 doi:10.20944/preprints201810.0679.v1

The relevance of foreshocks in earthquake triggering: A statistical study

E. Lippiello [1], **C. Godano** [1], **L. de Arcangelis** [2]

[1]Department of Mathematics and Physics, University of Campania "L. Vanvitelli", Viale Lincoln 5, 81100 Caserta, Italy

[2]Department of Engineering, University of Campania "L. Vanvitelli', Via Roma 29, 81031 Aversa (CE), Italy

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

Abstract

An increase of seismic activity is often observed before large earthquakes. Events responsible for this increase are usually named foreshock and their occurrence probably represents the most reliable precursory pattern. Many foreshocks statistical features can be interpreted in terms of the standard mainshock-to-aftershock triggering process and are recovered in the Epidemic Type Aftershock Sequence ETAS model. Here we present a statistical study of instrumental seismic catalogs from four different geographic regions. We focus on some common features of foreshocks in the four catalogs which cannot be reproduced by the ETAS model. In particular we find in instrumental catalogs a significantly larger number of foreshocks than the one predicted by the ETAS model. We show that this foreshock excess cannot be attributed to catalog incompleteness. We therefore propose a generalized formulation of the ETAS model, the ETAFS model, which explicitly includes foreshock occurrence. Statistical features of aftershocks and foreshocks in the ETAFS model are in very good agreement with instrumental results.

1 Introduction

The epidemic-type-aftershock sequence ETAS model *Ogata* [1985, 1988a, b, 1989] is nowadays considered " a de facto standard model, or null hypotheses, for other models and ideas to be compared to" *Huang et al.* [2016] The model assumes that two classes of earthquakes exist: Independent background and triggered earthquakes. An epidemic organization of events arises under the assumption that each earthquake can trigger its own descendents leading to a branching organization. From a physical point of view, background seismicity can be thought as the effect of the slow tectonic drive whereas triggered earthquakes are induced by stress redistribution after previous shocks. In the ETAS model the occurrence rate of triggered events is obtained on the basis of well established empirical laws controlling the spatio-temporal clustering of aftershocks. As a consequence, by construction, the model is very efficient in reproducing statistical features of aftershock organization observed in experimental catalogs. At the same time, in the ETAS model an event can trigger also a larger shock. In this situation the triggering event is often named "foreshock" and the triggered earthquake, if it is the largest event in the sequence, is named "mainshock".

In this study we adopt the standard definition of mainshocks as events sufficiently isolated in time and space from other larger events. Foreshocks (aftershocks) are then all events occurring close in space and in time before (after) the mainshock. We wish to stress that, within

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

the ETAS framework, this classification of events does not reflect different physical properties since, as anticipated, only two kinds (independent or triggered) earthquakes are assumed and, for instance, a mainshock can be either an independent or a triggered earthquake. On the other hand, according to a nucleation theory *Ohnaka* [1992, 1993]; *Dodge et al.* [1996], the nucleation phase can be characterized by the occurrence of smaller earthquakes inside the region involved in the fracture process of the subsequent incoming larger shock. This pre-shock seismicity is not implemented in the ETAS model and the main question addressed in this study is if its inclusion, within the ETAS modeling, gives a more accurate description of of foreshock organization in instrumental catalogs. In the last five years several studies have shown a lack of foreshocks in ETAS with respect to instrumental catalogs *Brodsky* [2011]; *Lippiello et al.* [2012c]; *Shearer* [2012a, b]; *Hainzl* [2013]; *Shearer* [2013]; *Bouchon et al.* [2013]; *Mignan* [2014]; *Chen and Shearer* [2013]; *Brodsky and Lay* [2014]; *Ogata and Katsura* [2014]; *Felzer et al.* [2015]; *Bouchon and Marsan* [2015]; *de Arcangelis et al.* [2016]; *Lippiello et al.* [2017]. Nevertheless, the deficit of foreshocks in ETAS catalogs has been attributed, at least partially, to the deficit of aftershocks in instrumental catalogs caused by spurious incompleteness. This point will be discussed in the following sections where we provide evidence that it cannot justify the excess of foreshocks in instrumental data sets.

2 Data sets and definitions

We perform a systematic analysis of four different instrumental catalogs: The relocated Southern California earthquake catalog (RSCEC) *Hauksson et al.* [2012] (from 01/01/1981 to 12/31/2013), the relocated Northern California earthquake catalog (RNCEC) *Waldhauser and Schaff* [2008] (from 01/01/1981 to 06/30/2011), the Italian earthquake catalog (ItEC) [ISI] (from 01/01/2002 to 12/31/2012) and the Japanese earthquake catalog (JaEC) [NIE] (from 01/01/1966 to 01/30/2011). We use the same definition of mainshock, aftershocks and foreshocks adopted in *Lippiello et al.* [2017]. More precisely, we define an event as "mainshock" if a larger earthquake does not occur in the previous y days and within a distance L. In addition a larger earthquake must not occur in the selected area in the following y_2 days. We then associate to each mainshock its own "aftershocks" and "foreshocks" defined as all earthquakes recorded in the subsequent or in the preceding time interval $T = 12$ h, respectively, and within a circle of radius $R \leq R_M$ centered in the mainshock epicenter. We use different R_M for different catalogs: $R_M = 2$ km for RSCEC and RNEC, $R_M = 5$ km for ItEC and $R_M = 10$ km for JaEC.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

Once mainshocks are identified, they are grouped in classes according to their magnitude $m \in [m_M, m_M+1)$ and, for each catalog, we evaluate the total number of mainshocks belonging to the given class $n_{main}(m_M)$, the total number of associated aftershocks $n_{aft}(m_M)$ and foreshocks $n_{fore}(m_M)$. We also evaluate the epicentral distance Δr between each main-aftershock and main-foreshock couple and construct the aftershock and foreshock epicentral distance distributions, $\rho_a(\Delta r, m_M)$ and $\rho_f(\Delta r, m_M)$. Their precise definition is given in Sec.4.1.

The choice of parameters has been deeply investigated in previous studies [*Felzer and Brodsky*, 2006; *Lippiello et al.*, 2009a, 2012c, 2017] and here we implement typical values, $L = 100$ km, $y = 3$ and $y_2 = 0.5$. The value of R_M is fixed imposing that, for each instrumental catalog, different choices of $T \leq 12$ h produces similar results $\rho_a(\Delta r, m_M)$ when $\Delta r < R_M$. This leads to $R_M = 2, 2, 5, 10$ km, for RSCEC, NSCEC; ItEC and JMAC, respectively.

