The matrix profile: a fast and sensitive template matching method for seismic event detection that does not require templates

Nader Shakibay Senobari[1], Gareth Funning[1], Zachary Zimmermann[2], Yan Zhu[2], Peter M. Shearer[3], Phillip Brisk[1], and Eamonn Keogh[1]

[1]University of California, Riverside
[2]Google
[3]U.C. San Diego

May 22, 2023

The matrix profile in seismology: template matching of everything with everything

Nader Shabikay Senobari[1], Peter M. Shearer[2], Gareth J. Funning[3], Zachary Zimmerman[1,4], Yan Zhu[1,4], , Philip Brisk[1]and Eamonn Keogh[1]

[1]Department of Computer Science and Engineering, University of California, Riverside

[2]Scripps Institution of Oceanography, University of California, San Diego

[3]Department of Earth and Planetary Sciences, University of California, Riverside

[4]Now at Google

Key Points:

- The matrix profile finds the maximum autocorrelation of every subsequence in a time series
- Peaks of high similarity in the matrix profile can be used to identify seismic events and similar event pairs
- The matrix profile method has similar sensitivity to template matching but does not require *a priori* templates

Corresponding author: Nader Shakibay Senobari, `nshak006@ucr.edu`

Abstract

Template matching has proven to be an effective method for seismic event detection, but is biased toward identifying events similar to previously known events, and thus is ineffective at discovering events with non-matching waveforms (e.g., those dissimilar to existing catalog events). In principle, this limitation could be overcome by cross-correlating every segment (possible template) of a seismogram with every other segment to identify all similar event pairs, but doing so would be computationally infeasible for long time series. Here we describe a method, called the 'Matrix Profile' (MP), a "correlate everything with everything" calculation that can be efficiently and scalably computed. The MP returns the maximum value of the correlation coefficient of every sub-window of continuous data with every other sub-window, as well as the best-correlated sub-window location. Here we show how MP methods can obtain valuable results when applied to months and years of continuous seismic data in both local and global case studies. We find that the MP can identify many new events in Parkfield, California seismicity that are not contained in existing event catalogs and that it can efficiently find clusters of similar earthquakes in global seismic data. Either used by itself, or as a starting point for subsequent template matching calculations, the MP is likely to provide a useful new tool for seismology research.

Plain Language Summary

Detecting and cataloguing earthquakes through analysis of seismic data—the shapes of seismic waves, recorded by seismometers—is foundational to our understanding of Earth's interior structure and processes, as well as geological hazards such as earthquakes. Methods to improve the efficiency and sensitivity of earthquake detection while maintaining accuracy, are critical, as seismic data volumes have grown exponentially in recent decades. Recently, methods that use recorded earthquakes as template patterns to identify in seismic data, have proven capable of detecting several times more earthquakes than traditional methods. However, such methods require knowledge of the template earthquakes ahead of time, and are best at identifying earthquakes whose waveforms are similar to the templates.

We present here a new method, called the 'Matrix Profile' (MP), that takes short windows of data from a seismic data stream, and compares them to all other parts of that data stream, identifying parts of the data that are highly similar. The MP effec-

tively identifies earthquakes, even those which are hidden within noise, does not require any templates to be provided upfront, and can be efficiently calculated. We demonstrate its success at detecting earthquakes in different types of seismic data, and provide practical guidelines for applying the method and interpreting the results.

1 Introduction

Event detection from seismograms has been a key part of seismology research and applications for over a century. Most often the focus has been on earthquakes, but explosions, volcanic eruptions, landslides and other sources of seismic waves can also be studied. At one time, this was a task largely assigned to human experts—analysts who specialized in manually picking the arrivals of specific seismic phases. Smaller-scale studies and local seismic networks may still make use of such expertise but as the potential applications and quantity of seismic data have grown in tandem, it is often not possible or economically viable to rely on manual picking. Thus, there is a growing and presently unmet need for automated methods and algorithms that can efficiently mine large data volumes for seismic events.

