Electronic Seismologist

BPMF: A Backprojection and Matched-Filtering Workflow for Automated Earthquake Detection and Location

Eric Beaucé*[1], William B. Frank[2], Léonard Seydoux[3], Piero Poli[4], Nathan Groebner[5], Robert D. van der Hilst[2], and Michel Campillo[6]

Abstract

We introduce BPMF (backprojection and matched filtering)—a complete and fully automated workflow designed for earthquake detection and location, and distributed in a Python package. This workflow enables the creation of comprehensive earthquake catalogs with low magnitudes of completeness using no or little prior knowledge of the study region. BPMF uses the seismic wavefield backprojection method to construct an initial earthquake catalog that is then densified with matched filtering. BPMF integrates recent machine learning tools to complement physics-based techniques, and improve the detection and location of earthquakes. In particular, BPMF offers a flexible framework in which machine learning detectors and backprojection can be harmoniously combined, effectively transforming single-station detectors into multistation detectors. The modularity of BPMF grants users the ability to control the contribution of machine learning tools within the workflow. The computation-intensive tasks (backprojection and matched filtering) are executed with C and CUDA-C routines wrapped in Python code. This leveraging of low-level, fast programming languages and graphic processing unit acceleration enables BPMF to efficiently handle large datasets. Here, we first summarize the methodology and describe the application programming interface. We then illustrate BPMF's capabilities to characterize microseismicity with a 10 yr long application in the Ridgecrest, California area. Finally, we discuss the workflow's runtime scaling with numerical resources and its versatility across various tectonic environments and different problems.

Cite this article as Beaucé, E., W. B. Frank, L. Seydoux, P. Poli, N. Groebner, R. D. van der Hilst, and M. Campillo (2023). BPMF: A Backprojection and Matched-Filtering Workflow for Automated Earthquake Detection and Location, *Seismol. Res. Lett.* **95**, 1030–1042, doi: 10.1785/0220230230.

Supplemental Material

Introduction

Small-magnitude earthquakes occur much more frequently than large ones (Gutenberg–Richter law; Gutenberg and Richter, 1941), and therefore bring essential information about fault and crustal processes either in tectonic (e.g., Rubin *et al.*, 1999; Shelly *et al.*, 2007; Sánchez-Reyes *et al.*, 2021), volcanic (e.g., Scholz *et al.*, 2019; Shelly and Thelen, 2019), or induced seismicity (e.g., De Barros *et al.*, 2021; Kwiatek *et al.*, 2022) settings. Seismic noise and wave scattering within the crust may hide the signals of small earthquakes in the seismic records. Studying these small events is therefore limited by our capability to detect and locate them, which requires both good-quality data and modern detection techniques that go beyond the visual identification of earthquake signals by analysts. Thus, automating the earthquake catalog building process is a necessary step toward fully using the information encoded in the large and ever-growing volumes of seismic data (e.g., Arrowsmith *et al.*, 2022). This information will feed the next advances in Earth sciences from gaining detailed knowledge of physical processes at the local scale to understanding natural systems over long timescales. Here, we present BPMF (backprojection and matched filtering)—a new community tool to automatically build low magnitude of completeness

1. Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, U.S.A., https://orcid.org/0000-0003-3138-9082 (EB); 2. Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A., https://orcid.org/0000-0001-7892-3081 (WBF); https://orcid.org/0000-0003-1650-6818 (RDvdH); 3. Institut de physique du globe de Paris, Université Paris cité, Paris, France, https://orcid.org/0000-0002-6596-5896 (LS); 4. Dipartimento di Geoscienze, Università di Padova, Padova, Italy; 5. Strabo Analytics, Inc, New Ulm, Minnesota, U.S.A., https://orcid.org/0000-0002-6535-1739 (NG); 6. Institut des Sciences de la Terre, Université Grenoble Alpes, Grenoble, France, https://orcid.org/0000-0001-6971-4499 (MC)

*Corresponding author: ebeauce@ldeo.columbia.edu

© Seismological Society of America

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

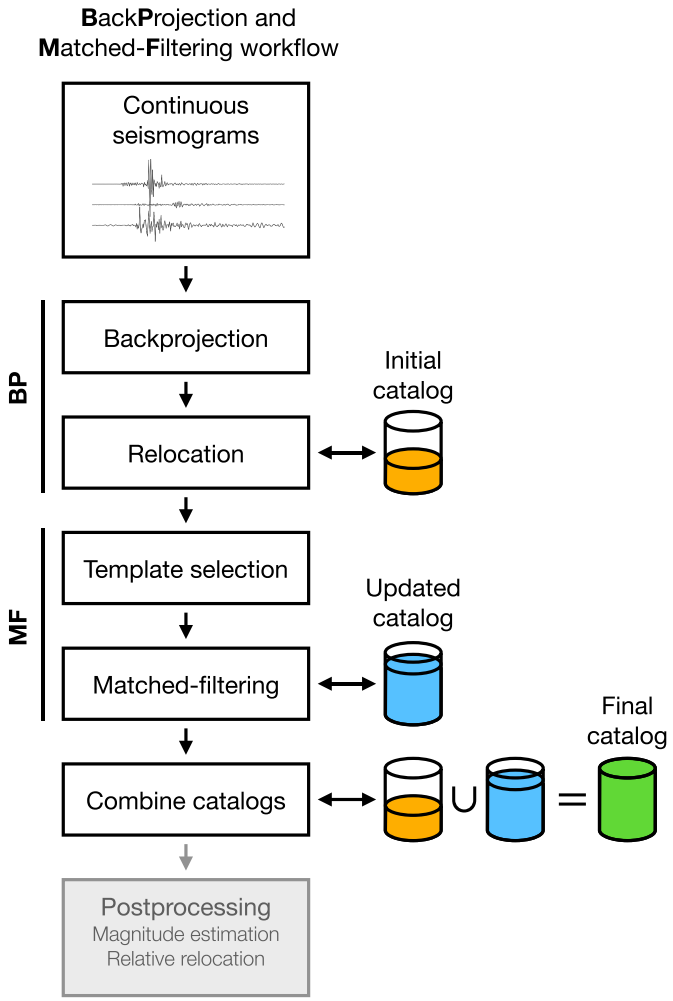


Figure 1. BPMF (backprojection and matched filtering) workflow. Multistation, multichannel continuous seismic data is first processed with a backprojection algorithm to build an initial catalog (backprojection catalog). Some of the detected events are selected to serve as template events that are then used in a matched-filter search to produce an updated catalog (matched filtering catalog). The backprojection and matched filtering catalogs are combined to make the final catalog. Additional postprocessing, such as magnitude estimation or relative relocation, may be achieved outside of the BPMF main workflow. The color version of this figure is available only in the electronic edition.

and short interevent time-resolution earthquake catalogs using large seismic datasets.

Our automated earthquake detection and location workflow, BPMF (Fig. 1), is based upon two seismic array processing techniques: backprojection (also called beamforming, e.g., Capon, 1970; Ringdal and Husebye, 1982; Rost and Thomas, 2002; Fan and Shearer, 2016) and matched filtering (also called template matching, e.g., Gibbons and Ringdal, 2006; Shelly *et al.*, 2007; Beaucé *et al.*, 2018); it also includes the recent phase-picking machine learning technologies (e.g., Zhu and Beroza, 2019; Mousavi *et al.*, 2020; Zhou *et al.*, 2022). Backprojection emerged in the second-half of the twentieth century as a powerful tool to detect, locate, and characterize earthquake sources (Capon, 1970; Ringdal and Husebye, 1982) in the context of nuclear test ban monitoring. More recently, it has been used to find aftershocks hidden in the coda of large earthquakes (Kiser and Ishii, 2013; Fan and Shearer, 2016) that conventional methods like the short-term average/long-term average (STA/LTA; Allen, 1982) failed to detect. It is now used to detect small, local earthquakes in continuous seismic data from local and regional arrays (Frank and Shapiro, 2014; Beaucé *et al.*, 2019; Feng *et al.*, 2020). The matched filtering technique consists of detecting near-repeating earthquakes using known events and waveform correlation to overcome the signal-to-noise ratio limitations that power detectors such as backprojection suffer from. It became popular in seismology at the beginning of the century (Gibbons and Ringdal, 2006; Shelly *et al.*, 2007) and is now commonly used to lower the detection threshold in existing catalogs (Frank *et al.*, 2013; Ross, Trugman, *et al.*, 2019; Beaucé *et al.*, 2022). In the past five years, several machine learning models have proven more effective at identifying seismic phases than traditional algorithms like STA/LTA or the Akaike information criterion (Takanami and Kitagawa, 1988). These models either use three-component single-station seismograms (generalized phase detection [Ross *et al.*, 2018], PhaseNet [Zhu and Beroza, 2019], and EQTransformer [Mousavi *et al.*, 2020]) or, recently, multistation seismograms (EdgePhase [Feng *et al.*, 2022]).

The BPMF workflow leverages the complementary nature of backprojection and matched filtering, and proposes a novel approach to enhancing these techniques with machine learning. The workflow is distributed with a user-friendly, flexible Python package that harnesses the computational efficiency of low-level C and CUDA-C routines. This new open-source software aims at fostering the use of modern detection techniques and comprehensive workflows within the seismology community. The workflow is applicable with no or little prior knowledge about the study region, although a more realistic model of the velocity structure ensures increased detection capability and more accurate locations. Because machine learning models do not perform well on data outside of the training distribution (e.g., Münchmeyer *et al.*, 2022), for which they require fine-tuning of their constituent parameters (e.g., Jozinović *et al.*, 2022), BPMF also offers a machine-learning-free alternative workflow. Furthermore, machine learning tools are usually built for regular earthquakes and, therefore, are not directly applicable to the monitoring of atypical earthquakes such as tectonic low-frequency earthquakes (e.g., Frank *et al.*, 2013, 2014; Poiata *et al.*, 2021), mid-ocean ridge earthquakes (e.g., Ekström, 2006; Fan, 2023; Poli, 2023), or volcanic long-period earthquakes (e.g., Shapiro *et al.*, 2017).

