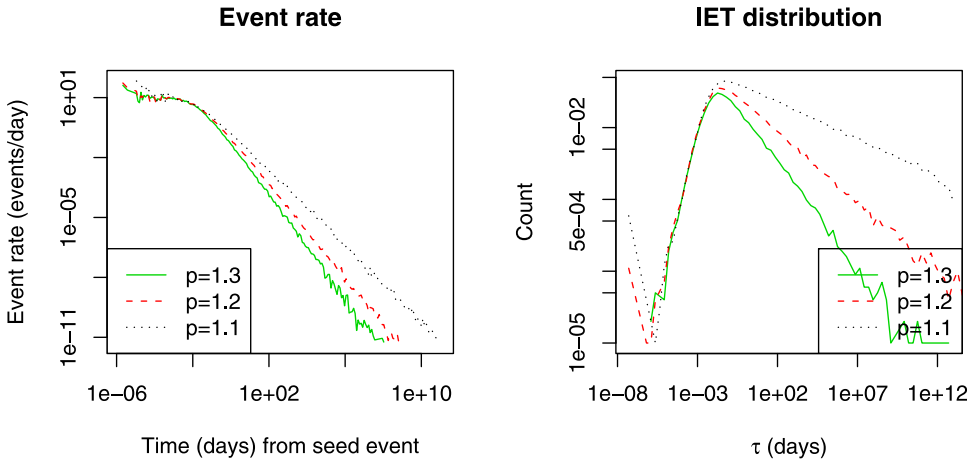

Figure 7. The effect of varying the ETAS parameter p on (left) the average global aftershock sequence (event rate against time) and (right) the distribution of interevent times τ within the sequence (average histogram). Other parameters were held fixed at $\mu = 0.01$, $\alpha = 1.1$, $c = 0.01$ and $A = 5$. IET, interevent time.

histogram of intervals labeled as same-sequence. This further confirms that our interpretation of the interevent time distribution and the origins of bimodality are relevant for real data. A possible practical application of the technique would be to obtain the power law exponent in the histogram as a means of inverting p through its relationship with this exponent (see *Helmstetter and Sornette* [2002] for details of the relationship).

[27] Having established the effect of the spontaneous rate μ and its connection with region size, we now proceed to explore the effects of the triggering parameters and the interaction between this and the seeding rate.

5. Distinguishability as a Function of Seeding Rate

[28] While μ determines the occurrence rate of "global" aftershock sequences as whole entities and clearly plays an important role in the temporal statistics, different aftershock generation parameters also make an impact on the statistics. The relative proportions of same-sequence and intersequence intervals in the series, and the bimodality or otherwise of the interevent time distribution, depend on both these aspects.

[29] The global aftershock sequence has been shown analytically to exhibit a double power law decay in the event rate, with a smooth transition between these two regimes [*Helmstetter and Sornette*, 2002], as opposed to the single power law for the "local" aftershock sequence (the nonrecursive Omori law). The exponents of the two power law segments are $1 - \theta$ for early times and $1 + \theta$ for later times, where $\theta = p - 1$. The effect of changes in p on the event rate of the global aftershock sequence is quite dramatic as shown in Figure 7 (left), which is produced by stacking simulated aftershock sequences.

[30] This effect on the power law exponent propagates through to the interevent times, altering the exponent in the same-sequence component of the interevent time distribution. Figure 7 also shows a histogram of the interevent times generated from the simulated aftershock sequences in Figure 7 (right). Note that in the absence of further seeding events, the power law decay in this distribution component has, in principle, no upper cutoff; in practice, it is truncated at a point dependent on the temporal length of each simulation run, which was arbitrarily set. We can see here that the requirement for p to be small in order to observe an apparently universal gamma distribution [from *Saichev and Sornette*, 2007] works by ensuring a shallow power law and thus reducing the appearance of bimodality.

[31] When ETAS parameters are inverted for a real catalog, events below the completeness threshold are first removed, and then the unjustified assumption is implicitly made that the completeness threshold m_c coincides with the smallest possible event m_0. Triggering relationships that exist between events on either side of the detection threshold are thus hidden by the removal of the small events. *Sornette and Werner* [2005] have shown that this unavoidable false assumption results in a renormalized set of parameter values and an underestimation of the true branching ratio. The upshot of this is that the magnitude threshold m_c can be thought of as a tuning dial for the effective branching parameters; the higher it is set, the lower the apparent branching ratio is. (It of course alters the effective spontaneous event rate also.)

[32] We now examine the effects of these two aspects, seeding and aftershock generation, together. We combine a selection of different branching parameterizations with first, a low μ value and, second, a high μ value, and we simulate full catalogs to determine the combined effect.

[33] The histograms of interevent times for the low-μ catalogs (Figure 8, left) are all bimodal. We see that the variability between global aftershock sequences carries through to the full catalogs, causing the shapes of the curves for low μ to differ significantly from each other. Perhaps counter intuitively, the position and height of the second (exponential) peak also varies with the branching parameters, even for constant μ; this is because it reflects the event rate within the average sequence at the time where overlap occurs, and so does not depend only on the spontaneous rate.