Many automated and semi-automated seismic event detection methods have been developed during the last few decades (Vassallo et al., 2012). Perhaps the most common approach uses the ratio between the short-term average (STA) and the long-term average (LTA) of the absolute seismogram amplitude (e.g., Allen, 1982; Earle & Shearer, 1994). Immediately following the arrival of seismic wave, we expect a sharp increase in the STA relative to the LTA, with the latter a measure of the pre-event noise in the seismogram. Many earthquake monitoring networks rely on the STA/LTA method, despite numerous efforts toward developing more sensitive and advanced approaches. Seismic networks produce catalogs that are often the basis of other products relating to seismic hazards, making the accuracy and reliability of catalogs crucial. Considering the power law increase of the volume of seismic data in recent years (e.g., Hutko et al., 2017), there is an urgent need for improved algorithms to process seismograms for event detection.

One widely used approach for event detection is template matching (also known as ‘matched filtering’ in the seismological literature and ‘query search’ in the computational data mining literature). In recent years, the seismological community has adopted template matching methods in order to detect smaller or more emergent events than those found in standard catalogs. The method uses the waveforms of known events as templates ('motifs' in the data mining literature), which are cross-correlated with continuous seismic waveforms in order to identify other similar events, indicated by peaks in the cross-correlation function (Shelly et al., 2006, 2007; Gibbons & Ringdal, 2006). Each earthquake recording represents a convolution between the earthquake source, the path taken by seismic waves, and the instrument that recorded it. Thus, this method is particularly effective at identifying events with similar mechanisms and/or locations to the template event, and can reveal similarly shaped events with disparate amplitudes, sometimes at or below the noise level.

In a recent study, template matching using cataloged earthquakes as the templates identified over 10 times more earthquakes than were present in the original catalog (Ross et al., 2019). While template matching is computationally intensive, several template matching software packages have become available, the most advanced of which exploit multi-core CPU or GPU parallelism to achieve high performance (e.g., Beaucé et al., 2018; Chamberlain et al., 2018; Shakibay Senobari et al., 2019). These codes can calculate the exact cross-correlation function for multiple known templates simultaneously, with the amount of available RAM and number of available cores as the main factors that limit performance. Conceptually, template matching can only detect new events that are similar in shape to the templates themselves, which are known cataloged events. They are far less effective at detecting new events whose waveforms are dissimilar to the templates. Consequently, the newly detected events tend to cluser near existing events, and events in locations that lack previously cataloged events tend to remain sparse; the resulting earthquake catalogs tend to exhibit high spatial variance in terms of completeness.

Developments in the field of machine learning offer another potential means to increase the number of seismic events detected in continuous seismic waveforms. In particular, Deep Learning (DL) approaches using Convolutional Neural Ntworks (CNNs) have been successfully trained to identify and pick P- and S-wave onsets (e.g., Perol et al., 2018; Ross et al., 2018; Dokht et al., 2019; Zhu et al., 2019). DL approach can identify more events than were originally detected in network catalogs (Mousavi et al., 2020), suggesting opportunities to improve event detection. In contrast to methods such as template matching, DL approaches often suffer from the 'black-box problem' (Alzubaidi et al., 2021; Sarker, 2021), wherein the results lack interpretability and cannot be traced back to how they were obtained. Unlike other machine learning systems which make decisions based on logical rules, DL models make decisions based on a learned representation of the data in a neural network, which is not very interpretable for humans (Chakraborty et al., 2017). In fields such as healthcare and medicine where interpretability and the uncertainty scale (statistics of output results) are crucial, lack of interpretability can be more problematic (Chakraborty et al., 2017; Sarker, 2021). In seismology, reliability and uncertainty scaling play an important role, as the final results are sometimes used for hazard and disaster assessments. Further, DL models require considerably more training data than other machine learning approaches (e.g., Alzubaidi et al., 2021; Sarker, 2021): to ensure robustness, a large training data set must include a wide variety of representations (classes) and numerous representative examples (templates) for each class. When seismic data characteristics change (e.g., network changes, noise characteristics at different stations), the versatility of an existing trained model may be compromised. While all these issues with DL have been identified and discussed in many different fields, pragmatic and computationally efficient solutions remain an open research problem (Alzubaidi et al., 2021; Sarker, 2021).

