EARTHQUAKES

Earthquake forecasting during the complex Amatrice-Norcia seismic sequence

Warner Marzocchi,* Matteo Taroni, Giuseppe Falcone

Earthquake forecasting is the ultimate challenge for seismologists, because it condenses the scientific knowledge about the earthquake occurrence process, and it is an essential component of any sound risk mitigation planning. It is commonly assumed that, in the short term, trustworthy earthquake forecasts are possible only for typical aftershock sequences, where the largest shock is followed by many smaller earthquakes that decay with time according to the Omori power law. We show that the current Italian operational earthquake forecasting system issued statistically reliable and skillful space-time-magnitude forecasts of the largest earthquakes during the complex 2016–2017 Amatrice-Norcia sequence, which is characterized by several bursts of seismicity and a significant deviation from the Omori law. This capability to deliver statistically reliable forecasts is an essential component of any program to assist public decision-makers and citizens in the challenging risk management of complex seismic sequences.

INTRODUCTION

On 24 August 2016, a magnitude 6 (M6) earthquake hit close to the city of Amatrice, causing severe damage and killing about 300 inhabitants. This earthquake was not anticipated by any increase in seismic activity, and it initiated a very complex seismic sequence that had several bursts of seismicity and a complex spatial pattern (Fig. 1), with more than 50,000 earthquakes recorded. After 2 months of a typical aftershock sequence characterized by many smaller earthquakes, in late October, the northern portion of the sequence was hit by M5.9 (on October 26) and M6.5 (on October 30) earthquakes that caused significant damage to a vast area around the city of Norcia, destroying the famous 14th century basilica of San Benedetto that had withstood earthquakes for more than six centuries. Later in January, four M5+ earthquakes occurred close in time and space in the southern part of the sequence, raising concerns for the potential impact on the nearby Campotosto dam that forms the second largest man-made lake in Europe.

This complex behavior is significantly different from the typical mainshock-aftershock sequences, where one large earthquake is followed by many smaller aftershocks. The statistics of typical mainshock-aftershock sequences is successfully described by the model of Reasenberg and Jones (R&J) (1), which has been widely used in aftershock forecasting since the 1989 Loma Prieta earthquake (1, 2). The R&J model is rooted in the idea that aftershocks from a single mainshock are distributed according to the Omori-Utsu (3, 4) and Gutenberg-Richter (5) empirical laws, thus contemplating the possibility that aftershocks can be larger than the mainshock. Although the R&J model gained wide acceptance in the seismological community, it cannot forecast the evolution of more complex seismic sequences, such as the Amatrice-Norcia sequence, because it does not include spatial information, and it assumes that there is only one major quake that triggers the whole aftershock sequence. The limited applicability of the R&J model was illustrated during the recent sequence that hit Kumamoto region, Japan, in 2016. After the first M6.5 earthquake on April 14, the Japan Meteorological Agency (JMA) released aftershock forecasts for smaller earthquakes. A few days later, on April 16, an M7.3 earthquake hit the same region, killing more than 50 inhabitants, and JMA stopped issuance of

forecasts because the sequence did not conform to typical mainshock-aftershock sequence (6).

More recent clustering models, such as the epidemic-type aftershock sequence (ETAS) (7) and the short-term earthquake probability (STEP) (8) models, are based on the same empirical laws used by R&J, but they introduce two important improvements: the spatial component of the forecasts and the assumption that every earthquake above a certain magnitude can generate other earthquakes, not only the mainshock. The latter precludes any meaningful distinction among foreshocks, mainshocks, and aftershocks (9, 10), because they can be attributed, at best, only a posteriori when the sequence is over. Despite these conceptual improvements, the increase in forecasting skill with respect to R&J and the capability of these models to deliver statistically reliable real-time forecasts during complex seismic sequences are still mostly unknown. With this specific aim, we analyze here the prospective real-time forecasting performance of ETAS and STEP models during the Amatrice-Norcia complex seismic sequence.

In the last few years, the Seismic Hazard Center at Istituto Nazionale di Geofisica e Vulcanologia (INGV) introduced an operational earthquake forecasting (OEF) system (OEF_Italy) (11, 12) that provides continuous authoritative information about time-dependent seismic hazards to the Italian Civil Protection. Currently, the system is still in a pilot phase. OEF_Italy consists of an ensemble of three different clustering models, two different flavors of ETAS (13, 14), and one STEP model (15). The ensemble modeling weighs the forecasts of each single model according to its past forecasting performance (16) and allows proper estimation of the epistemic uncertainty that is of paramount importance for testing (17) and for communicating uncertainties to the decision-makers. Here, we analyze the weekly earthquake forecasts for the complex seismic sequence still ongoing in Central Italy (Fig. 1). The weekly forecasts are updated every week or after any earthquake of M4.5 or above. Figure 2 shows the forecasts after (Fig. 2A) and before (Fig. 2, B to D) the most important earthquakes of the sequence. Notably, although each earthquake has an isotropic triggering spatial kernel in OEF_Italy models, the sum of the spatial triggering of all earthquakes yields a more heterogeneous spatial pattern. (All weekly forecasts used in this analysis are available upon request to the corresponding author.)