[34] The histograms in Figure 8 (right) are derived from the high-μ simulations and represent extreme examples where the power law is no longer visible. This is a consequence of the significant overlapping of aftershock sequences as discussed in section 3. By contrast with the low-μ case, it is hard

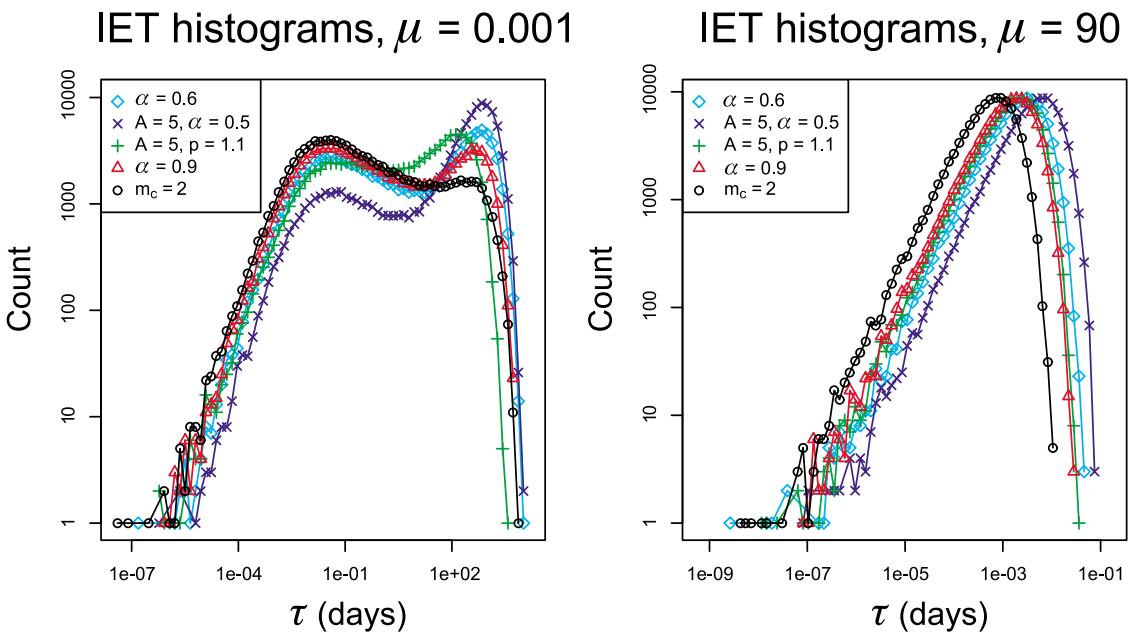

Figure 8. Histograms of interevent times τ from ETAS simulations of 100,000 events each, created using five different parameterizations, coupled with (left) a low μ and (right) a high μ; different shapes at low μ become very similar shapes at high μ. Other ETAS parameter values are as stated in legend, with unstated values defaulting to $A = 10$, $\alpha = 1.1$, $c = 0.01$, $p = 1.2$, $b = 1$ and $m_c = 0$.

to distinguish between the sets of parameters used for the high-μ simulations besides a shift of the distribution along the τ axis. Each curve could conceivably be created by a wide range of sets of parameter values, because the effect of the branching is "hidden."

[35] This point is underlined when we look at normalized and rescaled versions of these histograms (Figure 9), as is typically done when looking for universal data collapse [e.g., *Corral*, 2003]. Strong data collapse is observed at high μ, confirming that in that regime, the different parameterizations of the model result in interevent time distributions that are described by the same underlying function (an exponential). No such data collapse is observed at low μ, though visually this may seem an attractive interpretation from noisier (real) data.

[36] This strongly suggests that the uncertainty in all parameters varies systematically with the actual value the parameter μ takes relative to the others. We will now proceed to confirm this transition toward indistinguishability in branching characteristics as μ increases, through measurements of the accuracy of ETAS parameter inversions.

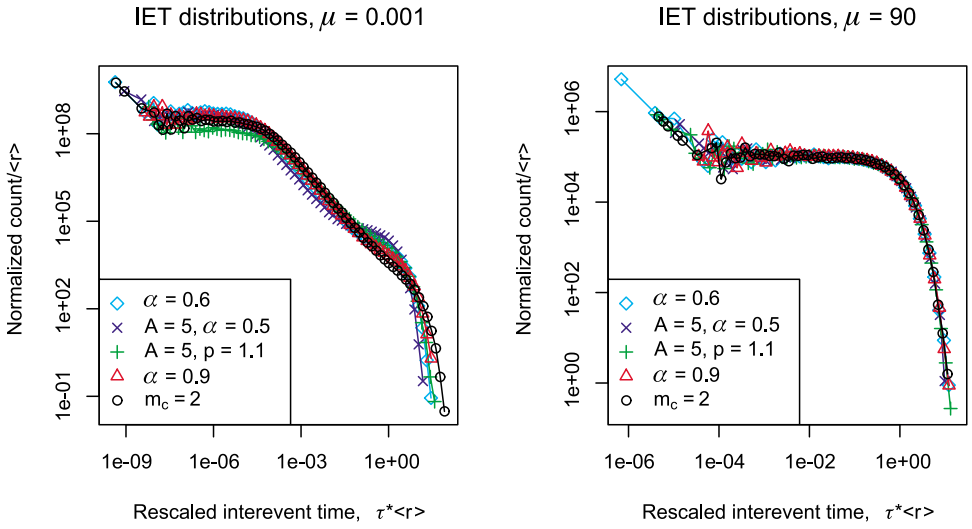

Figure 9. Normalized PDFs rescaled by the mean event rate, using the data from Figure 8. Variation at low μ becomes a data collapse at high μ.