Another potential approach to is to identify patterns (motifs) in seismic waveforms, with the recognition that earthquakes are more similar to other earthquakes than they are to noise. In a recent series of studies, a fast motif discovery methodology was developed that made use of 'fingerprinting' – converting seismic time series to small and dense proxies, or 'fingerprints' and then performing Locality-Sensitive Hashing (LSH) on them (e.g., Yoon et al., 2015; Bergen & Beroza, 2018). LSH is a fast approximate nearest neighbor search method that reduces the similarity search dimensions by mapping the domain of the similarity search to smaller domains containing similar objects with a high probability (Indyk & Motwani, 1998; Gionis et al., 1999). Although the LSH can speed up similarity search, it can generate both false positive and false negative results. LSH also requires careful selection of a number of tuning parameters that strongly influence the success of the search, and whose values may vary for different regions, data sets and applications. The tuning parameter selection process requires visual inspection and validation against the results of other methods, which we consider to be a significant drawback.

In this study we describe and illustrate an alternative method for seismic event detection, the 'Matrix Profile' (MP), which overcomes the shortcomings of template matching, DL approaches, and LSH. The MP offers sensitivity comparable to template matching, but does not require or utilize *a priori* event templates. Further, the MP is interpretable, does not generate false positive or negative results, and relies on a single parameter. The MP is similar to the autocorrelation method (e.g., Brown et al., 2009), which we further describe below, in that each sub-window of continuous waveform data becomes a potential template, but it has a substantially lower computational overhead enabling the rapid analysis of very large seismic datasets. Even for signals below the noise level, the MP can highlight repeated or near-repeated events in continuous data sets. Using data from both Parkfield, California, and global seismic network stations, we show how the information provided by MP processing can be used for event detection and other seismological applications.

2 The Similarity Matrix and the Matrix Profile

2.1 The Similarity Matrix

In the absence of *a priori* templates, we can attempt to identify repeating patterns in seismic waveform data by autocorrelation, i.e., by comparing subsets of the waveform to other parts of the waveform. This approach was introduced by Brown et al. (2008) and is mainly used to search for low frequency earthquakes (Brown et al., 2009; Royer & Bostock, 2014). Since the autocorrelation method is computationally intensive, it has only been applied to short periods of continuous seismic data (e.g., one hour; Royer & Bostock, 2014).

The results of such ‘all subsequences comparisons’ within a single continuous waveform can be represented as a ‘Similarity Matrix’ of correlation coefficients (CCs), or any other similarity measure such as Euclidean distance, between all pairs of subwindows in a data vector. In other words, any subwindow of length m in a data set of length n (where $n >> m$) is compared with all other length-m subwindows yielding a matrix

$$M_{ij} = \mathrm{CC}(x_i, x_j)_m \tag{1}$$

where CC indicates the correlation coefficient, and i and j are the locations of subsequences within the data vector x.

The Similarity Matrix can be computed at maximal resolution by sliding the subwindow by a single data point, or at a lower resolution by skipping over some data points.

In the former case, there will be approximately n individual subwindows that will be compared with approximately $n-1$ other subwindows, and the resulting Similarity Matrix will comprise $\sim n(n-1)$ CCs. While the resulting Similarity Matrix will have complete information about the self-similarity within a data vector, much of the information will be redundant or unimportant. For example, the similarities between parts of the data that only contain background seismic noise with other parts of the dataset do not have useful information for event detection purposes, but may comprise the majority of the computed CCs.

Another drawback is that the Similarity Matrix grows in size with the square of the length of the data vector, requiring very large amounts of memory and storage capacity, even for modest data vector lengths. For a brute force time domain method for computing CCs, the time complexity for computation of the similarity matrix is $\mathrm{O}(n^2m)$ and the memory complexity (memory required) is $\mathrm{O}(n^2)$. A time series data vector for 24 hours of data from a single seismometer component sampled at 20 Hz comprises $n = 24 \times 60 \times 60 \times 20 = 1,728,000$ data points. The corresponding similarity matrix will have $n^2 \approx 2.986 \times 10^{12}$ elements, and assuming it is stored as single precision floating point numbers, will require $4 \times 2.986 \times 10^{12} \approx 1.194 \times 10^{13}$ bytes of storage, equivalent to $\sim 11,200$ GB of RAM. Thus, computing the Similarity Matrix for a time series longer than just a few hours rapidly becomes impractical, despite the success of similar approaches at detecting new events (e.g., Brown et al., 2009).