To evaluate the scientific reliability of the forecasts, we test whether the forecasts satisfactorily describe the space-time-magnitude distribution of the 40 target earthquakes (1 of $M \geq 6$, 6 of $5 \leq M < 6$, and

Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy.
*Corresponding author. Email: warner.marzocchi@ingv.it
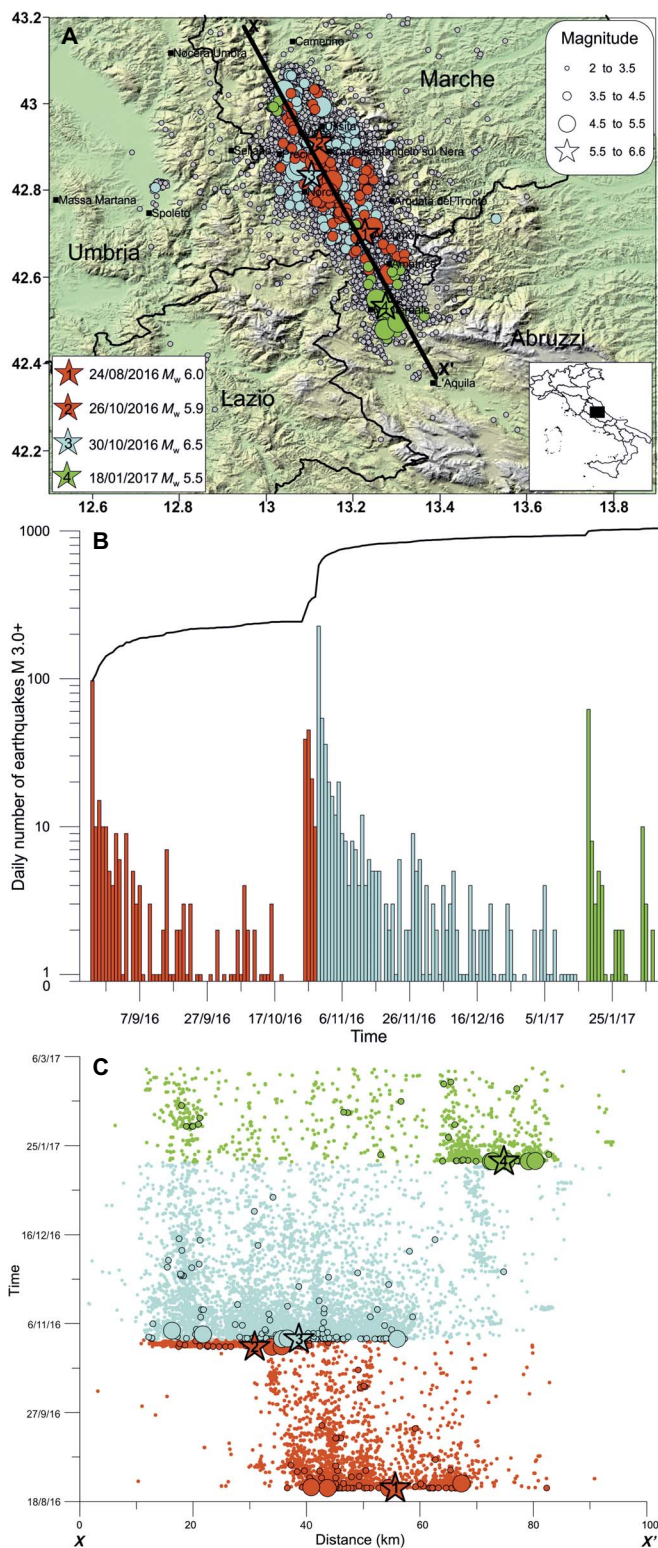
Fig. 1. Space-time evolution of the Amatrice-Norcia seismic sequence. (**A**) Seismicity above M3.5 observed during the sequence. Red, seismicity from 24 August to 29 October 2016; light blue, seismicity from 30 October 2016 to 17 January 2017; green, seismicity from 18 January to 31 January 2017. (**B**) Time evolution of the daily number of earthquakes of M3 or larger in the whole region. (**C**) Plot of the space-time evolution of the M2.5+ seismicity projected along the axis X-X' shown in Fig. 1A.

33 of $4 \leq M < 5$), which are the largest earthquakes that occurred in the period between 24 August 2016 to 31 January 2017. We do not consider as target earthquakes the few events with $M \geq M4$ that occurred in the first few hours after the two largest Amatrice and Norcia earthquakes. (We come back to this point later.) To test the space and time distribution, we use the S- and N-tests proposed in the Collaboratory for the Study of Earthquake Predictability (CSEP) framework (18, 19). The S-test (20) verifies whether the spatial occurrence of target earthquakes is consistent with the forecasts of the model, that is, if the empirical spatial distribution of the target earthquakes is not statistically different from the spatial distribution of the forecasts; the N-test (21) verifies whether the overall number of earthquakes predicted by the model is consistent with the number of earthquakes observed. More technical details on the S- and N-tests are provided in Materials and Methods. For magnitude distribution, we test whether the magnitude of the observed target earthquakes fits the Gutenberg-Richter law that is used by all models in OEF_Italy. Specifically, we test whether the magnitudes have an exponential distribution, that is, if $x \sim \exp(\Lambda)$, where $x = M - M_t + \varepsilon$, with M as the magnitude of the observed target earthquakes, $M_t = 3.95$ as the minimum magnitude for the target earthquakes, Λ as the average of x, and ε as a random uniform noise to avoid the binning problem of the magnitudes. This goodness of fit is carried out using the Lilliefors test (22).