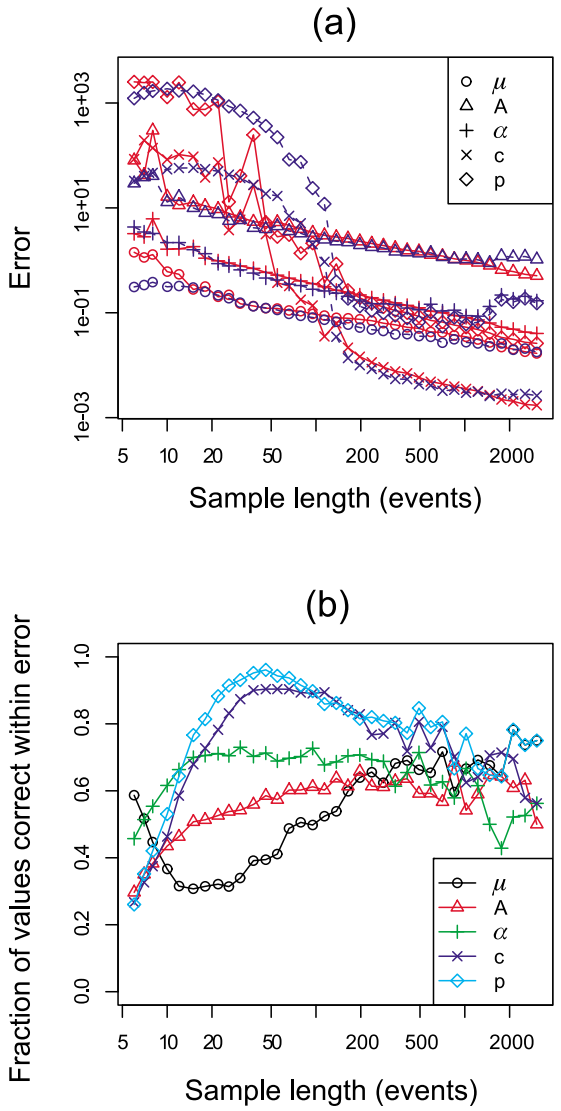

Figure 10. (a) Comparison of mean standard error e (red, solid line) with mean absolute difference between the true value x and its inverted estimate x^* (blue, dashed line) for each parameter x as a function of sample size, for inversions of ETAS parameters from a synthetic ETAS catalog. (b) Fraction of inverted values for which the true parameter value is within the bounds of the inverted value plus/minus the standard error, as a function of sample size. The ETAS parameters used to create the catalog were $\mu = 0.07$, $A = 5$, $\alpha = 1.1$, $c = 0.01$, $p = 1.1$.

[37] It is common to analyze earthquake catalogs by inverting ETAS parameters from the catalog using the maximum likelihood method. The log likelihood for a point process is given by

$$\log L = \sum_i \log \lambda(t_i|H_t) - \int \lambda(t|H_t)dt, \qquad (7)$$

where λ is the conditional intensity function, t_i refer to times of events, and H_t is the history of the process at the time t [*Harte*, 1998]. Optimization is typically performed by minimizing the negative log likelihood with a Newton-type algorithm.

[38] In this study, we invert ETAS parameters for synthetic catalogs generated by the ETAS model; unlike with real data, we can use our knowledge of the input parameter values to evaluate how well the inversion algorithm performs under various circumstances. We first test the algorithm on a synthetic catalog of 50,000 events created using typical ETAS parameters. Rather than invert parameters for the whole catalog, we divide it into samples of a particular number of successive events, and invert for each sample. This can then be repeated using different sample sizes, giving a spread of values and corresponding standard errors from which we can determine the sample size required to give an accurate inversion.

[39] Figure 10a shows, for each sample size, two types of mean error: first, the standard error, which is the square root of the appropriate diagonal entry in the approximate covariance matrix, averaged across all inversions; and secondly, the mean absolute difference between the inverted value and the true value, $|x - x^*|$ (where x represents a parameter). The mean of $|x - x^*|$ is generally just below the mean standard error, indicating that on average the standard error is a reasonable estimate of the actual error. Throughout much of the range of sample sizes, a power law relationship between error and sample size is apparent. The exponent is similar for all parameters; the curves are vertically offset in proportion to the values of the parameters. However, for c and p, both types of error are highly elevated for small sample sizes. The inversion procedure is failing to converge for some few samples and getting "lost," which is greatly influencing the average error. In practice these inversion results would of course be rejected, but we include them here to show that they exist, and clearly *don't* exist at larger sample sizes. This is presumably because, in a small sample of events, the temporal decay of the aftershock rate is highly uncertain and could be fit by a wide range of pairs of values in these parameters which have a clear covariance. At large sample sizes, variability in the error curve increases, simply due to obtaining fewer samples from the catalog over which to compute the average.