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

TABLE 1
Main Object Classes in the Backprojection and Matched-Filtering (BPMF) Python Package

Class (Object) Name	Description
BPMF.dataset.Network	Seismic station metadata
BPMF.dataset.Data	Seismic data and metadata
BPMF.template_search.TravelTimes	Travel times
BPMF.dataset.Event	Event data, metadata with methods for phase picking, relocation, and other useful operations
BPMF.dataset.Template	Subclass of the Event class with extra methods tailored for matched filtering
BPMF.dataset.TemplateGroup	Collections of templates with methods to format inputs for matched filtering
BPMF.template_search.Beamformer	Backprojection (or beamforming) given network, data, and travel-time information
BPMF.similarity_search.MatchedFilter	Matched filtering given network, data, and templates information

Past studies have applied earlier versions of BPMF and described in detail the methodology upon which it is based (Frank and Shapiro, 2014; Frank *et al.*, 2014; Beaucé *et al.*, 2019, 2022). Here, we introduce the first official release of BPMF (v.2.0.0; see Data and Resources), describe the workflow, and decribe the object-oriented application programming interface (API). A detailed online tutorial provides more information on the API (see Data and Resources). Finally, we present a decade-long application to the seismicity of the Ridgecrest, California area, where the BPMF workflow enabled the discovery of increasing modulation of the rate of seismicity by the solid Earth tides in the years preceding the 2019 M_w 7.1 Ridgecrest earthquake (Beaucé *et al.*, 2023).

Workflow

Overview

BPMF's workflow (Fig. 1) is divided into multiple sequential steps. The final earthquake catalog is the product of two detection methods: backprojection (BP) and matched filtering (MF).

BPMF is object-oriented software in which data are manipulated via classes, or structures, that provide a flexible user experience. The eight main classes are listed in Table 1 and will be described in this section.

Backprojection

Backprojecting the seismic wavefield recorded across a network of seismic stations, which involves migrating the network seismic information as a function of a potential source location at depth, has extensively been used to detect and locate seismic events (e.g., Kiser and Ishii, 2013; Frank *et al.*, 2014; Fan and Shearer, 2016; Poiata *et al.*, 2016; Beaucé *et al.*, 2019; Soubestre *et al.*, 2019; Feng *et al.*, 2020). A sum of time-shifted seismograms is called a beam (Capon, 1970; Ringdal and Husebye, 1982; Rost and Thomas, 2002), which quantifies how well the seismic wavefield focuses at a given point in space and time (see Fig. 2):

$$b_k(t) = \sum_{s,c,p} u_{s,c}\left(t + \tau_{s,p}^{(k)}\right), \tag{1}$$

in which $b_k(t)$ is the beam at time t that backprojects the wavefield at the source location indexed by k, and $u_{s,c}(t + \tau_{s,p}^{(k)})$ is the seismogram at seismometer s, channel c, and time shifted by the time delay (moveout) $\tau_{s,p}^{(k)}$ between source k and seismometer s for the seismic phase indexed by p (e.g., P or S wave). $\tau_{s,p}^{(k)}$ can model an arbitrarily complicated wave front geometry.

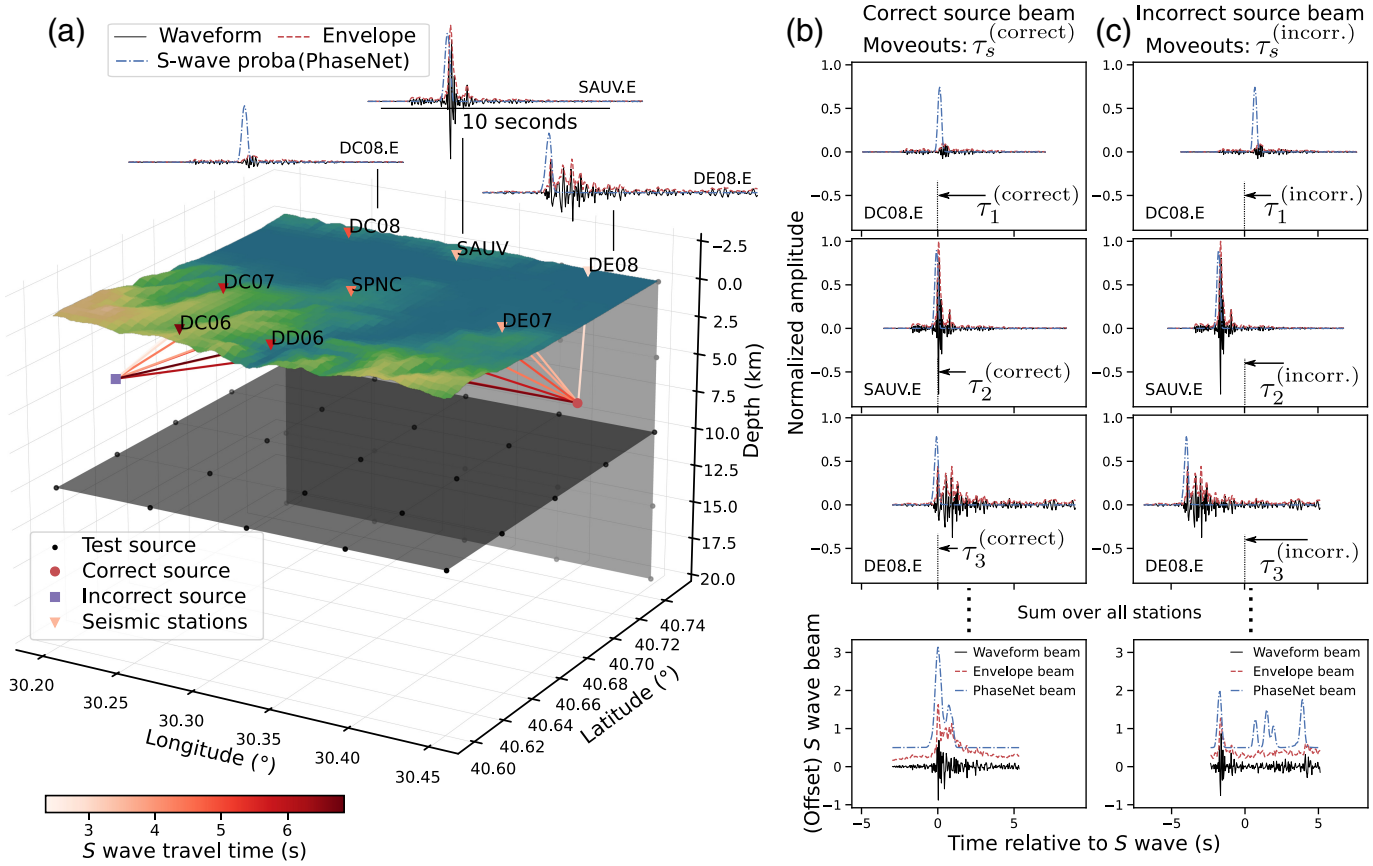
The opposite slip direction on either side of a fault generates waves with opposite polarity (Aki and Richards, 2002), potentially causing the beam to sum up to zero. Furthermore, coherently stacking the high-frequency waveform content requires accurate moveouts, involving the rarely known detailed velocity structure and computationally expensive fine grid spacing. Thus, equation (1) is typically modified to sum the radiated energy or any other smooth positive-valued transform $\mathcal{U}_{s,c}$ of the original seismogram:

$$\tilde{b}_k(t) = \sum_{s,c,p} \mathcal{U}_{s,c}(t + \tau_{s,p}^{(k)}), \tag{2}$$

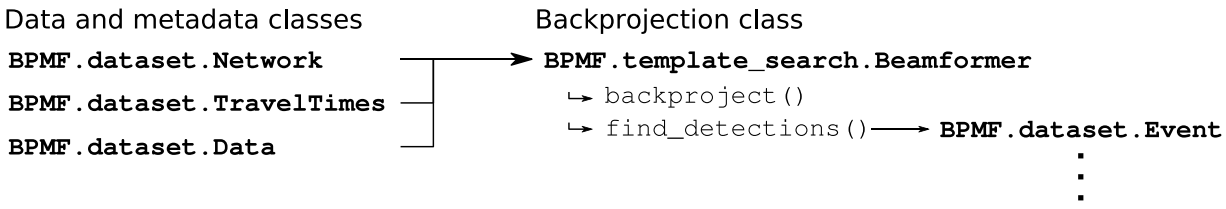
in which $\tilde{b}_k$ represents the beam power. For large arrays, BPMF allows to restrict the computation of $\tilde{b}_k(t)$ to the X closest stations with X to be defined by the user.

The choice of the waveform transform $\mathcal{U}_{s,c}$ depends on the nature of the target signals and the user's goals. The envelope of the waveforms (see Fig. 2) can be used to compute an energy-based beam that constitutes an efficient detector for both regular earthquakes (e.g., Beaucé *et al.*, 2019) as well as low-frequency earthquakes (e.g., Frank and Shapiro, 2014). In the application presented in the Application to the Ridgecrest, California Area section, we leveraged the recent developments in machine learning earthquake detectors (Zhu and Beroza, 2019; Mousavi *et al.*, 2020) and used PhaseNet's output probability time series (Zhu and Beroza, 2019) as the waveform transform $\mathcal{U}_{s,c}$ (we discuss other options in the Tuning the workflow and applicability to other tectonic contexts section). PhaseNet computes P- and S-wave arrival probabilities at any given time of a three-component seismic recording (see Fig. 2). Thus, backprojection is an efficient, physics-based method to turn single-station machine learning detectors into multistation detectors (see also Shi *et al.*, 2022)

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user



(d) Application programming interface for backprojection



and effectively solve the phase association problem (McBrearty *et al.*, 2019).