2.1 The ETAS model

The ETAS model is specified by the conditional intensity function, which represents the expected seismicity rate in a given space position conditioned to a given observational history. The conditional intensity function $\Lambda(m, \vec{x}, t)$, which represents the occurrence probability of events with magnitude $m \geq m_c$ in the position $\vec{x}$ at time t, can be written in the following form:

$$\Lambda(m, \vec{r}, t) = \left[\mu w(\vec{r}) + Q\left(|\vec{r}_i - \vec{r}|, t - t_i, m_i\right)\right] \frac{1}{b \log(10)} 10^{-b(m-m_c)} \tag{1}$$

and

$$Q(\Delta r_i, t - t_i, m_i) = \frac{A(p-1)}{c} \sum_{i:t_i<t} 10^{\alpha(m_i - m_c)} \left(1 + \frac{t - t_i}{c}\right)^{-p} G(\Delta r_i, m_i) \tag{2}$$

where $\Delta r_i = |\vec{r}_i - \vec{r}|$ and the sum extends over all events with magnitude m_i, epicentral coordinates $\vec{x}_i$ and occurrence time $t_i < t$. The function $G(\Delta r_i, m_i)$ is a spatial kernel which explicitly depends on the triggering magnitude m_i and $\mu w(\vec{x})$ is a time independent contribution due to background seismicity. The form of the spatio-temporal kernel (Eq.(2) implements three well established laws for aftershock triggering:

- A1: The number of aftershocks n_a depends on the mainshock magnitude class, according to the productivity law $n_a = K_a 10^{\alpha_a m_M}$;
- A2: The aftershock number decays as function of the time from the mainshock, consistently with the Omori law $n_a(\Delta t) \sim \Delta t^{-p}$ with $p \simeq 0.8$;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

- A3: The linear density distribution of epicentral distances between mainshock and aftershocks (Δr) clearly depends on the mainshock magnitude class m_M with the average distance $L(m_M) \propto 10^{\gamma m_M}$ with $\gamma \simeq 0.4$.

  In the following we present results from numerical simulations of the ETAS model performed according to the algorithm discussed in ref. *Lippiello et al.* [2012c, 2017]; *de Arcangelis et al.* [2016, 2018]. In particular, the spatial function $w(\vec{x})$ is obtained from the smoothed seismicity whereas the functional form of $G(\Delta r_i, m_i)$ is tuned in order to collpase the aftershock epicentral distribution on the instrumental one, for all values of m_M.

3 Results

3.1 Previous Results

The statistical features of foreshocks in instrumental catalogs has been recently investigated in ref.*Lippiello et al.* [2017]. This study has shown that foreshocks follow empirical laws like (A1-A3) of aftershocks but with important differences. More precisely:

- F1: The number of foreshocks n_f depends on the mainshock magnitude class according to a productivity law $n_f = K_f 10^{\alpha_f m_M}$;
- F1b: The number of foreshocks is systematically smaller than the aftershock number and $\alpha_f \simeq 0.7\alpha_a$;
- F2: The foreshock number increases approaching the mainshock occurrence time, consistently with an inverse Omori law $n_f(\Delta t) \sim |\Delta t|^{-p}$ with $p \simeq 0.8$;
- F3: The linear density distribution of epicentral distances fore-mainshock $\rho_f(\Delta r, m_M)$ depends on the mainshock magnitude class with a roughly symmetrical behavior between spatial distribution before and after the mainshocks $\rho_f(\Delta r, m_M) \simeq \rho_a(\Delta r, m_M)$;
- F3b: The foreshock linear density distribution $\rho_f(\Delta r, m_M)$ does not depend on the value of the lower magnitude cut-off m_{th}.

The comparison between instrumental and ETAS catalogs has shown (*Lippiello et al.* [2017]) that it is possible to generate ETAS catalogs which reproduce at quantitative level the statistical features (A1-A3) of instrumental aftershocks. At the same time ETAS catalogs can reproduce foreshock features F1 and F2. Conversely it is not possible to generate ETAS catalogs with foreshocks obeying features F3 and F3b. Furthermore, ETAS catalogs always present a deficit of foreshocks with respect to instrumental catalogs. In the following section we will better stress these differences between instrumental and ETAS catalogs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

3.2 The aftershock and foreshock number

In Fig.1 we plot the ratio between aftershock and mainshock number $n_{aft}(m_M)/n_{main}(m_M)$ for different mainshock classes m_M and for the different instrumental catalogs. We also plot the ratio between foreshock and mainshock number $n_{fore}(m_M)/n_{main}(m_M)$. We only consider events with magnitude $m > m_{th} = 2$. The lower threshold m_{th} must not be confused with m_c in Eq.(2). Indeed, m_c is a fixed parameter of the ETAS model and synthetic catalogs contain only events with $m \geq m_c$. The lower magnitude m_{th}, conversely, is a parameter implemented in the data analysis and it can be arbitrarily varied with $m_{th} \geq m_c$.

Results in Fig.1 show that the aftershock number is systematically larger than the foreshock number and this difference increases for increasing m_M. The aftershock number is consistent with the Utsu-productivity law (A1)

$$n_{aft}(m_M)/n_{main}(m_M) = K_a 10^{\alpha m_M} \tag{3}$$

and a similar law is also observed for foreshocks $n_{fore}(m_M)/n_{main}(m_M) = K_f 10^{\alpha_f m_M}$ (F1).

In Fig.1 we compare results for $n_{aft}(m_M)/n_{main}(m_M)$ and $n_{fore}(m_M)/n_{main}(m_M)$ for the RSCEC with the results obtained applying the same definition of aftershocks, mainshocks and foreshocks to simulated ETAS catalogs. The values of both $n_{aft}(m_M)/n_{main}(m_M)$ and $n_{fore}(m_M)/n_{main}(m_M)$ depend on the parameters A, p, c, α (Eq.(2) of the numerical model. In particular, in first approximation the value of $n_{aft}(m_M)/n_{main}(m_M)$ is given by

$$n_{aft}(m_M) \simeq \overline{\mu(R_M)}T + n_{main}(m_M) B_R 10^{\alpha(m_M - m_c)} \left(1 - (T/c+1)^{1-p}\right) H(R_m). \tag{4}$$

In this equation $\overline{\mu(R_M)}$ is the average background rate inside a circle of radius R_M, $n_{main}(m)dm$ represents the number of mainshocks in the range $[m, m+dm)$, $B_R = \frac{A}{log(10)(b-\alpha)}$ is the branching ratio, i.e. the average number of events triggered by any earthquake, and $H(R_m) = \int_0^{R_m} d\Delta r G(\Delta r, m_M)$. Neglecting the background contribution, Eq.(4) indicates that the ETAS model can reproduce the experimental result Eq.(3) with $\alpha_f = \alpha$. In numerical simulations we have explored a wide range of ETAS parameters A, p, c, α and verified that there exists a set of parameters leading to ETAS catalogs with the same behavior of $n_{aft}(m_M)/n_{main}(m_M)$ of the instrumental ones. In all cases the agreement between ETAS and instrumental catalogs is always recovered for values of $\alpha \gtrsim \alpha_f$. We wish to stress the difference between α and α_f: α is the model parameter which controls the productivity law in numerical simulations whereas α_f is the value obtained applying our definition of mainshock and aftershock to ETAS

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

catalogs and then performing a fit according to Eq.(3). The small discrepancies between α and α_f can be attributed to the background contribution which weakly affects data at small m_M whereas it can be neglected for increasing m_M.