[40] Figure 10b shows the fraction of inverted values that are correct within their standard error as a function of sample size. This is best understood as a measure of how good the standard error estimates are. At large sample sizes, for all parameters, around two thirds of the inversions produce values with error bounds which contain the true value. The algorithm is therefore producing a standard error comparable to one standard deviation, which is the 67% confidence limits. At smaller sample sizes, however, the curves separate out, with the error estimate being particularly incorrect for μ (too low) and c and p (too high). This separation of the parameters is consistent across similar inversions carried out with other synthetic catalogs (not shown), and can be understood in terms of under sampling of background events (for the inversion of μ) and of limited duration aftershock sampling (since c and p concern the temporal dimension of aftershock occurrence).

[41] Figures 10a and 10b indicate that a sample size of 1000 events is able to support the inversion of parameters

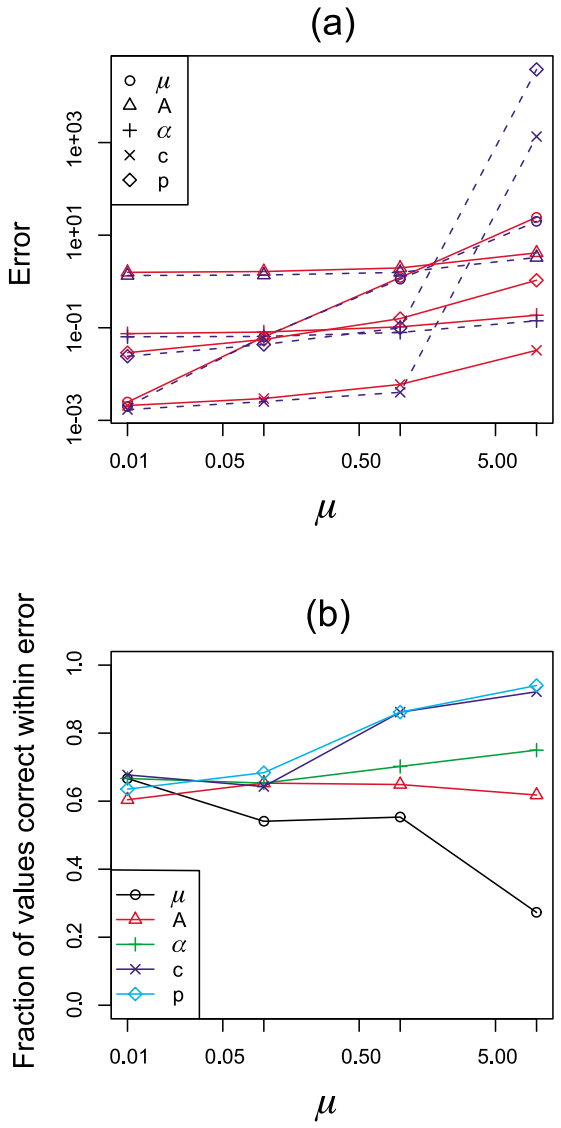

Figure 11. (a) Comparison of mean standard error e (red, solid line) with mean absolute difference between the true value x and its inverted estimate x^* (blue, dashed line) for each parameter x as a function of the true u value, for inversions of ETAS parameters from four synthetic ETAS catalogs of different μ (see the corresponding interevent time distributions in Figure 1). (b) Fraction of inverted values for which the true parameter value is within the bounds of the inverted value plus/minus the standard error, as a function of the true μ value. Other ETAS parameters were held fixed at $A = 10$, $\alpha = 1$, $c = 0.01$, $p = 1.2$. Inversions were started at the true solution.

with acceptable error estimates. We use this sample size to now explore the effect of the bimodality or otherwise of the interevent time distribution on the invertibility of parameters, by performing inversions on synthetic catalogs with different μ values. Each catalog is created with 100000 events, which are then divided into 100 samples of 1000 successive events each; parameter values are inverted for each sample. Rather than start each inversion off at some common set of initial parameter values, we eliminate the possible effect of initial proximity to the solution, which would not otherwise be constant, by starting each inversion off at the true solution.

[42] Figure 11a shows the two types of error again, this time as a function of μ, at constant sample size. They are again averaged over the 100 samples for each catalog. Both types of error increase with μ, particularly steeply in the case of the error in inverted μ. The error in c and p also increases dramatically for larger μ, similarly to the case for small sample sizes in Figure 10.

[43] In Figure 11b the fraction of inverted values that are correct within the standard error, for all parameters, is around two thirds for low μ as our previous inversion exercise led us to expect for a sample size of 1000. As μ is increased, however, the fraction of values correct within error starts to deviate from this value, indicating that just as for small sample sizes, large μ values produce inversions with inaccurate error estimates. The pattern of separation across the parameters is identical to the small sample sizes case, with the estimated error being too low for μ and too high for c and p.

[44] In order to establish beyond doubt that the parameter estimations are truly converging, we recorded the values of the estimates at each iteration, along with the gradients of the log likelihood function, for 12 of the samples at each value of μ. Figure 12a shows an example for the lowest value of μ, which is representative of the pattern observed in every inversion for that μ value. The log likelihood gradients go toward zero and the parameters, which start out at their true values, do not wander far from those values. This confirms that the iteration procedure has converged.