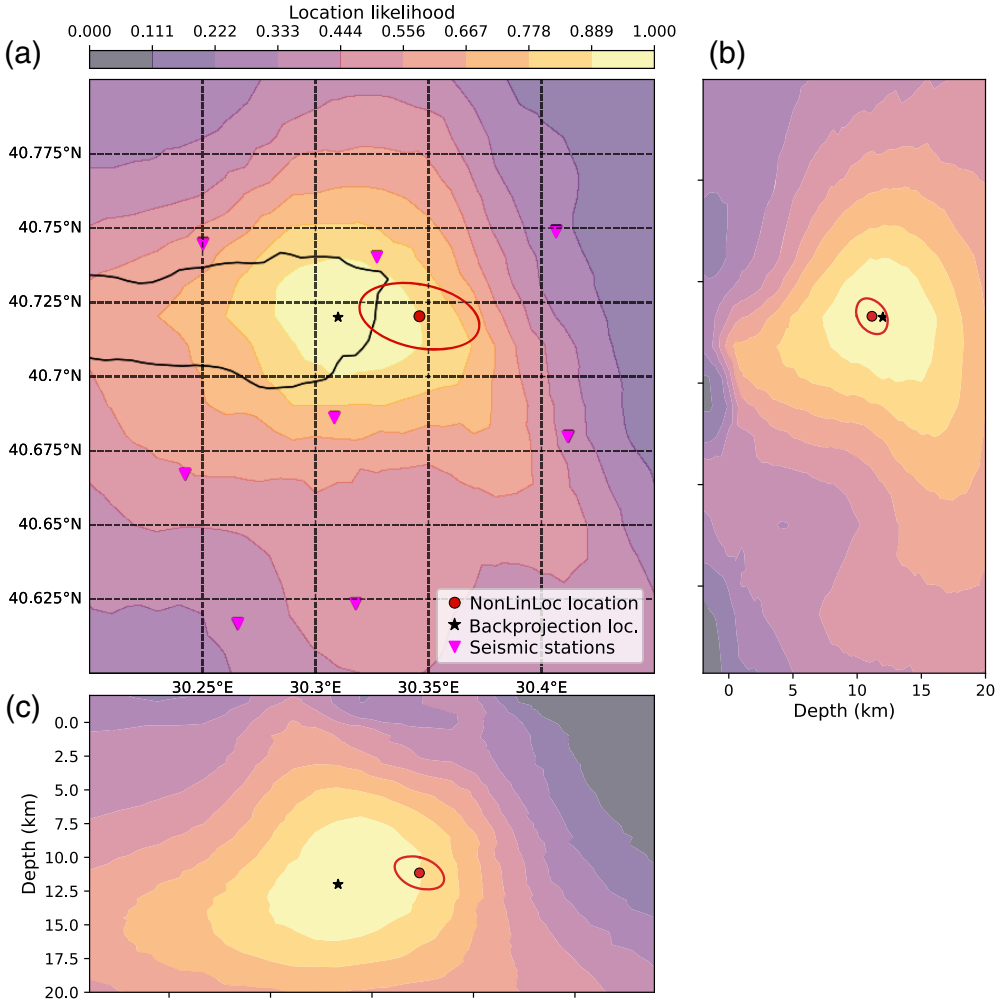
For continuous earthquake detection, the beam power is computed at every source k at each time t, and the maximum beam power $\tilde{b}_{k^*}(t)$ and the corresponding source index $k^*(t)$ are found at each time:

$$\tilde{b}_{k^*}(t) = \max_k\{\tilde{b}_k(t)\}, \quad \text{with} \quad k^*(t) = \text{argmax}_k\{\tilde{b}_k(t)\}. \qquad (3)$$

Peaks above a given threshold define the detection of new events. This threshold can be based on the statistics of $\tilde{b}_{k^*}(t)$ or set entirely empirically (as in the Application to the Ridgecrest, California Area section). The maximum pooling operation in equation (3) prevents the detection of multiple events simultaneously, for which the use of subarrays is necessary. We observe that, because of the trade-off between hypocentral time and location caused by uncertainties, the backprojection method cannot resolve short interevent times (specific numbers

Figure 2. Illustration of the travel-time-based shift-and-stack operation in the backprojection method. (a) The wavefield is continuously recorded at multiple seismic stations, and, when an earthquake occurs, the *P* and *S* waves are observed with a station-dependent time delay (moveout) determined by the source–receiver geometry. Given a velocity model, the *P*- or *S*-wave travel times are computed at discrete points (test sources, black dots). Test sources that are closer to the real source location also have closer moveouts to the true ones. Not all sources of the 3D grid are shown for the sake of clarity. (b) The recorded waveforms, the waveform envelopes, and PhaseNet's output (see Backprojection section) are time shifted using the correct *S*-wave moveouts and stacked. (c) The same as panel (b) but for incorrect *S*-wave moveouts. The difference in peak height between panels (b) and (c) increases as the number of stacked recordings increases. (d) Overview of BPMF's classes involved in the execution of backprojection. The Network, TravelTimes and Data classes contain the information required by the Beamformer class to backproject the wavefield. The detected events are identified on the maximum beam, $\tilde{b}_{k^*}(t)$ in equation (3), given a user-defined threshold, and are represented by instances of the Event class. The color version of this figure is available only in the electronic edition.

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user



(d) Application programming interface for (Re)location

```
BPMF.dataset.Event
  ↪ pick_PS_phases() + relocate(method="NLLoc")
  ↪ OR relocate(method="beam")
```

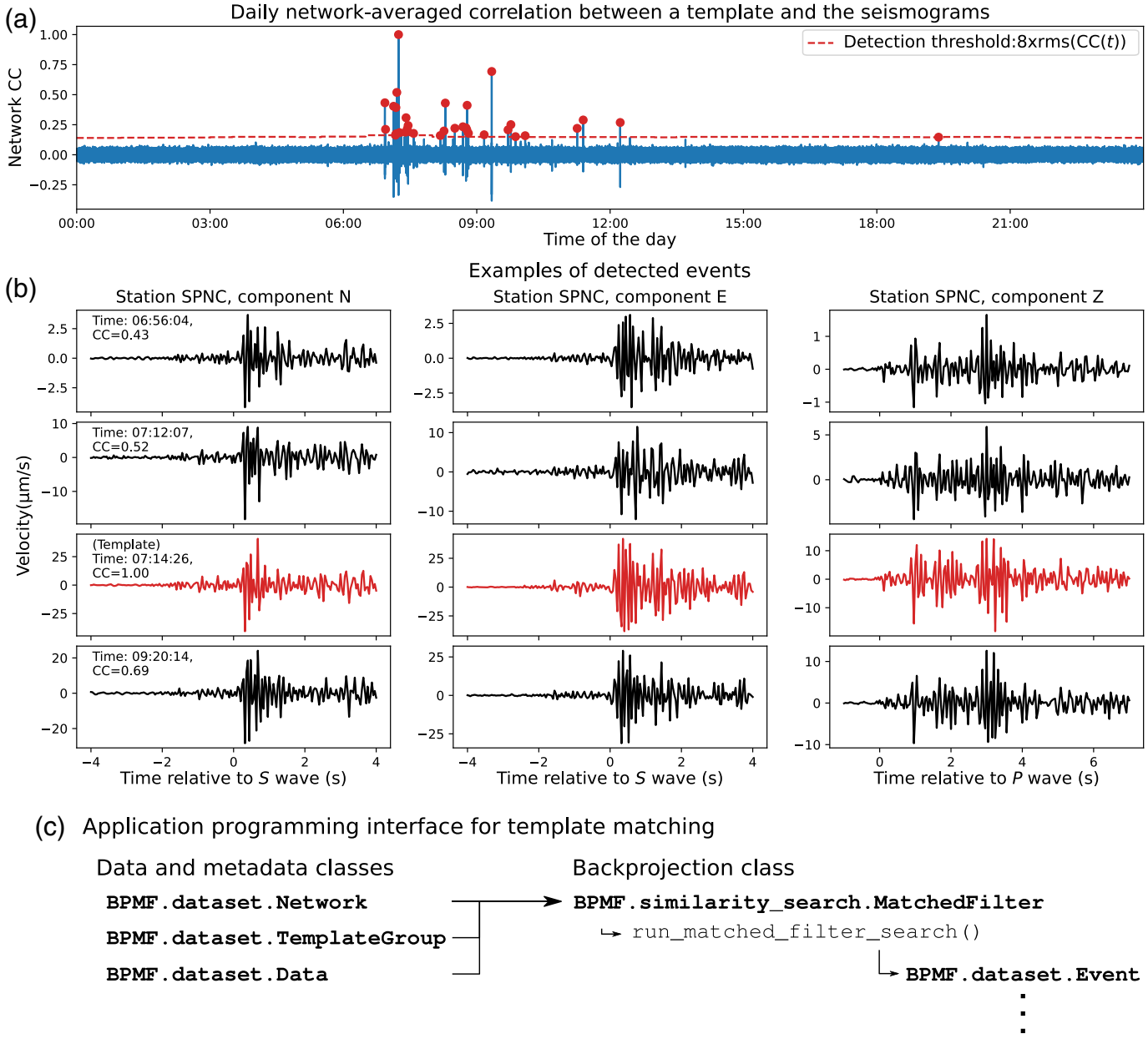
Figure 3. (a–c) Event location with the two methods implemented in BPMF: backprojection (black star and colored contours) and NonLinLoc (red dot). The NonLinLoc earthquake location software uses the *P*- and *S*-wave picks given by the machine learning phase picker. The uncertainty ellipse computed by NonLinLoc represents the 1-sigma error from the covariance matrix derived in the Gaussian approximation about the solution. See the Relocation and Location Uncertainties section for comments on the two location methods. (d) The NonLinLoc and backprojection location methods are implemented in an Event class method. The color version of this figure is available only in the electronic edition.

depend on the grid geometry). We leveraged the embarrassingly parallel structure of evaluating equations (2) and (3) in time (over t) and in space (over k) to design an efficient backprojection algorithm for central processing units (CPUs; C language) and graphic processing units (GPUs; CUDA-C language). The backprojection routines are distributed in a separate Python package, Beampower (see Data and Resources and Fig. S1, available in the supplemental material to this article), to allow usage outside of the BPMF workflow. In BPMF's API, the backprojection and event identification are done with the Beamformer class, which uses information formatted in instances of the Network, TravelTimes, and Data classes (see Fig. 2d and Table 1).

