Recognizing the waveform of a foreshock

Lippiello E.[1,*], Petrillo G.[1] & Godano C.[1,2]

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

[2] *Istituto Nazionale Geofisica e Vulcanologia, Napoli, Italy.*

The 2011 Mw9.1 Tohoku, Japan, earthquake is the paradigmatic example of an earthquake anticipated by a significant foreshock activity, with a Mw7.3 earthquake occurred two days before, within about 10 km [1]. Recent results [2] show that statistically relevant changes can be found in the magnitude distribution after the Mw7.3 foreshock but the discrimination between normal and foreshock activity still remains a scientific challenge [3]. Here we show that the envelope of the ground velocity recorded after the Mw7.3 foreshock presents an atypical *sawtooth* **profile very different from the one observed after other earthquakes [4]. We interpret this profile as the consequence of the locked state of the mainshock fault which reduces the possibility of the foreshock to trigger its own aftershocks. We find a similar sawtooth profile after other Mw6+ foreshocks followed within 10 days by a larger earthquake, as in the case of the 2014 Mw8.1 Iquique, Chile, sequence. This observation allows us to define a level of concern, simply extracted from the first 45 minutes of the recording waveform, associated to the occurrence of a larger earthquake. A test of the method for 47 Mw6+ worldwide earthquakes gives precise warning in time and space after all the 10 earthquakes followed by a**

larger one with only 2 **false alerts.**

The coseismic slip during a large earthquake causes a shear stress reduction in regions which have experienced large slips and, at the same time, concentrates residual shear stress near the slip zone margins [5]. This stress redistribution promotes the occurrence of aftershocks with an abrupt increase of the seismic rate. During normal activity the aftershock magnitudes get smaller for increasing time, but, occasionally, aftershocks larger than the mainshock are observed. In these cases the mainshock is relabeled foreshock, the largest earthquake becomes the mainshock and the key question becomes if it is possible to distinguish foreshocks from normal seismic activity. Focusing on moderate up to intermediate (Mw$<$5) mainshock magnitudes, after the original paper by Brodsky [6], several studies [7–12] have shown that the number of foreshocks in instrumental catalogs is larger than the one expected according to normal earthquake clustering models. This result is also in agreement with a recent study [13], before Mw4 mainshocks, which uses a high-resolution earthquake catalog. Statistically relevant deviations from normal seismicity has been also found [7] before Mw6+ mainshocks but the first clear proof of the relevance of foreshocks in improving the forecast of large (Mw6.5+) mainshocks has been only recently obtained by Gulia & Wiemer (GW) [2]. Indeed, GW demostrated that the b-value of the Gutenberg-Richter (GR) law decreases during foreshock activity whereas previous studies [14] have shown that it increases during normal aftershock sequences. This pattern has been observed in the temporal interval of two days separating the Mw7.3 foreshock and the 2011 Mw9.1 Tohoku mainshock. During the same interval, episodic slow slip events have been observed [15–17] and they have been interpreted as precursors according to the pre-slip model [18,19]. Other observations [20], conversely, support the cascade model [21,22] where foreshocks are no different from other sets of clustered earthquakes.

In this article we show that it is possible to discriminate between foreshocks and normal seismic sequences from the profile of the envelope $\mu(t)$, defined as the logarithm of the envelope of the ground velocity (see Methods). Immediately before an earthquake, $\mu(t)$ starts from the background level μ_B and rapidly raises up to the time t_M, when it reaches its maximum value μ_M. This value corresponds to the perceived magnitude close to the recording station and represents the mainshock magnitude apart from an additive term. The presence of aftershocks is clearly visible [4,23–26] in the decay of the envelope function $\mu(t)$, at times $t > t_M$. We present two limit cases in Fig.1[1a]: The 2019/11/27 Mw6.1 Platanos earthquake followed by no aftershock and the 2017/07/20 Mw6.6 Kos earthquake with many aftershocks identified in the first hours [27]. After the Mw6.1 Platanos earthquake, for times $t > t_M$, $\mu(t)$ fast decays and then remains stationary around $\mu(t) \simeq \mu_B$. Conversely, after the Kos earthquake, $\mu(t)$ does not go back to μ_B but fluctuates around a plateau with a minimum value μ_L significantly larger than μ_B. To understand the origin of the plateau, we must take into account that if an aftershock has occurred at time t_1, with percieved magnitude μ_1, it produces a peak $\mu(t_1) = \mu_1$ in the envelope. After this peak, $\mu(t)$ would decrease towards μ_B but if an other afteshock with perceived magnitude μ_2 occurrs at the time $t_2 > t_1$, the envelope raises again reaching a second peak $\mu(t_2) = \mu_2$. If the aftershock productivity is very high then the temporal distance $(t_2 - t_1)$ between two subsequent aftershocks is very short and $\mu(t)$ is not able to decay below a level of the plateau $\mu_L \simeq \mu_2$. In particular, in ref. [25] it was numerically shown that the number of aftershocks exponentially increases by increasing $\mu_L - \mu_B$. More precisely, a very qualitative estimate indicates that the expected number of aftershocks (with $\mu_i > \mu_M - 2$) following a mainshock with perceived magnitude μ_M, is roughly proportional to $n_{aft} = 10^{-\delta\mu}$, with $\delta\mu = (\mu_M - \mu_B) - 2(\mu_L - \mu_B)$. A precise estimate of the aftershock occurrence probability from $\mu(t)$, in the first minutes after t_M, can be found in Ref. [4].