[45] At high μ, there were two distinct scenarios in the examples. The first and most common scenario, occurring in 9 out of 12 inversions, was similar to the situation for low μ and is shown in Figure 12b. The log likelihood gradients similarly tend toward zero, and the parameter values take only small steps toward the end of the inversion, although they have moved far from their starting points in some cases. This situation also represents convergence, albeit toward a false solution. The second scenario is where the inversion terminates due to the maximum step size being exceeded on one parameter (Figure 12c), although the log likelihood gradients are also very low. We can infer that the likelihood surface is nearly flat in one dimension, leading that parameter rapidly away from the true value. This cannot be classed as convergence; however, it is a situation that only occurs at high μ, which is an important observation.

[46] The differences in the inversion outcomes as μ is varied are consistent with the fact that as μ is increased, the overlapping of aftershock sequences causes earthquake occurrence in the catalog to look more like a Poisson process in time (Figure 1), which hides the effect of the branching parameters on the time series (Figure 8). We suggest that this makes the origin of each event more ambiguous, and profoundly, we see here that this is reflected in poor inversion success. Even with a catalog produced by, and thus perfectly described by, the ETAS model, and with the ability to start the inversion at the true solution, luxuries that real data does not afford us, the error is still significant for high μ and the error estimates produced by the inversion are significantly incorrect.

[47] The situation may be improved with new more accurate inversion methods such as the expectation maximization (EM) method of *Veen and Schoenberg* [2008], but our

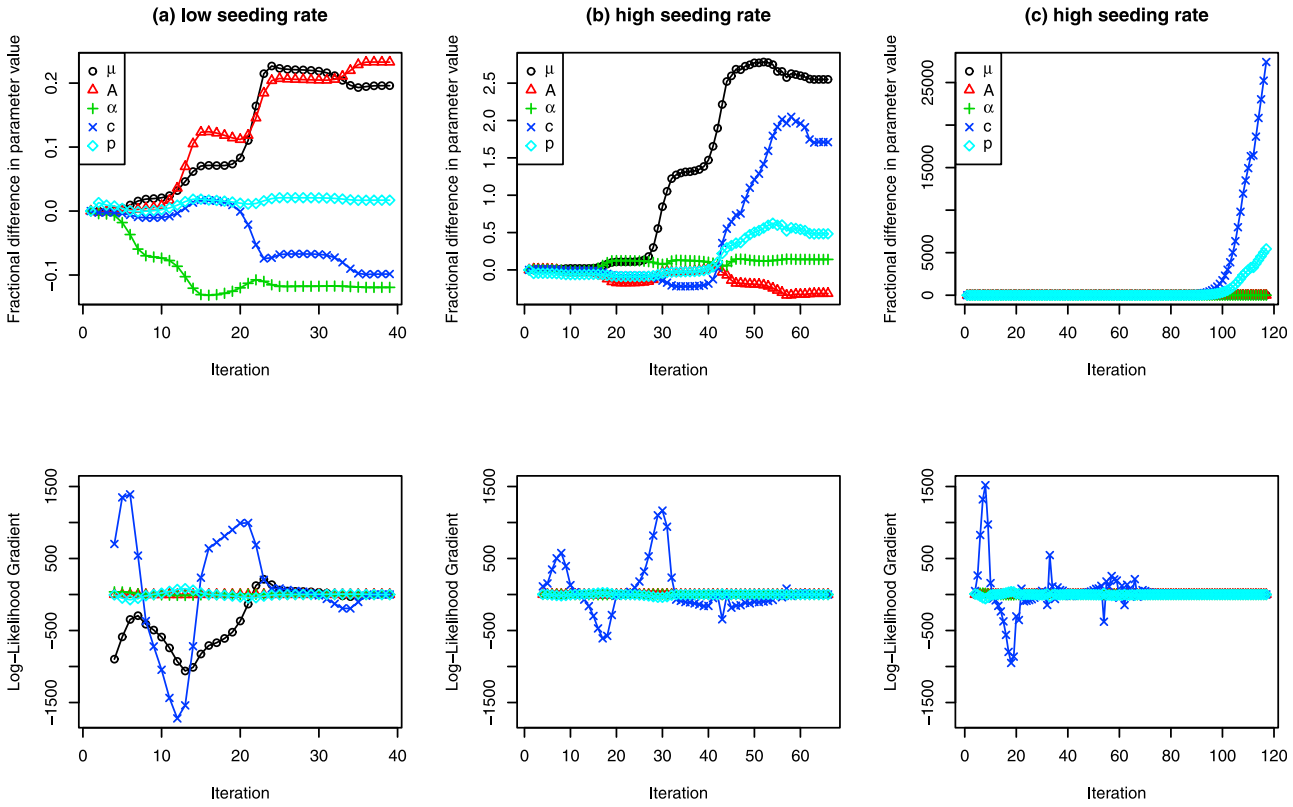

Figure 12. Parameter estimates (as a fractional difference from the initial, true value) and log likelihood function gradient at each iteration of an inversion for (a) one of the samples of the lowest μ value from Figure 11, and (b and c) one of the samples of the highest μ value. Likelihood gradient is not shown for the first three iterations where it fluctuates strongly. Toward the final iterations, parameter estimates are stable and gradients go to zero for Figures 12a and 12b, although in Figure 12b the parameter values have moved further from their (true) starting values. In Figure 12c the inversion terminated with excessively large steps being taken in one parameter, so this inversion failed to converge, although the final gradients are near zero.

primary concern is to test current standard practice and point out where and why it fails. These issues may also apply to alternate techniques, since, as we have illustrated, in principle, finding independent events in data where many aftershock sequences overlap is inherently problematic, so any improvement is unlikely to be first order.