Relocation and location uncertainties

Backprojection provides preliminary locations for the detected events, and the spatial beam field $b_k(t)$ may be used to quantify location uncertainties (Fig. 3). However, the beam-based location uncertainties have multiple drawbacks: they depend on grid size and grid resolution, require expensive computation for large grids, and are sensitive to the properties of the waveform transform $\mathcal{U}_{s,c}$ (equation 2). To overcome the aforementioned issues, BPMF proposes a relocation routine that is based on the nonlinear earthquake location software NonLinLoc (see Data and Resources and Lomax *et al.*, 2000) and a machine-learning-based phase-picking routine (Zhu and Beroza, 2019; Mousavi *et al.*, 2020; Woollam *et al.*, 2022). NonLinLoc combines grid search and sampling methods (OCT-Tree algorithm) to efficiently find the global solution to the location problem and estimate location uncertainties (see Fig. 3). Furthermore, the automated workflow highly benefits from NonLinLoc's robust formulation of the location problem (i.e., its robust loss function) that strongly mitigates the effects of outliers in the *P*- and *S*-wave picks. In summary, two location routines are implemented in BPMF: one based on machine learning phase pickers and NonLinLoc, and a beam-based location routine that does not require any machine learning tools. The latter may be preferred with datasets or signals (e.g., low-frequency earthquakes) with which machine learning phase pickers perform poorly. Both the location methods are implemented within an Event class method (see Fig. 3d

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user



and Table 1). Users can choose to use every station with *P*- and *S*-wave picks or focus on a subset of trusted stations.

Matched filtering and template selection

The earthquakes detected and located at this stage of the workflow (Fig. 1) are used in a matched-filter search to densify, in time and space, the backprojection catalog. Matched filtering (also called template matching) consists in scanning the continuous recordings in search of events with similar waveforms and moveouts than that of a reference event (template, see Fig. 4) at one or multiple stations (Gibbons and Ringdal, 2006; Shelly *et al.*, 2007; Beaucé *et al.*, 2022; Cabrera *et al.*, 2022). Waveform and moveout similarity implies the colocation and mechanism similarity of the sources. The similarity between the template event's waveforms $e_{s,c}(t)$ and the continuous recordings $u_{s,c}(t)$ at time t is quantified with the network-averaged correlation coefficient (CC; Fig. 4a):

Figure 4. Detecting new events with matched filtering (also called template matching). (a) Similarity between the template event and the continuous recordings measured by the network-averaged correlation coefficient (CC; equation 4). Peaks above the user-defined detection threshold are detections of new events. Here, the threshold is set to eight times the running root mean square of the CCs. (b) Examples of detected events based on the network-averaged CC shown in panel (a). The template event (red waveforms) detects itself with CC = 1. We are showing the waveforms on a single three-component station, but the CC is computed across the entire network (8 stations). (c) Overview of BPMF's classes involved in the execution of matched filtering. The Network, TemplateGroup, and Data classes contain the information required by the MatchedFilter class to run the matched-filter search. The detected events are identified on the network-averaged CC (panel (a), given a user-defined threshold, and are represented by instances of the Event class. The color version of this figure is available only in the electronic edition.

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

$$\mathrm{CC}(t) = \sum_{s,c} w_{s,c} \sum_{i=1}^{N} \frac{(e_{s,c}(i\Delta t) - \bar{e}_{s,c})(u_{s,c}(t + i\Delta t + \tau_{s,c}) - \bar{u}_{s,c}(t + \tau_{s,c}))}{\sqrt{\sum_{i=1}^{N}(e_{s,c}(i\Delta t) - \bar{e}_{s,c})^2 \sum_{i=1}^{N} \sum (u_{s,c}(t + i\Delta t + \tau_{s,c}) - \bar{u}_{s,c}(t + \tau_{s,c}))^2}}, \quad (4)$$

in which $\tau_{s,c}$ is the moveout at station s, and channel c used to align the continuous recordings with the template waveforms, $w_{s,c}$ is a user-prescribed weight (with $\sum_{s,c} w_{s,c} = 1$ to compute an average CC and all weights equal to compute the arithmetic mean), N is the template waveform duration in samples, Δt is the sampling time in seconds, and $\bar{e}_{s,c}$ and $\bar{u}_{s,c}$ are the template and windowed continuous data means. The evaluation of equation (4) over large timescales and seismic networks is a computationally intensive task but also embarrassingly parallel, for which we have developed the Fast Matched Filter GPU-accelerated algorithm (Beaucé *et al.*, 2018). In BPMF's API, the matched-filter search and event identification are done with the MatchedFilter class, which uses information stored in Network, TemplateGroup, and Data class instances (see Fig. 4c and Table 1).

Because nearby events with similar focal mechanisms share similar waveforms and moveouts, they detect the same events when used as templates in a matched-filter search (see Fig. S2). Thus, to avoid redundant computation and reduce the computation footprint of matched filtering, we identify groups of similar events and select only one event per group as a template. Choosing the similarity threshold above which earthquakes are grouped together trades off a reduction in computation time against increasing the number of detected events. Templates are further selected based on their location uncertainties. Poorly located events are associated with low signal-to-noise ratio signals and poor knowledge of the *P*- and *S*-wave arrivals, both reducing the chances of successful matched-filter searches.

The matched filtering method is able to detect very low signal-to-noise ratio events, and improves magnitude and interevent time completeness with respect to the initial catalog. However, because of the nature of similarity-based detection, the spatial completeness of a matched filtering catalog is entirely set by the initial catalog.

Compiling the final catalog and postprocessing

The backprojection and matched filtering catalogs are complementary. Even though the backprojection catalog is used to build the template database, not all events are used as templates (see the Matched Filtering and Template Selection section), and, therefore, some events of the backprojection catalog may not be included in the matched filtering catalog. User-defined criteria on interevent time and distance allow to find the events of the backprojection catalog that are not included in the matched filtering catalog to complement the latter and build the final catalog.

Before merging the two catalogs, the matched filtering catalog needs to be cleaned to remove multiple detections of unique events. Neighboring template earthquakes with similar mechanisms tend to detect similar events. The identification of these shared detections is achieved through user-defined criteria on intertemplate similarity, distance, and interevent times. Events detected within ΔT seconds of one another represent redundant detections of the same event if they were detected with template earthquakes that are separated by less than ΔR kilometers and with an average interevent CC above S. This cleaning procedure is essential for the interpretability of the final catalog (e.g., interevent time or frequency–magnitude analyses).

BPMF limits its core purpose to the detection and location of seismic events. However, BPMF provides tools to facilitate the estimation of event magnitudes (e.g., waveform extraction, displacement spectrum estimation). Postprocessing of the catalog may also commonly include the relative event relocation with the double-difference method (e.g., Poupinet *et al.*, 1984; Waldhauser and Ellsworth, 2000).

Application to the Ridgecrest, California Area

We present an application of the BPMF workflow to the Ridgecrest, California area in the decade preceding the M_w 6.4 and 7.1 earthquake sequences that struck the region in July 2019 (e.g., Shelly, 2020, and see Fig. 5). The catalog presented here (available online; see Data and Resources) was initially developed to study the statistical properties of small-magnitude seismicity in the region (Beaucé *et al.*, 2023). We compare our BPMF catalog to the Southern California Seismological Network (SCSN; see Data and Resources) catalog and the Quake Template Matching (QTM; see Data and Resources) catalog (Ross, Trugman, *et al.*, 2019). The QTM catalog is a matched filtering extension of the SCSN catalog. We stress that these catalogs were built in different contexts using more extended seismic arrays and for different purposes than our BPMF catalog. Thus, we here acknowledge the limitations of the following comparison.

