Forecasting of the first hour aftershocks by means of the perceived magnitude

E. Lippiello¹, G. Petrillo¹, C. Godano¹, A. Tramelli², E. Papadimitriou³ & V. Karakostas³

The majority of strong earthquakes takes place a few hours after a mainshock, promoting the interest for real-time post-seismic forecasting, which is, however, very inefficient because of the incompleteness of available catalogs. Here we present a novel method that uses, as only information, the ground velocity recorded during the first 30 min after the mainshock and does not require that signals are transferred and elaborated by operational units. The method considers the logarithm of the mainshock ground velocity, its peak value defined as the perceived magnitude and the subsequent temporal decay. We conduct a forecast test on the nine \( M \geq 6 \) mainshocks that have occurred since 2013 in the Aegean area. We are able to forecast the number of aftershocks recorded during the first 3 days after each mainshock with an accuracy smaller than 18% in all cases but one with an accuracy of 36%. 

¹ Department of Mathematics and Physics, University of Campania “L. Vanvitelli”, Viale Lincoln, 5, 81100 Caserta, Italy. ² Istituto Nazionale Geofisica e Vulcanologia, Sez. Napoli, Osservatorio Vesuviano, Via Diocleziano 328, Napoli 80124, Italy. ³ Geophysics Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece. Correspondence and requests for materials should be addressed to E.L. (email: eugenio.lippiello@unicampania.it)
Short-term aftershock forecasting (STAF) models\textsuperscript{1–3} are based on the Omori–Utsu (OU) law\textsuperscript{4} which states that the rate of aftershocks with magnitude $M$ above a threshold $M_0$, at time $t$ after the mainshock, is given by

$$\lambda(t) = \frac{K 10^{BM}}{(t + c)}$$ \hspace{1cm} (1)

where $\Delta M = M_\text{max} - M$, and $M_\text{max}$ is the mainshock magnitude. Equation (1) may be used in STAF models once the four model parameters $b$, $p$, $K$, and $c$ are known. The parameters $b$ and $p$ are similar for different sequences and setting $b = p = 1$ almost always provides reasonable results. Conversely the parameters $K$ and $c$ exhibit huge fluctuations from one sequence to another\textsuperscript{5} and, therefore, they must be fitted as soon as the ongoing earthquake sequence produces a sufficient number of aftershocks\textsuperscript{6}. This fitting procedure requires the identification of all aftershocks above a sufficiently small magnitude threshold. This, however, is a very difficult task because of the overlapping of coda-waves, among close in time aftershocks, which obscures the recordings of smaller events. As a consequence small aftershocks are not recorded in the first part of seismic sequences and the evaluation of $K$ and $c$ is strongly biased\textsuperscript{6–8}. Some fitting procedures\textsuperscript{11,13,16–20} have been recently developed to take into account this incompleteness of data sets, but they become efficient only when a sufficient number of aftershocks has been recorded. Another problem of STAF methods is that, once $\lambda(t)$ has been estimated, the predicted aftershock rate must be converted in terms of the probability of ground shaking at a given site. This requires accurate empirical attenuation functions\textsuperscript{21,22} which are usually only approximately known and often ignore site effects caused by local site conditions. Here we present a novel method based on a fitting procedure applied to the ground velocity recorded at a site of interest. The method does not require the identification of aftershocks and provides the probability of strong ground shaking without attenuation relations. This idea is, to a certain extent, similar to the proposal\textsuperscript{23} of extrapolating the strong ground shaking without attenuation relations. This idea is, therefore, they must be

We indicate with $N_1$ and $N_2$ the number of aftershocks with $M > M_\text{max}$ – 3.0 recorded within 1 and 3 months, respectively, after the mainshock occurrence. $\delta r$ is the distance in km between the mainshock epicenter and the recording station.
Fig. 1 Epicentral distribution of considered mainshocks. The spatial distribution of the $M \geq 6.0$ mainshocks that occurred in the Aegean area since October 2013 along with their fault plane solutions shown as lower hemisphere equal area projections. On the top of each ball the occurrence year is given, whereas the numbers inside the yellow stars representing the mainshock epicenters designate their sequential occurrence in time: (1) Crete, (2) Lixouri1, (3) Lixouri2, (4) North Aegean, (5) Karpathos, (6) Lefkada, (7) Lesvos, (8) Kos, (9) Zakynthos. The map is generated using the Generic Mapping Tool (http://www.soest.hawaii.edu/gmt/). Fault plane solutions were taken from http://ldeo.columbia.edu/gcmt.