[48] To illustrate the potential consequences of underestimating the error, we consider the following example. *Hainzl and Ogata* [2005] inferred fluid signals in Vogtland earthquake data in terms of a varying seeding rate, by performing a series of ETAS inversions on the data using a moving time window. Here we test the alternate hypothesis that an apparent variation in μ may be observed in synthetic ETAS catalogs with constant seeding rate, purely as a consequence of overlapping aftershock sequences.

[49] We repeat their analysis using synthetic catalogs, comparing the lowest and highest μ values used in our inversion accuracy analysis. Like *Hainzl and Ogata* [2005], we invert only for μ while holding all other parameters fixed at their true values, a procedure which might reasonably be expected to produce very reliable estimates of μ. We use, in each case, a window size (in days) calculated to give an average of 400 events per window, which is the approximate mean number of events in Hainzl and Ogata's windows; although the full seismic history is included in $\lambda\,(t|H_t)$ in equation (7), the sum and integration are performed only over the time window. Hainzl and Ogata do not specify the step size, but we make the step size equal to the window size so that each point in the graph represents a completely different sample with no overlap. Figure 13 shows the results of this series of inversions. The error bars shown are, in this case, two standard errors (i.e., two standard deviations, or 95% confidence limits).

[50] At both values of μ, a variation in the inverted μ values occurs, which mirrors the variation in the overall event rate as in Figure 7 of *Hainzl and Ogata* [2005]. The actual number of spontaneous background events in each time window, divided by the window length, is shown as a dotted line, which (unsurprisingly) remains in the vicinity of the true μ value and does not vary to the same extent as the inverted values. We would therefore interpret these type of inversion results with caution, and accept that while there may be good reason to reject a single ETAS parameterization for the Vogtland swarm data based on the additional analysis of Hainzl and Ogata, particular parameter values for an ETAS inversion cannot be regarded as too reliable.