We used the seismograms from eight broadband seismic stations (see Fig. 5) from 1 January 2009 to 3 July 2019, bandpass filtered them between 1 and 12 Hz, downsampled them to 50 Hz, and applied the BPMF workflow (Fig. 1). We used the deep-learning model PhaseNet to transform the three-component seismograms ($u_{s,c}$ in equation 1) into time series of *P*-wave and *S*-wave arrival probability ($\mathcal{U}_{s,c}$ in equation 2; see also Fig. 2). The model used here was trained on seismic data from northern California (Zhu and Beroza, 2019). The *P*- and *S*-wave travel times were computed by solving the Eikonal equation with PyKonal (White *et al.*, 2020; see Data and Resources) in the 3D velocity model from Zhang and Lin (2014). We empirically set the backprojection detection threshold to

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

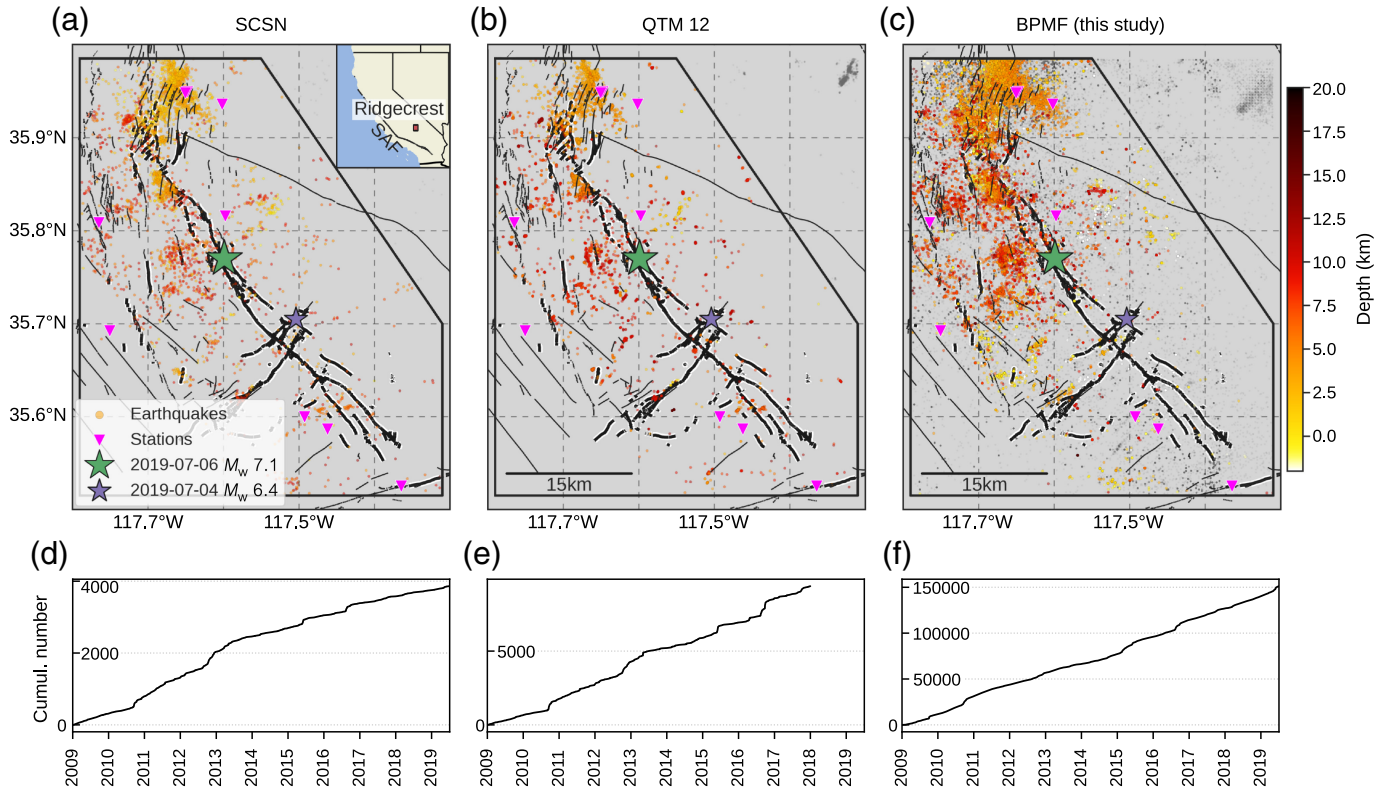


Figure 5. Comparison of the Southern California Seismological Network (SCSN), Quake Template Matching (QTM; with a detection threshold of 12× the median absolute deviation of CCs; Ross, Trugman, *et al.*, 2019), and BPMF (this study) seismicity catalogs in the Ridgecrest, California area. (a–c) Map of the earthquake epicenters with color-coded hypocentral depths. The inset at the top right corner in panel (a) shows the location of the Ridgecrest study region (red square) in California. For reference, the San Andreas fault (SAF) is shown with a black line. In panel (c), earthquake epicenters with uncertainties lower than 2.5 km are colored, and all others are shown with gray dots. Black lines are known faults, and the thick lines show the surface rupture of the M_w 6.4 and 7.1 2019 earthquakes. (d–f) Cumulative number of earthquakes as a function of time. The QTM catalog ends on 1 January 2018. The BPMF catalog (our study) reports 39 × more events than the SCSN catalog (over the entire time span) and 13 × more than the QTM catalog (between 2009 and 2018). The color version of this figure is available only in the electronic edition.

1.25 after a series of tests with manual checking of the detected events. We set the matched filtering detection threshold to 8 times the CC root mean square (rms), which compares with QTM's detection threshold of 12 times the CC median absolute deviation (MAD; 8 × rms = 11.86 × MAD). We focused the catalog comparison on the earthquakes located nearer to the stations (see black polygon in Fig. 5).

We detected and located 44,717 events with the backprojection and PhaseNet methods, out of which we selected 13,594 templates. Templates were made of 8 s waveforms, starting one second before the *P*-wave arrival on the vertical components and four seconds before the *S*-wave arrival on the horizontal components. Matched filtering tripled the number of earthquakes with a total of 151,229 events over the ≈10.5 yr (and inside the black polygon in Fig. 5f), which is 39 times more than the SCSN catalog (3863 events, Fig. 5d) and 13 times more than the QTM catalog (9403 events, Fig. 5e; Ross, Trugman, *et al.*, 2019; stopping the comparison in 2018 when the QTM catalog ends). Because of the nature of matched filtering (see the Matched Filtering and Template Selection section), the spatial completeness of the QTM catalog is restricted to that of the SCSN catalog (see Fig. 5a vs. 5b). Differences in earthquake locations between these two catalogs arise because events in the QTM catalog were relocated with the double-difference method (Ross, Trugman, *et al.*, 2019). In contrast, the BPMF catalog shows a more scattered seismicity (Fig. 5c), with small-magnitude events (a quantitative magnitude analysis follows) missing from the SCSN catalog and thus undetected in the QTM catalog. The scatter is also due to location uncertainties that are higher than that of the relative locations given in the QTM catalog.

Double-difference relative relocation techniques (e.g., Poupinet *et al.*, 1984; Waldhauser and Ellsworth, 2000) can be applied to the BPMF catalog to reduce uncertainties related to our limited knowledge of the velocity structure and build high-resolution maps of active faults. We applied the double-difference algorithm HypoDD (Waldhauser and Ellsworth, 2000, and see Data and Resources) to PhaseNet's *P*- and *S*-wave picks to refine BPMF's earthquake locations in selected subareas of the study region (Fig. 6). We only used picks that were within 10% of the predicted arrival times and for events that were located

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

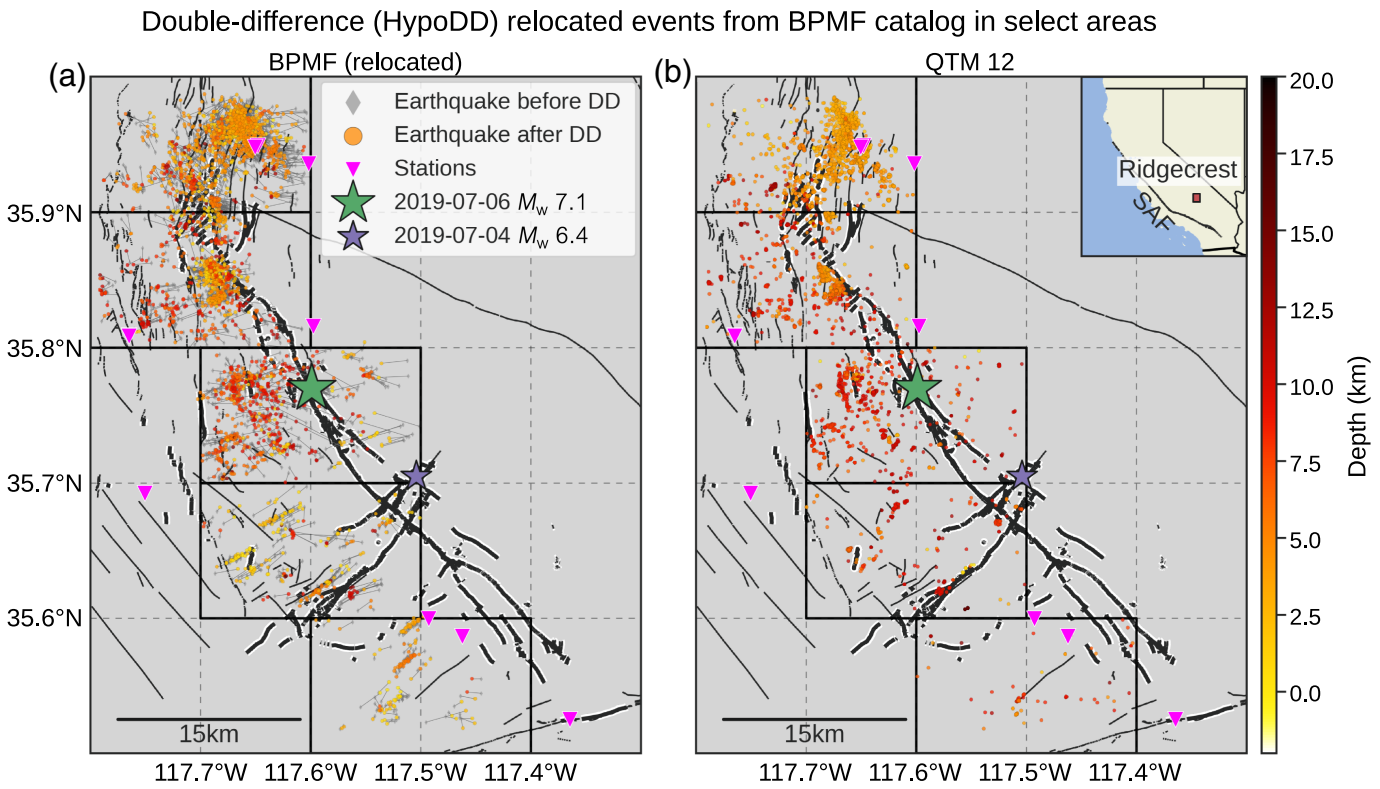


Figure 6. (a) Double-difference (DD) relative relocation of the BPMF catalog earthquake hypocenters with the HypoDD algorithm (Waldhauser and Ellsworth, 2000). For computational and performance concerns, we applied HypoDD to subsets of the catalog (black boxes). To limit the propagation of PhaseNet's pick errors into the relocation problem, we discarded outlier picks by selecting picks that were within 10% of the initial hypocenter's predicted time. Gray lines connect the initial and relocated epicenters. In each box from north to south, the number of selected versus relocated events was: box 1: 1433 versus 1279, box 2: 909 versus 802, box 3: 743 versus 612, box 4: 395 versus 223, and box 5: 154 versus 122. (b) For comparison, the QTM catalog is shown in the same selected boxes. The inset at the top right corner shows the location of the Ridgecrest study region (red square) in California. For reference, the San Andreas fault (SAF) is shown with a black line. The color version of this figure is available only in the electronic edition.