RESULTS

Figure 3A shows the results of the S-test for the different forecasts. The spatial probability of each target earthquake is given by the probability of the last time window that begins before the occurrence time of the target earthquake. The time of each forecast is reported in Table 1. The length of these forecasts is not an issue for the S-test, because the spatial probability of each target earthquake is normalized to the number of earthquakes that occurred in the time window. The results show good agreement between spatial forecasts and locations of the target earthquakes; the quantile score ξ, which mimics the P value of the test (Eq. 1), is always larger than 0.2 and often close to 1, which indicates optimal spatial forecasting.

Figure 3B shows the expected number of target earthquakes for each forecast. Owing to the well-known fact that clustering models based on ETAS severely underestimate the number of earthquakes in the hours immediately after a large shock because of the significant incompleteness of the seismic catalog (23), the statistical tests include neither the forecasts nor any target earthquakes immediately after the M6 Amatrice earthquake and after the M6.5 Norcia earthquake (forecast numbers 1, 2, and 19). These forecasts were corrected in real time with empirical factors before sending them to the Italian Civil Protection. From the figure, we note that there is one significant underestimate for forecast 18, which corresponds to the first forecast after the M5.9 earthquake in Visso on October 26. Being less than M6, we did not correct the OEF_Italy forecast; therefore, an underestimate is not a surprise, for the same reasons discussed above. The forecasts of Fig. 3B overlap in time, posing challenges for statistical testing. To avoid this problem, we rescale each weekly forecast for the real-time length of the forecast using a function proportional to $1/(t + c)$ (see Materials and Methods for more details). Figure 3C shows the results of the N-test, that is, that the cumulative number of observed target earthquakes (40) lies within the 95% confidence interval (7 to 46). Figure 3D shows the agreement between the magnitudes of the target earthquakes and the Gutenberg-Richter frequency-magnitude distribution of the forecasts; the P value
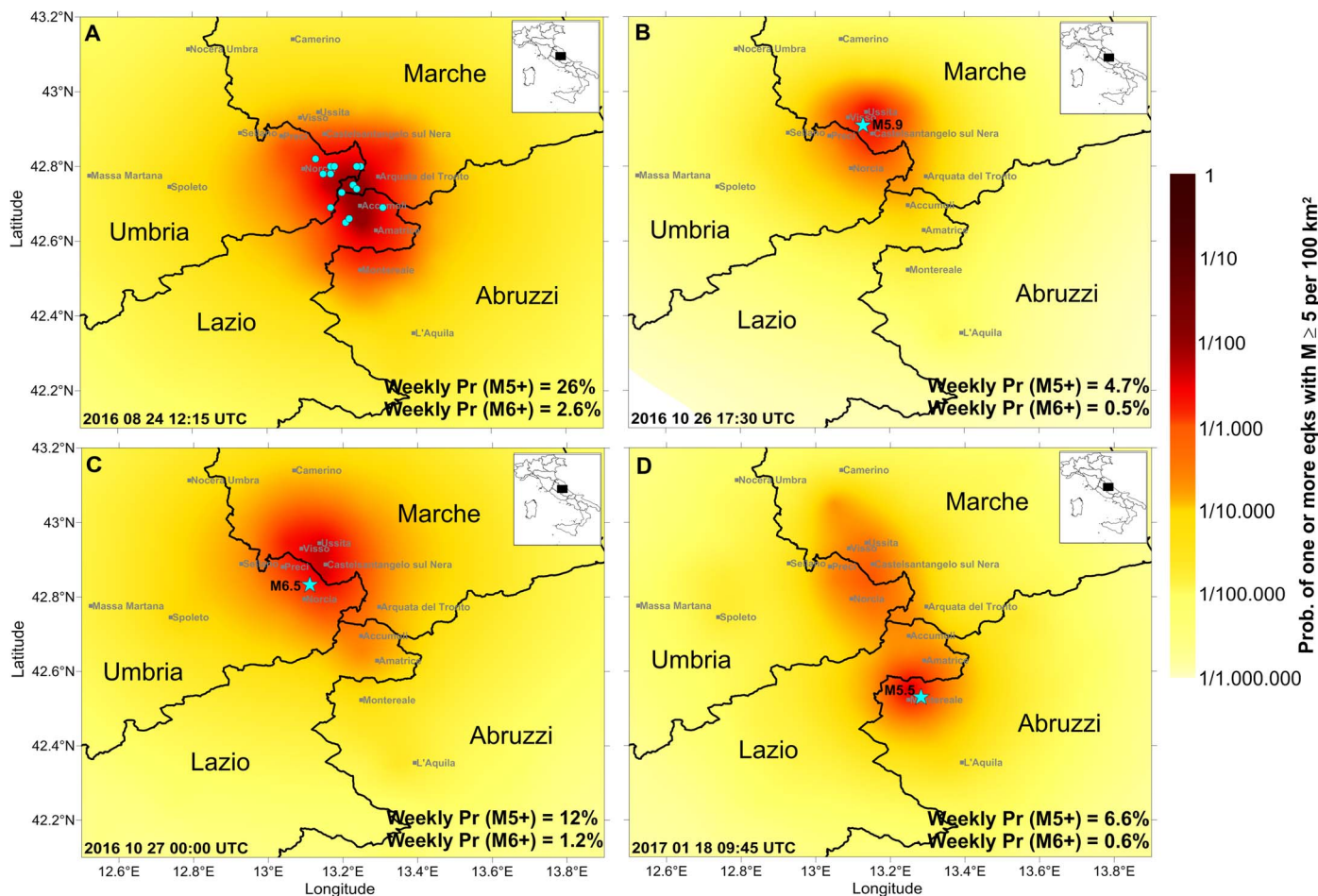
Fig. 2. Some examples of weekly forecasts (the number of the forecasts are reported on Table 1). (**A**) Forecast number 3, a few hours after the Amatrice earthquake and the M3.5+ earthquakes (blue-green circles) that occurred in the forecasting time window. (**B**) Forecast number 15, before the M5.9 earthquake (blue-green star) that occurred on October 26. (**C**) Forecast number 18, before the Norcia M6.5 earthquake (blue-green star) that occurred on October 30. (**D**) Forecast number 35, before the Campotosto M5.5 earthquake (blue-green star) that occurred on January 18.