Fig. 2 The average envelope function of the Kos, Lesvos, and Ischia earthquakes. a The average envelope function $\mu(t)$ is plotted as a function of the time divided by the characteristic time $\tau_M = 4.3$ s for Ischia, $\tau_M = 22.6$ s for Kos, and $\tau_M = 25.7$ s for Lesvos. Different open symbols and colors are used for the three mainshocks (see legend). Filled symbols represent the theoretical envelope $\mu_{th}(t)$ obtained setting $b = 1$, $p = 1.1$ and implementing different choices of $K$ and $c$ for the three different mainshocks. b The same data of panel (a) without time rescaling. c A zoom inside the orange dotted rectangle of the main panel. The green line is the scaling function $F(x) = 6 + \log_{10}(x) - 3.5\log_{10}(x + 0.43)$.
distance from the recording station (see Table 1) but also to soil conditions at the seismic station location. Concerning the functional dependence of the scaling function $F(x)$ in Eq. (3), the shape of the envelope function is dominated by different mechanisms such as source properties, attenuation, scattering effects, etc. The scaling collapse of Fig. 2c suggests that these features are, at a first approximation, incorporated in two seismic source-dependent parameters: the perceived magnitude $\mu_M$ and the horizontal time rescaling $\tau_M$. This result is confirmed by the scaling collapse also obtained for the other mainshocks, when plotting $\bar{\mu}(t) - \mu_M$ as a function of $(t - t_0)/\tau_M$ (Fig. 3). In all cases we find that, for $x < 10$, $F(x) \approx \log_{10}(x) - \log_{10}(x_0)$ with $q \approx 3.5$ (continuous curve in Fig. 2c) and $x_0 = (q - 1)\log(1 - \delta) \approx 0.43$ such that the maximum value of $F(x)$ is equal to zero. We assume the same functional dependence of $F(x)$ for all mainshocks in order to reduce the number of fitting parameters in our method. This fit corresponds to a power law decay of the amplitude of coda-waves of $V(t)$ which anticipates the exponential decay at longer timescales. We wish to emphasize that our results are weakly affected by the specific choice of $F(x)$ and we expect that other choices for $F(x)$ lead to similar results.

Figure 4 shows $\bar{\mu}(t)$ for the same mainshock recorded at seismic stations at different epicentral distances $\delta r$. We found that, as expected, attenuation effects lead to a perceivable magnitude $\mu_M$ which is a decreasing function of $\delta r$. At the same time, we observe that $\tau_M$ increases for increasing $\delta r$ and this can be probably attributed to scattering effects. Nevertheless, the plot of $\bar{\mu}(t) - \mu_M$ as a function of $(t - t_0)/\tau_M$ (right panels of Fig. 4) still exhibits a good scaling collapse which is consistent with Eq. (3) for $(t - t_0)/\tau_M < 10$. The collapse extends up to $(t - t_0)/\tau_M = x_B > 10$, whereas for $(t - t_0)/\tau_M > x_B$ the envelope presents an almost flat behavior $\bar{\mu}(t) \approx \mu_B$ related to the contribution of background seismicity which is expected to be quite stationary in time. Figure 4 presents results for only three earthquakes: The more recent one (Zakynthos), the one with the largest $\mu_M(t) - \mu_M$ for $t > \tau_M$ (Kos) and the one with the smallest one (Crete). A similar behavior is observed in the other earthquakes.