with a horizontal uncertainty of less than 5 km. The scatter in earthquake hypocenters decreased, and hypocenter alignments revealed fault structures that compare to the QTM catalog (Fig. 6). Better spatial resolution could be achieved by measuring differential arrival times with correlation techniques (e.g., Schaff *et al.*, 2004). The relocated BPMF catalog reveals active structures missing from the SCSN and QTM catalogs, like the northeast-southwest striking faults, mostly below 35.70° (Fig. 6a). Similarly striking faults were illuminated by the 2019 Ridgecrest aftershocks (Ross, Idini, *et al.*, 2019).

We characterized the shape of the magnitude distributions by fitting the parameters a and b of the Gutenberg–Richter law (Gutenberg and Richter, 1941) to the data: $\log(N(M)) = a - bM$, in which $N(M)$ is the number of earthquakes exceeding magnitude M during the observation period. The b-value relates to fault zone attributes such as stress (Scholz, 2015) or structural heterogeneity (Mogi, 1967). We computed a local earthquake magnitude M_L for each of our 151,229 cataloged events (see Fig. S3 and the supplemental section Magnitude Estimation), and compared the magnitude distribution against that of the SCSN and QTM catalogs (see Fig. 7). For this comparison, we only considered the BPMF events with horizontal location uncertainties (h_{max}) smaller than 2.5 km to exclude mislocated events that are actually further outside the study region. However, it is not possible to fully mitigate selection effects caused by the northern edge crossing a highly seismically active area (see Fig. 5). We computed the magnitude of completeness M_c with the maximum curvature method (Wiemer and Wyss, 2000). We obtained $M_{c,BPMF} = -0.34$ for the BPMF catalog and $M_{c,BP} = -0.14$ for the backprojection-only catalog (Fig. 7a) with a factor ≈3 increase in both the total number of detected events and the number of events with $h_{max} < 2.5$ km in the BPMF catalog. This is a reminder that the magnitude of completeness is a function of time and space, and that a scalar value does not fully describe catalog completeness. The small difference in event numbers indicates that backprojection and PhaseNet have already detected most of the events that matched filtering could find. Such good performance may not always be achieved for other datasets if the deep-learning model is not retrained. Matched filtering, importantly, improves the smallest resolved interevent time with respect to the backprojection catalog.

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

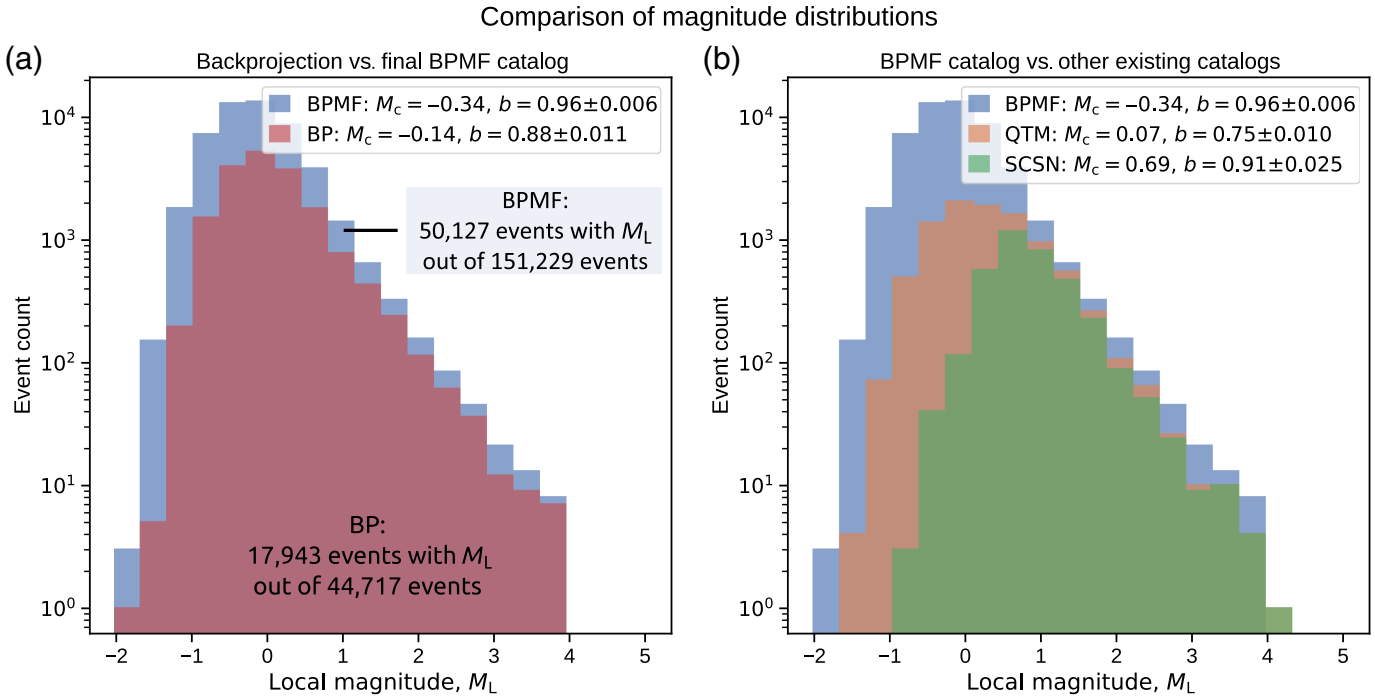


Figure 7. Comparison of magnitude distributions. (a) The backprojection-only catalog (abbreviated as BP, red histogram) from the first part of the workflow (see Fig. 1) is compared against the final catalog (BPMF, blue histogram). Only M_L from events with $h_{max} < 2.5$ km is considered here. There is only a factor ≈3 difference in both the total number of detected earthquakes, suggesting that backprojection with PhaseNet (see the Backprojection and Application to the Ridgecrest, California Area sections) is particularly efficient for the Ridgecrest data. Because the matched-filtering-only catalog is indistinguishable from the BPMF catalog, indicating that all events in the BP catalog were detected with matched filtering, it is not plotted here. (b) The final BPMF catalog is compared against the QTM and SCSN catalogs (see also Fig. 5). All catalogs were cut before 2018 for fair comparison with the QTM catalog. The BPMF catalog shows a magnitude of completeness M_c that is 0.41 and 1.03 units smaller than that of the QTM and SCSN catalogs, respectively. The color version of this figure is available only in the electronic edition.

The BPMF catalog magnitude of completeness is 0.41 and 1.03 units smaller than that of the QTM and SCSN catalogs, respectively. We conclude that the M_c difference is less than the 1 unit difference that one could expect to explain a 10-fold event number difference, despite the factor ≈10 difference in the number of events between the BPMF and QTM catalogs. Thus, we conclude that BPMF's improvement over the QTM catalog primarily lies in its better spatial coverage, that is, its more homogeneous magnitude of completeness than that of the QTM catalog, which is limited by the sparse event distribution of the SCSN catalog.

We computed the maximum-likelihood estimate of the b-value (Aki, 1965) and its standard error following Shi and Bolt (1982). We found that the BPMF and SCSN catalogs share similar estimates ($b_{BPMF} = 0.96 \pm 0.006$ vs. $b_{SCSN} = 0.91 \pm 0.025$, respectively; see Fig. 7), whereas the QTM catalog yielded a significantly lower b-value ($b_{QTM} = 0.75 \pm 0.010$). When matched filtering significantly decreases the magnitude of completeness with respect to the template catalog, as in the QTM catalog, it may result in a more spatially heterogeneous M_c if the template spatial coverage was sparse and, thus, bias the estimate of the b-value. In the geometrical interpretation of the b-value, in which b is related to the spatial hypocenter distribution (the fractal spatial dimension, e.g., Aki, 1981), one might expect that failing to detect isolated events artificially lowers the b-value.

Discussion

Runtime and runtime scaling

We built the catalog presented here in about one month using 48 Intel Xeon Silver 4310 2.10 GHz CPUs and 1 Nvidia A100 GPU, spending three days on the backprojection catalog and three weeks on the matched filtering catalog (13,594 templates). During the early development stages of the study in Beaucé *et al.* (2023), we ran the entire matched filtering process in 12 hr using four nodes of the Frontera supercomputer (see Data and Resources), each equipped with 32 Intel Xeon E5-2620 v.4 CPUs and 4 NVIDIA Quadro RTX 5000 GPUs. Also using Frontera's computing resources, we conducted runtime scaling tests for the two core routines of BPMF that run backprojection (Beampower) and matched filtering (Fast Matched Filter). The results (Fig. S4) show that the runtime τ of Beampower's and Fast Matched Filter's CPU implementations scales inversely proportionally to the CPU count

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

N ($\tau \sim 1/N$), meaning that parallelization makes optimal use of the resources. Beampower's and Fast Matched Filter's GPU implementations offer great speed-up with respect to their CPU counterparts. We note that, although the runtime scaling $\tau \sim 1/N$ is preserved for large-scale problems, it may be hampered for small-scale problems because of overhead computation that accounts for a nonnegligible portion of the total runtime for very short runtimes. Therefore, users might consider running multiple single-GPU jobs to process many small-scale problems at the same time. We note that the runtime scaling of the overall workflow with CPU and GPU resources depends on user choices about assembling the BPMF modules. For example, a user may choose or not to parallelize phase picking and event location, which is easily achieved because of BPMF's modularity but does not come as a built-in feature.