Clearly, the Similarity Matrix and related methods have the capability to identify events for which we have no prior templates. The probability of detecting more events will only increase with the length of the time series being analyzed, but without some means of reducing the volume of the output, the approach cannot scale. The Matrix Profile, a product derived from the Similarity Matrix, which returns self-similarity information about a time series in a much more efficient format, overcomes many of these limitations.

2.2 The Matrix Profile

The Matrix Profile (hereafter, 'MP') is an approach developed in the computer science community in recent years to extract exact self-similarity information from a time series in an efficient manner (e.g., Yeh et al., 2016; Zhu et al., 2016). All of the subsequence comparisons necessary to assemble the full Similarity Matrix are evaluated, but to minimize storage requirements, only the maximum CC value for each subsequence is retained, along with indexing information which indicates the position of the most similar subsequence. Using this method, the subwindow is shifted by a single data point for each subsequence comparison. The immediate vicinity of the subsequence – points in the time series that are less than one subsequence length away – are excluded from the results, as we would expect them to be very similar. The name 'Matrix Profile' originates in the idea that this output is a derived product, i.e., a 'profile', of the full Similarity Matrix – the row vector that one would obtain by searching for the maximum value of each column of the Similarity Matrix, excluding the diagonal and its immediate neighbors.

The main advantage of the MP compared to the Similarity Matrix is storage efficiency, as both intermediate and final outputs have a linear, rather than quadratic, storage requirement. SCAlable Matrix Profile (SCAMP; Zimmerman et al., 2019) is the fastest algorithm to compute the MP, and can be deployed at the data center scale (multi-node CPU or GPU). Using SCAMP, MP values can be computed using CC (hereafter r value), which is the same similarity measure used for standard template matching studies.

In essence, the MP indicates the extent to which each subsequence in a larger time series is repeated at least once. Here we explore several applications of this feature in seismology, as it can be a high-quality indicator to distinguish meaningful seismic events from noise and to identify pairs or clusters of repeating events. We demonstrate in several examples below how the MP can be used to mine seismic data to detect new and undiscovered events.

3 The 2004 Parkfield earthquake and aftershocks

For a local-scale dataset, we use data from the Parkfield region in central California, one of the best instrumented places in the world due to the Parkfield Earthquake Prediction Experiment (Bakun & Lindh, 1985). Specifically, we use data from the Parkfield High Resolution Seismic Network (HRSN), a dense array of borehole seismometers, due to their high sensitivity and low noise levels. We band-pass filter the data between 2 and 8 Hz, a frequency range suitable for detecting both local earthquakes and low frequency earthquakes (LFEs) and resample the waveforms to 20 Hz. We set the MP query length at 100 samples, equivalent to 5 seconds of data—sufficiently small to detect events with magnitudes below zero as well as low-frequency earthquakes (Shelly et al., 2009), but also capable of detecting aftershocks up to $M_w \sim 5.0$. We computed MP results for two different time periods:

1. We use 90 days of continuous HRSN data from stations LCCB, SMNB, and VARB and one Northern California network station (PGH) from 2004-08-25 to 2004-11-22. This includes about 1 month before the 09-28 M6 Parkfield mainshock and the first two months of aftershocks. The experiments were conducted on a system with two NVIDIA P100 GPUs and took approximately 12 hours for each station for 155,520,000 samples.
2. We use 580 days of horizontal component seismic data from HRSN station VCAB starting on 2003-11-30, thus including about 9 months of Parkfield aftershocks. For this data set, the experiment took 375.6 GPU hours on 40 NVIDIA V100 GPUs on an AWS EC2 spot instance fleet. Spot jobs for this particular data set took 2.5 days, since the instances at the time were in high demand. When instances were not in high demand, the spot job time was around 10 hours for a similar experiment (Zimmerman et al., 2019). In order to allow for further exploration, we have placed this data set in a public repository (`https://github.com/Naderss/MP4Seismo`).