In Fig.1[1b] we plot $\mu(t)$ after the Mw7.3 2011/03/09 foreshock. According to the cascade model [21,22] the occurrence of an aftershock larger than the mainshock is a rare event which is more probable to occur during an intense aftershock activity. Accordingly, the normal behavior of the envelope function (Fig.1[1a]) would have suggested a high plateau level with a large value of n_{aft}. Instrumental data (black line in Fig.1[1b]), conversely, show exactly the opposite trend with $\mu(t)$ reaching values close to μ_B about 8 minutes after the mainshock, similarly to the no-aftershock pattern observed after the Mw6.1 Platanos earthquake (Fig.1[1a]). At variance with the Platanos earthquake, in the case of the Mw7.3 Tohoku foreshock, after reaching the minimum value, the envelope $\mu(t)$ abruptly raises and then drops again to $\mu(t) \simeq \mu_B$ producing an anomalous *sawtooth* profile. The presence of large peaks corresponds to the occurrence of large foreshocks ($\mu_i \simeq \mu_B + 4$) whereas the presence of valleys with $\mu(t) \simeq \mu_B$ corresponds to temporal periods with very few $\mu_i > \mu_B$ earthquakes. Therefore $\mu(t)$ shows the existence of temporal periods of some minutes with zero events (valleys) interrupted by large earthquakes (peaks). This is incompatible with the GR law which predicts thousands of events with $\mu_i \simeq \mu_B + 1$ for each event with $\mu_i \simeq \mu_B + 4$. Looking at the envelope $\mu(t)$ after the Mw9.1 mainshock (red curve in Fig.1[1b]), conversely, we find the normal profile with a high plateau level μ_L, as during standard aftershock triggering. The same behavior is observed at different seismic stations (Suppl. Fig.3).

Summarizing, the sawtooth profile of $\mu(t)$ after the Mw7.3 Tohoku foreshock reveals a very limited capability in triggering small aftershocks. A physical interpretation of this behavior, consistent with the pre-slip model, is qualitatively illustrated in the cartoon of Fig.2a. The Tohoku sequence has occurred in a region of the plate boundary which has not experienced a large earthquake for over a century and, according to geodetic data, the area of maximum coseismic slip was probably locked for a period of several years before the foreshock [29]. It is then reasonable, consistently with the asperity model [30], that the accumulated elastic strain leads to the existence of a vast and strongly correlated region along the plate (the red region in Fig.2a): A slip of a sub-region, even small in size, inside the red area will produce the syncronized rupture of several asperities with the global failure of the whole region. Nevertheless, because of frictional heterogeneities, it is reasonable to expect the existence of weaker regions within the red area (blue regions in Fig.2a). These regions are less locked and will achieve slip instabilities before the red one. Foreshocks are caused by the coseismic slip of these (blue) regions leading to a stress concentration at their periphery, i.e. inside the red area. Within this interpretation, therefore, the occurrence of a foreshock either will cause the global failure of the whole (red) area or can trigger aftershocks only inside another blue region or outside the red region (green area in Fig.2a). In presence of a slow drive process, other blue regions are brought close to failure and the envelope takes the form of isolated peaks (foreshocks) separated by temporal periods of quasi-zero seismicity (valleys). This should affect the peak distribution $P(\mu)$ (see Methods) which typically follows the Ishimoto-Iida law [28] $P(\mu_i) \sim 10^{-\beta\mu_i}$, where β roughly coincides with the b-value of the GR law. Since the presence of isolated foreshocks corresponds to a deficit of small events we should expect a smaller β-value. We indeed measure a β-value ($\beta = 0.4 \pm 0.1$) after the Mw7.3 foreshock significantly smaller than the value $\beta = 1.0 \pm 0.1$ after the Mw9.1 mainshock. The observed change in the β-value is consistent with the one found by GW in the b-value [2].