Tuning the workflow and applicability to other tectonic contexts

The BPMF workflow is applicable in various tectonic contexts (including shear zones, subduction zones, and stable continental regions), diverse environments (onland and ocean bottom seismometers), and even in the context of human-induced seismicity. In each specific scenario, achieving optimal results entails fine-tuning the workflow. This fine-tuning involves adapting the frequency band of the filtered seismograms, carefully selecting the number of stations over which a beam (equation 1) or a CC (equation 4) is computed, and choosing the most suitable transform $\mathcal{U}_{s,c}$ for backprojection (equation 2). Among these considerations, designing $\mathcal{U}_{s,c}$ is one of the most important choices.

Exploring multiple deep-learning phase pickers, such as generalized phase detection (Ross *et al.*, 2018), PhaseNet (Zhu and Beroza, 2019), or EQTransformer (Mousavi *et al.*, 2020), is a valuable approach that can optimize detection performances. For the application presented in the Application to the Ridgecrest, California Area section, we used PhaseNet, because it was trained on Californian data and runs fast thanks to its relatively few model parameters. These machine learning tools may be compared against conventional earthquake detectors like STA/LTA (Allen, 1982) or the waveform envelope, which may be more performant for detecting unconventional earthquakes such as low-frequency earthquakes.

Finally, it is crucial to establish appropriate detection thresholds for backprojection and matched filtering. These thresholds should align with the specific scientific question that the seismic catalog aims to address, taking into consideration the acceptable balance between false positive and false negative detections. Ultimately, this decision relies on whether your task is more adversely affected by false detections and mislocated earthquakes, or by missing seismic events.

Summary and Concluding Remarks

We have presented a new open-source community tool, BPMF (backprojection and matched filtering), for automatic earthquake detection and location. The workflow (Fig. 1) is based on the association of two modern array detection methods—backprojection and matched filtering—to build low magnitude of completeness and short interevent time resolution catalogs. BPMF integrates the recent machine learning tools for phase picking (PhaseNet; Zhu and Beroza, 2019), which enables automatic earthquake location with NonLinLoc (Lomax *et al.*, 2000). We have also demonstrated here (see the Backprojection and Application to the Ridgecrest, California Area sections) that backprojection and existing machine learning detectors work efficiently together. Alternatively, the BPMF workflow can run without any machine learning tool to avoid the common shortcomings of learning-based algorithms. The BPMF API is introduced in detail in the online tutorial (see Data and Resources).

We presented an application of BPMF to the Ridgecrest, California seismicity (see the Application to the Ridgecrest, California Area section) and showed that we detected 151,229 between 1 January 2009 and 3 July 2019 with a magnitude of completeness M_c – 0.34. The BPMF catalog contains 39 times more events than the SCSN catalog and 13 times more than the QTM catalog, with a magnitude of completeness that is 1.03 and 0.41 units smaller than each catalog's, respectively. As an open-source tool, BPMF aims to encourage the analysis of large-scale seismic datasets to promote new discoveries and answer longstanding questions.

Data and Resources

This article describes the backprojection and matched-filter (BPMF) Python package v.2.0.0beta (https://zenodo.org/badge/latestdoi/346391253, last accessed July 2023). The latest version of BPMF can be accessed at https://github.com/ebeauce/Seismic_BPMF (last accessed September 2023). The backprojection core routine is distributed in its own Python package, Beampower (https://github.com/ebeauce/beampower, v.1.0.2, last accessed September 2023), and the matched-filtering core routine is distributed in the Python package Fast Matched Filter (https://github.com/beridel/fast_matched_filter, v.1.5.0, last accessed September 2023). A detailed tutorial is available online at https://ebeauce.github.io/Seismic_BPMF/tutorial.html (last accessed July 2023). BPMF depends on the seisbench package (v.0.4.0, https://github.com/seisbench/seisbench, last accessed July 2023; Woollam *et al.*, 2022) to use the machine learning phase picker PhaseNet (Zhu and Beroza, 2019). NonLinLoc (Lomax *et al.*, 2000) v.7.0 can be downloaded at http://alomax.free.fr/nlloc/ (last accessed July 2023). PyKonal (White *et al.*, 2020) v.0.4.0 is available at https://github.com/malcolmw/pykonal (last accessed July 2023). We used HypoDD v.2.1b (Waldhauser and Ellsworth, 2000), which can be downloaded at ~http://www.ldeo.columbia.edu/~felixw/HYPODD2.1b/ (last accessed July 2023). The runtime scaling tests discussed in the Runtime and Runtime Scaling section, and shown in Figure S4 were run on the Frontera supercomputer, which is part of the Texas Advanced Computing Center (TACC). The earthquake catalog in the Ridgecrest, California area presented in the Application to the Ridgecrest, California Area section can be downloaded at DOI: 10.5281/zenodo.8393318 (v.2.0.1). The Southern California Seismological Network (SCSN)

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

catalog is available at https://service.scedc.caltech.edu/eq-catalogs/date_mag_loc.php (last accessed July 2023) and the Quake Template Matching (QTM) catalog is available at https://scedc.caltech.edu/data/qtm-catalog.html (last accessed July 2023). The supplemental material to this article includes additional figures (Figs. S1–S4) providing information on the Beampower algorithm (Fig. S1), the identification of similar templates (Fig. S2), the comparison of BPMF magnitudes with SCSN and QTM magnitudes (Fig. S3), and the runtime scaling with numerical resources (Fig. S4). Details on the magnitude calculation are provided in text S2.

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

Eric Beaucé was funded by the Brinson Foundation. This project has received funding from the European Research Council (ERC) under the European Union's Horizon H2020 research and innovation program (Grant Agreement Number 742335, F-IMAGE). Allocation on Frontera at the Texas Advanced Computing Center (TACC) was supported through the SCOPED project (NSF-OAC/2103741, 2103494). Piero Poli was supported by the European Union Horizon 2020 Research and Innovation Program (Grant Agreement Number 802777, MONIFAULTS).

References

Aki, K. (1965). Maximum likelihood estimate of b in the formula log n= a-bm and its confidence limits, *Bull. Earthq. Res. Inst. Tokyo Univ.* **43,** 237–239.

Aki, K. (1981). A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction: An International Review*, Wiley Online Library, Hoboken, New Jersey, Vol. 4, 566–574.

Aki, K., and P. G. Richards (2002). *Quantitative Seismology*, University Science Books, Melville, New York.

Allen, R. (1982). Automatic phase pickers: Their present use and future prospects, *Bull. Seismol. Soc. Am.* **72,** no. 6B, S225–S242.

Arrowsmith, S. J., D. T. Trugman, J. MacCarthy, K. J. Bergen, D. Lumley, and M. B. Magnani (2022). Big data seismology, *Rev. Geophys.* **60,** no. 2, e2021RG000769, doi: 10.1029/2021RG000769.

Beaucé, E., W. B. Frank, A. Paul, M. Campillo, and R. D. van der Hilst (2019). Systematic detection of clustered seismicity beneath the southwestern Alps, *J. Geophys. Res.* **124,** no. 11, 11,531–11,548.

Beaucé, E., W. B. Frank, and A. Romanenko (2018). Fast matched filter (FMF): An efficient seismic matched-filter search for both CPU and GPU architectures, *Seismol. Res. Lett.* **89,** no. 1, 165–172.

Beaucé, E., P. Poli, F. Waldhauser, B. Holtzman, and C. Scholz (2023). Enhanced tidal sensitivity of seismicity before the 2019 magnitude 7.1 Ridgecrest, California earthquake, *J. Geophys. Res.* **50,** no. 14, e2023GL104375, doi: 10.1029/2023GL104375.

Beaucé, E., R. D. van der Hilst, and M. Campillo (2022). Microseismic constraints on the mechanical state of the North Anatolian fault zone 13 years after the 1999 M7.4 Izmit earthquake, *J. Geophys. Res.* **127,** no. 9, e2022JB024416.

Cabrera, L., P. Poli, and W. B. Frank (2022). Tracking the spatio-temporal evolution of foreshocks preceding the Mw 6.1 2009 L'Aquila earthquake, *J. Geophys. Res.* **127,** no. 3, e2021JB023888, doi: 10.1029/2021JB023888.

Capon, J. (1970). Applications of detection and estimation theory to large array seismology, *Proc. IEEE* **58,** no. 5, 760–770.

De Barros, L., N. Wynants-Morel, F. Cappa, and P. Danré (2021). Migration of fluid-induced seismicity reveals the seismogenic state of faults, *J. Geophys. Res.* **126,** no. 11, e2021JB022767, doi: 10.1029/2021JB022767.

Ekström, G. (2006). Global detection and location of seismic sources by using surface waves, *Bull. Seismol. Soc. Am.* **96,** no. 4A, 1201–1212.

Fan, W. (2023). Combining love and Rayleigh waves in detecting and locating seismic sources, *Geophys. J. Int.* doi: 10.1093/gji/ggad250.

Fan, W., and P. M. Shearer (2016). Local near instantaneously dynamically triggered aftershocks of large earthquakes, *Science* **353,** no. 6304, 1133–1136.

Feng, T., L. Meng, and H. Huang (2020). Detecting offshore seismicity: Combining backprojection imaging and matched-filter detection, *J. Geophys. Res.* **125,** no. 8, e2020JB019599, doi: 10.1029/2020JB019599.