**Theoretical interpretation.** The main observation at the basis of our method concerns the clear separation of the envelope $\bar{\mu}(t)$ for the three mainshocks at times $t > 10\tau_M$ (Fig. 2), which we attribute to a different aftershock productivity, i.e. different $K$ and $c$ values. This temporal domain, as already observed in refs. 15,31, is indeed controlled by aftershock occurrence. If an aftershock with perceived magnitude $\mu_i$ has occurred at time $t_0$, the envelope $\bar{\mu}(t)$ cannot become smaller than $\mu_i(t) \approx \mu_i$ at times $t \approx t_0$. Therefore, larger values of $\bar{\mu}(t)$, for equal $\mu_M$, indicate a higher occurrence rate of big aftershocks and, consequently, a greater aftershock productivity. In particular, the low level of $\bar{\mu}(t)$ for $t \in [50, 120]$ in the case of the Ischia earthquake can be interpreted as a clear indication of a very small occurrence probability of $\mu_i \geq 4$ aftershocks in Ischia, in the hours following the mainshock. Indeed, only aftershocks with $\mu_i \leq 3$ have been subsequently recorded. Accordingly, from the low levels of $\bar{\mu}(t)$ recorded few minutes after the Ischia earthquake it would have been already possible to inform the authorities and issue a confirmatory announcement for mitigating the economic impact of the earthquake. The situation, conversely, was different after the Kos earthquake where $\bar{\mu}(t)$ has been stably above $\mu_i(t) > 3.5$ in the first minutes and indeed tens of $\mu_i > 4$ aftershocks occurred in the next 2 days (see Table 2). Similar indications can be obtained from the envelope function $\bar{\mu}(t)$ for all the $M \geq 6$ mainshocks occurred in the Aegean area after 2013. Figure 3, indeed, confirms that dividing the time by the appropriate $\tau_M$, $\bar{\mu}(t) - \mu_M$ reveals a quite universal behavior up to $t/\tau_M < 10$. For longer times the split of different envelopes is evident with important consequences for aftershock forecasting. In the following we present a forecasting tool which provides the quantitative estimation of the above considerations.

**The forecasting method.** Since aftershocks and mainshock mostly occur in the same tectonic environment, we expect that they present very similar attenuation and scattering properties. This is confirmed by the envelope function of the largest
aftershocks, with occurrence time \( t_i \) and perceived magnitude \( \mu_i \), which is well described by \( \tilde{\mu}_i(t_i - t) = \mu_i + F((t_i - t)/\tau_M) \) with the same scaling function \( F(x) \) and the same \( \tau_M \) of the mainshock. Accordingly, given a mainshock with perceived magnitude \( \mu_B = \mu_M \) occurred at time \( t_B = 0 \) and an aftershock sequence \( \{t_i, \mu_i\} (i \geq 1) \) we expect that the envelope function is given by

\[
\bar{\mu}_{ib}(t) = \max_{t_i < t} \tilde{\mu}_i(t_i - t) + \mu_B,
\]

where the maximum must be evaluated over all events, including the mainshock, occurred at times \( t_i < t \), and \( \mu_B \) corresponds to the background level of ground velocity during stationary periods.

Equation (4) just states that \( \bar{\mu}_{ib}(t) \) initially follows the mainshock envelope but as soon as a sufficiently strong aftershock occurs, \( \bar{\mu}_{ib}(t) \) increases up to \( \bar{\mu}_{ib}(t_i) = \mu_i \) and decays according to \( F((t - t_i)/\tau_M) \) unless another aftershock occurs and the signal can increase again. Equation (4) is supported by the scaling collapse (Fig. 4) of the envelope of the signal recorded at different stations which indicates that, even for distant stations, \( \bar{\mu}(t) \) is dominated by the aftershock occurrence up to the time \( t_B \). Our method is based on the idea to infer the statistical features of aftershock occurrence from \( \bar{\mu}(t) \) in the time interval \([t_M, t_B]\). In order to reduce the contaminations of background seismicity, in the following we restrict to the signal recorded at the closest station.

Figure 2 shows that it is possible to generate a synthetic aftershock sequence \( \{t_i, \mu_i\} \), for each one of the three mainshocks, which presents a \( \bar{\mu}_{ib}(t) \) such that \( \bar{\mu}_{ib}(t) \approx \bar{\mu}(t) \) at different times \( t \).

More specifically, the aftershock sequence is simulated assuming that aftershock occurrence follows a generalized Poisson process with occurrence rate \( \lambda(t) \) given by the OU law (Eq. 1) (see Methods section). The agreement between the theoretical \( \bar{\mu}_{ib}(t) \) and the instrumental \( \bar{\mu}(t) \) represents the key ingredient of our method. Indeed, since \( \bar{\mu}_{ib}(t) \) strongly depends on \( \lambda(t) \), we observe that different choices of the parameters \((b, p, c, K)\) in the synthetic sequence lead to a very different \( \bar{\mu}_{ib}(t) \). On the other hand, different synthetic sequences implementing the same value of the parameters \((b, p, c, K)\), correspond to a similar average envelope \( \bar{\mu}_{ib}(t) \). As a consequence, the values of the model parameters \((b, p, c, K)\) which represent the best description of the aftershock decay rate, are expected to correspond to those values which provide the best agreement between \( \bar{\mu}_{ib}(t) \) and \( \bar{\mu}(t) \).