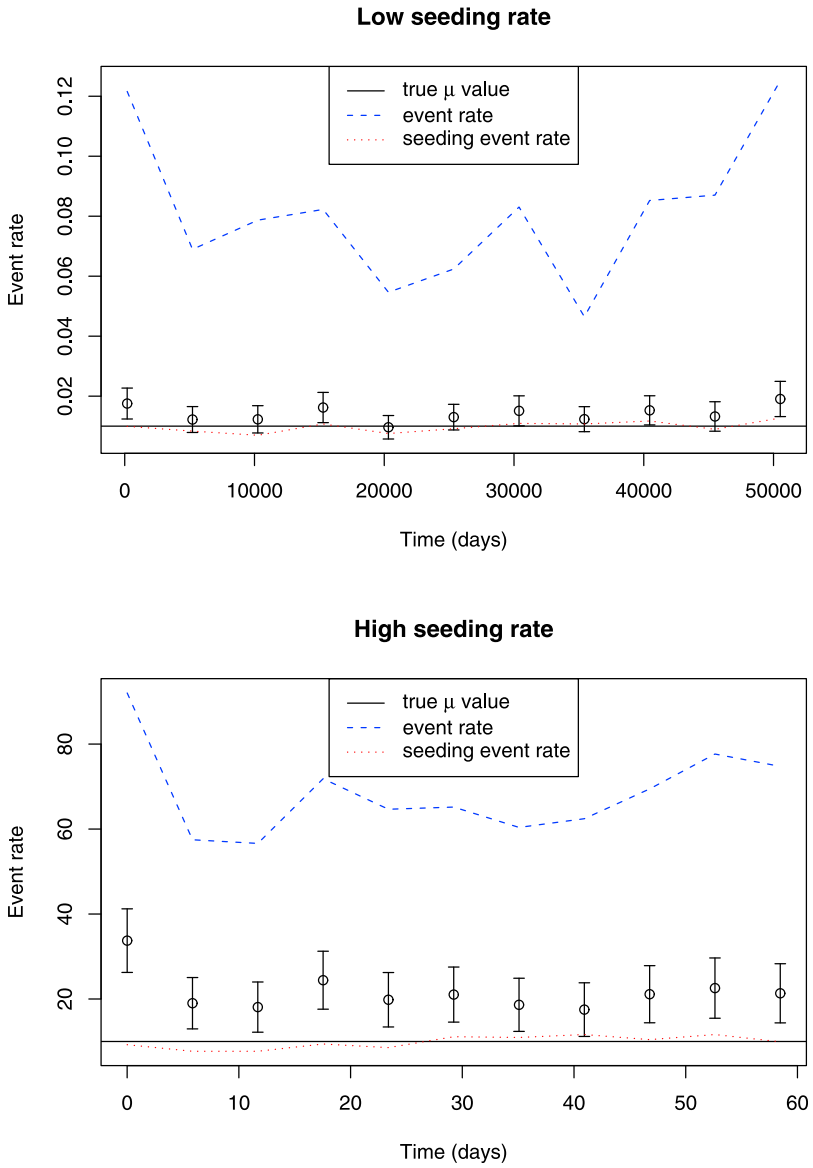

Figure 13. Circles with error bars show a series of inversions of μ, holding all other parameters fixed at their true values, from (top) a low-μ and (bottom) a high-μ synthetic catalog. In each case the window size was calculated to give an average of 400 events per window, and the windows do not overlap. Variation in the inverted values follows the variation in the event rate, demonstrating that an apparent variation in seeding rate can spuriously occur from a catalog with a constant seeding rate. Additionally, μ is overestimated when the true seeding rate is high due to temporal overlapping of aftershock sequences.

[51] In addition, we notice that for our high-μ simulation, the inverted value is consistently and significantly above the true value, showing that a bias is introduced in the inversion due to highly overlapping aftershock sequences. This is all the more surprising given that it is a one-dimensional inversion with all other parameters forced to be correct. We understand this as further confirmation that aftershocks are being mistaken for spontaneous seeding events due to the overlapping of sequences and the consequent destruction of their temporal signature. This effect naturally entails an underestimation of the branching ratio. In Figure 7 of *Hainzl and Ogata* [2005], the inverted μ values for the case where all parameters are allowed to vary is consistently higher than those values where only μ is inverted. This too is suggestive of the overestimation of μ we have shown can happen when aftershock sequences overlap.

[52] In summary, we have demonstrated clearly that the ETAS maximum likelihood inversion scheme, under specific (high) values of the seeding rate, produces a bias in the inverted seeding rate (and hence in the branching ratio also), and an underestimate of its uncertainty. For a finite recorded period, catalog size can be adjusted by varying the size of the included region; this, however, is akin to varying the effective μ, and so it must be borne in mind that while increasing the amount of data tends to improve the inversion success, the effect of increased temporal overlapping of aftershock

sequences has the opposite effect. We suggest that the absence of discernible bimodality in the interevent time distribution is likely to preclude accurate inversion of temporal ETAS parameters. Studies which seek to demonstrate a varying background rate must establish that the variation lies outwith the substantial error bounds caused by overlapping aftershock sequences; triggering is a difficult null hypothesis to reject.

6. Conclusions

[53] We have explored, within the framework of the ETAS model, the effects of the spontaneous event rate and the aftershock triggering parameters on the earthquake interevent time distribution. We have shown, using data from the PDE and SCEC catalogs, that this is fundamentally a bimodal distribution, and that the degree of observable bimodality depends on the region size of the data set. Our ETAS analysis demonstrates that the bimodality can be explained in terms of two populations of interevent times: those originating from same-sequence and intersequence event pairs, respectively. These populations depend on the extent of temporal overlap of separate aftershock sequences, and so the dependence of the distribution shape on region size is essentially a dependence on seeding rate. This observation forces us to reject the hypothesis proposed in earlier literature of a universal scaling law for earthquake interevent times: the small interevent times, which are dominated by same-sequence aftershocks, do not generally scale to the large ones, although the distribution may be seen to approximate a universal gamma form under the constraints of visually stationary periods, low p values and interevent times much greater than c. Our analysis thus represents a generalization of *Corral*'s [2003, 2004] and *Saichev and Sornette*'s [2006, 2007] work, taking it beyond these constraints.

[54] The statistics of the ETAS model are a realistic representation of real data, even for large region sizes, which are outside of the model's typical range of application, although inversion of parameters for large regions is problematic. Large regions occupy a high-spontaneous-rate subset of parameter space, in which the interevent time distribution becomes unimodal and exponential-like due to the temporal overlapping of aftershock sequences; parameterizations within this high-overlap regime have a high redundancy in terms of the shape of the distribution. As a result of this, maximum likelihood inversion of parameters from the simulated catalogs becomes less successful as the spontaneous event rate is increased, with mean parameter estimates across realizations becoming further away from the true value, and standard errors becoming less reliable and on average larger. The spontaneous rate is also systematically overestimated within this high-overlap regime. This introduces new challenges in determining both the reliability of parameter inversions from real data and the suitable range of region sizes for which temporal ETAS parameter inversion may be successful, as well as in the ability to detect significant temporal variations in parameters such as the spontaneous rate, for example in time-dependent hazard calculations.

[55] **Acknowledgments.** We would like to acknowledge the UK Engineering and Physical Sciences Research Council for funding S.T. on a Doctoral Training Award and M.N. on GR/T11753/01, the Royal Society of Edinburgh and the Scottish Government for supporting M.N. on a personal fellowship, and the EU TRIGS project and the NERIES project for funding S. T. R and SSLib were used in the data analysis. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF). (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk).

References

Bak, P., K. Christensen, L. Danon, and T. Scanlon (2002), Unified scaling law for earthquakes, *Phys. Rev. Lett.*, *88*, 178501, doi:10.1103/PhysRevLett.88.178501.