Figure 1 shows an example of the MP output for Parkfield station SMNB. At each time point t_i, the maximum value of the correlation coefficient (r) is returned for the cross-correlation of the window from t_i to t_{i+m} (where $m = 100$ for 5 s of data) with the entire time series (excluding times near t_i to avoid the peak in r at zero lag), together with the location (index) of the starting point for the window that gives the maximum correlation. The top panel of Figure 1 shows the r values, the middle panel is the filtered seismogram, and the bottom panel shows the index values for $r \geq 0.7$. The predicted P arrival times from 4 events in the Parkfield template-matched catalog of Neves et al. (2023) are shown as the vertical blue tick marks. Notice that the MP r values rise at the time of each event, indicating that the time series is correlated with some other part of the time series at those times. Often these index values change depending upon the exact location of the template window, indicating that there are multiple locations within the time series that are well-correlated with the template event.

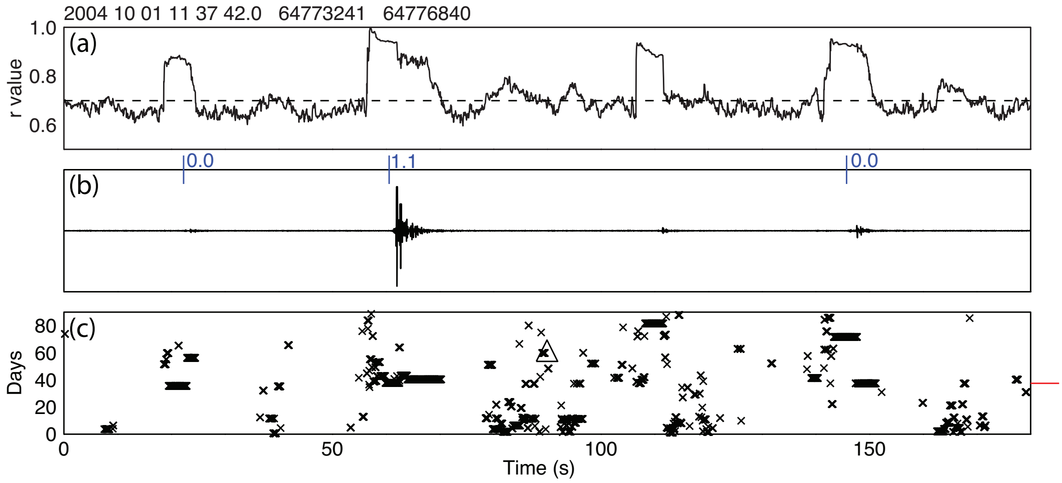

Figure 1. Matrix Profile (MP) output for 130 s of data from Parkfield station SMNB (starting at 2004-10-01 11:37:42, about 3 days after the M6 mainshock). (a) The MP correlation coefficient (r), showing the maximum correlation of a sliding 5 s-window with the rest of the selected 90-day-long seismogram. (b) The SMNB seismogram during the same time period. (c) The MP index values, showing when in the 90-day period (see y-axis) the maximum correlation with the sliding window in (a) occurred. The horizontal dashed line in (a) shows a reference r value of 0.7. The vertical tic marks at the top of panel (b) show the predicted P-wave arrival times and magnitudes of four events from the template-matched earthquake catalog of Neves et al. (2023). The horizontal line to the right of (c) indicates the time of window (a) within the full 90 days of data.

There is an apparent event at about 115 s in Figure 1 (seen in both the seismogram and the MP r values) that is not listed in either the NCSN or Neves et al. (2023) catalogs even though it is comparable in amplitude to the known event at about 25 s. This event could also have been detected using an STA/LTA filter but the MP output has the advantage of showing that the pulse is correlated with other signals in the data, which allows for additional processing options. This is illustrated in Figure 2, which shows MP results near a M 0.9 NCSN catalog earthquake (2004-10-02 11:01). For the value of t_i shown as the left red vertical tick mark in panel (a), the corresponding index value is shown as the triangle in panel (c), indicating the best waveform match occurs about 18 days later (2004-10-20). MP results for this index time are shown in panels (d)–(f). The time series from panel (a) is plotted in blue in order to compare with the correlated time series at the later time. The waveforms are nearly identical, as might be expected given the high r value (0.9627) of the correlation. Notice that the high correlation continues into the earthquake coda, beyond the time window used in the MP calculation to identify the correlated event.