A central observation is that all choices of parameters producing agreement in $n_{aft}(m_M)/n_{main}(m_M)$ between ETAS and instrumental catalogs give a value of $n_{fore}(m_M)/n_{main}(m_M)$ in ETAS catalogs systematically smaller than the instrumental value. It is difficult to obtain a simple approximated expression for the foreshock number as function of m_M as in Eq.(4). However, it is reasonable to expect that the ratio $n_{aft}(m_M)/n_{fore}(m_M)$ weakly depends on model parameters. This is supported by numerical simulations where we fix α in order to have the same value of α_f in ETAS catalogs. Results plotted in Fig.(2) show that different choices of B, p, c lead to similar results for $n_{aft}(m_M)/n_{fore}(m_M)$, significantly larger than the experimental value for any m_M.

4 Catalog incompleteness and the ETASI2 model

As anticipated in the introduction, the incompleteness of instrumental data sets can be responsible for the observed differences with ETAS catalogs. Indeed, because of the overlap of seismic coda waves, many aftershocks are not recorded in particular in the first temporal periods after large shocks *Kagan* [2004]; *Helmstetter et al.* [2006]; *Enescu et al.* [2007]; *Peng et al.* [2007]; *Peng and Zhao* [2009]; *Omi and Aihara* [2013]; *Lippiello et al.* [2016]; *Hainzl* [2016a, b]; *de Arcangelis et al.* [2018]. The direct inspection of seismic signals [*Lippiello et al.*, 2016; *de Arcangelis et al.*, 2018] has shown that, at a temporal distance τ after an event of magnitude m_0, there exists a lower magnitude level $m_x(\tau, m_0)$ such that it is impossible to detect events with $m \leq m_x(\tau, m_0)$. Results indicate a logarithmic decay of $m_x(\tau, m_0)$ in time

$$m_x(\tau, m_0) = m_0 - \phi \log(\tau) - \Delta m \tag{5}$$

with $\phi \simeq 1$ and $\Delta m \simeq 1$, if τ is measured in seconds. Accordingly earthquakes can be hidden by larger events occurring before them at small temporal distances. As a consequence incompleteness affects more strongly the aftershock than the foreshock number and could provide an explanation for the larger value of $n_{fore}(m_M)/n_{aft}(m_M)$ in instrumental catalogs.

In the following we take explicitly into account the aftershock incompleteness adopting the same procedure developed in ref.*de Arcangelis et al.* [2018] to reproduce both the non-trivial dependence of the c-value in the OU law on the manishock magnitude *Shcherbakov et al.* [2005]; *Lippiello et al.* [2007a]; *Bottiglieri et al.* [2009a, b, 2010, 2011]; *Lippiello et al.* [2012a]; *David-*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

sen and Baiesi [2016], as well as the non trivial magnitude correlations between subsequent earthquakes *Lippiello et al.* [2007b, 2008, 2009b, a]; *Sarlis et al.* [2010a, b]; *Sarlis* [2011]; *Lippiello et al.* [2012b, 2013]. The model, defined as ETASI2 model, implements aftershock incompleteness by multiplying the occurrence rate $Q(\Delta r, t-t_i, m_i)$ in Eq.(2) by a detection rate function of the magnitude $\Phi(m|q,\sigma)$ represented by the cumulative distribution function of the normal distribution. The function $\Phi(m|q,\sigma)$ depends on two parameters q and σ representing, respectively, the magnitude with a 50% detection rate and a partially detected magnitude range. In other words, $\Phi(q|q,\sigma) = 0.5$ whereas $\Phi(m|q,\sigma) \simeq 0$ when $m < q-\sigma$ and $\Phi(m|q,\sigma) \simeq 1$ when $m > q-\sigma$. In the ETASI2 model the q parameter depends on time according to Eq.(5), $q = m_x(\tau, m_i)$.

We have performed extended numerical simulations of the ETASI2 model exploring a wide range of model parameters and evaluated $n_{aft}(m_M)$ and $n_{fore}(m_M)$. Restricting to parameters with $n_{aft}(m_M)/n_{main}(m_M)$ in agreement with the instrumental RSCEC catalog, we find (Fig.(2)) that, as expected, the incompleteness increases the value of $n_{fore}(m_M)/n_{aft}(m_M)$ which, however, still remains systematically smaller than the value found in instrumental catalogs. For fixed ϕ and Δm in Eq.(5), the ratio $n_{fore}(m_M)/n_{aft}(m_M)$ does not strongly depend on different choices of σ, as well as on different values of A, p, c and is always significant smaller than the numerical one. We also find that $n_{fore}(m_M)/n_{aft}(m_M)$ slightly increases for decreasing Δm and becomes approximately Δm independent for $\Delta m \lesssim 0$. However, also in this case the value $n_{fore}(m_M)/n_{aft}(m_M)$ is significantly smaller than the one measured in the instrumental catalogs. The origin of this discrepancy is that incompleteness also affects the foreshock number. Indeed, considering a mainshock of magnitude m_2 triggered by an event (a foreshock) with magnitude $m_1 < m_2$, incompleteness does not only affect the identification of both m_1 and m_2 but it can hide foreshocks with magnitude $m < m_1$ occurring between them. Therefore, if the parameters are tuned in order to produce a higher aftershock incompleteness this also reduces the foreshock number and the experimental result is never recovered. In Fig.1 we plot for each instrumental catalog the results of the ETASI2 model setting the model parameters in order to achieve the best agreement for $n_{aft}(m_M)/n_{main}(m_M)$ and minimizing the discrepancy for $n_{fore}(m_M)/n_{main}(m_M)$. Results are obtained assuming $\phi = 0.75$ and $\Delta m = 0.8$ for all catalogs, and keeping $\sigma = 0.3$. Furthermore we use $p = 1.2$ and $c = 0.01$ sec. The values of A and α producing the best agreement are listed in the caption of Fig.1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

4.1 Aftershock and Foreshock spatial distribution

In the previous section we have shown that incompleteness can only partially explain the deficit of foreshocks in ETAS catalogs. As discussed in Sec.3.1 an important difference, which has been observed in previous studies concerns the feature F2. This implies that even if the aftershock and foreshock number are different, their linear density distribution are similar $\rho_a(\Delta r, m_M) \simeq \rho_f(\Delta r, m_M)$ for different values of m_M.

In this study we focus on the aftershock and foreshock average epicentral distance $\zeta_a(\Delta r, m_M)$ and $\zeta_b(\Delta r, m_M)$ defined as $\zeta_a(\Delta r, m_M) \equiv \frac{1}{\Delta r}\int_0^{\Delta r} dx r \rho_a(r, m_M)$ and $\zeta_f(\Delta r, m_M) \equiv \frac{1}{\Delta r}\int_0^{\Delta r} dr r \rho_a(r, m_M)$. Here $\rho_a(\Delta r, m_M)$ is the aftershock linear density distribution defined as the number of aftershocks with epicentral distance from the mainshock in the interval $[\Delta r, 1.2\Delta r)$, divided by $0.2\Delta r$ and by the total number of aftershocks in the interval $[0, R]$. A similar definition applies to the foreshock linear density $\rho_f(\Delta r, m_M)$.

