


# SCIENTIFIC REPORTS



OPEN

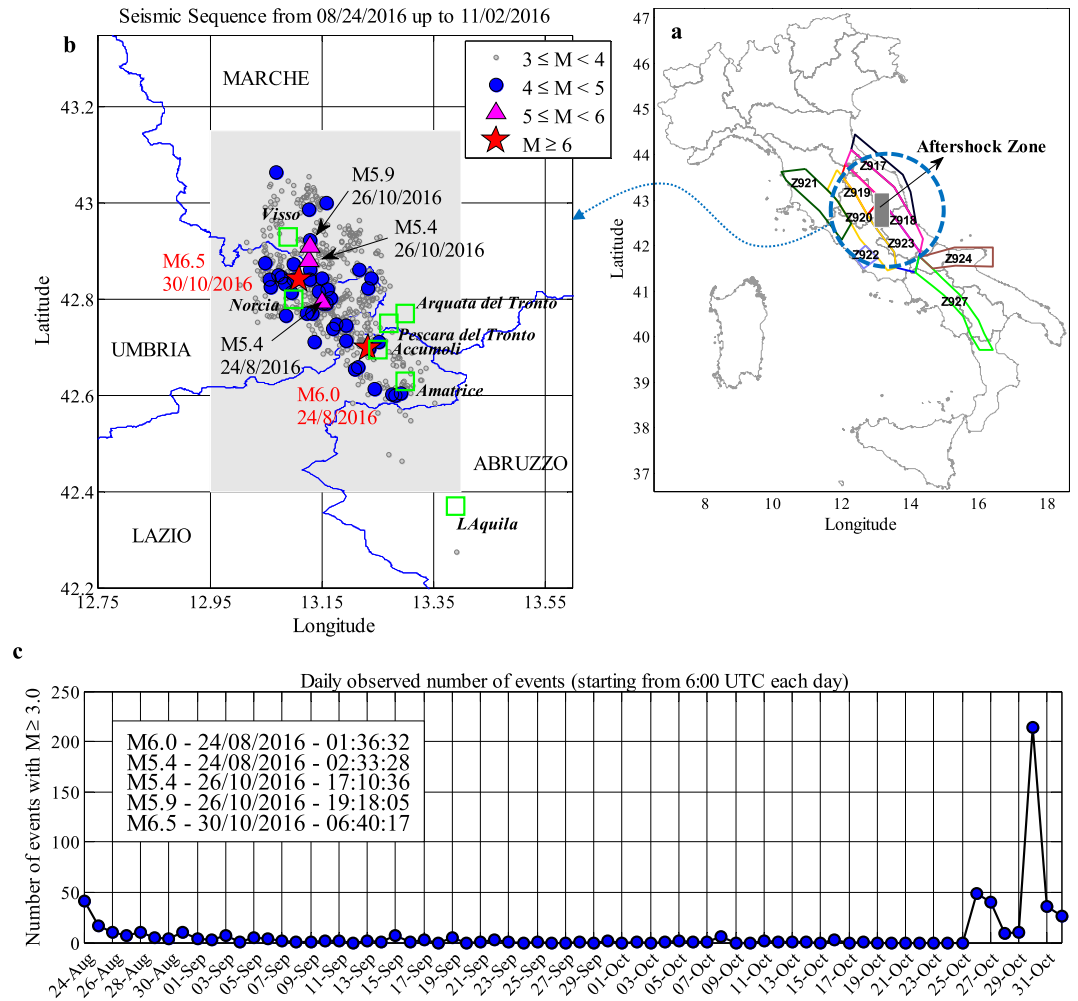
## Robust seismicity forecasting based on Bayesian parameter estimation for epidemiological spatio-temporal aftershock clustering models

Hossein Ebrahimian<sup>1</sup> & Fatemeh Jalayer<sup>2</sup> 

In the immediate aftermath of a strong earthquake and in the presence of an ongoing aftershock sequence, scientific advisories in terms of seismicity forecasts play quite a crucial role in emergency decision-making and risk mitigation. Epidemic Type Aftershock Sequence (ETAS) models are frequently used for forecasting the spatio-temporal evolution of seismicity in the short-term. We propose robust forecasting of seismicity based on ETAS model, by exploiting the link between Bayesian inference and Markov Chain Monte Carlo Simulation. The methodology considers the uncertainty not only in the model parameters, conditioned on the available catalogue of events occurred before the forecasting interval, but also the uncertainty in the sequence of events that are going to happen during the forecasting interval. We demonstrate the methodology by retrospective early forecasting of seismicity associated with the 2016 Amatrice seismic sequence activities in central Italy. We provide robust spatio-temporal short-term seismicity forecasts with various time intervals in the first few days elapsed after each of the three main events within the sequence, which can predict the seismicity within plus/minus two standard deviations from the mean estimate within the few hours elapsed after the main event.

Following a large earthquake and in the presence of a vast number of aftershocks, short-term operational seismicity forecasts (in the order of days to months) are of utmost importance for emergency decision-making and risk mitigation in the disaster area<sup>1–6</sup>. The aftershock activity is forecasted mainly based on the observed data of already registered (and mostly incomplete) recordings within the ongoing sequence<sup>7</sup>. The Epidemic Type Aftershock Sequence (ETAS) model<sup>8,9</sup> is the stochastic model most frequently used to describe earthquake occurrence within a seismic sequence<sup>10</sup>. It is an epidemic stochastic point process in which every earthquake within the sequence is a potential triggering event for subsequent earthquakes<sup>11</sup>, and therefore generates its own well-defined Modified Omori<sup>12,13</sup> (MO) aftershock decay<sup>14</sup>. Hence, it is capable of accounting for the triggering effect of all the events that have taken place before a desired time. The ETAS model performed quite well in operational seismic forecasting during the L'Aquila 2009 (central Italy) seismic sequence<sup>15</sup>. The model parameters are usually calibrated a priori based on the maximum likelihood criterion. The first effort on the calibration of temporal model parameters has been carried on by Ogata<sup>8</sup>, and extended later<sup>9,16–19</sup> to estimate the spatio-temporal model parameters. In addition, several attempts are made for developing improved algorithms to attain maximum likelihood estimates of ETAS parameters; e.g., the Expected-Maximization algorithm<sup>20</sup>, an improved maximum likelihood algorithm<sup>21</sup>, and a new algorithm based on Simulated Annealing optimization technique<sup>10</sup>, which allows for an automatic maximum likelihood estimation of the model parameters instead of fine-tuning of some algorithm parameters in advance. Adaptive model parameter estimation based on the events in the ongoing sequence (e.g., calibrating the parameters of MO and ETAS models based on the ongoing catalogue by employing Bayesian parameter estimation<sup>22–25</sup>) has the advantage of both tuning a sequence-specific model and also

<sup>1</sup>Post-doctoral Researcher, Department of Structures for Engineering and Architecture, University of Naples Federico II, Via Claudio 21, 80125, Naples, Italy. <sup>2</sup>Associate Professor, Department of Structures for Engineering and Architecture, University of Naples Federico II, Via Claudio 21, 80125, Naples, Italy. Correspondence and requests for materials should be addressed to F.J. (email: [fatemeh.jalayer@unina.it](mailto:fatemeh.jalayer@unina.it))



**Figure 1.** Amatrice 2016 seismic sequence. (a) The aftershock zone indicated by the grey-coloured box in perspective with the surrounding Italian seismogenic zonation. (b) The spatial distribution of aftershock events based on Catalogue 2 from August 24, 2016 (01:36 UTC) up to November 2, 2016 (10:32 UTC) bordering four neighbouring provinces in Italy. The grey-coloured box defines the considered aftershock zone and the most damaged towns are highlighted with green boxes. The main seismic events are illustrated as follows: M6.0 and M6.5 with red stars; M5.4 (24/08/16), M5.4 (26/10/16), M5.9 (26/10/16) with magenta triangles; aftershocks  $M \geq 3.0$  with grey circles. (c) The number of events (with  $M \geq 3.0$ ) in Catalogue 2 occurred within a 24-hour interval starting from 6:00 UTC of the desired day (MATLAB 2016b, <http://software.unina.it/matlab/> is used to create this figure).

capturing possible time variations of the model parameters. As the original purpose of the present paper, we propose a fully simulation-based method to provide a *robust estimate*<sup>24, 26, 27</sup> for the spatial distribution of the events in a prescribed forecasting time interval after the main event. In the context of this robust estimate, the uncertainty in the ETAS model parameters is taken into account as the posterior joint probability distribution for the model parameters conditioned on the events that have already occurred (i.e., registered events in the ongoing seismic sequence before the beginning of the forecasting interval). The Markov Chain Monte Carlo (MCMC) simulation scheme<sup>24, 25, 27</sup> is used to sample directly from the posterior probability distribution for ETAS model parameters (i.e., conditioned on the registered events in the ongoing sequence). Moreover, this robust estimate also considers the sequence of events that is going to occur during the forecasting interval (and hence affect the seismicity in an epidemic type model like ETAS). Although this sequence is unknown at the time of forecasting, we propose a stochastic procedure to generate it. The procedure leads to the stochastic spatio-temporal distribution of the forecasted events and consequently to the uncertainty in the estimated number of events, corresponding to a given forecasting interval. The resulting robust forecasts are directly applicable in adaptive daily aftershock hazard and risk assessment procedures<sup>23, 28, 29, 30</sup>.

The proposed methodology is applied to provide retrospective forecasting for seismic activities of the 2016 Amatrice sequence by analysing the registered data of quasi real-time catalogues from INGV (Istituto Nazionale di Geofisica E Vulcanologia). The corresponding aftershock zone, as shown in Fig. 1a by the gray-colored area, is located mostly within the seismic zone 923 based on the ZS9 Italian Seismogenic Zonation<sup>31</sup>. Fig. 1a shows also the seismogenic zonation surrounding the aftershock zone. It is to note that based on the Italian seismic

zonations, the upper-bound magnitude for seismic zone Z923 is  $M_{\max} = 7.06$ . On the 24th of August 2016 at 01:36 UTC, a Mw 6.0 earthquake struck the Central Italy between towns of Norcia and Amatrice, devastating Amatrice, Accumoli and several surrounding small towns and villages, causing almost 300 fatalities and leaving almost 30,000 homeless. The seismic sequence, including a Mw 5.4 aftershock (occurred almost one hour after the main shock at 02:33 UTC), triggered hundreds of earthquakes per day until the mid-September. Two months after, on the 26th of October, a Mw 5.4 followed within a two-hour delay by a Mw 5.9 earthquake (at 17:10 and 19:18 UTC, respectively) took place in the east of town Visso, and preceded the largest event of the sequence, a Mw 6.5 on October 30 at 06:40 UTC, North of Norcia<sup>32</sup>. This one is the largest earthquake recorded in Italy since the Mw 6.9 1980 Irpinia event. Fig. 1b and c illustrate the seismic activities within the aftershock zone during the first two months highlighting the key events taken place.

In the adaptive procedure presented in Section Method, the spatial maps can provide indications with regards to automatic re-sizing of the aftershock zone (if required) in case the spatial extension of the seismicity is not captured adequately inside a designated aftershock zone. For each forecasting interval, the spatial distribution of aftershock events is estimated based on equation (1) by calculating the (average) number of events  $N(x, y, m|\text{seq}, M_l)$  in the cell unit centered at Cartesian coordinate  $(x, y)$  with magnitude greater than or equal to  $m$ . For a prescribed aftershock zone extension, the boundaries of the zone would need to be adjusted if  $N(x, y, m|\text{seq}, M_l)$  is significantly greater than zero along the border cell units. This is a very important issue especially in case of Amatrice (multiple) seismic sequence 2016, where the seismicity was relocated towards north-west after two months from the main event on August 24. In this study, the aftershock zone is assigned a priori and the adequacy of its extension is monitored for each forecast.