The possibility to extrapolate the parameters \((b, p, c, K)\) from the decay of the envelope function has been already recognized in

Table 2 The comparison between the observed \( n_{obs} \) and predicted \( n_{pred} \) (T) aftershocks

| Code name | \( n_{obs} \) | \( n_{pred} \) (20 min) | \( n_{pred} \) (1 h) | \( n_{pred} \) (2 h) |
|-----------|---------------|------------------------|---------------------|---------------------|
| Crete     | 11            | 11                     | 9                   | 10                  |
| Lixouri1  | 455           | 157                    | 390                 | 396                 |
| Lixouri2  | 277           | 172                    | 252                 | 317                 |
| North Aegean | 94          | 93                     | 102                 | 81                  |
| Karpathos | 123           | 85                     | 108                 | 109                 |
| Lefkada   | 161           | 65                     | 102                 | 120                 |
| Lesvos    | 119           | 58                     | 100                 | 111                 |
| Kos       | 550           | 301                    | 621                 | 758                 |
| Zakynthos | 480           | 291                    | 472                 | 515                 |

We consider aftershocks with \( p > p_M = 3 \) in three learning periods \( T = 20\, min, 1\, h, and 2\, h \) and for the testing period \( \{t_B, t_f\} \), with \( t_B = 2\, h \) and \( t_f = 72\, h \), except for Crete where \( t_B = 48\, h \) and for Karpathos where \( t_B = 54\, h \). The quantity \( n_{pred}(T) \) is obtained with the values of \( K \) and \( c \) listed in Table 3, expressing \( t_B, t_f \) and \( c \) in seconds and using \( b = 1 \) and \( p = 1.1 \).

Fig. 4 The average envelope function at different epicentral distances. a–c The average envelope function \( \mu(t) \) obtained from the signal recorded at seismic stations at different epicentral distances \( \delta r \) from the mainshock after the Zakynthos earthquake (a), Crete earthquake (b), and Kos earthquake (c). The station name and epicentral distances are reported in the figure legend. d–f The same quantity of the corresponding left panel plotted after the rescaling \( \mu(t) - \mu_M \) and \( (t - t_B)/\tau_M \) with \( \tau_M = 11.4, 27.7, 16.1, 57.7, and 50.8\, s \) for the stations RTZL, ANKY, LKD2, ALN, and CHOS, respectively, for the Zakynthos earthquake (d), \( \tau_M = 16.5, 57.7, 55.4, and 18.5\, s \) for the stations IMMV, HORT, AGG, and GVD, respectively, for the Crete earthquake (e), and \( \tau_M = 22.6, 23.1, 25.6, and 41.8\, s \) for the stations DAT, ARG, KARP, and PRK, respectively, for the KDS earthquake (f). The green continuous line is the scaling function \( F(x) = \log_{10}(x) - 3.5\log_{10}(x + 0.43) \).
Lippiello et al.\textsuperscript{15}. The central assumption was that $\dot{\mu}(t)$, for times $t-t_0>10\tau_M$, obeys a logarithmic decay, $\dot{\mu}(t) = \mu_M - \Delta \mu - \phi \log(t)$ and that the parameters $(\Delta \mu, \phi)$ are somehow related to $K$ and $c$. The novel and more efficient method developed and applied in this study does not rely on assumptions on the functional form of $\dot{\mu}(t)$. In particular, it can be implemented in an automatic and almost real-time procedure which provides the occurrence probability of strong ground shaking in a few minutes after the mainshock. The procedure is detailed in the Methods section and here we outline the main steps. In the first step we obtain the best fit of $\mu_M$ and $\tau_M$ considering only data for $t \leq 100$ s. In the second step we consider the learning period $t \in (0, T)$ and, for a given set of parameters $K, c, b, p$, we simulate an aftershock sequence $\{t_i, \mu_i\}$ assuming that the aftershock occurrence probability in time is given by Eq. (1) with $\Delta M = \mu_M - \mu_i$. We then assume the Ishimoto and Iida law\textsuperscript{23} for the peak ground velocity which corresponds to an exponential law for the frequency distribution of the aftershock perceived magnitude $P(\mu_i) \propto 10^{-\beta \mu_i}$. This is related\textsuperscript{23} to the Gutenberg–Richter law for the frequency distribution of earthquake magnitude and, in order to reduce the number of fitting parameters, we always consider $\beta = b = 1$ which is a quite realistic value for all the considered mainshocks. For the same reason we also consider a fixed $p$ value, $p = 1.1$. We have explicitly verified that routines with variable $b$ and $p$ values do not produce any improvement.