We next look for the anomalous behavior of $\mu(t)$, observed after the Mw7.3 Tohoku foreshock, after other earthquakes. We consider all Mw6+ events, occurred after 2010, followed within 10 days by a larger earthquake and recorded by a seismic station at distance smaller than ~ 100 km, with a level of the background signal $\mu_B \lesssim 1$, which is sufficiently low to indicate that the envelope is weakly influenced by previous seismicity. There are 10 mainshocks, highlighted in Suppl. Table 1, that match our criterion. We have investigated $\mu(t)$ after each mainshock and its most relevant foreshocks (all shown in Sec.2 in the Supporting Information). We find that $\delta\mu$ after the foreshocks is always larger than the same quantity after the corresponding mainshock. In particular, the anomalous sawtooth profile, observed after the Mw7.3 Tohoku foreshock, is also found in other foreshocks when the foreshock hypocenter is located very close to the relative mainshock one and therefore reasonably well inside the locked (red) area, as in the case of the 2014/03/23 Mw6.2 Iquique foreshock [31]. The evolution of $\mu(t)$ then supports our interpretation (Fig.2a) and confirms the sawtooth profile as a distinctive feature of foreshock activity.

Fig.2b schematically describes a different situation where the foreshock occurs on the border of the locked (red) area. The 2014/03/14 Mw6.7 Iquique foreshock, the 2016/08/24 Mw6.2 Amatrice foreshock and the 2016/10/26 Mw6.1 Visso foreshock belong to this situation [31,34]. In this case a fraction of the stress is redistributed within the "no-aftershock" red zone but the remaining stress is concentrated within the green zone, leading to normal aftershock activity. In presence of a slow driving process we therefore expect that foreshock activity within blue regions is superimposed to standard aftershock triggering outside the locked portion of the fault. We indeed observe (green curve in Fig.1[1c]) that $\mu(t) - \mu_B$ drops to a small but significantly larger than zero value and this can be attributed to a low but non null normal aftershock triggering. At the same time we can easily identify the presence of some large isolated peak μ_i, which can be associated to the foreshock occurrence. Fig.2c schematically describes a further different situation where the foreshock occurs on a secondary fault, close in space but different from the mainshock one, as for the 2019/07/04 Mw6.4 Ridgecrest foreshock [35]. Also in this case, as in the situation of Fig.2b, we expect the coexistence of normal aftershock activity along the secondary fault with foreshock activity along the main fault. This should produce an hybrid behavior of $\mu(t)$ with a small but non null value of $\mu_L - \mu_B$ and, at the same time, the presence of large quite isolated peaks μ_i, as confirmed by instrumental data (green curve in Fig.1[1d]). Also the 2016/04/14 Mw6.2 Kumamoto foreshock occurred on a secondary fault [36] and $\mu(t)$ exhibits an hybrid behavior (Supp. Fig. 27). At the same time, the Ridgecrest sequence presents a Mw5.4 foreshock occurred very close in space (epicentral distance $\sim$ 2 km) only 16 hours before the mainshock. This foreshock exhibits a profile of $\mu(t)$ more similar to the Mw7.3 Tohoku foreshock (black curve in Fig.1[1d]).

According to previous observations, foreshock activity exhibits a much larger number of events n_{obs} with high peak values ($\mu_i > \mu_M - 2$) compared to its expected number n_{aft}. The quantity n_{obs}/n_{aft} should be therefore exhibit abnormal large values during foreshock sequences and $Q = (n_{obs}/n_{aft})10^{-\beta\mu_M}$ can be used to define the level of concern associated to the occurrence of a subsequent earthquake with peak magnitude larger than μ_M (see Methods). We consider as target earthquakes all Mw6+ earthquakes which, in the last decade, have been followed, within 10 days, by an earthquake with a larger value of μ_M. For comparison, we also take into account all Mw6+ earthquakes occurred after 2015 in geographic regions with dense seimsic stations (in-land Japan, Central Europe and North America). Imposing a constraint on the quality of the recorded waveform (see Methods) we collect a sample of 47 earthquakes listed in Suppl. Table 1 and including 10 target earthquakes. To statistically validate our method we use the receiver operating characteristic (ROC) diagram adopting a binary code which switches on the alarm when Q is larger than a given threshold Q_{th}. By changing the value of Q_{th} we evaluate the fraction of true positive rate (TPR), i.e. the number of events with $Q > Q_{th}$ and followed by a larger earthquake divided by the total number of target earthquakes. We also evaluate the fraction of false alarm rate (FAR), defined as the fraction of events with $Q > Q_{th}$ and NOT followed by a larger earthquake. By changing Q_{th} we obtain a diagram (Fig.3) with points close to the perfect prediction ($TPR = 1, FAR = 0$) and well above the random prediction, which can be discarded with a confidence level above the 99.99999%.