The present study strives to perform robust forecasts for the spatio-temporal evolution of the events in specific time intervals within the very complex sequence described above that is distinguished by three main events (“mainshocks”) of moment magnitudes 6.0, 5.9, and 6.5, respectively (as illustrated in Fig. 1b). We divide the sequence into three parts: (a) from 24-August to 25-October, (b) from 26-October to 29-October, and finally (c) from 30-October to 1-November. We have used two different catalogues herein in order to gather data backwards in time. The two catalogues are distinguished by the date of access to the source database and share the time of origin (01:36 UTC, August 24 2016). The first one (labelled as Catalog 1) lasts until September 9, 2016 (07:23 UTC, date of first access) when the aftershock activities are reduced considerably in the zone. The second one (labelled as Catalog 2 whose registered data are illustrated in Fig. 1b and c) covers until November 2, 2016 (10:32 UTC, date of second access).

## Results

**Providing daily forecasts of seismicity from August 24 up to October 25.** The prediction time window  $[T_{start}, T_{end}]$  indicates a 24-hour interval where  $T_{start}$  is 6:00 UTC of the following day. Each forecast uses available information at the time when the forecast is issued, i.e., the sequence (denoted as **seq** in Section Methods) comprised of events registered in the Catalogue 1 including the main event up to the time  $T_{start}$ . To issue the first forecast, the observation history, **seq**, comprises the main event with Mw 6.0 at 01:36 UTC and the triggered events up to 6:00 UTC of 24 August 2016, where the lower cut-off magnitude,  $M_b$ , of Catalogue 1 is equal to 3.0 based on the two methods discussed in Ebrahimian *et al.*<sup>23</sup>. It is to note that the completeness threshold of the catalog of aftershocks in subsequent forecasting intervals is demonstrated to be even lower than 3.0 (the procedure(s) adopted for evaluating the completeness magnitude  $M_c$  throughout the various phases of this multiple seismic sequence is described in detail in the Supplementary Information in a Section entitled “Discussion on the Completeness Magnitude  $M_c$ ”). In any case,  $M_l = 3.0$  is considered as the cut-off threshold for the computation of the aftershock rates for the upcoming days.

The first step towards providing seismicity forecasts (with reference to equation 9) is sampling from the distribution of modal parameters  $\theta$  based on posterior (target) probability distribution  $p(\theta|\text{seq}, M_l)$ . The vector  $\theta = [\beta, K, K_p, K_R, c, p, d, q]$  is updated on a daily basis by applying the Bayesian updating routine illustrated in equation (17) and considering that parameters  $K, K_p, K_R$  are derived as function of other parameters within vector  $\theta$  (see equations 4–7). As prior information, we assigned a normal distribution to the five model parameters  $[\beta, c, p, d, q]$ , with a coefficient of variation (COV) equal to 0.30. The prior mean values for  $[\beta, c]$  are assigned equal to those provided by the MO Italian generic model parameters<sup>33</sup>, while prior mean values assigned to  $[p, d, q]$  are equal to the ETAS model parameters calibrated for the L’Aquila aftershock sequence<sup>15</sup>. The prior COV for each model parameter is set equal to 0.3 –in some cases larger than the reported prior COV’s in the above-mentioned references– to avoid using an over-informative prior distribution (i.e., a prior with a very low COV). Samples for  $\theta$  are generated as a Markov Chain sequence directly from  $p(\theta|\text{seq}, M_l)$ , as noted in Section Methods. It should be noted that sampled  $p$  and  $q$  values smaller than one are rejected according to equations (4) and (5) with the constraint  $p > 1$  and  $q > 1$ . One key restraint on performing operational forecasting during an ongoing sequence is that the procedure should be performed in a reasonably small amount of time. To address this issue, an MCMC procedure for updating the model parameters is carried out adaptively (see the Section Methods).

The evolution in the statistics (mean and COV) of model parameters  $\theta = [\beta, K, K_p, K_R, c, p, d, q]$  are summarized in Table 1. We use Catalogue 1 for constructing **seq** in order to provide forecasts within the first two weeks after the main event (i.e., from August 24 up to September 06). The first row of Table 1 corresponds to the statistics for the prior marginal PDF’s for  $\theta$  while the subsequent rows indicate the statistics for the posterior distributions. In addition, Supplementary Fig. S1 illustrates the sampled histograms representing the marginal prior and posterior PDF’s corresponding to the six model parameters  $[\beta, K, c, p, d, q]$ . Moreover, we have illustrated the PDF for  $K$ , derived as a function of other model parameters, in Supplementary Fig. S1. The marginal PDF’s for the other two parameters  $K_p$  and  $K_R$  are not shown in the figure since they have a very straightforward relationship to other model parameters (see equations 4 and 5). The marginal distributions shown correspond to a 24 hours (1 day) forecasting interval and a magnitude 3.0 lower cut-off.

	$\beta$		$c$ [day]		$p$		$d$ [km]		$q$		$K$		$K_t$		$K_r$	
	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV
Prior	2.21	0.30	0.03	0.30	1.10	0.30	1.00	0.30	1.50	0.30						
24/08/2016	1.38	0.08	0.03	0.28	1.47	0.14	1.48	0.15	1.63	0.07	0.60	0.30	0.08	0.25	0.35	0.36
25/08/2016	1.53	0.07	0.03	0.27	1.44	0.08	1.60	0.11	1.70	0.06	0.35	0.17	0.08	0.14	0.47	0.43
26/08/2016	1.54	0.06	0.02	0.29	1.40	0.09	1.60	0.12	1.74	0.05	0.36	0.17	0.09	0.11	0.49	0.32
27/08/2016	1.57	0.06	0.02	0.28	1.40	0.08	1.49	0.11	1.72	0.05	0.33	0.15	0.09	0.11	0.43	0.31
28/08/2016	1.60	0.06	0.03	0.27	1.38	0.06	1.55	0.17	1.74	0.07	0.33	0.16	0.09	0.09	0.52	0.55
29/08/2016	1.59	0.06	0.03	0.30	1.34	0.10	1.54	0.12	1.75	0.05	0.36	0.19	0.09	0.08	0.48	0.35
30/08/2016	1.61	0.05	0.02	0.42	1.27	0.08	1.41	0.16	1.67	0.07	0.39	0.22	0.09	0.13	0.37	0.46
31/08/2016	1.64	0.05	0.02	0.30	1.24	0.06	1.53	0.14	1.74	0.06	0.38	0.21	0.09	0.11	0.48	0.39
01/09/2016	1.63	0.05	0.02	0.29	1.19	0.06	1.57	0.11	1.75	0.05	0.51	0.19	0.09	0.15	0.49	0.36
02/09/2016	1.64	0.04	0.02	0.33	1.18	0.06	1.60	0.15	1.76	0.06	0.49	0.24	0.08	0.21	0.54	0.42
03/09/2016	1.64	0.05	0.02	0.33	1.17	0.05	1.62	0.11	1.78	0.05	0.48	0.23	0.09	0.19	0.56	0.34
04/09/2016	1.63	0.05	0.02	0.24	1.19	0.05	1.52	0.11	1.73	0.04	0.46	0.20	0.09	0.14	0.45	0.31
05/09/2016	1.63	0.05	0.02	0.27	1.19	0.05	1.59	0.10	1.80	0.06	0.45	0.19	0.09	0.14	0.57	0.35
06/09/2016	1.65	0.05	0.02	0.29	1.17	0.05	1.57	0.09	1.75	0.04	0.46	0.23	0.09	0.16	0.49	0.27

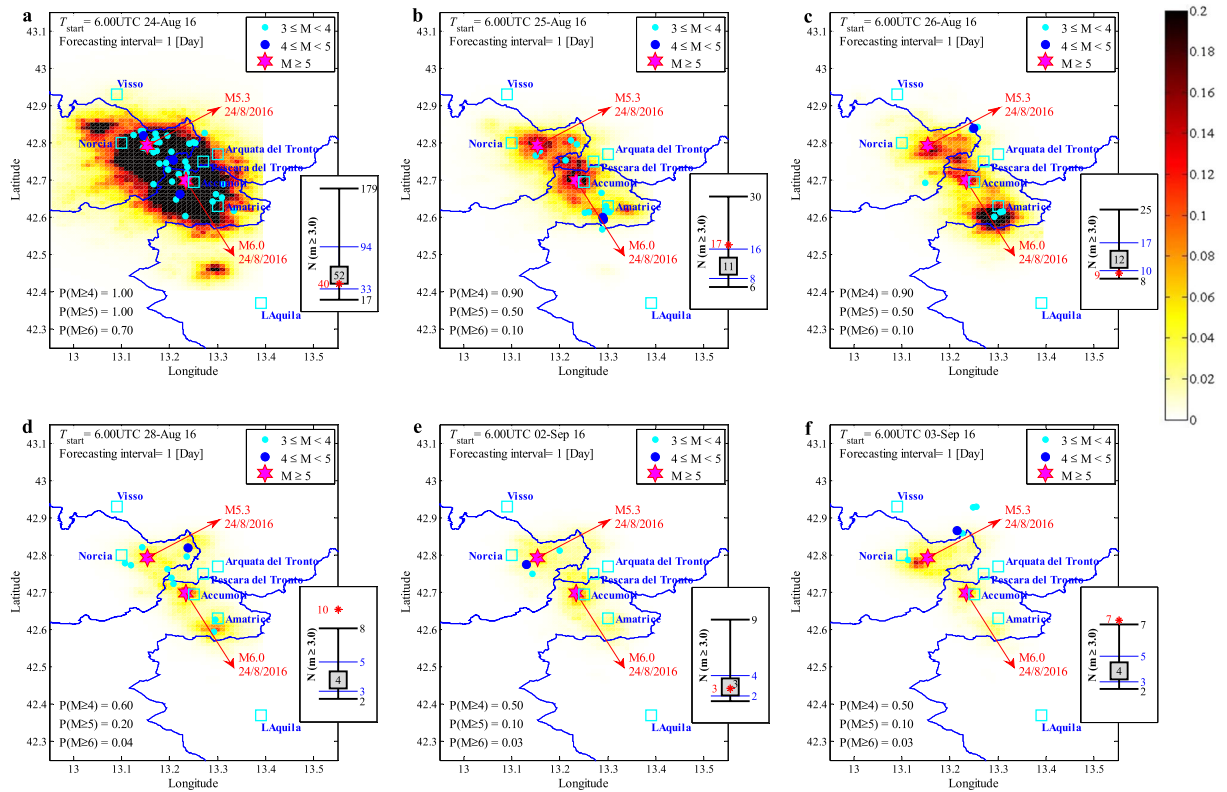
**Table 1.** Statistics (mean and COV) of ETAS model parameters  $\theta$  for estimating the daily seismicity in the first 14 days elapsed after Mw 6.0 at 24-Aug. 2016 (Supplementary Fig. S1 also illustrates the sampled histograms representing the marginal posterior PDF's corresponding to the six model parameters [ $\beta$ ,  $K$ ,  $c$ ,  $p$ ,  $d$ ,  $q$ ]).