The subsequent step is the evaluation of $\hat{\mu}_{tb}(t)$ along with the root mean square deviation $\chi(T) = \int_0^T dt (\hat{\mu}(t) - \hat{\mu}_{tb}(t))^2$ between the instrumental and theoretical envelope functions. By means of a simulated annealing procedure (see Methods section) we find the best parameter values which correspond to the minimum value of $\chi(T)$. This is a fast procedure, of the order of few minutes on standard CPU, which can be repeated as soon as time goes on and new data of $\dot{\mu}(t)$ become available. In this way we can obtain the best set of parameters in the OU law at different times $T$ after the mainshock.

Testing procedure. To test the method efficiency, the parameters $K, c$, and $\mu_M$ obtained as best fit in the learning period $(0, T)$ are used to forecast the occurrence of aftershocks with perceived magnitude $\mu_i > \mu_M - 3$ in the testing period $[t_{in} = 2\, h, t_f = 72\, h]$. The choice $\mu_i > \mu_M - 3$ represents a compromise between a sufficient number of aftershocks for robust statistical tests and considerable ground shaking induced by each aftershock. Aftershocks are identified according to the procedure outlined by Peng et al.\textsuperscript{8} selecting among all peaks of the envelope $\mu(t)$ with $\mu_{peak} \geq \mu_M - 3$. This procedure allows us to identify many more aftershocks than those reported in the official Greek catalog. Figure 5 presents with filled symbols the hourly rate $\lambda(t)$ of $\mu_i > \mu_M - 3$ aftershocks identified starting from 1 h after the mainshock. Different colors are used for the different mainshocks. The Kos sequence produces roughly 30 times more aftershocks than the Ischia one. In Kos and Lesvos sequences, $\lambda(t)$ is consistent with the OU law with $p \approx 1$, whereas the Ischia sequence comprises only a few aftershocks and the OU decay is not observed. In the same figure open symbols are used for the hourly rate of events with $\mu_i > \mu_M - 3$ predicted by implementing in the OU law (Eq. (1)) the best-fit parameters $K, c$ obtained in the learning period $(0, T)$. Different colors correspond to different mainshock sequences whereas different symbols, for each sequence, correspond to different values of $T$. Results show that for the Kos and Lesvos mainshocks an accurate prediction of subsequent aftershock occurrence could be available only ten minutes after the mainshock. In the Ischia earthquake, the forecasting after 10 min is not so efficient but for $T \geq 15\, \text{min}$ the agreement between predicted and recorded $\lambda(t)$ is satisfactory.

In order to better verify the efficiency of our method, we apply the same procedure to estimate the best OU parameters $K$ and $c$ for all the $M \geq 6$ mainshocks occurred after 2013 in the Aegean area (see Table 1). The best values of $K$ and $c$ obtained by our procedure for different learning periods $T$ and different mainshocks are listed in Table 3. In Fig. 6 we plot with continuous

---

**Fig. 5** Test of the method for the Kos, Lesvos, and Ischia earthquakes. We use filled large symbols for the hourly rate $\lambda(t)$ of aftershocks with $\mu_i > \mu_M - 3$ with different symbols and colors for different mainshocks (see legend). Small empty symbols are used for the predicted aftershock rate obtained using in the OU law Eq. (1) the values of $K$ and $c$ provided by our model. Different colors correspond to different mainshocks whereas different symbols (circles, squares, diamonds, triangles up, triangles left) correspond to different learning periods ($T = 1/6, 1/4, 1/2, 1$, and $6\, \text{h}$), respectively.
In the testing period 20 min, 1 h, 2 h. The time \( c \) is expressed in seconds and the parameter \( K \) is in units of seconds\(^{-1}\).