Feng, T., S. Mohanna, and L. Meng (2022). Edgephase: A deep learning model for multi-station seismic phase picking, *Geochem. Geophys. Geosys.* **23,** no. 11, e2022GC010453, doi: 10.1029/2022GC010453.

Frank, W., and N. Shapiro (2014). Automatic detection of low-frequency earthquakes (lfes) based on a beamformed network response, *Geophys. J. Int.* **197,** no. 2, 1215–1223.

Frank, W. B., N. M. Shapiro, A. L. Husker, V. Kostoglodov, A. Romanenko, and M. Campillo (2014). Using systematically characterized low-frequency earthquakes as a fault probe in Guerrero, Mexico, *J. Geophys. Res.* **119,** no. 10, 7686–7700.

Frank, W. B., N. M. Shapiro, V. Kostoglodov, A. L. Husker, M. Campillo, J. S. Payero, and G. A. Prieto (2013). Low-frequency earthquakes in the Mexican sweet spot, *Geophys. Res. Lett.* **40,** no. 11, 2661–2666.

Gibbons, S. J., and F. Ringdal (2006). The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.* **165,** no. 1, 149–166.

Gutenberg, B., and C. Richter (1941). *Seismicity of the Earth*, Vol. 34, Geological Society of America, Boulder, Colorado.

Jozinović, D., A. Lomax, I. Stajduhar, and A. Michelini (2022). Transfer learning: Improving neural network based prediction of earthquake ground shaking for an area with insufficient training data, *Geophys. J. Int.* **229,** no. 1, 704–718.

Kiser, E., and M. Ishii (2013). Hidden aftershocks of the 2011 mw 9.0 Tohoku, Japan earthquake imaged with the backprojection method, *J. Geophys. Res.* **118,** no. 10, 5564–5576.

Kwiatek, G., P. Martínez-Garzón, J. Davidsen, P. Malin, A. Karjalainen, M. Bohnhoff, and G. Dresen (2022). Limited earthquake interaction during a geothermal hydraulic stimulation in Helsinki, Finland, *J. Geophys. Res.* **127,** e2022JB024354, doi: 10.1029/2022JB024354.

Lomax, A., J. Virieux, P. Volant, and C. Berge-Thierry (2000). Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, in *Advances in Seismic Event Location*, Springer, Dordrecht, The Netherlands, 101–134.

McBrearty, I. W., J. Gomberg, A. A. Delorey, and P. A. Johnson (2019). Earthquake arrival association with backprojection and graph theory, *Bull. Seismol. Soc. Am.* **109,** no. 6, 2510–2531.

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user

Mogi, K. (1967). Earthquakes and fractures, *Tectonics* **5,** no. 1, 35–55.
Mousavi, S. M., W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza (2020). Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking, *Nat. Commun.* **11,** no. 1, 3952.
Münchmeyer, J., J. Woollam, A. Rietbrock, F. Tilmann, D. Lange, T. Bornstein, T. Diehl, C. Giunchi, F. Haslinger, D. Jozinović, *et al.* (2022). Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers, *J. Geophys. Res.* **127,** no. 1, e2021JB023499, doi: 10.1029/2021JB023499.
Poiata, N., C. Satriano, J.-P. Vilotte, P. Bernard, and K. Obara (2016). Multiband array detection and location of seismic sources recorded by dense seismic networks, *Geophys. J. Int.* **205,** no. 3, 1548–1573.
Poiata, N., J.-P. Vilotte, N. M. Shapiro, M. Supino, and K. Obara (2021). Complexity of deep low-frequency earthquake activity in Shikoku (Japan) imaged from the analysis of continuous seismic data, *J. Geophys. Res.* **126,** no. 11, e2021JB022138, doi: 10.1029/2021JB022138.
Poli, P. (2023). Continuation of events detection with global long-period seismic data: An analysis from 2010 to 2022, *Seismol. Res. Lett.* doi: 10.1785/0220230148.
Poupinet, G., W. Ellsworth, and J. Frechet (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California, *J. Geophys. Res.* **89,** no. B7, 5719–5731.
Ringdal, F., and E. Husebye (1982). Application of arrays in the detection, location, and identification of seismic events, *Bull. Seismol. Soc. Am.* **72,** no. 6B, S201–S224.
Ross, Z. E., B. Idini, Z. Jia, O. L. Stephenson, M. Zhong, X. Wang, Z. Zhan, M. Simons, E. J. Fielding, S.-H. Yun, *et al.* (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence, *Science* **366,** no. 6463, 346–351.
Ross, Z. E., M.-A. Meier, E. Hauksson, and T. H. Heaton (2018). Generalized seismic phase detection with deep learning, *Bull. Seismol. Soc. Am.* **108,** no. 5A, 2894–2901.
Ross, Z. E., D. T. Trugman, E. Hauksson, and P. M. Shearer (2019). Searching for hidden earthquakes in southern California, *Science* **364,** no. 6442, 767–771.
Rost, S., and C. Thomas (2002). Array seismology: Methods and applications, *Rev. Geophys.* **40,** no. 3, 2–1.
Rubin, A. M., D. Gillard, and J.-L. Got (1999). Streaks of microearthquakes along creeping faults, *Nature* **400,** no. 6745, 635–641.
Sánchez-Reyes, H., D. Essing, E. Beaucé, and P. Poli (2021). The imbricated foreshock and aftershock activities of the Balsorano (Italy) mw 4.4 normal fault earthquake and implications for earthquake initiation, *Seismol. Res. Lett.* **92,** no. 3, 1926–1936.
Schaff, D. P., G. H. Bokelmann, W. L. Ellsworth, E. Zanzerkia, F. Waldhauser, and G. C. Beroza (2004). Optimizing correlation techniques for improved earthquake location, *Bull. Seismol. Soc. Am.* **94,** no. 2, 705–721.
Scholz, C. H. (2015). On the stress dependence of the earthquake b value, *Geophys. Res. Lett.* **42,** no. 5, 1399–1402.
Scholz, C. H., Y. J. Tan, and F. Albino (2019). The mechanism of tidal triggering of earthquakes at mid-ocean ridges, *Nat. Commun.* **10,** no. 1, 2526.
Shapiro, N. M., D. Droznin, S. Y. Droznina, S. Senyukov, A. Gusev, and E. Gordeev (2017). Deep and shallow long-period volcanic seismicity linked by fluid-pressure transfer, *Nat. Geosci.* **10,** no. 6, 442–445.
Shelly, D. R. (2020). A high-resolution seismic catalog for the initial 2019 ridgecrest earthquake sequence: Foreshocks, aftershocks, and faulting complexity, *Seismol. Res. Lett.* **91,** no. 4, 1971–1978.
Shelly, D. R., and W. A. Thelen (2019). Anatomy of a caldera collapse: Kīlauea 2018 summit seismicity sequence in high resolution, *Geophys. Res. Lett.* **46,** no. 24, 14,395–14,403.
Shelly, D. R., G. C. Beroza, and S. Ide (2007). Non-volcanic tremor and low-frequency earthquake swarms, *Nature* **446,** no. 7133, 305–307.
Shi, P., F. Grigoli, F. Lanza, G. C. Beroza, L. Scarabello, and S. Wiemer (2022). Malmi: An automated earthquake detection and location workflow based on machine learning and waveform migration, *Seismol. Res. Lett.* **93,** no. 5, 2467–2483.
Shi, Y., and B. A. Bolt (1982). The standard error of the magnitude-frequency b value, *Bull. Seismol. Soc. Am.* **72,** no. 5, 1677–1687.
Soubestre, J., L. Seydoux, N. Shapiro, J. De Rosny, D. Droznin, S. Y. Droznina, S. Senyukov, and E. Gordeev (2019). Depth migration of seismovolcanic tremor sources below the Klyuchevskoy volcanic group (Kamchatka) determined from a network-based analysis, *Geophys. Res. Lett.* **46,** no. 14, 8018–8030.
Takanami, T., and G. Kitagawa (1988). A new efficient procedure for the estimation of onset times of seismic waves, *J. Phys. Earth* **36,** no. 6, 267–290.
Waldhauser, F., and W. L. Ellsworth (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.* **90,** no. 6, 1353–1368.
White, M. C., H. Fang, N. Nakata, and Y. Ben-Zion (2020). Pykonal: A Python package for solving the Eikonal equation in spherical and Cartesian coordinates using the fast marching method, *Seismol. Res. Lett.* **91,** no. 4, 2378–2389.
Wiemer, S., and M. Wyss (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan, *Bull. Seismol. Soc. Am.* **90,** no. 4, 859–869.
Woollam, J., J. Münchmeyer, F. Tilmann, A. Rietbrock, D. Lange, T. Bornstein, T. Diehl, C. Giunchi, F. Haslinger, D. Jozinović, *et al.* (2022). Seisbench—A toolbox for machine learning in seismology, *Seismol. Res. Lett.* **93,** no. 3, 1695–1709.
Zhang, Q., and G. Lin (2014). Three-dimensional vp and vp/vs models in the coso geothermal area, California: Seismic characterization of the magmatic system, *J. Geophys. Res.* **119,** no. 6, 4907–4922.
Zhou, Y., H. Yue, L. Fang, S. Zhou, L. Zhao, and A. Ghosh (2022) An earthquake detection and location architecture for continuous seismograms: Phase picking, association, location, and matched filter (palm), *Seismol. Res. Lett.* **93,** no. 1, 413–425.
Zhu, W., and G. C. Beroza (2019). PhaseNet: A deep-neural-network-based seismic arrival-time picking method, *Geophys. J. Int.* **216,** no. 1, 261–273.

Manuscript received 12 July 2023
Published online 4 December 2023

Downloaded from http://pubs.geoscienceworld.org/ssa/srl/article-pdf/95/2A/1030/6257418/srl-2023230.1.pdf
by Columbia University user