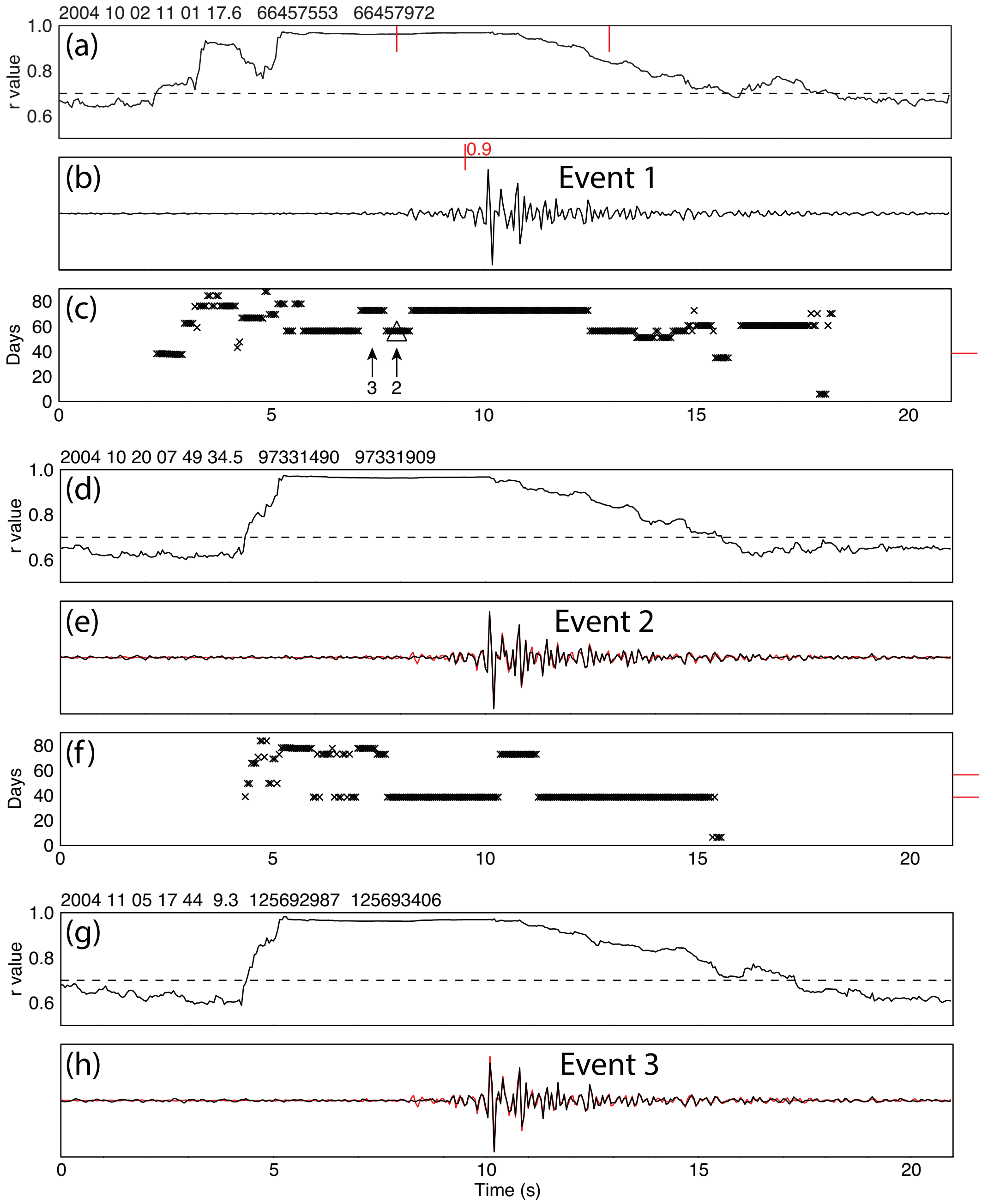

Figure 2. An example of how similar earthquakes can be extracted from MP output. (a) MP r values for about 21 s of data from Parkfield station SMNB (starting at 2004-10-02 11:01:18, about 4 days after the M6 mainshock), which includes an NCSN catalog earthquake of M 0.9. The vertical red tick marks show the specific 5-s window we use to extract the index value used in panels d–f. (b) The SMNB seismogram during the same time period. (c) The MP index values, showing when in the 90-day period (see y-axis) the maximum correlation with the sliding window in panel (a) occurred. The triangle (and arrow labeled with 2) indicates the index value at the start of the time window shown in panel a and used for panels d–f for Event 2. Panels (d) – (f) are similar to panels a–c, but show the results at the time defined by the index value in panel c, which resolves an earthquake on 2004-10-20 (Event 2) that is very similar to the 2004-10-02 event (Event 1). For comparison, the seismogram from panel a is plotted in red in panel e, showing that it is nearly identical to the later record (plotted in black). Panels (g) – (h) are similar to panels e–f, but show the results at the time defined by the index value marked by the arrow and number 3 in panel a, which point to another similar earthquake occurring on 2004-11-05 (Event 3).

This figure illustrates several properties of the MP results for the Parkfield data. As is known from previous studies of repeating earthquakes (e.g., Nadeau et al., 1995; Nadeau & Johnson, 1998; K. H. Chen et al., 2010) and template-matched catalogs (Peng & Zhao, 2009; Meng et al., 2013; Neves et al., 2023), there are many earthquake pairs in the Parkfield seismicity with very similar waveforms. Often, multiple matching events can be extracted from the MP results from a single starting event. For example, if we move the time window in panel (a) to the left, then the maximum correlation ($r = 0.9618$) is found for a different matching event, which occurred on 2004-11-05. The waveform match to the panel (a) event is shown in panel (h). As shown in panel (f), additional similar events can be obtained by checking the index values from the MP results for the matching events. In this way, a cluster of similar events can be obtained, even if the MP results do not directly connect every pair of events in the cluster.