The sequence of events taking place during the forecasting interval,  $\mathbf{seq}_g$ , is generated (see equation 11 and Section Methods) given that  $\theta$  and  $\mathbf{seq}$  are known. Finally, substituting the sampled values for  $\theta$  and  $\mathbf{seq}_g$  in equation (9), the robust estimate for the number of events with  $M \geq m$  in a spatial cell units centered at  $(x, y)$  within the aftershock zone is obtained. This robust estimate is calculated as the expected number of events considering the uncertainties in the spatio-temporal distribution of the sequence of events. Fig. 2 shows the forecasted seismicity maps in terms of the mean plus two logarithmic standard deviation (98% confidence interval) for the number of events with  $M \geq 3.0$  within each spatial cell unit issued for the 24-hour time forecasting intervals. The earthquakes of interest occurred within the corresponding forecasting interval are illustrated as coloured dots (distinguished by magnitude). The two main events of the sequence with  $M \geq 5.0$  (see also Fig. 1) are identified with coloured stars (these events are shown for reference only and they did not necessarily take place in the illustrated map's corresponding forecasting interval). We also report the forecasted daily probabilities of having earthquakes of magnitude equal to or larger than  $m = 4, 5$  and 6 in the whole aftershock zone. These probabilities are calculated from the equation  $1 - e^{-N(m|\mathbf{seq}, M_l)}$  where  $N(m|\mathbf{seq}, M_l)$  is the sum over the whole aftershock zone of the expected number of events  $N(x, y, m|\mathbf{seq}, M_l)$  from equation (9).

At the right-hand side of each sub-figure, the observed (shown as a red star) vs. forecasted number of events (shown in an error-bar format) is illustrated for events with  $M_l = 3.0$  for the entire aftershock zone. The error-bar for the forecasted number of events features: the median value (the 50<sup>th</sup> percentile, equivalent of the logarithmic mean in the arithmetic scale) inside a grey-filled square, the (logarithmic) mean plus/minus one (logarithmic) standard deviation indicating the interval between 16<sup>th</sup> and 84<sup>th</sup> percentiles (marked with blue horizontal lines), and the (logarithmic) mean plus/minus two (logarithmic) standard deviations indicating the interval between 2<sup>nd</sup> and 98<sup>th</sup> percentiles (marked with black horizontal lines). This is done to help in locating the observed number of events within plus or minus certain number of standard deviations from the mean estimate. It can be seen that the observed number of events lies within plus/minus one standard deviation of the mean estimate except for 28<sup>th</sup> of August and 3<sup>rd</sup> of September (for this one lies within two standard deviations away from the mean estimate).

**Providing daily forecasts of seismicity from October 26 up to October 29.** As mentioned before, on 26<sup>th</sup> of October, a Mw 5.4, followed within a two-hour delay by a Mw 5.9 earthquake (at 17:10 and 19:18 UTC, respectively), took place in the east of town Visso (located in the north-western part of the aftershock zone, see Fig. 1b). This triggered a new aftershock sequence within the ongoing one. At this stage, given the time elapsed from the occurrence of the mainshock (i.e., around two months), it seemed quite tedious to consider all the events of interest up to the time of origin (i.e., 24<sup>th</sup> of August) for each forecasting interval. To achieve this (i.e., avoid considering all the events back to 24<sup>th</sup> of August), we performed a shift in the time of origin  $T_o$  from August 24<sup>th</sup> to 17:10 UTC of October 26<sup>th</sup> (time of occurrence of the Mw 5.4 earthquake, see Fig. 1). In the following, we describe the details of the forecasting procedure for the period from 26<sup>th</sup> of October to 29<sup>th</sup>.

**24-hour forecasting from 6:00 UTC of 26/10/2016.** We performed the forecasting in this period considering as the sequence  $\mathbf{seq}$  all the events with magnitude greater than or equal to 3.0 that took place after 01:36 UTC of 24/08/2016 up to  $T_{start}$  which is 6:00 UTC of 26/10/2016 (including the main event of August 24<sup>th</sup>). To facilitate the ETAS model parameter  $\theta$  estimation, we employed as prior distribution (instead of the priors considered in the previous section) the updated distribution of model parameters on 06-September (see the last row of Table 1); assuming that the model parameters did not undergo significant changes from 6<sup>th</sup> of September to 26<sup>th</sup> of October (the sequence activity has decayed significantly in this period). The MCMC procedure is carried out as described

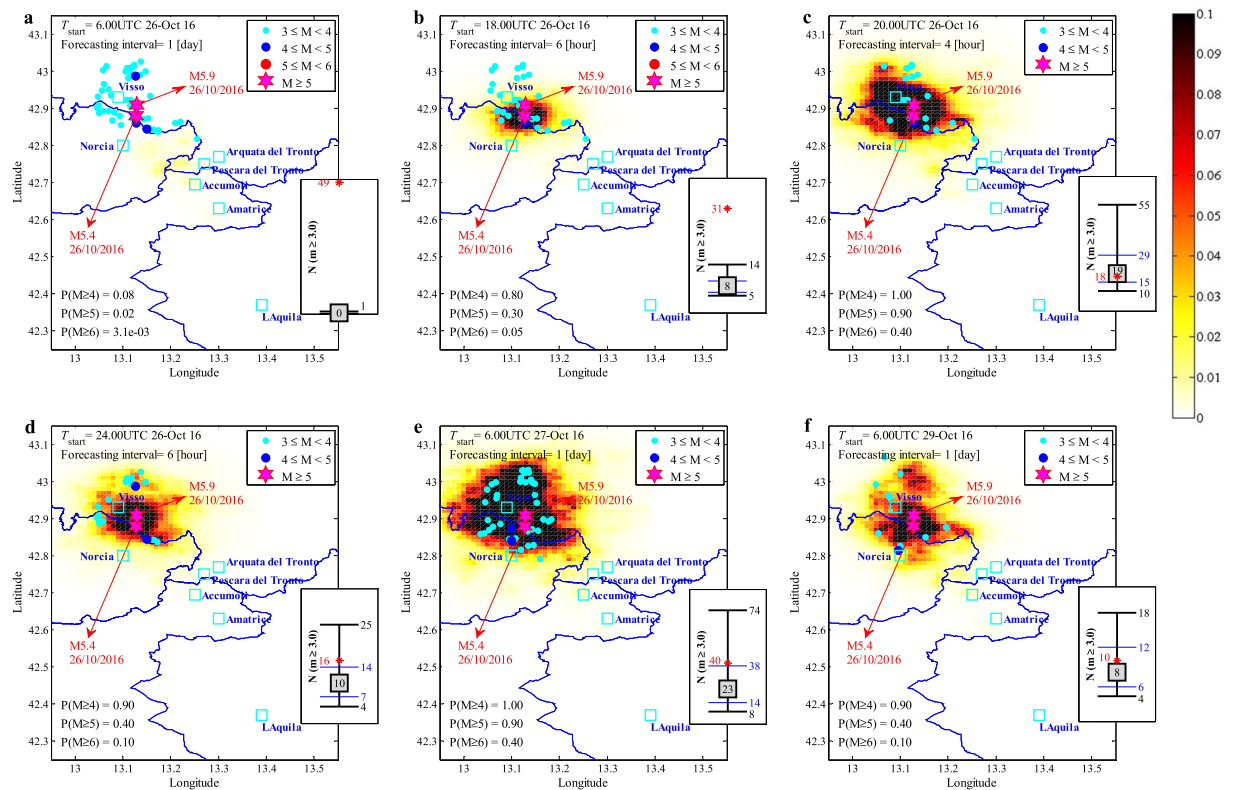


**Figure 2.** Forecasted vs. observed seismicity distribution in the aftershock zone, the maps report the mean +2 standard deviation confidence interval for the number of events per [km<sup>2</sup>] (latitude/longitude cells of a 0.01° × 0.01° grid) equal to or greater than magnitude  $M_i = 3$  in the indicated 24-hour forecasting time window. In the lower left corner, the daily probabilities of having earthquakes with magnitudes 4, 5, and 6 or larger are reported. Each sub-plot also features the earthquakes (coloured dots) that occurred during the corresponding forecasting time window together with the two main events with  $M \geq 5.0$  (magenta stars). The sub-figures illustrate the observed (plotted in red star) vs. the error-bar for the forecasted number of events with  $m \geq M_i$  corresponding to the forecasting time interval: the median value (the 50<sup>th</sup> percentile, equivalent of the logarithmic mean in the arithmetic scale) inside a grey-filled square, the (logarithmic) mean plus/minus one (logarithmic) standard deviation indicating the interval between 16<sup>th</sup> and 84<sup>th</sup> percentiles (marked with blue horizontal lines), and the (logarithmic) mean plus/minus two (logarithmic) standard deviations indicating the interval between 2<sup>nd</sup> and 98<sup>th</sup> percentiles (marked with black horizontal lines). (MATLAB 2016b, <http://softwaresso.unina.it/matlab/> is used to create this figure).

	$\beta$		$c$ [day]		$p$		$d$ [km]		$q$		$K$		$K_i$		$K_R$	
	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV
26/10/2016 <sup>a</sup>	1.67	0.05	0.02	0.36	1.17	0.06	1.57	0.12	1.74	0.06	0.56	0.26	0.08	0.23	0.47	0.31
26/10/2016 <sup>b</sup>	1.34	0.17	0.03	0.27	1.20	0.14	1.39	0.17	1.55	0.09	9.48	0.20	0.08	0.46	0.28	0.43
26/10/2016 <sup>c</sup>	1.53	0.13	0.03	0.27	1.38	0.16	1.53	0.15	1.58	0.08	0.70	0.39	0.08	0.29	0.32	0.46
27/10/2016	1.35	0.10	0.03	0.26	1.38	0.14	1.51	0.13	1.69	0.08	1.33	0.19	0.09	0.23	0.43	0.41
29/10/2016	1.49	0.07	0.03	0.27	1.27	0.09	1.58	0.13	1.71	0.06	0.57	0.34	0.10	0.14	0.47	0.37

**Table 2.** Statistics (mean and COV) of ETAS model parameters  $\theta$  for estimating the daily seismicity rate of the desired dates for the second part of the sequence after Mw 5.4 at 26-October 2016 (Supplementary Fig. S2 also illustrates the sampled histograms representing the marginal posterior PDF's corresponding to the six model parameters [ $\beta, K, c, p, d, q$ ]). <sup>a</sup>Corresponding to Fig. 3(a,b); <sup>b</sup>corresponding to Fig. 3c; <sup>c</sup>corresponding to Fig. 3d.