In Fig.(3) we plot $\zeta_a(\Delta r, m_M)$ and $\zeta_f(\Delta r, m_M)$ for different catalogs. It is evident that, for all catalogs, data corresponding to different m_M are well separated and in all cases $\zeta_a(\Delta r, m_M) \simeq \zeta_f(\Delta r, m_M)$. The latter result is a direct consequence of the similarity $\rho_a(\Delta r, m_M) \simeq \rho_f(\Delta r, m_M)$ (F2) which cannot be reproduced by the ETAS model. Indeed, neglecting the contribution of background seismicity, we indicate by $p_a(m|m_0)$ the probability that an event with magnitude m_0 triggers a smaller earthquake ($m < m_0$) inside a temporal window T. The aftershock epicentral distribution is approximatively given by $\rho_a(\Delta r, m_0) \simeq p_a(m|m_0)G(\Delta r, m_0)$. Considering triggering events in the range $[m_M, m_M + dM)$ we therefore obtain

$$\rho_a(\Delta r, m_M) \simeq \frac{\int_{m_M}^{m_M+dM} dm_0 \int_{m_{th}}^{m_0} dm G(\Delta r, m_0) p_a(m|m_0)}{\int_{m_M}^{m_M+dM} dm_0 \int_{m_{th}}^{m_0} dm p_a(m|m_0)}. \quad (6)$$

At the same time we indicate with $p_f(m|m_0)$ the probability that an event with magnitude m triggers a larger earthquake ($m < m_0$) inside a temporal window T. In this case, the foreshock epicentral distance is approximatively given by $\rho_f(\Delta r, m_0) \simeq p_f(m|m_0)G(\Delta r, m)$ and considering triggered events in the range $[m_M, m_M + dM)$ we find

$$\rho_f(\Delta r, m_M) \simeq \frac{\int_{m_M}^{m_M+dM} dm_0 \int_{m_{th}}^{m_0} dm G(\Delta r, m) p_f(m|m_0)}{\int_{m_M}^{m_M+dM} dm_0 \int_{m_{th}}^{m_0} dm p_f(m|m_0)}. \quad (7)$$

We wish to stress the fundamental difference between Eq. (6) and Eq. (7). In Eq. (6) the spatial distance is controlled by the kernel $G(\Delta r, m_0)$ which depends on $m_0 \in [m_M, m_M + dM)$ and if $dM \to 0$ $\rho_a(\Delta r, m_M) \simeq G(\Delta r, m_M)$. Conversely, in Eq. (7) the spatial kernel is $G(\Delta r, m)$ with m ranging in $m \in [m_{th}, m_M)$. In this case if $dM \to 0$, since $p_f(m|m_0)$ is an exponential decreasing function of m, the integral in Eq. (7) is mainly controlled by the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

C

contribution from $m \simeq m_{th}$ which leads to $\rho_f(\Delta r, m_M) \simeq G(\Delta r, m_{th})$. As a consequence, in the ETAS model, we expect that $\rho_f(\Delta r, m_M)$ only weakly depends on m_M differently from ρ_a and strongly depends on m_{th}. The comparison between Eq.s(6,7) therefore shows that, independently of the value of the model parameters, the condition $\rho_a(\Delta r, m_M) \simeq \rho_f(\Delta r, m_M)$ can never be observed. This is confirmed by the results of numerical simulations (Fig.(4)) which show that, even if one can generate ETAS catalogs with $\zeta_a(\Delta r, m_M)$ in good agreement with instrumental catalogs (Fig.(4)a), significant differences are observed between the numeric and the experimental $\zeta_f(\Delta r, m_M)$ (Fig.(4)b). This difference becomes more pronounced for increasing m_M and can simply attributed to the nature of foreshocks in the ETAS model which are tipical events that have triggered a larger shock.

We finally stress that the observed difference in $\zeta_f(\Delta r, m_M)$ between instrumental and ETAS catalogs cannot be related to catalog incompleteness. Indeed, incompleteness typically affects the number of aftershocks and foreshocks but not their spatial distribution. We have explicitly verified this point by means of numerical simulations evidencing that the dependence on Δr and m_M of $\zeta_a(\Delta r, m_M)$ and $\zeta_f(\Delta r, m_M)$ is substantially indistinguishable for the ETAS and the ETASI2 model (not shown).

5 The ETAFS model

According to the nucleation theory [*Ohnaka*, 1992, 1993; *Dodge et al.*, 1996] one expects a concentration of seismicity inside the nucleation zone during the preparatory phase. This kind of seismic activity is not considered in the ETAS model and can justify the excess of foreshocks observed in instrumental catalogs (Fig.(2)). Furthermore, since the size of the nucleation zone scales with the magnitude of the incoming mainshock[*Ohnaka*, 1992; *Dodge et al.*, 1996], the pre-seismic activity can also explain the result $\zeta_a(\Delta r, m_M) \simeq \zeta_f(\Delta r, m_M)$ (Fig.(3)). We then propose a novel model, the Epidemic Type Aftershocks and Foreshock Sequence (ETAFS) model, which implements together with the standard aftershock triggering also additional earthquakes corresponding to the foreshocks expected according to a nucleation scenario. In the ETAFS model each earthquake can trigger its own aftershocks with a probability Q (Eq.(2)) as in the ETAS model. The new ingredient is that each earthquake can be also anticipated by a number of foreshocks according to a probability

$$Q_f(\Delta r, t - t_i, m_i) = \frac{B(p-1)}{c'} \sum_{i:t_i<t} 10^{\alpha'(m_i - m_c)} \left(1 + \frac{t - t_i}{c'}\right)^{-p} G_f(\Delta r_i, m_i). \qquad (8)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

Table 1. Best parameters of the ETAFS model

Catalog	A	α	B	α_f	$\mu\ s^{-1}$
RSCEC	0.084	0.9	0.050	0.54	5.84E-4
RNCEC	0.082	0.88	0.033	0.59	4.98E-4
ItEC	0.086	0.88	0.052	0.60	5.21E-4
JapEC	0.234	0.6	0.160	0.36	5.92E-3

Because of the experimental result $\rho_a(\Delta r, m_M) \simeq \rho_f(\Delta r, m_M)$ we set $G_f(\Delta r_i, m_i) = G(\Delta r_i, m_i)$ in Eq.(2). We have also implemented an inverse-Omori law with the same p as for aftershock occurrence, to reduce the number of model parameters. There is no physical justification for it and we expect that similar results can be recovered with other functional forms of temporal clustering. We fix $c = 100$ s and chose the parameters B, c', α', for each instrumental catalog, in order to reproduce the value $n_{fore}(m_M)/n_{main}(m_M)$, for different m_M (see Tab. 1). We also take explicitly into account aftershock incompleteness, implemented as in the ETASI2 model, and finally generate synthetic catalogs with the same number of $n_{aft}(m_M)/n_{main}(m_M)$ and $n_{fore}(m_M)/n_{main}(m_M)$ of instrumental ones. Fig.2 shows that for all values of m_M ETASI2 catalogs contains the same number of both aftershocks and foreshocks of instrumental data sets. In (Fig.4) we compare the average fore-mainshock distance $\zeta_f(\Delta r, m_M)$ between the ETAFS and the RSCEC catalog. Results show good agreement for all values of m_M.