In Suppl. Fig.1 we show that there exists a specific $Q_{th} = 0.18$ which allows us to score all the 10 successful alerts, with no missed events, two false alerts, and 35 correct negatives. Interesting, this value of Q_{th} was identified at the time of the first submission and the two foreshocks among the 7 earthquakes occurred later (last 7 events in Suppl. Table 1) are correctly discriminated by the comparison between Q and $Q_{th} = 0.18$.

This result is obtained using $\beta = 1$ in the definition of Q, which leads to an expression of Q only in terms of the quantities μ_B, μ_L and n_{obs} which can be easily obtained from the first minutes of the envelope $\mu(t)$ (see Methods). This gives some advantages with respect to the GW method (see Supplementary Materials) and allows us to overcome all the problems related to the estimate of the b-value [3]. Indeed, the quantity Q is analogously defined for all earthquakes whereas the choice of the normal b-value requests a decision-making specific for each earthquake. Furthermore, our results are not affected by problems of aftershock completeness and we can therefore test the model over an ensemble including up to 10 true positive cases. However, we wish to remark that our method can provide information of the stress state of a fault only at the occurrence time of a big foreshock. Indeed, the earthquake must be sufficiently large to distribute stress over a wide region of the fault so to discriminate locked from unlocked regions (red and blue areas in Fig.2). The evaluation of the b-value, conversely, provides the opportunity to monitor the evolution of the stress on the fault over sufficiently long temporal periods as for instance in the case of the Tohoku earthquake when a decreasing trend in the b-value has been documented for a period of several years before the Mw7.3 foreshock [37]. At the same time the b-value can be used to identify the stress change induced by big foreshocks with results in substantial agreement with the evaluation of Coulomb stress changes [38]. As a consequence it appears reasonable that one can achieve more accurate forecasting by combining the two approaches: the Q value can be adopted as an initial discriminant and the b-value can be used to identify higher stressed regions, which will be probably host the subsequent mainshock. Nevertheless, in the Kumamoto sequence the b-value has increased in the surrounding of the future mainshock hypocenter [39]. As an example of combining the two approaches we observe that, consistently with the GW prediction of small b-values during foreshock activity, the peak distribution $P(\mu)$ (upper panels of Fig.1[2a-2d] and Suppl. Table.1) presents atypical small values of $\beta \leq 0.74$ after the foreshocks compared to larger values of β observed after the mainshocks. Implementing in the definition of Q the β value extracted from $P(\mu)$, we find an even clearer discrimination of foreshock from normal aftershock activity (Fig.3 and Suppl. Fig.2) with just one false alert.

Summarizing, we present a simple procedure which can be used to characterize the stress state of a fault and can help local authorithies in the management of post-seismic risk.

Methods

The envelope function

We filter the vertical component of the ground velocity in the range $[2, 10]$ Hz, we apply a Hilbert transformation and we take the logarithm to base 10. We finally define the envelope function $\mu(t)$ after applying a smoothing procedure on a moving window including 5 points. We find very similar results using the two horizontal components and different choices of the frequency range. Only for graphical purposes, in Fig.1 we extend the smoothing procedure to a moving window of 200 points.

The evaluation of Q

We define the quantity t_M as the time such that $\mu(t)$ takes its maximum value and the perceived magnitude is $\mu_M = \mu(t_M)$. We define the quantity μ_B as the minimum value of $\mu(t)$ in the temporal window $[t_M - \Delta t_B, t_M]$ and μ_L as the minimum value of $\mu(t)$ in the temporal window $[t_M, t_M + \Delta t_L]$. In our study we consider $\Delta t_B = 400$ sec and $\Delta t_L = 45$ min. We have verified that similar results are obtained for $\Delta t_B \in [2, 10]$ min and $\Delta t_L \in [30, 60]$ min. We take the signal recorded at the station with the smallest value of μ_B and in all cases we always consider waveforms with $\mu_B \leq 1$. At the same time we only take into account waveforms with a perceived magnitude $\mu_M \geq 5$. Since μ_M is the decreasing function of the epicentral distance of the recording station we usually consider the closest station compatible with the above constraint on μ_B. We have also verified that the value of μ_M is not affected by staturation problems.

The selection of earthquakes

We adopt the searching criterion of the USGS earthquake hazard program to obtain the occurrence time and epicentral coordinates of an Mw6+ earthquake. We next consider regional networks to verify this information and to identify the stations closest to the earthquake epicenter. Once the waveform with the ground velocity has been downloaded, the procedures to obtain the envelope and to evaluate Q, illustrated above, are automatically implemented.