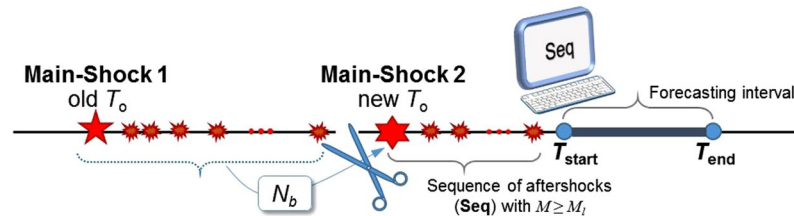
in the previous section. The statistics of the updated model parameters (mean and COV) are reported in the first row of Table 2. It can be observed (comparing with the prior distribution in the last row of Table 1) that model parameters remain more-or-less invariant (we could have directly used the ETAS model parameters estimated for 6<sup>th</sup> of September). Fig. 3a illustrates the forecasted seismicity (mean plus two standard deviations) and the events of interest that occurred during the 24-hour interval of interest. The estimated number of events within the aftershock zone (shown in the error-bar plot reported in the right-hand side of the figure) shows a substantial



**Figure 3.** Forecasted vs. observed seismicity distribution in the aftershock zone, the maps report the mean  $\pm 2$  standard deviation confidence interval for the number of events per  $[km^2]$  (latitude/longitude cells of a  $0.01^\circ \times 0.01^\circ$  grid) equal to or greater than magnitude  $M_i = 3$  in the indicated forecasting time window. In the lower left corner, the daily probabilities of having earthquakes with magnitudes 4, 5, and 6 or larger are reported. Each sub-plot also features the earthquakes (colored dots) that occurred during the corresponding forecasting time window together with the two main events with  $M \geq 5.0$  (magenta stars). The sub-figures illustrate the observed (plotted in red star) vs. the error-bar for the forecasted number of events with  $m \geq M_i$  corresponding to the forecasting time interval: the median value (the 50<sup>th</sup> percentile, equivalent of the logarithmic mean in the arithmetic scale) inside a grey-filled square, the (logarithmic) mean plus/minus one (logarithmic) standard deviation indicating the interval between 16<sup>th</sup> and 84<sup>th</sup> percentiles (marked with blue horizontal lines), and the (logarithmic) mean plus/minus two (logarithmic) standard deviations indicating the interval between 2<sup>nd</sup> and 98<sup>th</sup> percentiles (marked with black horizontal lines). (MATLAB 2016b, <http://software.unina.it/matlab/> is used to create this figure).

reduction in the seismicity compared to the first days elapsed after the main event in August 24. The exceedance probabilities are calculated as  $P(M \geq 4) = 0.08$ ,  $P(M \geq 5) = 0.02$ , and  $P(M \geq 6) = 0.003$ . It is interesting to note that these probabilities are substantially higher than those provided from the calculation of the long-term seismicity level based on the ZS9 Italian Seismogenic Zonation<sup>31</sup> data, estimating a daily probability of around  $3.83 \times 10^{-4}$  for  $M \geq 4.76$ . Hence, the level of (forecasted) seismicity at the desired date is more than 40 times higher than the base seismicity level for  $M \geq 5$ . The distribution of seismicity forecasted for 26<sup>th</sup> of October (shown in Fig. 3a) is used next as an estimate of the background (base) seismicity denoted as  $N_b(x, y, m|M_i)$  in Section Methods.

**6-hour forecasting from 18:00 UTC of 26/10/2016.** After the occurrence of the event with Mw 5.4 at 17:10 UTC of 26/10/2016, we provide a 6-hour prediction of seismicity for the forecasting interval starting from  $T_{start}$  set to 18:00 UTC of 26/10/2016 (i.e., 50 minutes after the occurrence of Mw 5.4 event). At this point, we performed a shift in the time of origin by setting  $T_0$  to 17:10 UTC of 26<sup>th</sup> of October (see Fig. 4). The sequence **seq** includes all the triggered events with  $M \geq 3$  occurred after 17:10 UTC of 26/10/2016 (including the main Mw 5.4 event). It should be noted that the event Mw 5.4 was not preceded by any foreshocks (i.e., no  $M \geq 3$  events took place between 06:00 UTC and 17:10 UTC of 26 of October). Given the very low seismic activity prior to the major event and given the presence of very few events in **seq**, we did not perform Bayesian updating on the model parameter  $\theta$  and used the statistics provided in the first row of Table 2 issued for 26/10/2016 (the parameters updated in the previous step). It is important to note that the forecasted seismicity for the 24-hour interval elapsed after 06:00 UTC of October 26 in the previous step (shown in Fig. 3a) is used herein (after proportioning it to a 6-hour forecasting interval) as the background seismicity  $N_b(x, y, m|M_i)$ . The background seismicity usually considers the long-term seismicity in the calculations and was assumed to be equal to zero in our previous calculations for the first part of the sequence starting from 24<sup>th</sup> of August. Herein, we use this background seismicity to conservatively approximate the triggering effect of the events occurred in the first part of the sequence (from August 24<sup>th</sup> to



**Figure 4.** Schematic sketch of the shift in the time of origin  $T_o$ .

October 26<sup>th</sup>). The background seismicity is added as a constant term to the contribution of the triggering events (see Section Method). The forecasted seismicity map in terms of the mean plus two standard deviation for the number of events with  $M \geq 3.0$  is shown in Fig. 3b. Observed events with  $M \geq 3.0$  (coloured dots) occurred within the corresponding 6-hour forecasting interval are also highlighted on the map. The main two events with Mw 5.4 at 17:10 UTC assigned as the mainshock and the Mw 5.9 event at 19:18 UTC (which lies within the 6-hour forecasting interval) are shown with magenta stars. According to the right-hand side error-bar plot of Fig. 3b, the total number of registered events within the 6-hour forecasting interval (red star) is significantly higher than the forecasted values. This can be attributed to very few number of observed input data in **seq** for performing the robust estimation and to the fact that model parameters were not tuned to the newly triggered sequence. Although less successful in predicting the number of events, the model predicts exceedance probabilities  $P(M \geq 5)$  and  $P(M \geq 6)$  to be more than 15 times compared to the initial estimates in Fig. 3a.

**4-hour forecasting from 20:00 UTC of 26/10/2016.** After the occurrence of the event with Mw 5.9 at 19:18 UTC, the seismicity forecast is provided again for the interval starting from 20:00 UTC (42 minutes after the Mw 5.9 event) up to 24:00 UTC of 26/10/2016 (i.e., a 4-hour time interval). The corresponding **seq** includes all the events with  $M \geq 3.0$  which occurred (including the main event of Mw 5.4 at 17:10 UTC and Mw 5.9 at 19:18 UTC) after 17:10 UTC up to the starting time (20:00 UTC) of October 26. The model parameters  $\theta$  are updated based on the information provided by the sequence **seq** with  $M_l$  set to 2.5 and reported in the second row of Table 2. Note that the cut-off magnitude lower than 3.0 is assigned only for model updating purposes to gain more data and the seismicity rate is later calculated with  $M_l = 3.0$ . Fig. 3c illustrates the forecasted seismicity map in terms of the mean plus two standard deviation for the number of events with  $M \geq 3.0$  within the 4-hour forecasting interval. Note that for the 4-hour time interval, the exceedance probabilities  $P(M \geq 5)$  and  $P(M \geq 6)$  increase considerably in Fig. 3c after the occurrence of the event with Mw 5.9 at 19:18 UTC (compared to Fig. 3a). In addition, the observed number of events within the 4-hour time interval (sub-figure) lies within the plus/minus one standard deviation confidence interval.

**6-hour forecasting from 24:00 UTC of October 26, and 24-hour forecasting from 6:00 UTC of 27<sup>th</sup> and 29<sup>th</sup> of October.** At this stage, seismicity forecasts are provided first for the 6-hour time interval starting from 24:00 UTC of 26/10/2016, and later extended for issuing 24-hour forecasts with  $T_{start} = 6:00$ UTC corresponding to the two consecutive days of October 27 and October 29. These predictions are issued based on the sequence of events **seq** including the mainshock of Mw 5.4 (considering 17:10 UTC 26<sup>th</sup> of October as time of origin  $T_o$ ) and the triggered events with  $M \geq 3.0$  up to the associated  $T_{start}$ . The last three rows of Table 2 illustrate the statistics of updated model parameters. The predicted seismicity distribution maps and the corresponding error-bar plots for the number of events (in sub-plots) are presented in Fig. 3(d–f), which manage to properly forecast the distribution/number of observed events.

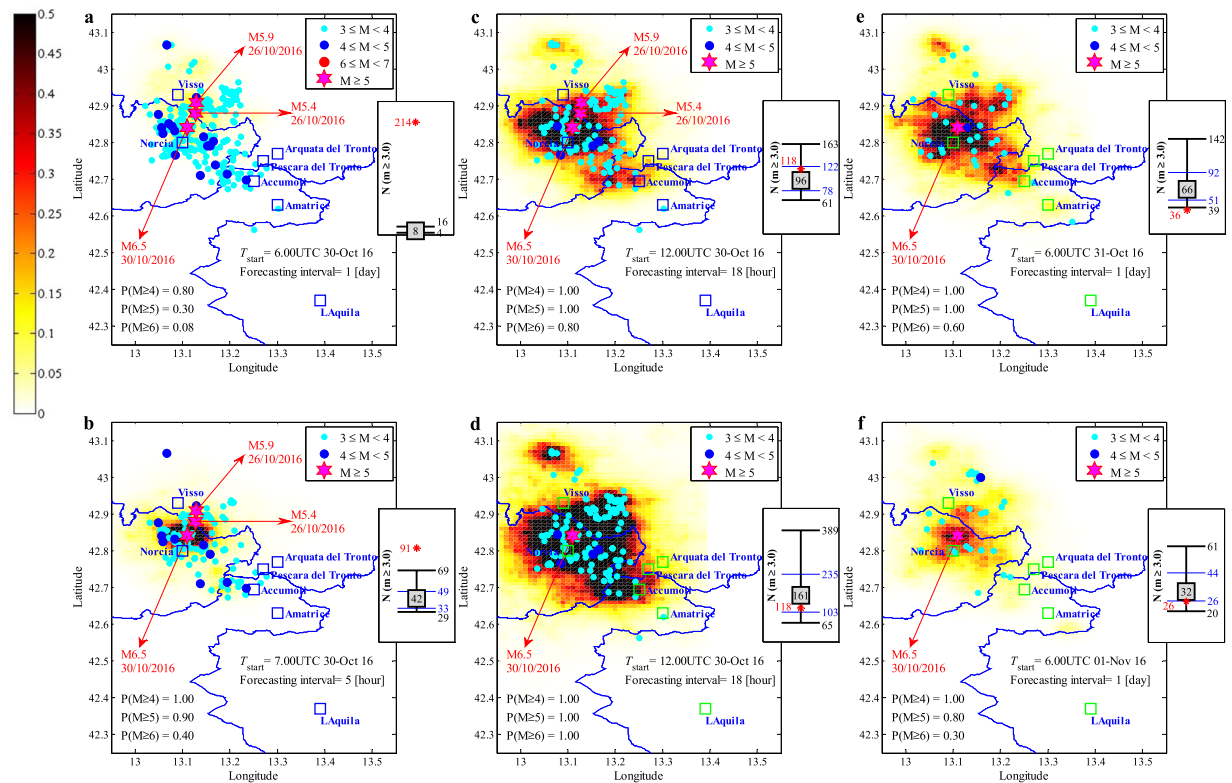
**Providing daily forecasts of seismicity from October 30 up to November 1.** As mentioned before, on 30<sup>th</sup> of October, a Mw 6.5 event occurred in the North of Norcia at 6:40 UTC (located in the north-western part of the aftershock zone, see Fig. 1b).

**24-hour forecasting from 6:00 UTC of 30<sup>th</sup> of October (starting time 40 minutes before the main event).** The forecasted seismicity is issued for a 24-hour time interval starting at 6:00UTC of 30/10/2016 based on the sequence of events **seq** comprising of events occurring after the time of origin  $T_o$  set to 17:10 UTC of 26<sup>th</sup> of October (time of occurrence of the Mw5.4 event). The statistics of the updated model parameters are shown in the first row of Table 3. The forecasted seismicity distribution for the next 24 hours for  $M \geq 3$  is mapped in Fig. 5a. The background seismicity is set to the time-invariant distributed seismicity equal to  $N_b(x, y, m|M_l)$  as defined in the previous section (shown in Fig. 3a). According to Fig. 5a, our blind prediction of the exceedance probability  $P(M \geq 6)$  is a value around 10% which is quite high as compared to the daily probability equal to  $3.83 \times 10^{-4}$  for  $M \geq 4.76$  calculated based on long-term seismicity in the previous section. It is also interesting to note that the forecasted value for  $P(M \geq 6)$  is equal to the forecast provided for October 29; this can be viewed as somewhat alarming (does not represent the expected decay with time). Nevertheless, the forecasted error-bar for the number of events (see right-hand side of Fig. 5a) is not able to properly predict the huge number of events triggered due to Mw 6.5 at 06:40UTC.

