nature COMMUNICATIONS

COMMENT

https://doi.org/10.1038/s41467-021-24952-6 OPEN

Machine learning and earthquake forecasting—next steps

Gregory C. Beroza[1]✉, Margarita Segou[2] & S. Mostafa Mousavi[1]

A new generation of earthquake catalogs developed through supervised machine-learning illuminates earthquake activity with unprecedented detail. Application of unsupervised machine learning to analyze the more complete expression of seismicity in these catalogs may be the fastest route to improving earthquake forecasting.

The past 5 years have seen a rapidly accelerating effort in applying machine learning to seismological problems. The serial components of earthquake monitoring workflows include: detection, arrival time measurement, phase association, location, and characterization. All of these tasks have seen rapid progress due to effective implementation of machine-learning approaches. They have proven opportune targets for machine learning in seismology mainly due to the large, labeled data sets, which are often publicly available, and that were constructed through decades of dedicated work by skilled analysts. These are the essential ingredient for building complex supervised models. Progress has been realized in research mode to analyze the details of seismicity well after the earthquakes being studied have occurred, and machine-learning techniques are poised to be implemented in operational mode for real-time monitoring. We will soon have a next generation of earthquake catalogs that contain much more information. How much more? These more complete catalogs typically feature at least a factor of ten more earthquakes (Fig. 1) and provide a higher-resolution picture of seismically active faults.

This next generation of earthquake catalogs will not be the single, static objects seismologists are accustomed to working with. For example, less than 2 years after the 2019 Ridgecrest, California earthquake sequence there already exist four next-generation catalogs, each of which were developed with different enhanced detection techniques. Now, and in the future, this will be the norm, and earthquake catalogs will be updated and improved—potentially dramatically—with time. Second-generation deep learning models[1] that are specifically designed based on earthquake signal characteristics and that mimic the manual processing by analysts, can lead to performance increases beyond those offered by earlier models that adapted neural network architectures from other fields. Those interested in using earthquake catalogs for forecasting can anticipate a shifting landscape with continuing improvements.

While these improvements are impressive, the value of the extra information they provide is less clear. What will we learn about earthquake behavior from these deeper catalogs and how might it improve the prospects for the stubbornly difficult problem of earthquake forecasting?

Short-term deterministic earthquake prediction remains elusive and is perhaps impossible; however, probabilistic earthquake forecasting is another matter. It remains the subject of focused and sustained attention and it informs earthquake hazard characterization[2] and thus both policy and earthquake risk reduction. A key assumption is that what we learn from the newly

[1] Department of Geophysics, Stanford University, Stanford, CA, USA. [2] British Geological Survey, Research Avenue South, Lyell Centre, Edinburgh, UK. ✉email: beroza@stanford.edu

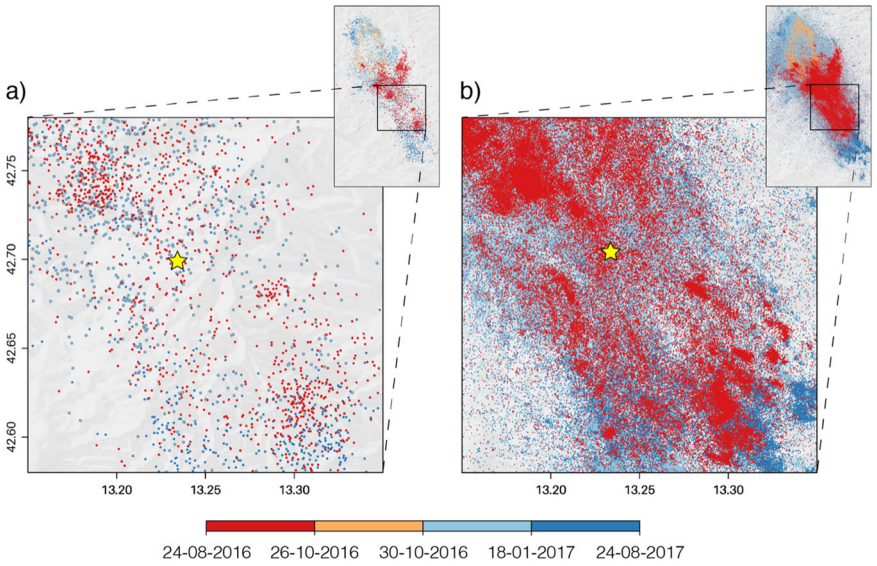

Fig. 1 A year of seismicity in the epicentral area of the 2016 $M = 6.0$ Amatrice earthquake (star) in Italy color coded by time of occurrence. a Real-time catalog, available at http://cnt.rm.ingv.it/ and (**b**) machine-learning catalog[16] are shown for event magnitudes above their respective magnitude of completeness[12,16] Mc = 2.2 and Mc = 0.5.