of the Lilliefors test is larger than 0.05, which is the significance level usually adopted by the CSEP experiments (*19–21*).

In summary, this prospective testing phase shows a good agreement between the forecasts and the observed space-time-magnitude distribution of the largest earthquakes that occurred during the sequence. The skill and reliability with which OEF_Italy tracks the spatial evolution of the sequence can be qualitatively grasped by looking at Fig. 1C; besides showing the spatial pattern of the evolution of the sequence, the plot shows that the largest earthquakes of the sequence (marked by stars) occurred in areas that are characterized by spatial clustering of the preceding seismicity.

DISCUSSION

Statistically reliable and skillful OEF forecasts are a prerequisite for any practical use or dissemination of this information (*12*). However, the conversion of OEF forecasts in terms of risk reduction strategies is challenging and is still a matter of discussion (*24, 25*). To illustrate this general problem for the Central Italy sequence in Fig. 4, we show the time evolution of the weekly exceedance probability of macroseismic intensities VII and VIII in a circle of 10 km around Norcia before the

M6.5 earthquake on October 30 that severely damaged this historical city. Figure 4 (top) shows the long-term background probability, including the increase in probability due to the nearby M6.2 L'Aquila earthquake that occurred in 2009 at about 55 km from Norcia. In Fig. 4 (bottom), we show that weekly earthquake probability increases to 100 to 1000 times that of background before the M6.5 earthquake, achieving weekly probabilities of about 5% (for intensity VII) and slightly less than 1% (for intensity VIII) a few days before the M6.5 event. Although the probability of such a ground shaking occurrence is a few percent, the associated risk can be large, very likely above any a priori–defined acceptable risk threshold (*26*). As is the case for any kind of low-probability high-impact events, management of this risk poses a great challenge to the wide range of possible decision-makers (*27*); but despite its unpredictability, the risk can hardly be considered negligible or irrelevant. For this reason, we foresee in the near future a clear urgency to improve communication strategies to make OEF messages comprehensible to any possible interested stakeholders.

Besides the practical implications of seismic risk reduction for society, the capability to accurately forecast the evolution of the natural phenomena is the distinctive feature of science with respect to all other
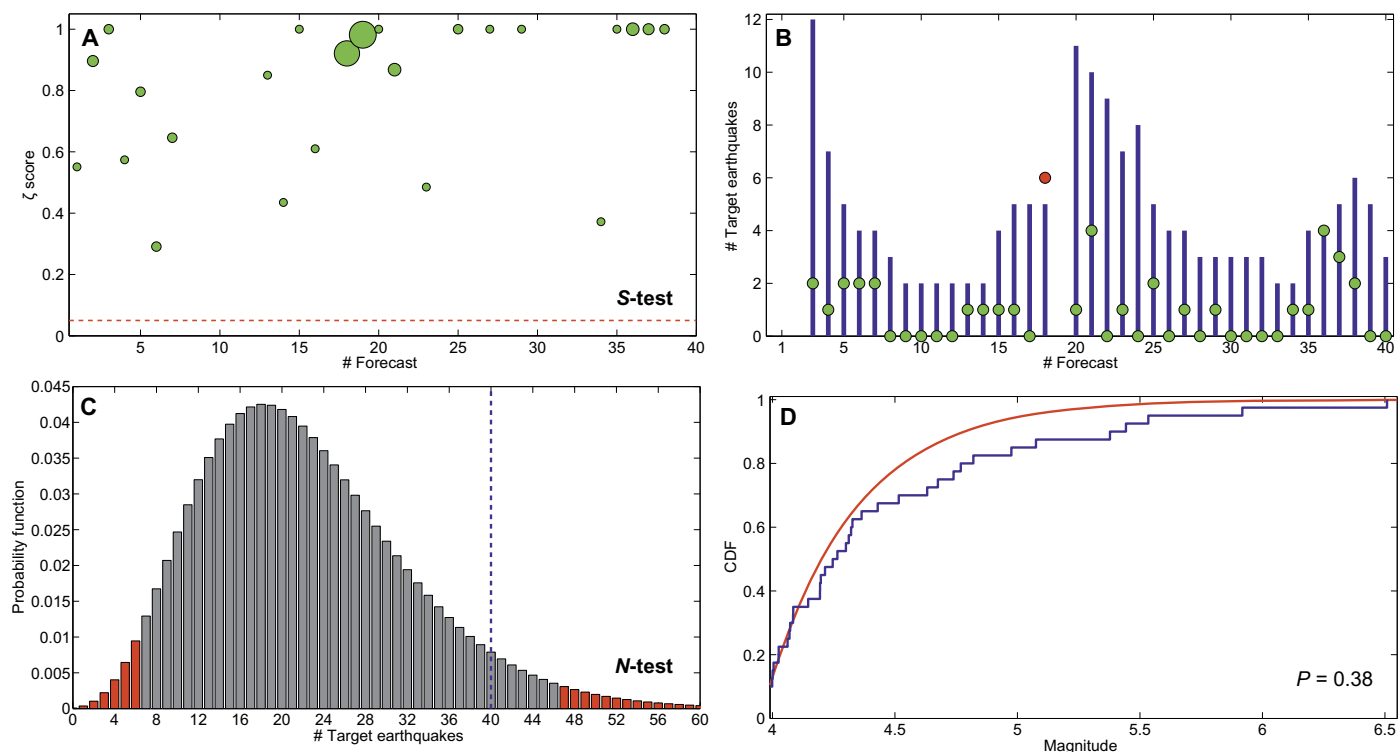
Fig. 3. Results of the statistical tests. (**A**) The quantile score ζ for each forecast of Table 1. Dot size is proportional to the number of target earthquakes that occurred during the forecasting time window; the red dashed horizontal line is the critical value ($\zeta = 0.05$) for rejection. (**B**) Expected number of target earthquakes for each forecast, calculated with negative binomial distribution (NBD) at 99% (see Materials and Methods). The green and red dots represent the following observations: green if inside the expected range and red if otherwise. (**C**) Probability distribution of the expected number of target earthquakes with a 95% confidence interval (red tails) and the observed number of target earthquakes (vertical dashed blue line). (**D**) Plot of the forecast and observed frequency-magnitude distribution. The P value of the Lilliefors test is reported inside the plot. These results show that the space-time-magnitude distribution of the forecasts and of the target earthquakes is in agreement.