6 Conclusions

In conclusions, we have discussed the differences between statistical features of foreshocks in ETAS and instrumental catalogs. We have then introduced the novel ETAFS model which explicitly implements foreshocks in the ETAS model and is able to reproduce the entire ensemble of experimental observations. All properties investigated in this study are obtained by means of a stacking procedure. An interesting point is the behavior expected according to the ETAFS model for the seismic activity before a single large shock. The best-fit parameters of the ETAFS model (Tab.1) indicate a small coefficient $\alpha_f \simeq 1.1$ in the foreshock productivity law and accordingly the number of foreshock remains relatively small also before large m_M. As an example, we expect on average less than 8 ($m > 2$) foreshocks the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

day before a magnitude $m = 7$ mainshock, within a radius of 10 km. This small number implies that, for a single mainshock, foreshock activity can at most appear in the form of isolated bursts not leading to an evident systematic increase of the seismic rate. This is consistent with experimental observations, where the inverse Omori-law is obtained only after a stacking procedure and rarely observed inside isolated sequences [*Papadopoulos et al.*, 2010; *Daskalaki et al.*, 2016].

The agreement with experimental data suggests that the ETAFS model can contribute to a significant improvement of pre-seismic forecasting. A rigorous validation of this point, however, needs to be tested in prospective tests. Unfortunately, a main limitation of the model is that it is not immediately suitable to be implemented in this kind of analysis. Indeed, in order to forecast the occurrence of an earthquake, according to the ETAFS model it is necessary to distinguish foreshock from normal earthquake triggering. An attempt in this direction ([*Lippiello et al.*, 2012c]) multiplies the ETAS occurrence probability by an ad-hoc function, giving different weights to aftershock and foreshock clustering. This produces significant gain in the retrospective forecasting of $m > 6$ earthquakes. The nature of foreshocks implemented in the ETAFS model is consistent with this approach promoting further studies on the relevance of foreshocks in seismic forecasting.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

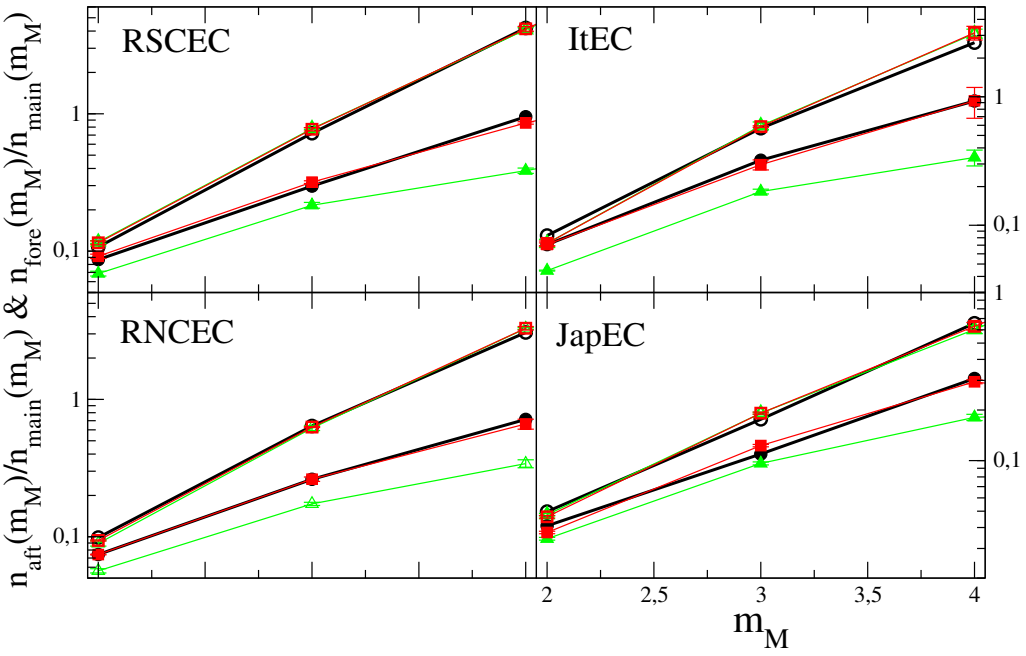

Figure 1. (Color online) The ratio $n_{aft}(m_M)/n_{main}(m_M)$ and $n_{fore}(m_M)/n_{main}(m_M)$ in instrumental and synthetic catalogs. Different panels correspond to different instrumental catalogs. We use open symbols for $n_{aft}(m_M)/n_{main}(m_M)$ and filled symbols for $n_{fore}(m_M)/n_{main}(m_M)$. Results from the instrumental data sets are indicated with black circles. Green triangles are results for the ETASI2 model and red squares for the ETAFS model. The error bars (of the same size of symbols) in numerical catalogs represent the standard deviation from 100 realization of synthetic catalogs. The best parameter of the ETAFS model are listed in Tab.1 whereas for the ETASI2 model the best agreement is obtained with $A = 0.084$, $\alpha = 0.9$ and $\mu = 5.8510^{-4}s^{-1}$ for RSCEC, $A = 0.082$, $\alpha = 0.88$ and $\mu = 4.9810^{-4}s^{-1}$ for RNCEC, $A = 0.082$, $\alpha = 0.88$ and $\mu = 5.2110^{-4}s^{-1}$ for ItEC and $A = 0.26$, $\alpha = 0.60$ and $\mu = 5.9210^{-3}s^{-1}$ for JapEC.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