**5-hour forecasting from 7:00 UTC of 30<sup>th</sup> of October (starting time 20 minutes after the main event of Mw 6.5).** In order to test the forecasting ability of the model right after the main event, we provided a forecast with starting

	$\beta$		$c$ [day]		$p$		$d$ [km]		$q$		$K$		$K_I$		$K_R$	
	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV	mean	COV
30/10/2016 <sup>*</sup>	1.52	0.06	0.03	0.27	1.22	0.08	1.67	0.13	1.76	0.06	0.70	0.22	0.09	0.16	0.58	0.40
30/10/2016 <sup>†</sup>	1.42	0.06	0.03	0.26	1.17	0.07	1.64	0.12	1.70	0.06	1.16	0.26	0.08	0.30	0.48	0.36
30/10/2016 <sup>‡</sup>	1.39	0.04	0.03	0.26	1.25	0.08	1.94	0.11	1.81	0.06	0.71	0.23	0.10	0.11	0.81	0.36
30/10/2016 <sup>††</sup>	1.29	0.06	0.03	0.23	1.24	0.14	1.61	0.13	1.63	0.06	3.09	0.39	0.08	0.38	0.38	0.35
31/10/2016	1.44	0.03	0.04	0.21	1.31	0.10	1.61	0.11	1.71	0.04	0.69	0.25	0.10	0.11	0.47	0.29
01/11/2016	1.50	0.03	0.04	0.20	1.33	0.08	1.73	0.09	1.73	0.04	0.54	0.20	0.11	0.08	0.54	0.27

**Table 3.** Statistics (mean and COV) of ETAS model parameters  $\theta$  for estimating the daily seismicity rate of the desired dates for the third part of the sequence after Mw 6.5 at 30-October 2016 (Supplementary Fig. S3 also illustrates the sampled histograms representing the marginal posterior PDF's corresponding to the six model parameters [ $\beta$ ,  $K$ ,  $c$ ,  $p$ ,  $d$ ,  $q$ ]). \*Corresponding to Fig. 5a; †corresponding to Fig. 5b; ‡corresponding to Fig. 5c; ††corresponding to Fig. 5d.



**Figure 5.** Forecasted vs. observed seismicity distribution in the aftershock zone, the maps report the mean + 2 standard deviation confidence interval for the number of events per [km<sup>2</sup>] (latitude/longitude cells of a 0.01° × 0.01° grid) equal to or greater than magnitude  $M_I = 3$  in the indicated forecasting time window. In the lower left corner, the daily probabilities of having earthquakes with magnitudes 4, 5, and 6 or larger are reported. Each sub-plot also features the earthquakes (colored dots) that occurred during the corresponding forecasting time window together with the main events with  $M \geq 5.0$  (magenta stars). The sub-figures illustrate the observed (plotted in red star) vs. the error-bar for the forecasted number of events with  $m \geq M_I$  corresponding to the forecasting time interval: the median value (the 50<sup>th</sup> percentile, equivalent of the logarithmic mean in the arithmetic scale) inside a grey-filled square, the (logarithmic) mean plus/minus one (logarithmic) standard deviation indicating the interval between 16<sup>th</sup> and 84<sup>th</sup> percentiles (marked with blue horizontal lines), and the (logarithmic) mean plus/minus two (logarithmic) standard deviations indicating the interval between 2<sup>nd</sup> and 98<sup>th</sup> percentiles (marked with black horizontal lines). (MATLAB 2016b, <http://software.unina.it/matlab/> is used to create this figure).

time  $T_{start}$  set to 7:00UTC (only 20 minutes after the main event). The same as the previous forecast, the time of origin is set to 17.10 UTC of 26<sup>th</sup> of October and the same invariant background seismicity  $N_0(x, y, m|M_I)$  is being adopted. The second row of Table 3 shows the statistics of the updated model parameters. The forecasted (5-hour) seismicity distribution is mapped in Fig. 5b. Moreover, the error-bar plot for the forecasted number of events is



reported against the observed number of events. Comparing Fig. 5b with Fig. 5a, we can appreciate the improvement in the forecasted seismicity distribution and number of events based on the few events that occurred in the one hour that separates the two starting times (i.e., 6:00 UTC and 7:00 UTC, respectively).

*18-hour forecasting from 12:00 UTC of 30<sup>th</sup> of October (starting time 5 hours and 20 minutes after the main event of Mw 6.5,  $T_o = 6:40$  UTC of 30<sup>th</sup> of October).* The next forecasting is performed for the same day of 30/10/2016 with  $T_{start}$  set to 12:00UTC and  $T_{end}$  set to 06:00 UTC of 31/10/2016 (i.e., 18-hour interval). At this stage, we performed a shift in the time of origin  $T_o$  from 17:10 UTC of 26<sup>th</sup> of October to 6:40 UTC of 30<sup>th</sup> of October. The background seismicity  $N_b(x, y, m|M_i)$  is set (and proportioned to an 18-hour interval) to that of 30<sup>th</sup> of October for a 24-hour interval with starting time set to 6:00 UTC (shown in Fig. 5a). The prior distributions for the model parameters are taken equal to those of the original priors reported in the first row of Table 1. The fourth row of Table 3 shows the statistics of the updated model parameters. Fig. 5d shows the map of forecasted seismicity with the back-drop of events occurred in this interval. The error-bar plot for the forecasted number of events manages to capture the observed number of events within one standard deviation confidence interval.

*18-hour forecasting from 12:00 UTC of 30<sup>th</sup> of October (starting time 5 hours and 20 minutes after the main event of Mw 6.5,  $T_o = 17:10$  UTC of 26<sup>th</sup> of October).* To measure the effect of the shift in the time of origin, the same forecasting presented in the previous step is performed with time of origin set to 17:40 UTC of 26<sup>th</sup> of October. The third row of Table 3 illustrates the statistics of updated model parameters. Fig. 5c shows the forecasted map of seismicity and the error-bar for the predicted number of events. The forecasted number of events are slightly lower than those predicted in the previous step in Fig. 5c (after shifting the time of origin). This is to be expected since the latter forecast employs a time-invariant background seismicity to consider the events of interest occurred in the time interval between 17:10 UTC of 26<sup>th</sup> of October and 6.40 UTC of 30<sup>th</sup> of October. This is while the former forecast explicitly considers the triggering contribution of these events and the associated time-decay. Overall, it is reassuring to note that the two forecasts provide essentially the same information.

*24-hour forecasting from 6:00 UTC of October 31<sup>st</sup> and November 1<sup>st</sup>.* The seismicity forecasts are provided next for the 24-hour time intervals starting from 6:00UTC of October 31 and November 1. These predictions are issued based on time of origin  $T_o$  set to 6:40 UTC of 30 October and the sequence of records **seq** including all the events of interest  $M \geq 3.0$  that occurred between the time of origin  $T_o$  (including the 6.5 Mw “mainshock”) and the time of start  $T_{start}$  of each forecasting interval (6:00 UTC of each of these two days). The prior probability distributions are set to the original priors (whose statistics are reported in the first row of Table 1). The last two rows of Table 3 illustrate the statistics of updated model parameters. The seismicity forecasts (i.e, mean estimate plus two standard deviations) are shown in Fig. 5(e,f). The observed number of events lies within two standard deviation from the forecasted mean estimate.

## Discussion

We have proposed a fully simulation-based procedure for both Bayesian model updating of an epidemiological-type aftershock spatio-temporal clustering model and robust operational forecasting of the number of events of interest expected to happen in each time frame. The adopted epidemiological model parameters are defined so that the model converges asymptotically (spatio-temporally speaking) to the long-term seismicity of the zone of interest. The forecasting is “robust” because it considers the uncertainty (i.e., the joint probability distribution) in the model parameters. Apart from being quite efficient (the most challenging forecasting we performed took 45 minutes on a normal PC), the model updating and forecasting procedure is carried on without human interference and use of expert judgement. The model is simply “tuning-in” automatically into the sequence of observed events. The choice of the recent Central Italy sequence of events as a demonstration of this procedure, albeit quite a natural choice, proved to be very challenging. This is because the sequence embedded three “sub-sequences” with different productive and decaying properties. We used the peculiarities of this sequence to test several different strategies for forecasting. For example, we performed a shift in the time of origin of the sequence by conservatively introducing a constant background seismicity (calculated by the procedure). This shift proved to be quite useful as it relieved us from the burden of summing up the triggering properties of all the events that took place in the previous “sub-sequence” (or the previous part of the sequence as we may wish to call it) at the small price of neglecting the time-decay in their triggering contribution. We observe that after an initial transition time (in the order of few hours, enough to accumulate sufficient events for updating the model parameters), the model quickly tunes into the sequence and provides forecasting that is reliable in most cases up to plus/minus one standard deviations. As expected, the procedure falls short of predicting the “mainshock” of 26<sup>th</sup> of October (17:10 UTC) as it happened when the sequence had decayed. The procedure, however, did a better job for forecasting the events occurred at 19:18 UTC of 26<sup>th</sup> of October and on 6.40 UTC of 30<sup>th</sup> October. This relative success can be attributed to the fact that these events took place at the initial stages of the newly triggered sequence of 26<sup>th</sup> of October when the seismic activity was still very high. The estimated model parameters present some time-dependent fluctuations but after a certain number of days elapsed after the main event, they seem to stabilize. In general, the first sub-sequence (Mw 6 “mainshock” occurred at 1:36 UTC 24<sup>th</sup> of August) seems to be the mildest one in terms of the time decay in seismicity and is the least active in terms of sequence’s productivity. The second sub-sequence (Mw 5.4 “mainshock” occurred at 17:10 UTC 26<sup>th</sup> of October) is intermediate both in terms of the rate of time-decay and the productivity. The third sub-sequence (Mw 6.5 “mainshock” occurred at 6:40 UTC 30<sup>th</sup> of October) has the steepest time-decay of seismicity and is the most active in terms of the productivity of the sequence. Last but not least, it is important to mention that the proposed procedure for robust forecasting is conditioned on the available catalogue of events and the epidemiological model adopted for capturing the spatio-temporal aftershock clustering.

### Methods

Let the aftershock occurrence be described by a non-homogenous Poisson point process over the two-dimensional space and time. Hence, the aftershock zone can be described as the set  $A$  in the Cartesian space discretized into mutually exclusive and collectively exhaustive (MECE) subsets or spatial cell units centered at  $(x, y) \in A$ . In this manner,  $\lambda(t, x, y, m|\mathbf{seq}, M_i)$  represents the rate of occurrence of events in the forecasting interval  $[T_{start}, T_{end}]$  at time  $t$  elapsed after the main event (a.k.a. mainshock) occurred at time of origin  $T_0$  with magnitude greater than or equal to  $m$  and in the cell unit centered at  $(x, y) \in A$ , given (a) the observation history  $\mathbf{seq}$  which is the sequence of  $N_0$  events (including mainshock and the sequence of aftershocks) taken place before the forecasting interval (i.e. in the interval  $[T_0, T_{start}]$ ), and (b) the lower cut-off magnitude  $M_i$ . Hence,  $\mathbf{seq}$  can be expressed as  $\mathbf{seq} = \{(t_i, x_i, y_i, m_i), t_i < T_{start}, m_i \geq M_i, i = 1: N_0\}$ , where  $t_i$  is the arrival time for the  $i$ th event with magnitude  $M_i$  and location  $(x_i, y_i) \in A$ . The average number of events in the spatial cell unit centered at  $(x, y)$  with magnitude greater than or equal to  $m$  in the forecasting interval  $[T_{start}, T_{end}]$  can then be calculated as:

$$N(x, y, m|\mathbf{seq}, M_i) = N_b(x, y, m|M_i) + \int_{T_{start}}^{T_{end}} \lambda(t, x, y, m|\mathbf{seq}, M_i) dt \tag{1}$$

where  $N_b(x, y, m|M_i)$  is a constant representing the background seismicity of the area. Let  $\theta$  denote the vector of model parameters for  $\lambda(t, x, y, m|\mathbf{seq}, M_i)$ . Given a particular space-time model and a realization of the vector of model parameters  $\theta$ , one can calculate a plausible value for the rate of occurrence denoted as  $\lambda(t, x, y, m|\theta, \mathbf{seq}, M_i)$ . Note that we have not included the conditioning on the model assumptions<sup>34</sup> for the sake of brevity. A robust estimate<sup>24, 26, 27, 35, 36</sup> of the average number of events in the spatial cell unit centered at  $(x, y)$  with magnitude greater than or equal to  $m$  in the forecasting interval  $[T_{start}, T_{end}]$ , and over the domain of the model parameters  $\Omega_\theta$  can be calculated as:

$$\mathbb{E}[N(x, y, m|\mathbf{seq}, M_i)] = N_b(x, y, m|M_i) + \int_{\Omega_\theta} \int_{T_{start}}^{T_{end}} \lambda(t, x, y, m|\theta, \mathbf{seq}, M_i) \cdot p(\theta|\mathbf{seq}, M_i) d\theta dt \tag{2}$$

where  $p(\theta|\mathbf{seq}, M_i)$  is the conditional probability distribution function (PDF) for  $\theta$  given the  $\mathbf{seq}$  and the lower cut-off magnitude  $M_i$ .

