Appendix: The set up of the forecasting ETAS model

The ETAS model estimates the perturbation of the seismicity rate λ in space and time induced by previous earthquakes through the equation

$$\lambda(t,x,y,m/\mathcal{H}_t) = \left[\mu u(x,y) + \sum_{T_i<t} \frac{Ke^{\alpha(M_i-M_{min})}}{(t-T_i+c)^p} \frac{c_{d,q}}{[r_i^2+(d\cdot e^{\gamma(M_i-M_{min})})^2]^q}\right]\beta e^{-\beta(m-M_{min})}. \quad (1)$$

where $\mathcal{H}_t = \{T_i, X_i, Y_i, M_i), \quad T_i < t\}$ is the observation history up time t, M_{min} is the minimum magnitude of catalog, K, c and p are the parameters of the modified Omori Law [*Utsu*, 1961] defining the decay in time of short-term triggering effect, α defines the dependence of triggering capability by magnitude of an earthquake, d and q define the spatial distribution of triggered events ($C_{d,q}$ is the normalization constant) as a function of the distance r_i between a general location (x,y) and the epicenter of the i-th earthquake (X_i, Y_i), γ takes into account the correlation between the aftershocks area and the magnitude of triggering event, and finally $\beta = b\cdot ln(10)$ is the parameter of the Gutenberg-Richter law [*Gutenberg and Richter*, 1944], defining the distribution of magnitudes.

The parameters $\mu, K, c, p, \alpha, d, b$ of equation 1 reported in the text were estimated through the iteration algorithm developed by *Zhuang et al.* [2002] that provides also an estimation of the spatial distribution $u(x,y)$ of the background events. We impose a priori the parameter $q = 1.5$ in order to have a spatial decay that mimics the co-seismic stress interaction [cf., *Hill et al.*, 2003]; this choice is also justified by the recognized trade-off between parameters q and d, that is different pairs of q and d values provide almost the same likelihood of the model [*Kagan and Jackson*, 2000]. The whole model was parameterized by using about 4 years (from April 16th 2005 to April 4th 2009) of seismicity, recorded into official INGV bulletin (http://iside.rm.ingv.it) and affecting a larger area [12.9-13.9E, 41.8-42.8N] respect to region struck by the 2009 L'Aquila sequence. The starting time of learning dataset (April 16th 2005) marks a strong improvement of Italian seismic network with a resulting decreasing of threshold detection magnitude up to $M_l = 1.5$. In this time period the catalog reports 640 events in the magnitude range 1.5-4.0. The estimated parameters are: $\mu = (0.18 \pm 0.01)$ day^{-1}, $k = (0.033 \pm 0.004)$ day^{p-1}, $p = 1.2 \pm 0.05$, $c = (0.03 \pm 0.01)$ day,

$\alpha = 1.5 \pm 0.1$, $d = (0.8 \pm 0.1)$ km, $b = 1.0 \pm 0.05$, $\gamma = 0.0 \pm 0.1$. All of them were kept fixed for whole duration of experiment together with the spatial distribution $u(x, y)$ of background events. Immediately after the M_w 6.3 event, the completeness magnitude increases for few days. We account for it using some empirical adjustment of the triggering effects of the missed events above M=1.5.

References

[1] Gutenberg B., C.F. Richter, (1944), Frequency of earthquakes in California, *Bull. Seismol. Soc. Am. 34*, 185-188.

[2] Hill D.P., *et al.* (1993), Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science 260*, 1617-1623.

[3] Utsu T. (1961), A statistical study on the occurrence of aftershocks, *Geophysical Magazine 30*, 521-605.

[4] Zhuang J., Y. Ogata, D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc. 97*, 369-380.