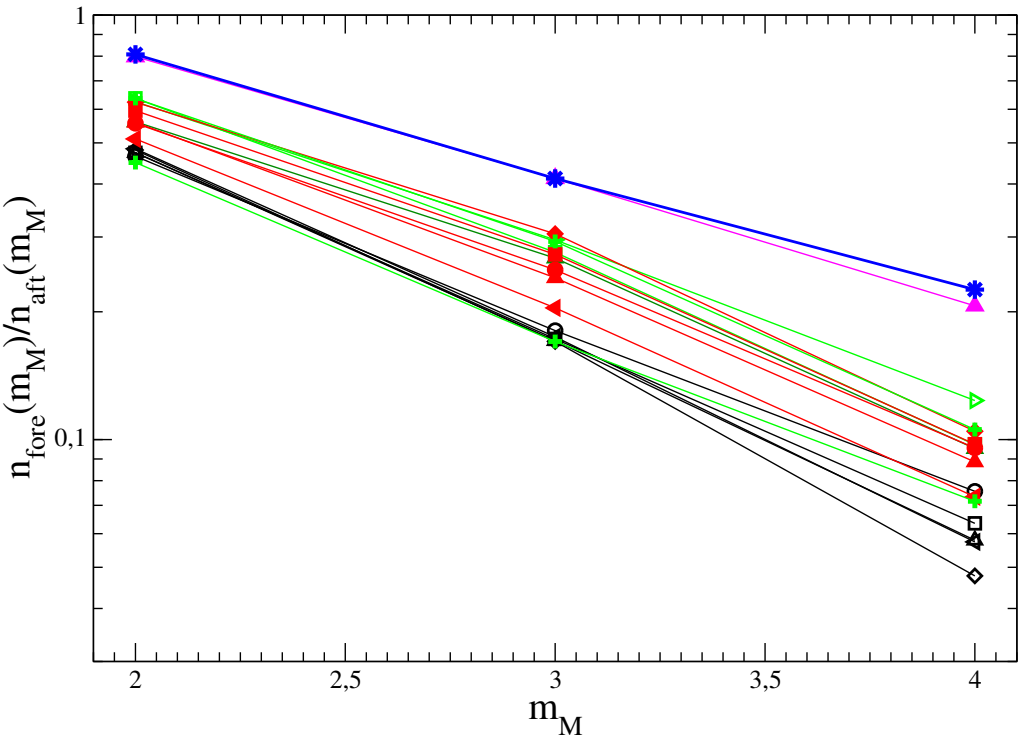

Figure 2. (Color online) The ratio $n_{fore}(m_M)/n_{aft}(m_M)$ in the RSCEC catalog (blue stars) is compared with the value obtained in synthetic ETAS and ETASI2 and ETAFS catalog. The black open symbols are results for the ETAS model for different choices of the parameters $A \in [0.05, 0.12]$, $p \in [1.1, 1.25]$ and $c \in [0.001, 0.1]$. The red filled symbols are results for the ETASI2 model implementing Eq.(5) with $\phi = 0.75$ and $\Delta m = 0.8$ and for different choices of the parameters $A \in [0.05, 0.12]$, $p \in [1.1, 1.25]$ and $c \in [0.001, 0.1]$. Green symbols correspond to $A = 0.084$, $c = 0.01$ and $p = 1.2$ used in Fig.1 for the RSCEC catalog and considering different values of ϕ, Δm and σ. The filled magenta up triangles are results of the ETAFS model with the best set of parameters listed in Tab.1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

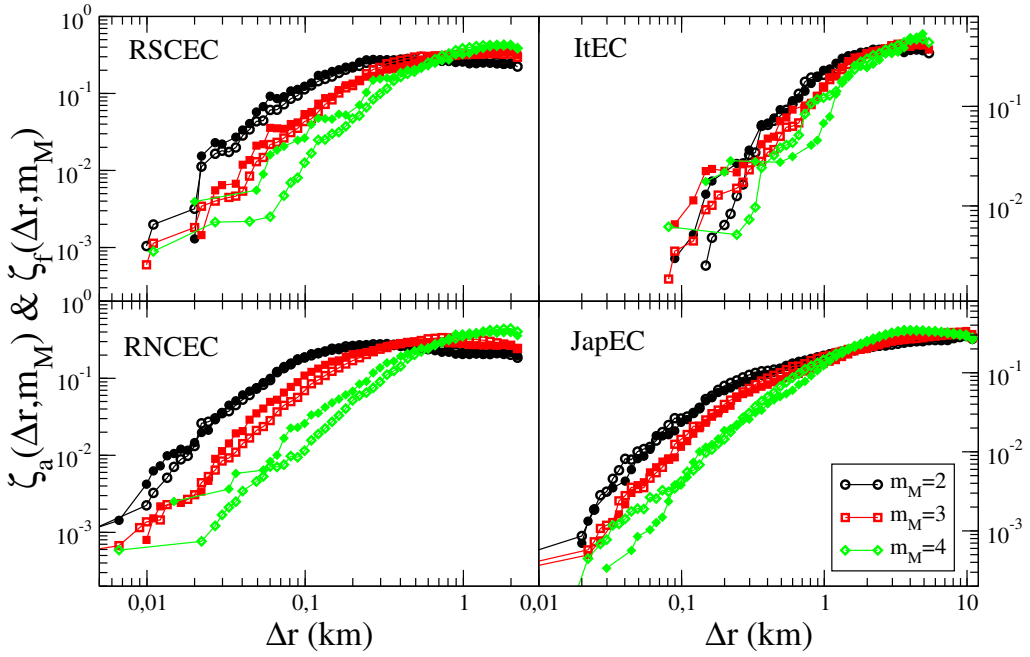

Figure 3. (Color online) The average distance $\zeta_a(\Delta r, m_M)$ of aftershocks (open symbols) and of foreshocks $\zeta_f(\Delta r, m_M)$ (filled symbols) is plotted as function of Δr for the different catalogs. Different mainshock magnitude classes are plotted with different colors and symbols.

Acknowledgments

We thank the National Research Institute for Earth Science and Disaster Prevention for the mainland Japan catalog.

References

From I.S.I.D.E. italian seismological instrumental and parametric data-base.

From Japan Meteorological Agency Earthquake Catalog, provided by N.I.E.D.

Bottiglieri, M., L. de Arcangelis, C. Godano, and E. Lippiello, The generalized omori law: magnitude incompleteness or magnitude clustering, *International Journal of Modern Physics B*, *23*(28n29), 5597–5608, doi:10.1142/S0217979209063882, 2009a.

Bottiglieri, M., E. Lippiello, C. Godano, and L. de Arcangelis, Identification and spatiotemporal organization of aftershocks, *Journal of Geophysical Research: Solid Earth*, *114*(B3), B03,303, doi:10.1029/2008JB005941, 2009b.

Bottiglieri, M., L. de Arcangelis, C. Godano, and E. Lippiello, Multiple-time scaling and universal behavior of the earthquake interevent time distribution, *Phys. Rev. Lett.*, *104*,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 October 2018 doi:10.20944/preprints201810.0679.v1