The peak distribution

We identify aftershocks from the envelope function $\mu(t)$ by means of the procedure developed by Peng et al. [23]. We define the quantity n_{obs} as the number of aftershocks producing a peak magnitude $\mu_i \geq \mu_M - 2$ in the temporal interval $[t_M + 300sec, t_M + \Delta t_L]$. We remove the first 300 seconds from this analysis since, in this time window, aftershocks can be hidden by the coda wave of the main event. We evaluate of the peak distribution $P(\mu)$ extending the temporal interval to 1 day, $t \in [t_M + 300sec, t_M + 1day]$ and the β-value is given by the best exponential fit $P(\mu) = A10^{-\beta\mu}$ for $\mu > \mu_M - 2$, according to the Ishimoto-Iida law.

1. Hirose, F., Miyaoka, K., Hayashimoto, N., Yamazaki, T. & Nakamura, M. Outline of the 2011 off the pacific coast of Tohoku earthquake (Mw 9.0) seismicity: foreshocks, mainshock, aftershocks, and induced activity. *Earth, Planets and Space* **63** (2011).

2. Gulia, L. & Wiemer, S. Real-time discrimination of earthquake foreshocks and aftershocks. *Nature* **574**, 193–199 (2019).

3. Brodsky, E. Predicting if the worst earthquake has passed. *Nature* **574**, 185–186 (2019).

4. Lippiello, E. *et al.* Forecasting of the first hour aftershocks by means of the perceived magnitude. *Nature Communications* **10**, 2953 (2019).

5. Wetzler, N., Lay, T., Brodsky, E. E. & Kanamori, H. Systematic deficiency of aftershocks in areas of high coseismic slip for large subduction zone earthquakes. *Science Advances* **4** (2018).

6. Brodsky, E. E. The spatial density of foreshocks. *Geophysical Research Letters* **38**, L10305 (2011). L10305.

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

8. Shearer, P. M. Self-similar earthquake triggering, Bath's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for Southern California. *Journal of Geophysical Research-Solid Earth* **117**, (2012).

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

10. de Arcangelis, L., Godano, C., Grasso, J. R. & Lippiello, E. Statistical physics approach to earthquake occurrence and forecasting. *Physics Reports* **628**, 1 – 91 (2016).

11. Lippiello, E., Giacco, F., Marzocchi, W., Godano, C. & Arcangelis, L. d. Statistical features of foreshocks in instrumental and ETAS catalogs. *Pure and Applied Geophysics* 1–19 (2017).

12. Seif, S., Zechar, J. D., Mignan, A., Nandan, S. & Wiemer, S. Foreshocks and Their Potential Deviation from General Seismicity. *Bulletin of the Seismological Society of America* **109**, 1–18 (2018).

13. Trugman, D. T. & Ross, Z. E. Pervasive foreshock activity across Southern California. *Geophysical Research Letters* **46**, 8772–8781 (2019).

14. Gulia, L. *et al.* The effect of a mainshock on the size distribution of the aftershocks. *Geophysical Research Letters* **45**, 13,277–13,287 (2018).

15. Ito, Y. *et al.* Episodic slow slip events in the Japan subduction zone before the 2011 Tohoku-Oki earthquake. *Tectonophysics* **600**, 14 – 26 (2013). Great Earthquakes along Subduction Zones.

16. Kato, A. *et al.* Propagation of slow slip leading up to the 2011 Mw 9.0 Tohoku-Oki earthquake. *Science* **335**, 705–708 (2012).

17. Brodsky, E. E. & Lay, T. Recognizing foreshocks from the 1 april 2014 Chile earthquake. *Science* **344**, 700–702 (2014).

18. Ellsworth, W. L. & Beroza, G. C. Seismic evidence for an earthquake nucleation phase. *Science* **268**, 851–855 (1995).

19. Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. Detailed observations of California foreshock sequences: Implications for the earthquake initiation process. *Journal of Geophysical Research: Solid Earth* **101**, 22371–22392 (1996).

20. Marsan, D. & Enescu, B. Modeling the foreshock sequence prior to the 2011, Mw9.0 Tohoku, Japan, earthquake. *Journal of Geophysical Research: Solid Earth* **117** (2012).

21. Fukao, Y. & Furumoto, M. Hierarchy in earthquake size distribution. *Physics of the Earth and Planetary Interiors* **37**, 149 – 168 (1985).

22. Helmstetter, A. & Sornette, D. Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research: Solid Earth* **108**, 2457 (2003).

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

24. Sawazaki, K. & Enescu, B. Imaging the highfrequency energy radiation process of a main shock and its early aftershock sequence: The case of the 2008 Iwate Miyagi Nairiku earthquake, Japan. *Journal of Geophysical Research: Solid Earth* **119**, 4729–4746 (2014).

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

26. Lippiello, E., Cirillo, A., Godano, C., Papadimitriou, E. & Karakostas, V. Post seismic catalog incompleteness and aftershock forecasting. *Geosciences* **9**, 355 (2019).