human enterprises (28). OEF advances a valuable scientific approach, because it describes what we really know about the earthquake occurrence process, and it represents a benchmark for measuring any scientific improvement through a prospective statistical testing phase. Prospective statistical testing is the scientific gold standard for validating models (17), showing where models can be improved, and allowing scientists to quantify the forecasting ability of a new model compared to that of models currently in use. This framework modifies the way in which the earthquake predictability problem has been approached by scientists and perceived by society, moving from the tacit dichotomy for which earthquakes either can be predicted at a probability close to 1 or cannot be predicted at all toward a more incremental continuum of probabilities between 0 and 1, traditionally termed forecasts.

Although the results reported here demonstrate that current OEF models can provide statistically reliable and skillful forecasts even during complex seismic sequences, OEF is still in a nascent stage. The steady accumulation of scientific knowledge, which is implicit in the OEF's brick-by-brick approach (18), may eventually pave the way to a "quiet revolution" in earthquake forecasting as in weather forecasting (29). Without pretending to be exhaustive, we foresee that significant improvement may come from including complete geodetic and/or geological information (30), a better physical description of the earthquake-generating process and fault interaction, and the identification of possible physical and/or empirical features that may shed light on

the so-far elusive distinction between the preparatory phase of small and large earthquakes.

MATERIALS AND METHODS
S- and N-tests
The S-test (20) is a likelihood test applied to a forecast that has been normalized to the observed number of target earthquakes, thereby isolating the spatial distribution of the forecast. After normalizing each forecast, the S-test is summarized by a quantile score

$$\zeta = \frac{n\{S_i | S_i \leq S, \ S_i \in S_s\}}{n\{S_s\}} \qquad (1)$$

where S_i is the ith simulated spatial likelihood according to the model, S_s is the whole set of simulated spatial likelihoods, S is the likelihood of the spatial forecast relative to the observed seismicity, and $n\{A\}$ indicates the number of elements in a set $\{A\}$. The score ζ indicates the percentage of times in which the simulated spatial likelihoods are less or equal to the spatial likelihood observed, mimicking a statistical P value. If ζ is below the critical threshold value, say 0.05, the spatial forecast is deemed inconsistent; values close to 1 indicate optimal spatial forecasts.

Table 1. Forecast number (as used in Figs. 2 and 3) and the starting date. UTC, Coordinated Universal Time.

Forecast	Starting date	Forecast	Starting date
1*	24 August 2016, 0200 UTC	21	30 October 2016, 1345 UTC
2*	24 August 2016, 0300 UTC	22	01 November 2016, 0815 UTC
3	24 August 2016, 1215 UTC	23	03 November 2016, 0000 UTC
4	25 August 2016, 0000 UTC	24	03 November 2016, 0100 UTC
5	25 August 2016, 0345 UTC	25	10 November 2016, 0000 UTC
6	26 August 2016, 0445 UTC	26	17 November 2016, 0000 UTC
7	01 September 2016, 0000 UTC	27	24 November 2016, 0000 UTC
8	08 September 2016, 0000 UTC	28	01 December 2016, 0000 UTC
9	15 September 2016, 0000 UTC	29	08 December 2016, 0000 UTC
10	22 September 2016, 0000 UTC	30	15 December 2016, 0000 UTC
11	29 September 2016, 0000 UTC	31	22 December 2016, 0000 UTC
12	06 October 2016, 0000 UTC	32	29 December 2016, 0000 UTC
13	13 October 2016, 0000 UTC	33	05 January 2017, 0000 UTC
14	20 October 2016, 0000 UTC	34	12 January 2017, 0000 UTC
15	26 October 2016, 1730 UTC	35	18 January 2017, 0945 UTC
16	26 October 2016, 1945 UTC	36	18 January 2017, 1015 UTC
17	26 October 2016, 2200 UTC	37	18 January 2017, 1030 UTC
18	27 October 2016, 0000 UTC	38	18 January 2017, 1345 UTC
19*	30 October 2016, 0700 UTC	39	19 January 2017, 0000 UTC
20	30 October 2016, 1215 UTC	40	26 January 2017, 0000 UTC