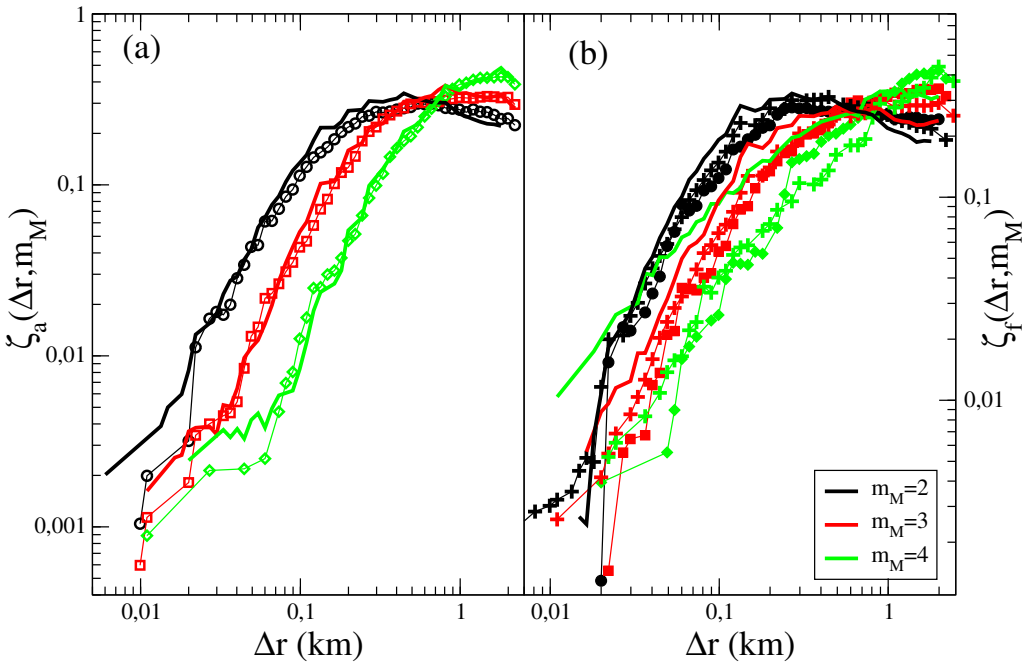

Figure 4. (Color online) (Left panel) The average distance of aftershocks $\zeta_a(\Delta r, m_M)$ in the RSCEC (open symbols) and in the synthetic ETASI2 catalogs (continuous lines) is plotted as function of Δr for different mainshock magnitude classes. (Right Panel) The average distance of foreshocks $\zeta_f(\Delta r, m_M)$ in the RSCEC (filled symbols) and in the synthetic ETASI2 catalog (continuous lines) is plotted as function of Δr for different mainshock magnitude classes. Results for the EATFS model, for the best set of model parameters listed in Tab.1, are plotted with crosses.

158,501, doi:10.1103/PhysRevLett.104.158501, 2010.

Bottiglieri, M., E. Lippiello, C. Godano, and L. de Arcangelis, Comparison of branching models for seismicity and likelihood maximization through simulated annealing, *Journal of Geophysical Research: Solid Earth*, *116*(B2), n/a–n/a, doi:10.1029/2009JB007060, b02303, 2011.

Bouchon, M., and D. Marsan, Reply to artificial seismic acceleration, *Nature Geosci.*, *8*, 83, 2015.

Bouchon, M., V. Durand, D. Marsan, H. Karabulut, and J. Schmittbuhl, The long precursory phase of most large interplate earthquakes, *Nature Geosci.*, *6*, 299302, 2013.

Brodsky, E. E., The spatial density of foreshocks, *Geophysical Research Letters*, *38*(10), L10,305, doi:10.1029/2011GL047253, l10305, 2011.

Brodsky, E. E., and T. Lay, Recognizing foreshocks from the 1 april 2014 chile earthquake, *Science*, *344*(6185), 700–702, 2014.

Chen, X. W., and P. M. Shearer, California foreshock sequences suggest aseismic triggering process, *Geophysical Research Letters*, *40*(11), 2602–2607, doi:10.1002/grl.50444, n/a, 2013.

Daskalaki, E., K. Spiliotis, C. Siettos, G. Minadakis, and G. A. Papadopoulos, Foreshocks and short-term hazard assessment to large earthquakes using complex networks: the case of the 2009 l'aquila earthquake, *Nonlinear Processes in Geophysics Discussions*, *2016*, 1–20, doi:10.5194/npg-2015-80, 2016.

Davidsen, J., and M. Baiesi, Self-similar aftershock rates, *Phys. Rev. E*, *94*, 022,314, doi:10.1103/PhysRevE.94.022314, 2016.

de Arcangelis, L., C. Godano, J. R. Grasso, and E. Lippiello, Statistical physics approach to earthquake occurrence and forecasting, *Physics Reports*, *628*, 1 – 91, doi:http://dx.doi.org/10.1016/j.physrep.2016.03.002, statistical physics approach to earthquake occurrence and forecasting, 2016.

de Arcangelis, L., C. Godano, and E. Lippiello, The overlap of aftershock coda waves and short-term postseismic forecasting, *Journal of Geophysical Research: Solid Earth*, *123*(7), 5661–5674, doi:10.1029/2018JB015518, 2018.

Dodge, D. A., G. C. Beroza, and W. L. Ellsworth, Detailed observations of california foreshock sequences: Implications for the earthquake initiation process, *Journal of Geophysical Research: Solid Earth*, *101*(B10), 22,371–22,392, doi:10.1029/96JB02269, 1996.

Enescu, B., J. Mori, and M. Miyazawa, Quantifying early aftershock activity of the 2004 mid-niigata prefecture earthquake (mw6.6), *Journal of Geophysical Research: Solid Earth*, *112*(B4), B04,310, doi:10.1029/2006JB004629, 2007.

Felzer, K. R., and E. E. Brodsky, Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature*, *441*, 735–738, 2006.

Felzer, K. R., M. T. Page, and A. J. Michael, Artificial seismic acceleration, *Nature Geosci.*, *8*, 82–83, 2015.

Hainzl, S., Comment on self-similar earthquake triggering, bth's law, and foreshock/aftershock magnitudes: Simulations, theory, and results forsouthern california by p. m. shearer, *Journal of Geophysical Research: Solid Earth*, *118*(3), 1188–1191, doi:10.1002/jgrb.50132, 2013.

Hainzl, S., Apparent triggering function of aftershocks resulting from rate-dependent incompleteness of earthquake catalogs, *Journal of Geophysical Research: Solid Earth*, *121*(9), 6499–6509, doi:10.1002/2016JB013319, 2016JB013319, 2016a.

Hainzl, S., Ratedependent incompleteness of earthquake catalogs, *Seismological Research Letters*, *87*(2A), 337–344, 2016b.

Hauksson, E., P. Shearer, , and W. Yang, Waveform relocated earthquake catalog for southern california (1981 to june 2011), *Bulletin of the Seismological Society of America*, *102*(5), 2239–2244, doi:10.1785/0120120010, 2012.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson, Comparison of short-term and time-independent earthquake forecast models for southern california, *Bulletin of the Seismological Society of America*, *96*(1), 90–106, doi:10.1785/0120050067, 2006.

Huang, Q., M. Gerstenberger, and J. Zhuang, Current challenges in statistical seismology, *Pure and Applied Geophysics*, *173*(1), 1–3, doi:10.1007/s00024-015-1222-7, 2016.

Kagan, Y. Y., Short-term properties of earthquake catalogs and models of earthquake source, *Bulletin of the Seismological Society of America*, *94*(4), 1207–1228, 2004.