27. Karakostas, V., Papadimitriou, E., Mesimeri, M., & Begum, C. The 2017 Kos sequence: Aftershocks relocation and implications for activated fault segments. *36th ESC General Assembly, Valetta, Malta* , ESC2018-S6-645 (2018).

28. Ishimoto, M. & Iida, K. Observations sur les seismes enregistres parle microsismographe construit dernierement. *Bulletin of the Earthquake Research Institute* **17**, 443–478 (1939).

29. Ikuta, R., Satomura, M., Fujita, A., Shimada, S. & Ando, M. A small persistent locked area associated with the 2011 Mw9.0 Tohoku-Oki earthquake, deduced from GPS data. *Journal of Geophysical Research: Solid Earth* **117** (2012).

30. Lay, T. & Kanamori, H. *An Asperity Model of Large Earthquake Sequences*, 579–592 (American Geophysical Union (AGU), 2013).

31. Ruiz, S. *et al.* Intense foreshocks and a slow slip event preceded the 2014 Iquique Mw 8.1 earthquake. *Science* **345**, 1165–1169 (2014).

32. Nealy, J. L. *et al.* 2017 Valparaso earthquake sequence and the megathrust patchwork of central Chile. *Geophysical Research Letters* **44**, 8865–8872 (2017).

33. Liu, S., Tang, C., Chen, C. & Xu, R. Spatiotemporal evolution of the 2018 Mw 6.4 Hualien Earthquake Sequence in Eastern Taiwan. *Seismological Research Letters* **90**, 1446–1456 (2019).

34. Improta, L. e. a. Multi-segment rupture of the 2016 Amatrice-Visso-Norcia seismic sequence (central Italy) constrained by the first high-quality catalog of early aftershocks. *Scientific Reports* **9** (2019).

35. Ross, Z. E. *et al.* Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. *Science* **366**, 346–351 (2019).

36. Yue, H. *et al.* The 2016 Kumamoto Mw = 7.0 earthquake: A significant event in a fault-volcano system. *Journal of Geophysical Research: Solid Earth* **122**, 9166–9183 (2017).

37. Nanjo, K. Z., Hirata, N., Obara, K. & Kasahara, K. Decade-scale decrease in b value prior to the M9-class 2011 Tohoku and 2004 Sumatra quakes. *Geophysical Research Letters* **39** (2012).

38. Nanjo, K. Were changes in stress state responsible for the 2019 Ridgecrest, California, earthquakes? *Nature Communications* **11**, 3082 (2020).

39. Nanjo, K. Z. & Yoshida, A. Anomalous decrease in relatively large shocks and increase in the p and b values preceding the April 16, 2016, M7.3 earthquake in Kumamoto, Japan. *Earth Planets Space* **69**, 13 (2017).

Author contributions E.L. conceived the study, data analysis has been performed by G.P and E.L. All authors have contributed to write the manuscript.

Competing Interests The Authors declare no competing interests.

Materials & Correspondence Correspondence and material requests can be addressed to E.L. (eugenio.lippiello@unicampania.it).

Acknowledgements We thank Eleftheria Papadimitriou and Vassilios Karakostas (Aristotle University of Thessaloniki) for data of Greek mainshocks. We thank the facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata, and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the National Science Foundation under Cooperative Agreement EAR-1261681. We thank the National Research Institute for Earth Science and Disaster Resilience (2019), NIED Hi-net, National Research Institute for Earth Science and Disaster Resilience, doi:10.17598/NIED.0003, for data of the Tohoku sequence. We thank the GEOFON Data Centre (1993): GEOFON Seismic Network. Deutsches GeoForschungsZentrum GFZ. Other/Seismic Network. doi:10.14470/TR560404, for data of the Iquique sequence. We thank the Institute of Earth Sciences, Academia Sinica, Taiwan, for data of the Hualien sequence. This research activity has been supported by the Program VAnviteLli pEr la RicErca: VALERE 2019, financed by the University of Campania ”L. Vanvitelli. We also acknowledge support from MIUR-PRIN project ”Coarse-grained description for non-equilibrium systems and transport phenomena (CO-NEST)” no. 201798CZL.