**An epidemiological model for space-time clustering of aftershocks.** The ETAS model is an epidemiological stochastic point process in which every earthquake is a potential triggering event for subsequent earthquakes<sup>8, 9, 16-19</sup>. According to the general ETAS model, we adopt the spatio-temporal triggering effect of a given sequence on the seismicity rate, denoted as  $\lambda_{ETAS}$ , as:

$$\lambda_{ETAS}(t, x, y, m|\theta, \mathbf{seq}_t, M_i) = e^{-\beta(m-M_i)} \sum_{t_j < t} K e^{\beta(M_j-M_i)} \cdot \frac{K_t}{(t - t_j + c)^p} \cdot \frac{K_R}{(r_j^2 + d^2)^q} \tag{3}$$

where  $\mathbf{seq}_t = \{(t_j, x_j, y_j, M_j), t_j < t, M_j \geq M_i\}$  is the observation history up to the time  $t$ ; parameter  $\beta$  is related to Gutenberg-Richter seismicity; parameters  $c$  and  $p$  are similar to those of the Modified Omori's Law<sup>12, 13</sup> defining the decay in time of short-term triggering effect;  $d$  and  $q$  characterize the spatial distribution of the triggered events;  $r_j$  is the distance between the location  $(x, y) \in A$  and the epicenter of the  $j$ th event  $(x_j, y_j)$ ; parameters  $K, K_t$  and  $K_R$  satisfy the achievement of asymptotic compatibility between ETAS predictions and the long-term seismicity. Thus, the vector of model parameters  $\lambda$  can be defined as  $\theta = [\beta, K, K_t, K_R, c, p, d, q]$ . The integral of  $\lambda_{ETAS}$  over infinite space and time needs to converge in limit to the number of events predicted by the Gutenberg and Richter model with magnitude greater than  $m$ , denoted generally as  $K_0 e^{-\beta m}$ ; note that  $K_0$  is different from  $K$ . Herein, such compatibility is achieved by making sure that the following three conditions are satisfied:

- 1) The normalizing coefficient  $K_t$  is obtained such that integrating the time-dependent term over infinite time will in limit be equal to unity (see also Lippiello *et al.*<sup>21, 37</sup>):

$$K_t \int_{t_j}^{+\infty} \frac{dt}{(t - t_j + c)^p} = 1 \therefore p > 1, K_t = (p - 1)c^{p-1} \tag{4}$$

- 2) The coefficient  $K_R$  is normalized such that integrating the spatial term over infinite space will in limit be equal to unity (see also Lippiello *et al.*<sup>21, 37</sup>):

$$\int_{x_j}^{+\infty} \int_{y_j}^{+\infty} \frac{K_R dx dy}{((x - x_j)^2 + (y - y_j)^2 + d^2)^q} \triangleq K_R \int_0^{+\infty} \frac{2\pi r dr}{[r^2 + d^2]^q} = 1 \therefore q > 1, K_R = \frac{(q - 1)}{\pi} d^{2(q-1)} \tag{5}$$

- 3) The coefficient  $K$  is calibrated such that the number of events with magnitude greater than or equal to  $M_i$  taking place in time interval  $t \in [T_0, T_{start}]$  over the whole aftershock zone  $A$  is equal to  $N_0$ :

$$\int_{T_0}^{T_{start}} \iint_{(x,y) \in A} \lambda_{ETAS}(t, x, y, M_i|\theta, \mathbf{seq}_t, M_i) dx dy dt = N_0 \tag{6}$$

where  $T_0$  defines the mainshock occurrence time; the term  $\lambda_{ETAS}(t, x, y, M_i|\theta, \mathbf{seq}_t, M_i)$  is obtained by substituting  $m = M_i$  in equation (3), and denotes the rate of events with magnitude greater than or equal to  $M_i$ . Since  $N_0$

denotes the events occurred within the aftershock zone  $A$ , the integration over the whole aftershock zone can be approximated with that over infinite space. Thus, according to equation (5),  $K$  can be derived as follows:

$$\begin{aligned}
 K &= \frac{N_o}{\int_{T_o}^{T_{start}} \left( \sum_{t_j < t} \frac{K e^{\beta(M_j - M_i)}}{(t - t_j + c)^p} \right) dt} \\
 &= \frac{N_o / K_t}{\sum_{i=2}^{N_o} \sum_{t_j < t_i} e^{\beta(M_j - M_i)} I_0(t_i, t_{i-1}, t_j) + \sum_{t_j < T_{start}} e^{\beta(M_j - M_i)} I_0(T_{start}, t_{N_o}, t_j)} \\
 I_0(t_i, t_{i-1}, t_j) &= \int_{t_{i-1}}^{t_i} \frac{dt}{(t - t_j + c)^p} \\
 &= \begin{cases} \left[ (t_i - t_j + c)^{1-p} - (t_{i-1} - t_j + c)^{1-p} \right] / (1 - p) & p \neq 1 \\ \ln \left[ (t_i - t_j + c) / (t_{i-1} - t_j + c) \right] & p = 1 \end{cases} \tag{7}
 \end{aligned}$$

Since the integral with respect to time in equation (7) cannot be calculated analytically over the interval  $[T_o, T_{start}]$ , we approximated it by summing over the sub-intervals  $[t_{i-1}, t_i]$  (where  $i = 2:N_o$ ) and the last interval  $[t_{N_o}, T_{start}]$  (where  $t_{N_o}$  is the arrival time of the  $N_o$ th event).

It is to note that parameters  $K$ ,  $K_p$ , and  $K_R$  are derived as a function of other model parameters in  $\theta$ ; therefore, the main parameters of the ETAS model include  $\{\beta, c, p, d, q\}$ . The rate of events in the ETAS model with magnitude (exactly) equal to  $m$ , denoted herein as  $\mu_{ETAS}$  herein, is calculated by taking the derivative of equation (3) with respect to magnitude  $m$ :

$$\mu_{ETAS}(t, x, y, m | \theta, \mathbf{seq}_t, M_i) = |\partial \lambda_{ETAS} / \partial m| = \beta e^{-\beta(m - M_i)} \lambda_{ETAS}(t, x, y, M_i | \theta, \mathbf{seq}_t, M_i) \tag{8}$$

**Robust estimation for the number of aftershock events.** As mentioned above,  $\mathbf{seq}$  denotes the sequence of events taking place before the beginning of the forecasting interval (i.e., in the interval  $[T_o, T_{start}]$ ). However, the triggering effect of the events taking place during the forecasting interval  $[T_{start}, T_{end}]$  is expected to play a major role. The sequence of events taking place during the forecasting interval denoted as  $\mathbf{seqg}$ , which is unknown at the time of forecasts, is simulated/generated herein. Let us assume that a plausible  $\mathbf{seqg}$  is defined as the events within the forecasting interval defined as  $\mathbf{seqg} = \{(LAT_i, x_i, y_i, m_i), T_{start} \leq t_i \leq T_{end}, m_i \geq M_i\}$ , where  $IAT_i = t_i - t_{i-1}$  stands for the inter-arrival time. The robust estimate for the number of aftershock events in equation (2) should also consider all the plausible sequences of events  $\mathbf{seqg}$  (i.e., the domain  $\Omega_{\mathbf{seqg}}$ ) that can happen during the forecasting time interval:

$$\begin{aligned}
 N(x, y, m | \mathbf{seqg}, M_i) &= N_b(x, y, m | M_i) \\
 &+ \int_{\Omega_\theta} \left[ \int_{\Omega_{\mathbf{seqg}}} \left( \int_{T_{start}}^{T_{end}} \lambda(t, x, y, m | \mathbf{seqg}, \theta, \mathbf{seq}, M_i) dt \right) \right. \\
 &\left. \times p(\mathbf{seqg} | \theta, \mathbf{seq}, M_i) d\mathbf{seqg} \right] p(\theta | \mathbf{seq}, M_i) d\theta \tag{9}
 \end{aligned}$$

where  $p(\mathbf{seqg} | \theta, \mathbf{seq}, M_i)$  is the PDF for the generated sequence  $\mathbf{seqg}$  given that  $\theta$  and  $\mathbf{seq}$  are known and  $\lambda(t, x, y, m | \mathbf{seqg}, \theta, \mathbf{seq}, M_i)$  is the space-time clustering ETAS model considering also the sequence of events taking place within the forecasting interval. The integral with respect to time in equation (9) cannot be calculated analytically over the entire interval  $[T_{start}, T_{end}]$ , and is approximated by summing over the sub-intervals  $[t_{i-1}, t_i]$  within  $\mathbf{seqg}$ :

$$\begin{aligned}
 \int_{T_{start}}^{T_{end}} \lambda(t, x, y, m | \mathbf{seqg}, \theta, \mathbf{seq}, M_i) dt &= \sum_{\Omega_{\mathbf{seqg}}} \int_{t_{i-1}}^{t_i} \lambda_{ETAS}(t, x, y, m \\
 &| \mathbf{seqg}_{i-1}, \theta, \mathbf{seq}, M_i) dt \\
 &= \sum_{\Omega_{\mathbf{seqg}}} \left[ K K_t K_R e^{-\beta(m - M_i)} \right. \\
 &\left. \sum_{t_j < t_i} \frac{e^{\beta(M_j - M_i)} I_0(t_i, t_{i-1}, t_j)}{(r_j^2 + d^2)^q} \right] \tag{10}
 \end{aligned}$$

where  $\lambda_{ETAS}$  has the functional form presented in equation (3), and  $\mathbf{seqg}_{i-1}$  is the previous  $(i - 1)$  events within the generated sequence. In the following sections, it is described first how sequence of events  $\mathbf{seqg}$  for the forecasting interval is sampled based on  $p(\mathbf{seqg} | \theta, \mathbf{seq}, M_i)$ . Later on, the method for sampling  $\theta$  from the distribution  $p(\theta | \mathbf{seq}, M_i)$  is discussed.