Bonnet, E., O. Bour, N. E. Odling, P. Davy, I. Main, P. Cowie, and B. Berkowitz (2001), Scaling of fracture systems in geological media, *Rev. Geophys.*, *39*, 347–383, doi:10.1029/1999RG000074.

Bottiglieri, M., L. Arcangelis, C. Godano, and E. Lippiello (2010), Multiple-time scaling and universal behavior of the earthquake interevent time distribution, *Phys. Rev. Lett.*, *104*, 158501, doi:10.1103/PhysRevLett.104.158501.

Corral, A. (2003), Local distributions and rate fluctuations in a unified scaling law for earthquakes, *Phys. Rev. E*, *68*, 035102, doi:10.1103/PhysRevE.68.035102.

Corral, A. (2004), Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes, *Phys. Rev. Lett.*, *92*, 108501, doi:10.1103/PhysRevLett.92.108501.

Davidsen, J., and C. Goltz (2004), Are seismic waiting time distributions universal?, *Geophys. Res. Lett.*, *31*, L21612, doi:10.1029/2004GL020892.

Davidsen, J., S. Stanchits, and G. Dresen (2007), Scaling and universality in rock fracture, *Phys. Rev. Lett.*, *98*, 125502, doi:10.1103/PhysRevLett.98.125502.

de Arcangelis, L., E. Lippiello, C. Godano, and M. Nicodemi (2008), Statistical properties and universality in earthquake and solar flare occurrence, *Eur. Phys. J. B*, *64*(3–4), 551–555, doi:10.1140/epjb/e2008-00057-5.

Hainzl, S., and Y. Ogata (2005), Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.*, *110*, B05S07, doi:10.1029/2004JB003247.

Hainzl, S., F. Scherbaum, and C. Beauval (2006), Estimating background activity based on interevent-time distribution, *Bull. Seismol. Soc. Am.*, *96*(1), 313–320, doi:10.1785/0120050053.

Hainzl, S., A. Christophersen, and B. Enescu (2008), Impact of earthquake rupture extensions on parameter estimations of point-process models, *Bull. Seismol. Soc. Am.*, *98*(4), 2066–2072, doi:10.1785/0120070256.

Harte, D. (1998), Documentation for the statistical seismology library, *Res. Rep. 98-10*, Sch. of Math. and Comput. Sci., Victoria Univ. of Wellington, Wellington, New Zealand.

Helmstetter, A., and D. Sornette (2002), Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, *J. Geophys. Res.*, *107*(B10), 2237, doi:10.1029/2001JB001580.

Huc, M., and I. G. Main (2003), Anomalous stress diffusion in earthquake triggering: Correlation length, time dependence, and directionality, *J. Geophys. Res.*, *108*(B7), 2324, doi:10.1029/2001JB001645.

Lennartz, S., V. N. Livina, A. Bunde, and S. Havlin (2008), Long-term memory in earthquakes and the distribution of interoccurrence times, *Europhys. Lett.*, *81*, 69001, doi:10.1209/0295-5075/81/69001.

Molchan, G. (2005), Interevent time distribution in seismicity: A theoretical approach, *Pure Appl. Geophys.*, *162*(6–7), 1135–1150, doi:10.1007/s00024-004-2664-5.

Saichev, A., and D. Sornette (2006), "Universal" distribution of interearthquake times explained, *Phys. Rev. Lett.*, *97*, 078501, doi:10.1103/PhysRevLett.97.078501.

Saichev, A., and D. Sornette (2007), Theory of earthquake recurrence times, *J. Geophys. Res.*, *112*, B04313, doi:10.1029/2006JB004536.

Shcherbakov, R., G. Yakovlev, D. L. Turcotte, and J. B. Rundle (2005), Model for the distribution of aftershock interoccurrence times, *Phys. Rev. Lett.*, *95*, 218501, doi:10.1103/PhysRevLett.95.218501.

Sornette, D., and M. J. Werner (2005), Apparent clustering and apparent background earthquakes biased by undetected seismicity, *J. Geophys. Res.*, *110*, B09303, doi:10.1029/2005JB003621.

Touati, S., M. Naylor, and I. G. Main (2009), Origin and nonuniversality of the earthquake interevent time distribution, *Phys. Rev. Lett.*, *102*, 168501, doi:10.1103/PhysRevLett.102.168501.

Veen, A., and F. P. Schoenberg (2008), Estimation of space-time branching process models in seismology using an EM-type algorithm, *J. Am. Stat. Assoc.*, *103*(482), 614–624, doi:10.1198/016214508000000148.

Zaliapin, I., A. Gabrielov, V. Keilis-Borok, and H. Wong (2008), Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.*, *101*, 018501, doi:10.1103/PhysRevLett.101.018501.

M. Christie, Institute of Petroleum Engineering, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK. (mike.christie@pet.hw.ac.uk)

I. G. Main, M. Naylor, and S. Touati, School of GeoSciences, University of Edinburgh, Grant Institute, West Mains Road, Edinburgh EH9 3JW, UK. (ian.main@ed.ac.uk; mark.naylor@ed.ac.uk; s.touati-2@sms.ed.ac.uk)