uncovered small earthquakes in AI-based catalogs, will inform earthquake forecasting for events of all magnitudes. The observed scale invariance of earthquake behavior suggests this is a reasonable expectation.

Empirical seismological relationships have played a key role in the development of earthquake forecasting. These include Omori's law[3] that describes the temporal decay of aftershock rate, the magnitude-frequency distribution, with the b-value describing the relative numbers of small vs. large earthquakes[4], and the Epidemic Type Aftershock Sequence (ETAS) model[5] in which earthquakes are treated as a self-exciting process governed by Omori's law for their frequency of occurrence and Gutenberg–Richter statistics for their magnitude. These empirical laws continue to prove their utility. Just in the past few years, the time dependence of the b-value has been used to try to anticipate the likelihood of large earthquakes during an ongoing earthquake sequence[6] and the ETAS model has been improved to better anticipate future large events[7]. So it appears that there is value in applying these longstanding relationships to improved earthquake catalogs, but our opinion is that much more needs to be done.

The relationships cited above date from 127, 77, and 33 years ago. The oldest of them, Omori's Law, was developed based on felt reports without the benefit of instrumental measurements. We suggest that a fresh approach using more powerful techniques is warranted. Earthquake catalogs are complex, high-dimensional objects and as Fig. 1 makes clear, that is even more true for the deeper catalogs that are being developed through machine learning. Their high dimensionality makes them challenging for seismologists to explore, and the conventional approaches noted above seem unlikely to be taking advantage of the wealth of new information available in the new generation of deeper catalogs. We suggest that, having first enabled the development of these catalogs, the statistical-learning techniques of data science are now poised to play an important role in uncovering new relationships within them. The obvious next step is to apply the techniques of machine learning in discovery mode[8] to discern new relationships encoded in the seismicity.

There are tantalizing indications that such an approach may lead to new insights. In double-direct-shear experiments, background signals that were thought to be uninformative random noise have instead been shown to encode information on the state of friction and the eventual time of failure of faults in a laboratory setting[9]. Well-controlled laboratory analogs to faults lack the geologic complexity of the Earth, yet, weak natural background vibrations of a similar sort, that again were thought to be random noise, have been shown to embody information that can be used to predict the onset time of slow slip events in the Cascadia subduction zone[10]. Finally, unsupervised deep learning, in which algorithms are used to discern patterns in data without the benefit of prior labels, applied to seismic waveform data uncovered precursory signals preceding the large and damaging 2017 landslide and tsunami in Greenland[11].

These examples are compelling but come with the caveat that they are not representative of the typical fast rupture velocity earthquakes on tectonic faults that are of societal concern. For such earthquakes, however, there are also indications from state-of-the-art forecasting approaches that next-generation earthquake catalogs may contain information that will lead to progress. Physics-based forecasting models, which account for changes in the Coulomb failure stress due to antecedent earthquakes that favor the occurrence of subsequent earthquakes, have shown increasing skill such that they are competitive with, and are beginning to outperform, statistical models. Coulomb failure models benefit particularly from deeper catalogs because they include many more small magnitude earthquakes. These small earthquakes add predictive power through their secondary triggering effects tracking the evolution of the fine-scale stress field that ultimately controls earthquake nucleation in foreshock and aftershock sequences. They can also be used to define the emerging active structures that comprise fault networks and by doing so clarify the relevant components of stress that would act to trigger earthquakes[12]. Secondary triggering and background stress heterogeneity were shown to improve stress triggering models[13] but were most effective when they incorporated near-real-time aftershock data from the sequence as it unfolded[14]. We note that there is no reason why more complete earthquake catalogs, developed with pre-trained neural network models, cannot be created in real time as an earthquake sequence unfolds. Finally, despite the disappointing history of the search for precursors, due diligence requires that seismologists consider the pursuit of signals that might be precursory.