Generating sequences according to  $p(\text{seqg}|\theta, \text{seq}, M_i)$ . The probability distribution  $p(\text{seqg}|\theta, \text{seq}, M_i)$  can be written as follows (based on the probability product rule, see e.g. Jaynes<sup>38</sup>):

$$p(\text{seqg}|\theta, \text{seq}, M_i) = \prod_i p(IAT_i, x_i, y_i, M_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) \tag{11}$$

where  $\text{seqg}_i$  is the generated sequence up to the  $i$ th event, where  $\text{seqg}_i = \{\text{seqg}_{i-1}, (IAT_i, x_i, y_i, m_i)\}$ , and the sequence of events that precede the  $i$ th generated event is  $\{\text{seq}, \text{seqg}_{i-1}\}$ . The probability distribution  $p(IAT_i, x_i, y_i, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i)$  can be further expanded (again using the probability product rule) as follows:

$$\begin{aligned} p(IAT_i, x_i, y_i, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) &= p(m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) \\ &= p(IAT_i|m_i, \text{seqg}_{i-1}, \theta, \text{seq}, M_i) \\ &= p(x_i, y_i|IAT_i, m_i, \text{seqg}_{i-1}, \theta, \text{seq}, M_i) \end{aligned} \tag{12}$$

where  $p(m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i)$  is the marginal PDF for the magnitude  $m_i$  given the sequence of events that precede it,  $\theta$ , and  $M_i$ ;  $p(IAT_i|m_i, \text{seqg}_{i-1}, \theta, \text{seq}, M_i)$  is the (conditional) marginal PDF for inter-arrival time  $IAT_i$  given that the value of magnitude is equal to  $m_i$ ; finally, the term  $p(x_i, y_i|IAT_i, m_i, \text{seqg}_{i-1}, \theta, \text{seq}, M_i)$  is the conditional joint PDF for the spatial position  $(x_i, y_i) \in A$  given that  $IAT_i$  and  $m_i$  are known. It should be noted that the break-down into the product of several conditional PDFs is necessary during the sequence generation process.

To generate a plausible sequence of events during the forecasting interval, the procedure, illustrated by the flowchart in Supplementary Fig. S4, is adopted. The  $i$ th event within the sequence  $\text{seqg}$  is generated through the following steps:

- 1) Generate the magnitude of the  $i$ th event,  $m_i$ , within  $\text{seqg}$  according to the following truncated Exponential PDF with rate  $\beta^{23}$  (see Phase 1 in Supplementary Fig. S4).

$$p(m|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) \cong p(m|\theta, \text{seq}, M_i) = \frac{\beta e^{-\beta m}}{e^{-\beta M_i} - e^{-\beta M_{\max}}} \tag{13}$$

where  $M_{\max}$  is the upper-bound magnitude of the site. This truncated Exponential PDF has a cumulative density function equal to  $\frac{1 - \exp(-\beta(m - M_i))}{1 - \exp(-\beta(M_{\max} - M_i))}$ ; hence, a generated/sampled  $m_i$  can

analytically be drawn as  $m_i = \frac{F_{M_{\max}} - F_{M_{\max}} \cdot r_{\text{rand}}}{\beta} + M_i$  where  $r_{\text{rand}}$  (as shown also in the flowchart of Supplementary Fig. S4) is a random number generated from a Uniform distribution in the interval (0, 1).

- 2) Generate the inter-arrival time of the  $i$ th event within  $\text{seqg}$  (given that its magnitude  $m_i$  is already known from previous step) by using the Thinning algorithm<sup>39,40</sup>. The algorithm, as shown in Phase 2 in Supplementary Fig. S4, consists of a two-stage process. First, the temporal Poisson rate over the entire aftershock zone is calculated for an event with magnitude equal to  $m_i$  at time  $t = t_i - 1$ , and denoted by  $\mu_{\max}$  using equation (8) as follows:

$$\mu_{\max} = \iint_{x,y \in A} \mu_{\text{ETAS}}(t_{i-1}, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) dx dy \tag{14}$$

Note that for generating the first event within the forecasting interval  $[T_{\text{start}}, T_{\text{end}}]$ ,  $t_{i-1}$  is set to  $T_{\text{start}}$ . In the next stage of Thinning algorithm, the inter-arrival time  $IAT_{\text{gen}}$  is sampled/generated from a homogeneous Exponential PDF with the form  $\mu_{\max} \exp(-\mu_{\max} IAT)$  which is equivalent to generating it as  $IAT_{\text{gen}} = \frac{-\log(1 - r_{\text{rand}})}{\mu_{\max}}$ . The Poisson rate at time  $t_{\text{gen}} = t_{i-1} + IAT_{\text{gen}}$  and denoted by  $\mu_{\text{gen}}$  can then be calculated from equation (14) by substituting  $t_{i-1}$  with  $t_{\text{gen}}$ . The generated inter-arrival time can then be (1) either accepted with probability  $p = \frac{\mu_{\text{gen}}}{\mu_{\max}}$  and thus  $t_i = t_{\text{gen}}$  and the procedures continues by generating the next inter-arrival time until  $t_i > T_{\text{end}}$  (2) or rejected with probability  $1-p$ . Let us denote the rejected  $t_{\text{gen}}$  as  $t_{\text{gen}}^{(-)}$  (in order to keep track of it for the next simulation). Also in the case of rejection, the procedure continues by sampling a new inter-arrival time  $IAT_{\text{gen}}$  from the homogeneous Exponential PDF with rate  $\mu_{\max}$ . The new generated arrival time is calculated as  $t_{\text{gen}} = t_{\text{gen}}^{(-)} + IAT_{\text{gen}}$ ,  $t_{\text{gen}} < T_{\text{end}}$ . The quantities  $\mu_{\text{gen}}$  and  $p$  are calculated again to test whether the newly generated interarrival time is accepted or rejected.

- 3) Generate/sample the Cartesian coordinates  $(x_i, y_i)$  for the  $i$ th event, given that the magnitude  $m_i$ , the time of occurrence  $t_i$ , and the previous  $(i - 1)$  events within the generated sequence  $\text{seqg}_{i-1}$  are known (see Supplementary Fig. S4, Phase 3), according the following joint PDF using the probability product rule:

$$p(x, y|IAT_i, m_i, \text{seqg}_{i-1}, \theta, \text{seq}, M_i) = \frac{p(IAT_i, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i)}{\iint_{x,y \in A} p(IAT_i, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) dx dy} \tag{15}$$

The nominator  $p(IAT_i, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i)$  in equation (15) can be calculated as:

$$\begin{aligned} p(IAT_i, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) &= \mu_{\text{ETAS}}(t_i, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) \\ &\times e^{-\int_{t_{i-1}}^{t_i} \mu_{\text{ETAS}}(t, x, y, m_i|\text{seqg}_{i-1}, \theta, \text{seq}, M_i) dt} \end{aligned} \tag{16}$$

where  $\mu_{ETAS}$  is calculated from equation (8), and the integral in the sub-interval  $[t_{i-1}, t_i]$  has the similar analytical expression shown in equation (10) being multiplied by  $\beta$  in order to account for  $\mu_{ETAS}$ .

*Sampling  $\theta$  from the distribution  $p(\theta|\mathbf{seq}, M_l)$ .* The probability distribution  $p(\theta|\mathbf{seq}, M_l)$  can be calculated using Bayesian parameter estimation:

$$p(\theta|\mathbf{seq}, M_l) = C^{-1}p(\mathbf{seq}|\theta, M_l)p(\theta|M_l) \tag{17}$$

where  $p(\mathbf{seq}|\theta, M_l)$  denotes the likelihood of the observed sequence given the vector of model parameters  $\theta$  and lower cut-off magnitude  $M_l$ ,  $p(\theta|M_l)$  is the prior distribution for the vector  $\theta$ , and  $C^{-1}$  is a normalizing constant. In lieu of additional information (e.g., statistics of regional model parameters), the prior joint distribution  $p(\theta|M_l)$  can be estimated as the product of marginal uniform probability distributions for each model parameter. In order to sample from  $p(\theta|\mathbf{seq}, M_l)$ , Markov Chain Monte Carlo (MCMC) simulation routine is employed which is particularly useful for cases where the sampling needs to be done from a probability distribution that is known up to a constant value<sup>27</sup> (herein,  $C^{-1}$ ). The MCMC routine uses the Metropolis-Hastings (MH) algorithm<sup>41, 42</sup> in order to generate samples as a Markov Chain sequence used first to sample from the target probability distribution  $p(\theta|\mathbf{seq}, M_l)$ , and later to estimate the robust reliability in equation (9). The MH routine, as shown in Supplementary Fig. S5, functions by generating a Markov chain that produces a sequence of samples  $[\theta_1 \rightarrow \theta_2 \rightarrow \dots \rightarrow \theta_n \rightarrow \dots]$ , where  $\theta_n$  represents the state of Markov chain at  $n$ th iteration (the first few samples are often discarded to reduce the initial transient effect). This procedure continues until the  $n_s$  Markov chain samples were simulated. It can be shown that the samples from the chain after the initial transient ones reflect samples from the target distribution  $p(\theta|\mathbf{seq}, M_l)$ . Given the Markovian nature of this simulation scheme, we limit ourselves to describe how the  $(n + 1)$ th sample  $\theta_{n+1}$  is generated given that the  $n$ th sample  $\theta_n$  is already known:

- Generate a candidate sample  $\theta^*$  from a proposal (candidate) distribution  $\zeta(\theta|\theta_n)$ . It is important to note that there are no specific restrictions about the choice of  $\zeta(\cdot)$  apart from the fact that it should be possible to calculate both  $\zeta(\theta_{i+1}|\theta_i)$  and  $\zeta(\theta_i|\theta_{i+1})$ .
- Accept the candidate sample with the probability  $p_{\text{accept}} = \min(1, r)$  (where  $r$  is defined in equation (18) as follows) and set  $\theta_{n+1} = \theta^*$ ; otherwise,  $\theta_{n+1} = \theta_n$ :

$$r = \frac{p(\theta^*|\mathbf{seq}, M_l)}{p(\theta_n|\mathbf{seq}, M_l)} \cdot \frac{\zeta(\theta_n|\theta^*)}{\zeta(\theta^*|\theta_n)} = \left( \frac{p(\mathbf{seq}|\theta^*, M_l)p(\theta^*|M_l)}{p(\mathbf{seq}|\theta_n, M_l)p(\theta_n|M_l)} \right) \cdot \frac{\zeta(\theta_n|\theta^*)}{\zeta(\theta^*|\theta_n)} \tag{18}$$

It can be shown<sup>27</sup> using the Total probability theorem that, if the current sample  $\theta_n$  is distributed as  $p(\cdot|\mathbf{seq}, M_l)$ , also the  $(n + 1)$ th sample  $\theta_{n+1}$  is distributed as  $p(\cdot|\mathbf{seq}, M_l)$ :

$$\begin{aligned} PDF(\theta_{n+1}|\mathbf{seq}, M_l) &= \int_{\Omega_{\theta_n}} p(\theta_{n+1}|\theta_n)p(\theta_n|\mathbf{seq}, M_l)d\theta_n \\ &= \int_{\Omega_{\theta_n}} \min(1, r)\zeta(\theta_{n+1}|\theta_n)p(\theta_n|\mathbf{seq}, M_l)d\theta_n \\ &= \int_{\Omega_{\theta_n}} \min(1, r^{-1})\zeta(\theta_n|\theta_{n+1})p(\theta_{n+1}|\mathbf{seq}, M_l)d\theta_n \\ &= p(\theta_{n+1}|\mathbf{seq}, M_l) \int_{\Omega_{\theta_n}} p(\theta_n|\theta_{n+1})d\theta_n = p(\theta_{n+1}|\mathbf{seq}, M_l) \end{aligned} \tag{19}$$

Equation (19) is derived using the arithmetic property that for any positive numbers  $A$  and  $B$ , the identity  $\min(1, A/B) \cdot B = \min(1, B/A) \cdot A$  holds. Having the proposal PDF  $\zeta$  centered around the current sample renders the MH algorithm similar to a local random walk that adaptively leads to the generation of the target PDF. In order to improve the rate of convergence of the simulation process, we have used an adaptive MH algorithm (as proposed by Beck and Au<sup>27</sup>) that introduces a sequence of intermediate evolutionary candidate PDF's that resemble more and more the target PDF.