*The forecasts that were modified in real time with empirical corrections because of severe seismic catalog underreporting immediately after large earthquakes.

The N-test (21) verifies whether $N_{obs} \sim F(N_{forecast})$, that is, if the observed number of target earthquakes (N_{obs}) over the entire region samples the distribution $F(N_{forecast})$ that describes the expected total number of target earthquakes predicted by the OEF model ($N_{forecast}$). We assumed that the target earthquakes follow a generalized Poisson distribution with the time-varying parameter λ that has a gamma

distribution to describe the variability among the rates produced by each one of the three models in OEF_Italy. This gamma-Poisson mixture produces an NBD for the expected number of target earthquakes.

NBD of the number of target earthquakes

The mass function of the NBD can be written as

$$f(n; r, p) = \frac{\Gamma(r + n)}{n!\Gamma(r)} p^n (1 - p)^r \quad (2)$$

where $\Gamma(r)$ is the gamma function, n is a positive integer random variable (the number of target earthquakes), and $r > 0$, $p \in (0, 1)$ are the two parameters of the distribution. NBD can be derived from the gamma-Poisson mixture distribution

$$f(n; r, p) = \int_0^\infty f_1(n; \lambda) f_2\left(\lambda; r, \frac{1 - p}{p}\right) d\lambda \quad (3)$$

where $f_1(n; \lambda)$ is the Poisson distribution of the positive integer random variable n, and the parameter λ of the Poisson distribution is assumed to follow a gamma distribution

$$f_2(\lambda; k, \theta) = \frac{\lambda^{k-1} e^{-\frac{\lambda}{\theta}}}{\theta^k \Gamma(k)} \quad (4)$$

where $k, \theta > 0$ are the parameters of the gamma distribution that are related to the parameters r and p of NBD through

$$r = k \quad (5)$$

and

$$p = \frac{\theta}{1 + \theta} \quad (6)$$

The models included in OEF_Italy provide a set of different rates for each forecast with different weights, describing the past forecasting performance of each model. From these rates, we estimated the parameters of the gamma distribution using the method of moments. In particular, we computed the weighted mean $\tilde{\mu}$ and the weighted variance $\tilde{\sigma}^2$ of the three rates for each forecast, and then we used them to calculate the parameters k and θ of the gamma distribution, that is