Lippiello, E., M. Bottiglieri, C. Godano, and L. de Arcangelis, Dynamical scaling and generalized omori law, *Geophysical Research Letters*, *34*(23), L23,301, doi:10.1029/2007GL030963, 2007a.

Lippiello, E., C. Godano, and L. de Arcangelis, Dynamical scaling in branching models for seismicity, *Phys. Rev. Lett.*, *98*, 098,501, doi:10.1103/PhysRevLett.98.098501, 2007b.

Lippiello, E., L. de Arcangelis, and C. Godano, Influence of time and space correlations on earthquake magnitude, *Phys. Rev. Lett.*, *100*, 038,501, doi:

10.1103/PhysRevLett.100.038501, 2008.

Lippiello, E., L. de Arcangelis, and C. Godano, Role of static stress diffusion in the spatiotemporal organization of aftershocks, *Phys. Rev. Lett.*, *103*, 038,501, doi: 10.1103/PhysRevLett.103.038501, 2009a.

Lippiello, E., L. de Arcangelis, and C. Godano, Time, space and magnitude correlations in earthquake occurrence, *International Journal of Modern Physics B*, *23*(28n29), 5583–5596, doi:10.1142/S0217979209063870, 2009b.

Lippiello, E., A. Corral, M. Bottiglieri, C. Godano, and L. de Arcangelis, Scaling behavior of the earthquake intertime distribution: Influence of large shocks and time scales in the omori law, *Phys. Rev. E*, *86*, 066,119, doi:10.1103/PhysRevE.86.066119, 2012a.

Lippiello, E., C. Godano, and L. de Arcangelis, The earthquake magnitude is influenced by previous seismicity, *Geophysical Research Letters*, *39*(5), L05,309, doi: 10.1029/2012GL051083, 2012b.

Lippiello, E., W. Marzocchi, L. de Arcangelis, and C. Godano, Spatial organization of foreshocks as a tool to forecast large earthquakes, *Sci. Rep.*, *2*, 1–6, 2012c.

Lippiello, E., C. Godano, and L. de Arcangelis, Magnitude correlations in the Olami-Feder-Christensen model, *Europhysics Letters*, *102*(5), 59,002, 2013.

Lippiello, E., A. Cirillo, G. Godano, E. Papadimitriou, and V. Karakostas, Real-time forecast of aftershocks from a single seismic station signal, *Geophysical Research Letters*, *43*(12), 6252–6258, doi:10.1002/2016GL069748, 2016GL069748, 2016.

Lippiello, E., F. Giacco, W. Marzocchi, C. Godano, and L. d. Arcangelis, Statistical features of foreshocks in instrumental and etas catalogs, *Pure and Applied Geophysics*, pp. 1–19, doi:10.1007/s00024-017-1502-5, 2017.

Mignan, A., The debate on the prognostic value of earthquake foreshocks: A meta-analysis, *Sci. Rep.*, *4*, 4099–4103, 2014.

Ogata, Y., Statistical models for earthquake occurrences and residual analysis for point processes, *Research Memo. Technical report Inst. Statist. Math., Tokyo.*, *288*, 1985.

Ogata, Y., Statistical models for earthquake occurrences and residual analysis for point processes, *J. Amer. Statist. Assoc.*, *83*, 9 27, 1988a.

Ogata, Y., Space-time point-process models for earthquake occurrences, *Ann. Inst. Math.Statist.*, *50*, 379402, 1988b.

Ogata, Y., A monte carlo method for high dimensional integration, *Numerische Mathematik*, *55*(2), 137–157, doi:10.1007/BF01406511, 1989.

Ogata, Y., and K. Katsura, Comparing foreshock characteristics and foreshock forecasting in observed and simulated earthquake catalogs, *Journal of Geophysical Research: Solid Earth*, *119*(11), 8457–8477, doi:10.1002/2014JB011250, 2014JB011250, 2014.

Ohnaka, M., Earthquake source nucleation: A physical model for short-term precursors, *Tectonophysics*, *211*(14), 149 – 178, 1992.

Ohnaka, M., Critical size of the nucleation zone of earthquake rupture inferred from immediate foreshock activity, *Journal of Physics of the Earth*, *41*(1), 45–56, doi: 10.4294/jpe1952.41.45, 1993.

Omi, Y. O. Y. H., T., and K. Aihara, Forecasting large aftershocks within one day after the main shock, *Scientific Report*, *3*(2218), 2013.

Papadopoulos, G. A., M. Charalampakis, A. Fokaefs, and G. Minadakis, Strong foreshock signal preceding the l'aquila (italy) earthquake (m_w 6.3) of 6 april 2009, *Natural Hazards and Earth System Science*, *10*(1), 19–24, doi:10.5194/nhess-10-19-2010, 2010.

Peng, Z., and P. Zhao, Migration of early aftershocks following the 2004 parkfield earthquake, *Nature Geosci*, *2*(12), 877–881, 2009.

Peng, Z., J. E. Vidale, M. Ishii, and A. Helmstetter, Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in japan, *Journal of Geophysical Research: Solid Earth*, *112*(B3), n/a–n/a, doi:10.1029/2006JB004386, b03306, 2007.

Sarlis, N. V., Magnitude correlations in global seismicity, *Phys. Rev. E*, *84*, 022,101, doi: 10.1103/PhysRevE.84.022101, 2011.

Sarlis, N. V., E. S. Skordas, and P. A. Varotsos, Nonextensivity and natural time: The case of seismicity, *Phys. Rev. E*, *82*, 021,110, doi:10.1103/PhysRevE.82.021110, 2010a.

Sarlis, N. V., E. S. Skordas, and P. A. Varotsos, Nonextensivity and natural time: The case of seismicity, *Phys. Rev. E*, *82*, 021,110, doi:10.1103/PhysRevE.82.021110, 2010b.

Shcherbakov, R., G. Yakovlev, D. L. Turcotte, and J. B. Rundle, Model for the distribution of aftershock interoccurrence times, *Phys. Rev. Lett.*, *95*, 218,501, doi: 10.1103/PhysRevLett.95.218501, 2005.

Shearer, P. M., Self-similar earthquake triggering, bath's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for southern california, *Journal of Geophysical Research-Solid Earth*, *117*, doi:10.1029/2011jb008957, n/a, 2012a.

Shearer, P. M., Space-time clustering of seismicity in california and the distance dependence of earthquake triggering, *Journal of Geophysical Research-Solid Earth*, *117*, n/a,

 doi:10.20944/preprints201810.0679.v1

2012b.

Shearer, P. M., Reply to comment by s. hainzl on self-similar earthquake triggering, bth's law, and foreshock/aftershock magnitudes: Simulations, theory and results for southern california, *Journal of Geophysical Research: Solid Earth*, *118*(3), 1192–1192, doi:10.1002/jgrb.50133, 2013.

Waldhauser, F., and D. P. Schaff, Large-scale relocation of two decades of northern california seismicity using cross-correlation and double-difference methods, *Journal of Geophysical Research: Solid Earth*, *113*(B8), B08,311, doi:10.1029/2007JB005479, 2008.