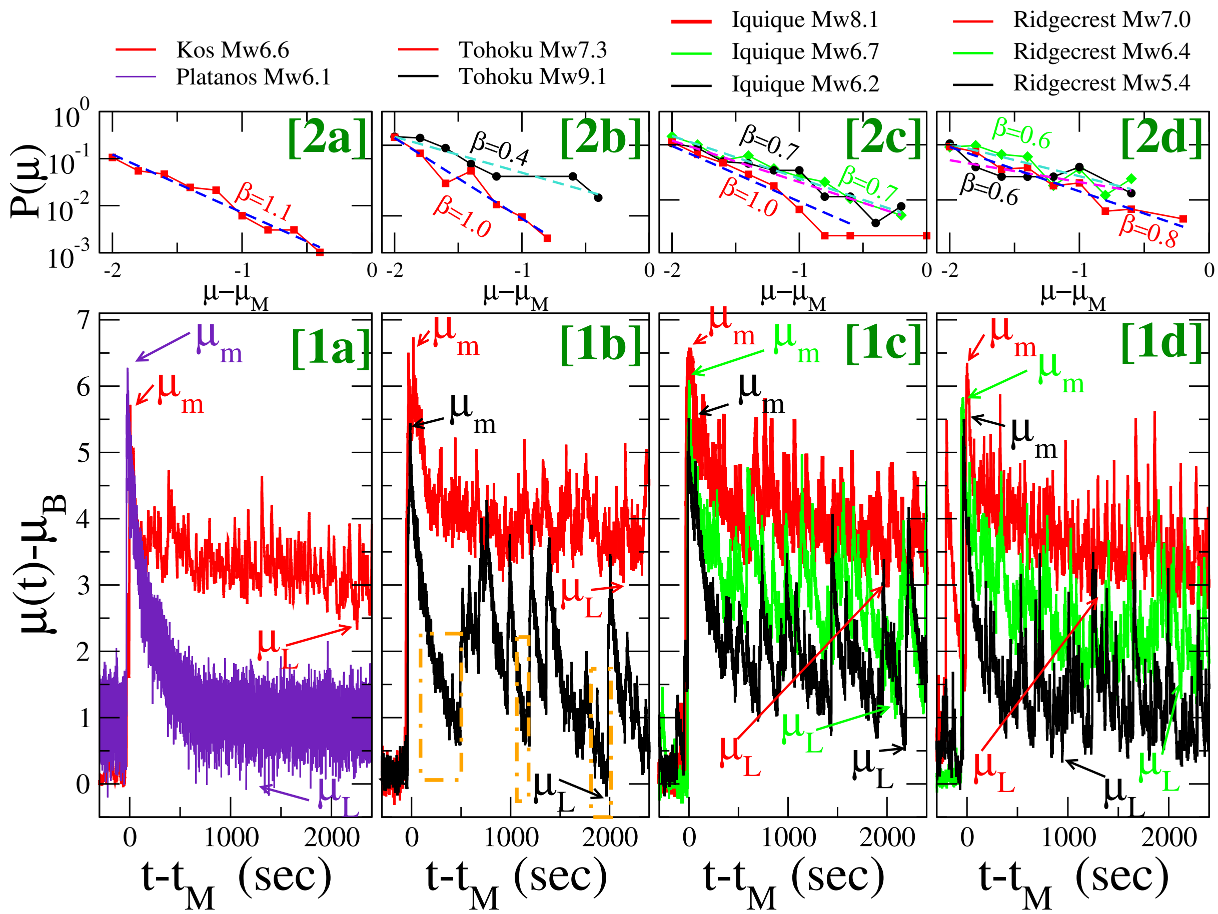

Figure 1: **The envelope function and the peak distribution.** The envelope function $\mu(t)$ is plotted as function of the time $t - t_M$ in the lower panels [1a-1d] and the corresponding peak distribution $P(\mu)$ is plotted in the upper panels [2a-2d]. Dashed lines in the upper panels are the exponential distribution $P(\mu) \propto 10^{-\beta\mu}$ and different colors correspond to the different best-fit β values. In panels [1a,2a] we consider two 'extreme' situations of normal aftershock triggering: The 2018 Mw6.6 Kos earthquake followed by a huge number of aftershocks (red curve [1a] and red squares [2a]) and the 2019 Mw6.1 Platanos earthquake followed by zero ($\mu_i > \mu_B$) aftershocks (violet curve [1a]). In panels [1b,2b] we consider the *2011 Tohoku* sequence. Red line (lower panel) and red squares (upper panel) are used for the Mw9.1 mainshock. Black line (lower panel) and black circles (upper panel) are used for the Mw7.3 foreshock. The orange dot-dashed rectangles identify temporal periods with no $\mu_i > \mu_B + 1$ aftershock .