We conclude that it is now possible to image the activity on active fault systems with unprecedented spatial resolution. This will enable experimentation with familiar hypotheses and enable the formulation of new hypotheses. It seems certain that the underlying processes that drive earthquake occurrence are encoded in this next generation of earthquake catalogs, but we may not find them unless we put new effort into searching for them. Unsupervised learning methods[15] are particularly well-suited tool for that effort.

Received: 28 May 2021; Accepted: 12 July 2021;
Published online: 06 August 2021

References

1. Mousavi, S. M. et al. Earthquake Transformer - an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* **11**, 3952, https://doi.org/10.1038/s41467-020-17591-w (2020).
2. Field, E. H. et al. A synoptic view of the third Uniform California Earthquake Rupture Forecast model. *Seismological Res. Lett.* **88**, 1259–1267, https://doi.org/10.1785/0220170045 (2017).
3. Omori, F. J. On the aftershocks of earthquakes. *Coll. Sci. Imp. Univ. Tokyo* **7**, 111–200 (1894).
4. Gutenberg, B. & Richter, C. F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **34**, 185–188 (1944).
5. Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* **83**, 9–27 (1988).
6. Gulia, L. & Wiemer, S. Real-time discrimination of earthquake foreshocks and aftershocks. *Nature* **574**, 193–199, https://doi.org/10.1038/s41586-019-1606-4 (2019).
7. Shcherbakov, R. et al. Forecasting the magnitude of the largest expected earthquake. *Nat. Commun.* **10**, 4051, https://doi.org/10.1038/s41467-019-11958-4 (2019).
8. Bergen, K. J. et al. Machine learning for data-driven discovery in solid earth geoscience. *Science* **363**, eaau0323, https://doi.org/10.1126/science.aau0323 (2019).
9. Rouet-Leduc, B. et al. Machine learning predicts laboratory earthquakes. *Geophys. Res. Lett.* **44**, 9276–9282 https://doi.org/10.1002/2017GL074677 (2017).
10. Hulbert, C. et al. An exponential build-up in seismic energy suggests a months-long nucleation of slow slip in Cascadia. *Nat. Commun.* **11**, 4139, https://doi.org/10.1038/s41467-020-17754-9 (2020).
11. Seydoux, L. et al. Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. *Nat. Commun.* **11**, 3972, https://doi.org/10.1038/s41467-020-17841-x (2020).
12. Mancini, S. et al. Improving physics-based aftershock forecasts during the 2016-2017 Central Italy earthquake cascade. *J. Geophys. Res. Solid Earth* **124**, https://doi.org/10.1029/2019JB017874 (2019).
13. Segou, M., & Parsons, T. A new technique to calculate earthquake stress transfer and to probe the physics of aftershocks. *Bull. Seismol. Soc. Am.* **110**, 863–873, https://doi.org/10.1785/0120190033 (2020).
14. Mancini, S. et al. The predictive skills of elastic Coulomb rate-and-state aftershock forecasts during the 2019 Ridgecrest, California, earthquake sequence. *Bull. Seismol. Soc. Am.* **110**, 1736–1751, https://doi.org/10.1785/0120200028 (2020).
15. Mousavi, S. M., et al. Unsupervised clustering of seismic signals using deep convolutional autoencoders. *IEEE Trans. Geosci. Remote Sens. Lett.* **16**, 1693–1697, https://doi.org/10.1109/LGRS.2019.2909218 (2019).
16. Tan, Y. J. et al. Machine-learning-based high-resolution catalog reveals how complex fault structures were activated during the 2016-2017 Central Italy sequence. *Seismic Rec.* **1**, 11–19, https://doi.org/10.1785/0320210001 (2021).

Acknowledgements

This work is supported by the NERC-NSFGEO 13 funded project The Central Apennines Earthquake Sequence under a New Microscrope (NE/R0000794/1). G.C.B. was supported by Department of Energy (Basic Energy Sciences; Award DE-SC0020445). Thanks to Dr. Simone Mancini for preparing the figure.

Author contributions

All authors contributed to the development of this comment, and all were involved in revising it.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.C.B.

Reprints and permission information is available at http://www.nature.com/reprints

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

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

© The Author(s) 2021