### Table 3 The estimated parameters \( K \) and \( c \)

| Code name | \( K \) (T = 20 min) | \( c \) (T=20 min) | \( K \) (T = 1 h) | \( c \) (T = 1 h) | \( K \) (T = 2 h) | \( c \) (T = 2 h) |
|-----------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Crete     | 0.0100            | 16                | 0.0084            | 3.75              | 0.0092            | 1.51              |
| Lixouri1  | 0.1286            | 285               | 0.3174            | 159               | 0.3227            | 215               |
| Lixouri2  | 0.1397            | 80                | 0.2045            | 155               | 0.2579            | 156               |
| North Aegean | 0.0757    | 281               | 0.0834            | 198               | 0.0663            | 218               |
| Karpathos | 0.0740            | 163               | 0.0952            | 283               | 0.0956            | 261               |
| Lefkada   | 0.0532            | 282               | 0.0830            | 238               | 0.0976            | 153               |
| Lesvos    | 0.0470            | 205               | 0.0824            | 365               | 0.0907            | 296               |
| Kos       | 0.2434            | 53                | 0.5020            | 38                | 0.6132            | 52                |
| Zakynthos | 0.2375            | 265               | 0.3850            | 240               | 0.4185            | 165               |

For each mainshock we consider three different learning periods \( T = 20 \text{ min}, 1 \text{ h}, 2 \text{ h} \). Results show huge differences (up to 6000%) in the number of observed aftershocks with a discrepancy typically smaller than 20%. Results also show that the best agreement is obtained for a learning period \( T \in (0.5 \text{ h}, 2 \text{ h}) \). The agreement tends to become worse for increasing \( T \) probably because of the contribution of higher order generation aftershocks, not considered in our study.

### Fig. 6 Short-term test of the method for the recent Aegean earthquakes. We compare the number of observed and predicted aftershocks with \( \mu_1 > \mu_M - 3 \) in the testing period \([t_{in}, t_T]\) with \( t_{in}=2 \text{ h} \) and \( t_T=72 \text{ h} \) except for Crete where \( t_T=48 \text{ h} \) and for Karpathos where \( t_T=54 \text{ h} \) Open symbols represent the value of \( n_{\text{pred}}(T) \) for different values of the testing period \( T \). The error bars correspond to statistical fluctuations in the number of expected aftershocks for fixed \( K \) and \( c \) values. Dashed horizontal lines represent \( n_{\text{obs}} \), i.e. the number of identified aftershocks according to the method described in ref. 8. Different symbols and colors are used for different mainshocks (see legend).

Indeed, the possibility of each aftershock to trigger its own aftershocks leads to deviations from the pure OU law and the estimate of parameters \((K, c)\) according to Eq. (1) becomes less and less accurate for increasing time\(^{32-34}\). Notice that for testing periods with \( T > 2 \text{ h} \), the testing and learning periods overlap. We make this choice in order to have a single value of \( n_{\text{obs}} \) in Fig. 6 and in Table 2 and this simplifies the presentation. At the same time the overlap is null in the more relevant temporal window \( T \leq 2 \text{ h} \) where our method produces the best agreement. Similar results are also found for a testing period with \( t_{in}=1 \text{ h} \), as illustrated in the Supplementary Figs. 2 and 3 and in the Supplementary Table 1.

The same comparison between observed and predicted aftershock number is performed in Fig. 7. In this case we plot the number \( n_{\text{pred}}(T) \) of \( \mu_1 > \mu_M - 3 \) aftershocks predicted by our model in the temporal range \([t_{in}=7, t_T=90] \text{ days}\) when short-term catalog incompleteness is not relevant. We therefore directly compare with the number \( n_{\text{at}} \) of \( m_1 > M_{in} - 3 \) aftershocks included in the Greek earthquake catalog in the same temporal
period. We still found a good agreement between prediction and observation, which shows that in only a few minutes after the mainshock it is already feasible to provide insights in the evolution of seismicity in the following months. There is an exception related to the 2014 North Aegean sequence since the model predicts about 70 aftershocks with $m_i > M_m - 3$ in the interval $[7, 90]$ days whereas no aftershock is included in the Greek catalog. In this case we observe an abrupt change of the seismic rate, in the mainshock area, which is not consistent with the OU law and, therefore, it is not predicted by our procedure. On the other hand, in the case of the Zakynthos earthquake we find $n_{pred}(T = 1 \text{ h}) = 220 \pm 30$ and we cannot compare with $n_{fit}$ in the Greek catalog, since the testing period is not yet finished. The same analysis of Fig. 7 for a testing period of $[30, 90]$ days is presented in Supplementary Fig. 4.