Figure 1: In panels [1c,2c] we consider the *2014 Iquique* sequence. Red line (lower panel) and red squares (upper panel) are used for the Mw8.1 mainshock. Green line (lower panel) and green diamonds (upper panel) are used for the Mw6.7 foreshock. Black line (lower panel) and black circles (upper panel) are used for the Mw6.2 foreshock. In panels [1d,2d] we consider the *2019 Ridgecrest* sequence. Red line (lower panel) and red squares (upper panel) are used for the Mw7.0 mainshock. Green line (lower panel) and green diamonds (upper panel) are used for the Mw6.4 foreshock. Black line (lower panel) and black circles (upper panel) are used for the Mw5.4 foreshock

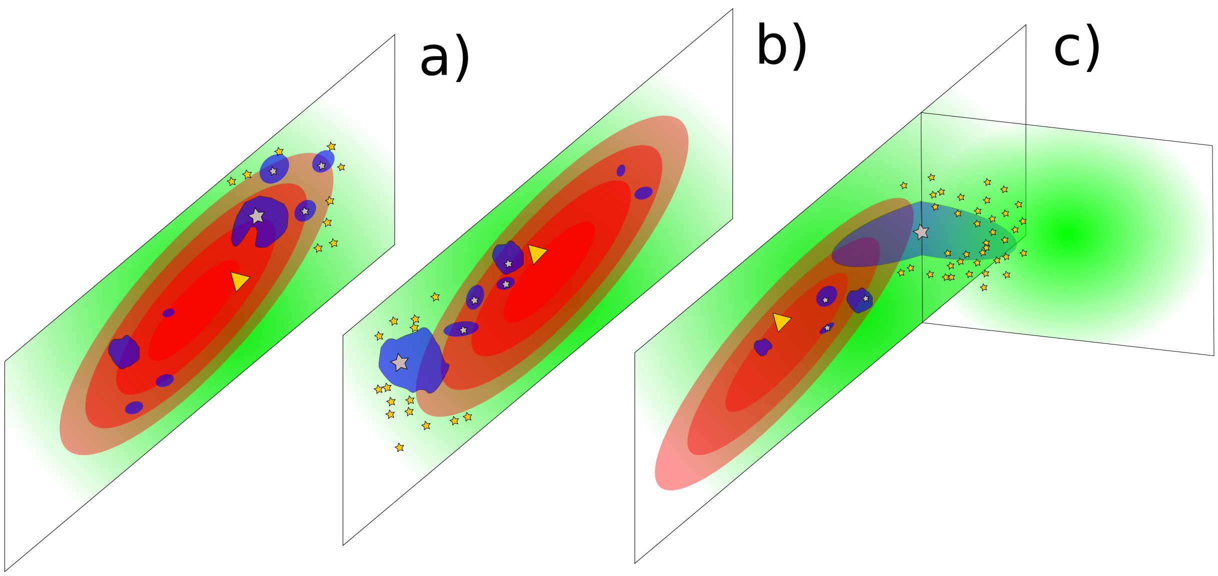

Figure 2: **Schematic description of the stress condition on the fault plane.** The color code indicates the degree of coupling of different regions on the fault plane, with deep red indicating very locked regions whereas light green indicates less coupled regions. The blue color indicates less locked region, inside the locked (red) area, where foreshock nucleation takes place. Grey stars, yellow stars and yellow triangles indicate the position of the epicenter of foreshocks, aftershocks and of the mainshock, respectively. Different panels schematically represent different scenarios, for the occurrence of the largest foreshock, qualitatively similar to the instrumental fore-mainshock sequences considered in Fig.1[1b-1d].

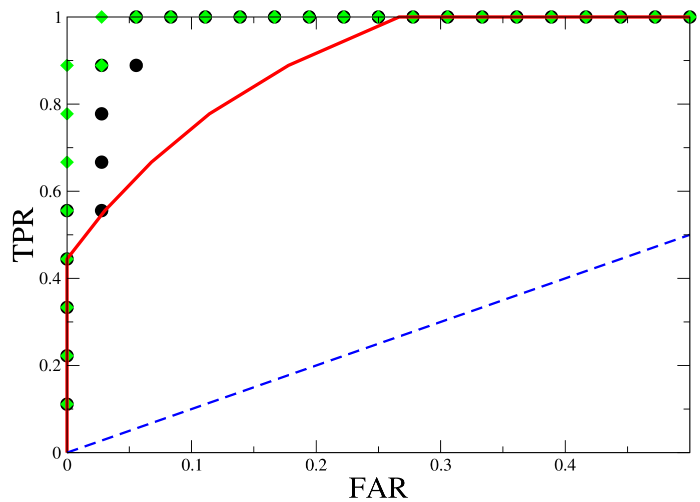

Figure 3: **The ROC diagram.** The true positive rate (TPR) is plotted versus the false alarm rate (FAR) for the method based on the Q-value with $\beta = 1$ (black circles) and with β extracted from data (green diamonds). The dashed blue diagonal represents random prediction (the null-hypothesis) whereas for points above the continuous red line, the null hypothesis can be rejected with a confidence level larger than 99.99999%.