Often the actual peak (maximum) in the MP r values occurs at the start of a much broader plateau with nearly constant r. This can occur when the leading edge of the cross-correlation window crosses the first arrival (usually the P-wave). The first few cycles of the P-wave arrival can be very similar for earthquake pairs that are not so well correlated when more of the wavetrain is included in the cross-correlation window. Thus more reliable results for identifying truly similar events are obtained from later parts of the MP plateau, which represent cross-correlations that include substantial parts of the earthquake signal, not just the initial arrivals. During these times, the index values tend to be relatively stable for several tenths of seconds or longer, whereas greater variability in the index values in typically seen when only the initial part of the P-wave is included in the cross-correlation window.

3.1 Extracting Parkfield similar event clusters from MP results

The MP returns an enormous amount of information about correlated waveforms throughout a time series, which can be used in a variety of ways. For our Parkfield data, one goal is to identify clusters of similar earthquakes, which could then be compared to existing earthquake catalogs. After some experimentation, we adopted the following approach:

1. We define a minimum correlation coefficient, $r_{\min}$ and consider only MP results with $r \geq r_{\min}$. We tried $r_{\min}$ values of 0.99, 0.95, and 0.90. As might be expected,

larger values of $r_{\min}$ lead to more highly correlated waveforms, but return smaller numbers of similar events.

2. Beginning with the first point ($t = 0$ and proceeding through the time series one point at a time, we save times, t_{peak}, from the MP r peaks that define similar waveforms that meet the following criteria: (a) there are at least 40 points (2 s) with $r \geq r_{\min}$ before t_{peak}, (b) the next r value is less than the current r value or the index value changes to a point at least 10 s away from the current index value, (c) there are no existing saved t_{peak} values in which the current time is within 10 s of t_{peak} and the corresponding index values are also within 10 s.
3. This list of times (points in the MP time series) defines a series of similar event pairs. We begin by defining each of these pairs as a cluster of two events. We then iteratively combine clusters in which the time of either event is within 3 s of the time of either of the events in the other cluster and continue iterating until no more clusters are combined. In this process, we also use the relative timing information to compute time shifts for the events within each cluster to align the waveforms.