**Discussion**

Our results have shown that, in only 30 min after a mainshock, it is possible to have an accurate forecasting of the aftershock occurrence rate in the subsequent days. We wish to emphasize that, to our knowledge, existing methods are not capable to provide such an efficient forecasting in a very short learning period of $T \sim 1 \text{ h}$, like the one here considered. The most refined way to take explicitly into account aftershock incompleteness is the Omi et al. method\textsuperscript{11,16,20}. This method is based on a Bayesian estimate of the completeness magnitude which, however, requires a sufficient number of identified aftershocks. In Omi et al. (2018) the method has been applied only to sequences with at least 30 events in the learning period. In the case of the mainshocks considered in our study, the number of identified aftershocks, at temporal distances smaller than $T \sim 2 \text{ h}$ reported in the Greek earthquake catalog, is typically smaller than 20 and this makes the prediction according to the Omi et al. method very unstable.

In conclusion, we have presented a novel method for aftershock forecasting where the parameters of the OU law are extracted from the ground velocity recorded some minutes after the mainshock. The accuracy of the method has been verified by means of retrospective tests on recent Greek mainshocks and future developments correspond to the validation by means of prospective tests. Refinements of the method can be obtained by incorporating higher order generation aftershocks which should contribute to improve the agreement for learning periods larger than 1h.

**Methods**

**The average envelope.** The quantity $\hat{\mu}(t)$ is obtained by means of a logarithmic smoothing procedure that corresponds to the average value of $\mu(t)$ over windows of increasing duration $\Delta t_\mu = \Delta t_{0}(1 + \chi^2)$ with $\Delta t_0 = 0.1 \text{ s}$ and $\chi = 0.005$.

**Algorithm to invert the OU parameters.** The algorithm is composed of four separate routines.

The first routine considers $\hat{\mu}(t)$ restricted to the time interval $[t_0, t_0 + 100 \text{ s}]$ and looks for the parameters $\mu_M$ and $\tau_M$ which minimize the quantity

$$\chi = \left( \hat{\mu}(t) - \mu_M - \log\left( \int (t-t_0) / \tau_M \right) + \log\left( (t + t_0) / \tau_M + x_0 \right) \right)^2$$

with $g = 3.5$.

The second routine generates, for a given set of parameters $(b, c, K)$, an aftershock sequence. More precisely, we assume that the number of aftershocks follows a Poisson distribution with average value $K10^{\Delta M}$ with $\Delta M = 3$. We randomly extract the occurrence time of each aftershock $t_i$ from the probability distribution $P(t_i) = \frac{\Delta t_{0}(1 + \chi^2)}{\tau_0}$. To each aftershock we associate a perceived magnitude $\mu_t = \mu_M - \Delta M + m_i$, with $m_i$ is extracted from the GR law $P(m_i) = \log(10)10^{-m_i}$, without any extra constraint, i.e. $\mu_t$ can be greater than $\mu_M$.

The third routine uses the information on $\tau_M$ and $\mu_M$ provided by the first routine, and starting from an aftershock sequence $\{t_i, \mu_t\}$ generated by the second routine, itgenerates a value $\hat{\mu}_M(t)$ according to Eq. (4). More precisely, the time is discretized in steps of $0.05 \text{ s}$ and at each time $t$, we evaluate the quantity

$$\bar{\mu}_M(t - t_i) = \mu_t + \log\left( (t-t_0) / \tau_M \right) - 3.3\log\left( (t-t_0) / \tau_M + x_0 \right)$$

for all aftershocks with $t_i < t$. The quantity $\hat{\mu}_M(t)$ is given by the maximum value of $\bar{\mu}_M(t - t_i)$ and we use the same logarithmic smoothing applied to the instrumental envelope to obtain $\hat{\mu}_M(t)$.

The fourth routines finds the best OU parameters $(b, c, K)$ via the minimization of the discrepancy between the theoretical and the instrumental envelope, up to a learning time $T$. This is achieved by means of the algorithm developed by Bottiglieri et al.\textsuperscript{35} for the log-likelihood maximization. More precisely, for a given parameter set $(b, c, K)$, we use $n_{real}$ times the third routine to obtain $n_{real}$ independent realizations of the envelope $\hat{\mu}_M^{(j)}(t)$ with $j = 1, ..., n_{real}$.