$$k = \frac{\tilde{\mu}^2}{\tilde{\sigma}^2} \quad (7)$$

and

$$\theta = \frac{\tilde{\sigma}^2}{\tilde{\mu}} \quad (8)$$

Once the k and θ parameters are obtained, it is possible to estimate the parameters r and p of the NBD through Eqs. 5 and
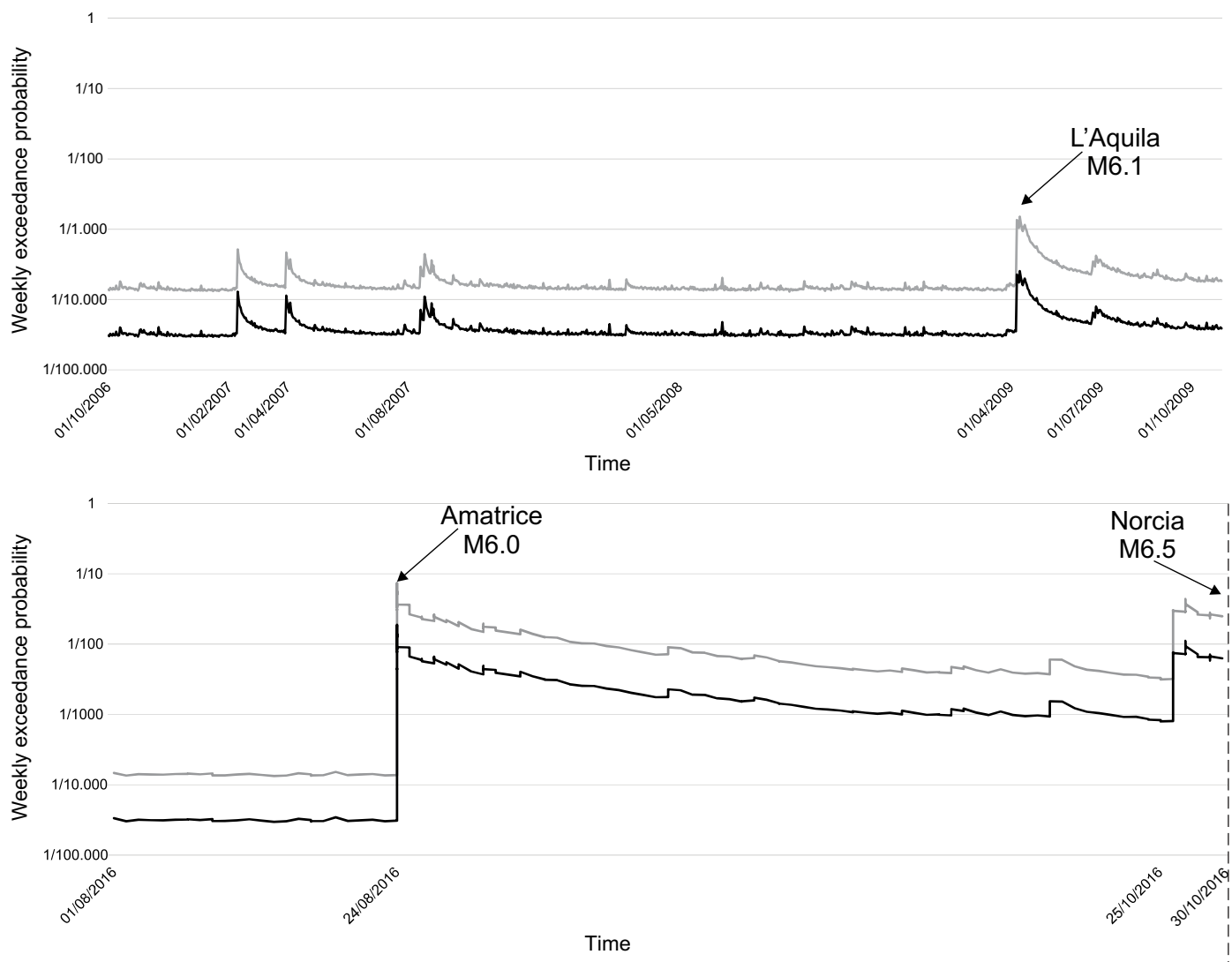
Fig. 4. Weekly exceedance probability for macrointensities VII (gray line) and VIII (black line) for a circular area around Norcia with a radius of 10 km. Top: A time period of 3 years including the L'Aquila earthquake (an M6.2 earthquake that occurred in 2009 at about 55 km of distance) (to give an idea about the background values). Bottom: A time period of few months, from 1 August to 29 October 2016, the day before the Norcia earthquake (M6.5). The increase in probability before the Norcia earthquake (bottom right) is caused by the increase in seismicity following the M5.4 earthquake that occurred on 26 October at the north of Norcia.

6. More information on the derivation of the NBD as mixture of Poisson distribution can be found in the study of Greenwood and Yule (*31*).

Correcting the weekly forecasts to account for shorter forecasting time windows

Weekly forecasts (that is, the expected number of target earthquakes in 1 week) have been updated every week or every earthquake of M4.5 or larger. This means that some weekly forecasts have been superseded by a new forecast before the end of the established 1-week forecasting time window (see Table 1). Thus, when testing the expected number of target earthquakes for the entire sequence (*N*-test, Fig. 3C), we had to correct each weekly forecast accounting for the real extent of the corresponding forecasting time window. For this purpose, we assumed that the earthquake rate λ decays with time within the forecasting window by

following a simple power law equation

$$\lambda = \frac{\text{constant}}{t + c} \tag{9}$$

where $c = 0.015$ day^{-1} (*13*). The expected number of target earthquakes in a generic forecasting time window T is

$$n_T = \int_0^T \lambda \, dt = \text{constant} \ln(1 + T/c) \tag{10}$$

Using Eq. 10, we can calculate the expected number of target earthquakes n_T for a generic forecasting time window of length T (shorter

than or equal to 7 days) from the weekly expected number of target earthquakes n_{week} given by the model

$$n_T = \frac{\ln(1 + T/c)}{\ln(1 + T/c)} \, n_{week} \qquad (11)$$

We have verified that replacing Eq. 9 with the more complex function

$$\lambda = \frac{\text{constant}}{(t + c)^p} \qquad (12)$$

did not change the results significantly for any realistic value of p estimated by Lombardi and Marzocchi (13).