*Calculating the likelihood of the observed sequence  $p(\mathbf{seq}|\theta, M_l)$ .* The likelihood for the observed sequence,  $\mathbf{seq} = \{(t_i, x_i, y_i, m_i), t_i < T_{start}, m_i \geq M_l, i = 1: N_o\}$ , with  $N_o$  events, including the mainshock (with  $i = 1$ ) and the sequence of aftershocks, can be calculated as:

$$\begin{aligned} p(\mathbf{seq}|\theta, M_l) &= \left( \prod_{i=2}^{N_o} p(IAT_i, x_i, y_i, m_i|\theta, \mathbf{seq}_{i-1}, M_l) \right) P(IAT_{N_o+1} > T_{start} - t_{N_o}|\theta, \mathbf{seq}, M_l) \\ &= \prod_{i=2}^{N_o} \left( \mu_{ETAS}(t_i, x_i, y_i, m_i|\theta, \mathbf{seq}_{i-1}, M_l) e^{-\int_{t_{i-1}}^{t_i} \mu_{ETAS}(t, x_i, y_i, m_i|\theta, \mathbf{seq}_{i-1}, M_l) dt} \right) \\ &\quad \times e^{-\int_{t_{N_o}}^{T_{start}} \int_{x,y \in A} \lambda_{ETAS}(t, x, y, M_l|\theta, \mathbf{seq}) dx dy dt} \end{aligned}$$

where  $\lambda(\cdot)$  and  $\mu(\cdot)$  are ETAS rates calculated from equations (3) and (8), respectively. The first term of the likelihood function in equation (20) is calculated by multiplying the probabilities that the  $i$ th arrival time ( $i = 2:N_o$ ) is

equal to the observed value,  $t_i = t_{i-1} + IAT_i$ , occurred at spatial position  $(x_i, y_i) \in A$  with magnitude  $m_i$  assuming a non-homogenous Poisson process (i.e., the inter-arrival times are independent) with a rate equal to  $\mu_{ETAS}$  (equation 8). Accordingly

$$\int_{t_{i-1}}^{t_i} \mu_{ETAS}(t, x_i, y_i, m_i | \theta, \text{seq}_{i-1}, M_i) dt = KK_i K_R \beta e^{-\beta(m-M_i)} \sum_{t_j < t_i} \frac{e^{\beta(M_j - M_i)} I_0(t_i, t_{i-1}, t_j)}{(r_{ji}^2 + d^2)^q} \quad (20)$$

where the term  $I_0$  is previously defined in equation (7) and  $r_{ji}$  indicates the distance of the  $i$ th event to the previously occurred  $j$ th event. The last probability term in equation (20) is the probability that no event with magnitude greater than the cut-off level  $M_i$  takes place over the entire aftershock zone  $A$  in the time interval between the  $N_0^{\text{th}}$  event at  $t_{N_0}$  and  $T_{start}$ .

## References

- Gerstenberger, M. C., Wiemer, S., Jones, L. M. & Reasenberg, P. A. Real-time forecasts of tomorrow's earthquakes in California. *Nature* **435**, 328–331 (2005).
- Jordan, T. H. & Jones, L. M. Operational earthquake forecasting: Some thoughts on why and how. *Seism. Res. Lett.* **81**(4), 571–574 (2010).
- Jordan, T. H. *et al.* Operational earthquake forecasting: State of knowledge and guidelines for implementation. *Annals of Geophysics* **54**(4), 315–391, doi:10.4401/ag-5350 (2011).
- Jordan, T. H., Marzocchi, W., Michael, A. J. & Gerstenberger, M. C. Operational earthquake forecasting can enhance earthquake preparedness. *Seism. Res. Lett.* **85**(5), 955–959 (2014).
- Marzocchi, W., Lombardi, A. M. & Casarotti, E. The establishment of an operational earthquake forecasting system in Italy. *Seism. Res. Lett.* **85**(5), 961–969 (2014).
- Zechar, J. D., Marzocchi, W. & Wiemer, S. Operational earthquake forecasting in Europe: progress, despite challenges. *Bull. Earthq. Eng.* **14**, 2459–2469 (2016).
- Omi, T., Ogata, Y., Hirata, Y. & Aihara, K. Forecasting large aftershocks within one day after the main shock. *Sci. Rep.* **3**, 2218, doi:10.1038/srep02218 (2013).
- Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* **83**, 9–27 (1988).
- Ogata, Y. Space-time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50**(2), 379–402 (1998).
- Lombardi, A. M. Estimation of the parameters of ETAS models by Simulated Annealing. *Sci. Rep.* **5**, 8417, doi:10.1038/srep08417 (2015).
- Lombardi, A. M. & Marzocchi, W. The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment. *Annals of Geophysics* **53**, 155–164 (2010).
- Utsu, T. A statistical study of the occurrence of aftershocks. *Geophys. Mag.* **30**, 521–605 (1961).
- Utsu, T., Ogata, Y. & Matsu'ura, R. S. The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth.* **43**, 1–33 (1995).
- Zhang, X. & Shcherbakov, R. Power-law rheology controls aftershock triggering and decay. *Sci. Rep.* **6**, 36668, doi:10.1038/srep36668 (2016).
- Marzocchi, W. & Lombardi, A. M. Real-time forecasting following a damaging earthquake. *Geophys. Res. Lett.* **36**, L21302, doi:10.1029/2009GL040233 (2009).
- Zhuang, J., Ogata, Y. & Vere-Jones, D. Stochastic declustering of space-time earthquake occurrences. *J. Am. Stat. Assoc.* **97**(458), 369–380 (2002).
- Zhuang, J., Ogata, Y. & Vere-Jones, D. Analyzing earthquake clustering features by using stochastic reconstruction. *J. Geophys. Res.* **109**, B05301, doi:10.1029/2003JB002879 (2004).
- Ogata, Y. & Zhuang, J. Space-time ETAS models and an improved extension. *Tectonophysics* **413**(1), 13–23 (2006).
- Zhuang, J. Next-day earthquake forecasts for the Japan region generated by the ETAS model. *Earth planets space* **63**(3), 207–216 (2011).
- Veen, A. & Schoenberg, F. P. Estimation of space-time branching process models in seismology using an EM-type algorithm. *J. Am. Stat. Assoc.* **103**, 614–624 (2008).
- Lippiello, E., Giacco, F., de Arcangelis, L., Marzocchi, W. & Godano, C. Parameter Estimation in the ETAS Model: Approximations and Novel Methods. *Bull. Seism. Soc. Am.* **104**(2), 985–994 (2014).
- Jalayer, F., Asprone, D., Prota, A. & Manfredi, G. A decision support system for post-earthquake reliability assessment of structures subjected to aftershocks: an application to L'Aquila earthquake, 2009. *Bull. Earthq. Eng.* **9**(4), 997–1014 (2011).
- Ebrahimian, H. *et al.* Adaptive daily forecasting of seismic aftershock hazard. *Bull. Seism. Soc. Am.* **104**(1), 145–161 (2014).
- Jalayer, F. & Ebrahimian, H. MCMC-based Updating of an Epidemiological Temporal Aftershock Forecasting Model. *Vulnerability, Uncertainty, and Risk*, 2093–2103, doi:10.1061/9780784413609.210 (2014).
- Omi, T., Ogata, Y., Hirata, Y. & Aihara, K. Intermediate-term forecasting of aftershocks from an early aftershock sequence: Bayesian and ensemble forecasting approaches. *J. Geophysical Res: Solid Earth* **120**(4), 2561–2578, doi:10.1002/2014JB011456 (2015).
- Papadimitriou, C., Beck, J. L. & Katafygiotis, L. S. Updating robust reliability using structural test data. *Probabilist. Eng. Mech.* **16**(2), 103–113 (2001).
- Beck, J. L. & Au, S. K. Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation. *J. Eng. Mech. ASCE* **128**(4), 380–391 (2002).
- Ebrahimian, H. *et al.* A performance-based framework for adaptive seismic aftershock risk assessment. *Earthq. Eng. Struct. Dyn.* **43**(14), 2179–2197 (2014).
- Jalayer, F. & Ebrahimian, H. Seismic risk assessment considering cumulative damage due to aftershocks. *Earthq. Eng. Struct. Dyn.* **46**(3), 369–389 (2016).
- Ebrahimian, H. *et al.* An outlook into time-dependent aftershock vulnerability assessment. ECCOMAS Thematic Conference (COMPADYN2013): 4th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, M. Papadrakakis, V. Papadopoulos, V. Plevris (eds.), Kos Island, Greece, pp 54–76 12–14 June (2013).
- Gruppo di Lavoro. *Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003*. Rapporto Conclusivo per il Dipartimento della Protezione Civile, INGV, Milano-Roma, April 2004: 65 pp. +5 appendixes (in Italian).
- Michele, M. *et al.* The Amatrice 2016 seismic sequence: a preliminary look at the mainshock and aftershocks distribution. *Annals of Geophysics* **59**, 5, doi:10.4401/ag-7227 (2016).
- Lolli, B. & Gasperini, P. Aftershocks hazard in Italy part I: Estimation of time-magnitude distribution model parameters and computation of probabilities of occurrence. *J. Seism.* **7**(2), 235–257 (2003).
- Muto, M. & Beck, J. L. Bayesian updating and model class selection for hysteretic structural models using stochastic simulation. *J. Vib. Control* **14**(1–2), 7–34 (2008).

35. Jalayer, F. & Beck, J. L. Effects of two alternative representations of ground-motion uncertainty on probabilistic seismic demand assessment of structures. *Earthq. Eng. Struct. Dyn.* **37**(1), 61–79 (2008).
36. Jalayer, F., Iervolino, I. & Manfredi, G. Structural modeling uncertainties and their influence on seismic assessment of existing RC structures. *Struct. Saf.* **32**(3), 220–228 (2010).
37. Lippiello, E., Marzocchi, W., de Arcangelis, L. & Godano, C. Spatial organization of foreshocks as a tool to forecast large earthquakes. *Sci. Rep.* **2**, 846, doi:[10.1038/srep00846](https://doi.org/10.1038/srep00846) (2012).
38. Jaynes, E. T. *Probability theory: The logic of science* (Cambridge University Press, London, 2003).
39. Lewis, P. A. W. & Shedler, G. S. Simulation methods for Poisson processes in nonstationary systems. *Proceedings of the 10<sup>th</sup> conference on Winter simulation* 155–163 (IEEE Press, 1978).
40. Ogata, Y. On Lewis' simulation method for point processes. *IEEE Trans. Inform. Theory* **27**(1), 23–31 (1981).
41. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21**(6), 1087–1092 (1953).
42. Hastings, W. K. Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1), 97–109 (1970).

## Acknowledgements

This work was supported in part by Project METROPOLIS (Metodologie e Tecnologie Integrate e Sostenibili Per Ladattamento e La Sicurezza di Sistemi Urbani, “Methodologies and sustainable innovative technologies to assess and manage natural and man-caused hazards in urban environment”). This support is gratefully acknowledged.

## Author Contributions

F.J. developed the framework and H.E. implemented the framework and performed the analyses. H.E. and F.J. both contributed to the writing of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-09962-z](https://doi.org/10.1038/s41598-017-09962-z)

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017