**Fig. 7** Long-term test of the method for the recent Aegean earthquakes. The same plot as in Fig. 6 for the number of predicted aftershocks $n_{pred}(T)$ (open symbols) with $m_i > M_m - 3$ in the testing period $[7, 90]$ days. Dashed horizontal lines is the number of aftershocks with $m_i > M_m - 3$ recorded in the Greek catalog in the interval $[7, 90]$ days. Aftershocks are de

| Country       | Lefkada | Kos   | Karpathos | N.Aegean | Lixouri1 | Lixouri2 | Crete |
|---------------|---------|-------|-----------|----------|----------|----------|-------|
| $n_{pred}$    | 100     | 10    | 80        | 100      | 100      | 100      | 100   |
| $n_{obs}$     | 100     | 10    | 80        | 100      | 100      | 100      | 100   |
For each realization we evaluate the quantity \(\chi(T) = \frac{1}{2} \int \left( \rho(t) - \langle \rho(t) \rangle \right)^2 dt\) and then define \(\chi(T)\) as the minimum value of \(\chi(T)\) for \(j = 1, \ldots, n_{\text{tot}}\). The parameters \((b, p, c, k)\) are then updated according to the algorithm of Bottiglieri et al.35 until we find the best values which correspond to the minimum of \(\chi(T)\). To reduce the number of parameters we fix \(b = 1\) and \(p = 1.1\). The code is available from the authors upon request.

In the limit \(n_{\text{tot}} \rightarrow \infty\) the procedure identifies the synthetic aftershock sequence which, according to Eq. (4), gives the best description of the instrumental signal and therefore the corresponding \((b, p, c, k)\) is considered the best parameter set. We have verified (see Supplementary Fig. S) that the procedure appears quite stable for \(n_{\text{tot}} \geq 10\). In order to reduce the computational time, results presented in this manuscript are obtained for \(n_{\text{tot}} = 10\). We have also explored two alternative definitions of \(\chi(T)\): The average value of \(\chi(T)\) for \(j = 1, \ldots, n_{\text{tot}}\) and its median value. We have found that the three different definitions lead, for \(n_{\text{tot}} \geq 10\), to a similar number of predicted aftershocks \(n_{\text{pred}}(T)\) within an uncertainty typically smaller than the 20%.

**Data Availability**

All relevant data are available from the corresponding author upon reasonable request.

Received: 11 September 2018 Accepted: 22 May 2019
Published online: 04 July 2019

---

**References**

1. Reasenberg, P. A. & Jones, L. M. Earthquake hazard after a mainshock in California. Science 243, 1173–1176 (1989).
2. Reasenberg, P. A. & Jones, L. M. Earthquake aftershocks: update. Science 265, 1251–1252 (1994).
3. 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).
4. Utsey, 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).
5. Wetzler, N., Brody, E. E. & Lay, T. Regional and stress drop effects on aftershock productivity of large megathrust earthquakes. Geophys. Res. Lett. 43, 12,012–12,020 (2016).
6. Kagan, Y. Y. Short-term properties of earthquake catalogs and models of earthquake source. Bull. Seismol. Soc. Am. 94, 1207–1228 (2004).
7. Helmstetter, A., Kagan, Y. Y. & Jackson, D. D. High-resolution time- and therefore the corresponding \((b, p, c, k)\) is considered the best parameter set. We have verified (see Supplementary Fig. S) that the procedure appears quite stable for \(n_{\text{tot}} \geq 10\). In order to reduce the computational time, results presented in this manuscript are obtained for \(n_{\text{tot}} = 10\). We have also explored two alternative definitions of \(\chi(T)\): The average value of \(\chi(T)\) for \(j = 1, \ldots, n_{\text{tot}}\) and its median value. We have found that the three different definitions lead, for \(n_{\text{tot}} \geq 10\), to a similar number of predicted aftershocks \(n_{\text{pred}}(T)\) within an uncertainty typically smaller than the 20%.

**Data Availability**

All relevant data are available from the corresponding author upon reasonable request.

Received: 11 September 2018 Accepted: 22 May 2019
Published online: 04 July 2019

---

**Acknowledgements**

This paper is dedicated to the memory of our dear friend and colleague Ferdinando Giacco, who passed away too prematurely. He made the time spent together special, in and out of the office, with his critical thinking, irony, and kindness. We acknowledge support of this work by the project HELP, Hellenic System for Lithosphere Monitoring (MIS 500269) which is implemented under the Action Reinforcement of the Research and Innovation Infrastructure, funded by the Operational Program “Competitiveness, Entrepreneurship and Innovation” (NSRF 2014–2020) and co-financed by Greece and the European Union (European Regional Development Fund). Geophysics Department Contribution 916/2019.

**Author contributions**

L.E. conceived the study and developed the algorithm. L.E., P.G., G.G., T.A., P.E., and K.V. contributed extensively to data analysis and to write the manuscript.

**Additional information**

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-10763-3

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information: Nature Communications thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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