REFERENCES AND NOTES

1. P. A. Reasenberg, L. M. Jones, Earthquake hazard after a mainshock in California. *Science* **243**, 1173–1176 (1989).
2. P. A. Reasenberg, L. M. Jones, California aftershock hazard forecasts. *Science* **247**, 345–346 (1990).
3. T. Utsu, A statistical study of the occurrence of aftershocks. *Geophys. Mag.* **30**, 521–605 (1961).
4. T. Utsu, Aftershocks and earthquake statistics (II)—Further investigation of aftershocks and other earthquake sequences based on a new classification of earthquake sequences. *J. Fac. Sci. Hokkaido Univ. Ser. 7* **3**, 197–266 (1970).
5. B. Gutenberg, C. F. Richter, Earthquake magnitude, intensity, energy and acceleration. *Bull. Seismol. Soc. Am.* **46**, 105–145 (1956).
6. N. Kamaya, N. Yamada, Y. Ishigaki, K. Takeda, H. Kuroki, S. Takahama, K. Moriwaki, M. Yamamoto, M. Ueda, T. Yamauchi, M. Tanaka, Y. Komatsu, K. Sakoda, N. Hirota, J.-. Suganomata, A. Kawai, Y. Morita, S. Annoura, Y. Nishimae, S. Aoki, N. Koja, K. Nakamura, G. Aoki, T. Hashimoto, "[MIS34-P01] Overview of the 2016 Kumamoto earthquake," Japan Geoscience Union Meeting, Chiba, Japan, 25 May 2016.
7. Y. Ogata, Space-time point-process models for earthquake occurrences. *Ann. Inst. Stat. Math.* **50**, 379–402 (1998).
8. M. C. Gerstenberger, S. Wiemer, L. M. Jones, P. A. Reasenberg, Real-time forecasts of tomorrow's earthquakes in California. *Nature* **435**, 328–331 (2005).
9. A. Helmstetter, D. Sornette, Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *J. Geophys. Res.* **107**, 2237 (2002).
10. K. R. Felzer, R. E. Abercrombie, G. Ekström, A common origin for aftershocks, foreshocks, and multiplets. *Bull. Seismol. Soc. Am.* **94**, 88–98 (2004).
11. W. Marzocchi, A. M. Lombardi, E. Casarotti, The establishment of an operational earthquake forecasting system in Italy. *Seismol. Res. Lett.* **85**, 961–969 (2014).
12. T. H. Jordan, Y.-T. Chen, P. Gasparini, R. Madariaga, I. Main, W. Marzocchi, G. Papadopoulos, G. Sobolev, K. Yamaoka, J. Zschau, Operational earthquake forecasting. State of knowledge and guidelines for implementation. *Ann. Geophys.* **54**, 315–391 (2011).
13. A. M. Lombardi, W. Marzocchi, The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment. *Ann. Geophys.* **53**, 155–164 (2010).
14. G. Falcone, R. Console, M. Murru, Short-term and long-term earthquake occurrence models for Italy: ETES, ERS and LTST. *Ann. Geophys.* **53**, 41–50 (2010).
15. J. Woessner, A. Christophersen, J. D. Zechar, D. Monelli, Building self-consistent, short-term earthquake probability (STEP) models: Improved strategies and calibration procedures. *Ann. Geophys.* **53**, 141–154 (2010).
16. W. Marzocchi, J. D. Zechar, T. H. Jordan, Bayesian forecast evaluation and ensemble earthquake forecasting. *Bull. Seismol. Soc. Am.* **102**, 2574–2584 (2012).
17. W. Marzocchi, T. H. Jordan, Testing for ontological errors in probabilistic forecasting models of natural systems. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11973–11978 (2014).
18. T. H. Jordan, Earthquake predictability, brick by brick. *Seismol. Res. Lett.* **77**, 3–6 (2006).
19. J. D. Zechar, D. Schorlemmer, M. Liukis, J. Yu, F. Euchner, P. J. Maechling, T. H. Jordan, The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science. *Concurr. Comput.* **22**, 1836–1847 (2010).
20. J. D. Zechar, M. C. Gerstenberger, D. A. Rhoades, Likelihood-based tests for evaluating space–rate–magnitude earthquake forecasts. *Bull. Seismol. Soc. Am.* **100**, 1184–1195 (2010).
21. D. Schorlemmer, M. C. Gerstenberger, S. Wiemer, D. D. Jackson, D. A. Rhoades, Earthquake likelihood model testing. *Seismol. Res. Lett.* **78**, 17–29 (2007).
22. H. W. Lilliefors, On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *J. Am. Stat. Assoc.* **64**, 387–389 (1969).
23. W. Marzocchi, A. M. Lombardi, Real-time forecasting following a damaging earthquake. *Geophys. Res. Lett.* **36**, L21302 (2009).
24. K. Wang, G. C. Rogers, Earthquake preparedness should not fluctuate on a daily or weekly basis. *Seismol. Res. Lett.* **85**, 569–571 (2014).
25. T. H. Jordan, W. Marzocchi, A. J. Michael, M. C. Gerstenberger, Operational earthquake forecasting can enhance earthquake preparedness. *Seismol. Res. Lett.* **85**, 955–959 (2014).
26. W. Marzocchi, I. Iervolino, M. Giorgio, G. Falcone, When is the probability of a large earthquake too small? *Seismol. Res. Lett.* **86**, 1674–1678 (2015).
27. E. H. Field, T. H. Jordan, L. M. Jones, A. J. Michael, M. L. Blanpied, The potential uses of operational earthquake forecasting. *Seismol. Res. Lett.* **87**, 313–322 (2016).
28. American Association for the Advancement of Science (AAAS), *Science for All Americans: A Project 2061 Report on Literacy Goals in Science, Mathematics, and Technology* (AAAS, 1989).
29. P. Bauer, A. Thorpe, G. Brunet, The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).
30. E. H. Field, K. R. Milner, J. L. Hardebeck, M. T. Page, N. van der Elst, T. H. Jordan, A. J. Michael, B. E. Shaw, M. J. Werner, A spatiotemporal clustering model for the third Uniform California Earthquake Rupture Forecast (UCERF3-ETAS): Toward an operational earthquake forecast. *Bull. Seismol. Soc. Am.* **107**, 1049–1081 (2017).
31. M. Greenwood, G. U. Yule, An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference of multiple attacks of disease or of repeated accidents. *J. R. Stat. Soc.* **83**, 255–279 (1920).