Figure 3 shows one of the similar event clusters extracted from the SMNB time series using $r_{\min} = 0.95$. In this case, one M 0.9 event was in the NCSN catalog and one smaller event was included in the Neves et al. (2023) catalog, but there are 5 additional events not contained in these catalogs. Note that the small bump in the 2004-10-02 event at about 0.7 s is not seen for the other earthquakes, an indication this is likely a small precursory event. This inference is only made possible by inspecting the group of similar events rather than the 2004-10-02 event by itself. MP methods provide an efficient way to extract these similar events.

Most of the similar event clusters that we have identified in the Parkfield MP results contain events that are already in existing catalogs but also often contain additional events. Presumably these events could have been found using template matching (but perhaps did not meet signal-to-noise or other requirements). However, we also find examples of similar clusters in which none of the events are in existing catalogs, as shown in Figure 4 from MP results for station VARB using $r_{\min} = 0.95$. None of these 7 events are in the NCSN or Neves et al. (2023) catalogs and thus would be very difficult to identify without using MP, which does not require any pre-defined template events.

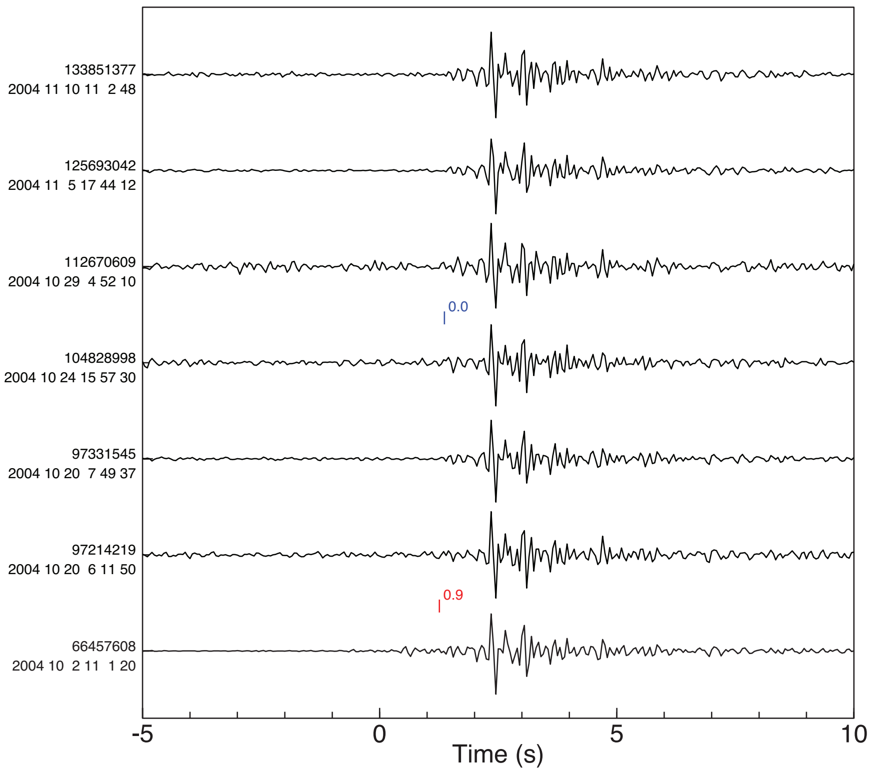

Figure 3. Seven earthquakes with similar waveforms obtained from MP results from 90 days of data from Parkfield station SMNB. Seismogram sample number and data/time for the plotted start times are shown to the left. Seismograms are self-scaled to the same maximum amplitude. Predicted P arrival times for a M 0.9 NCSN catalog earthquake and a M 0.0 earthquake from the template-matched catalog of Neves et al. (2023) are shown as red and blue ticks, respectively.

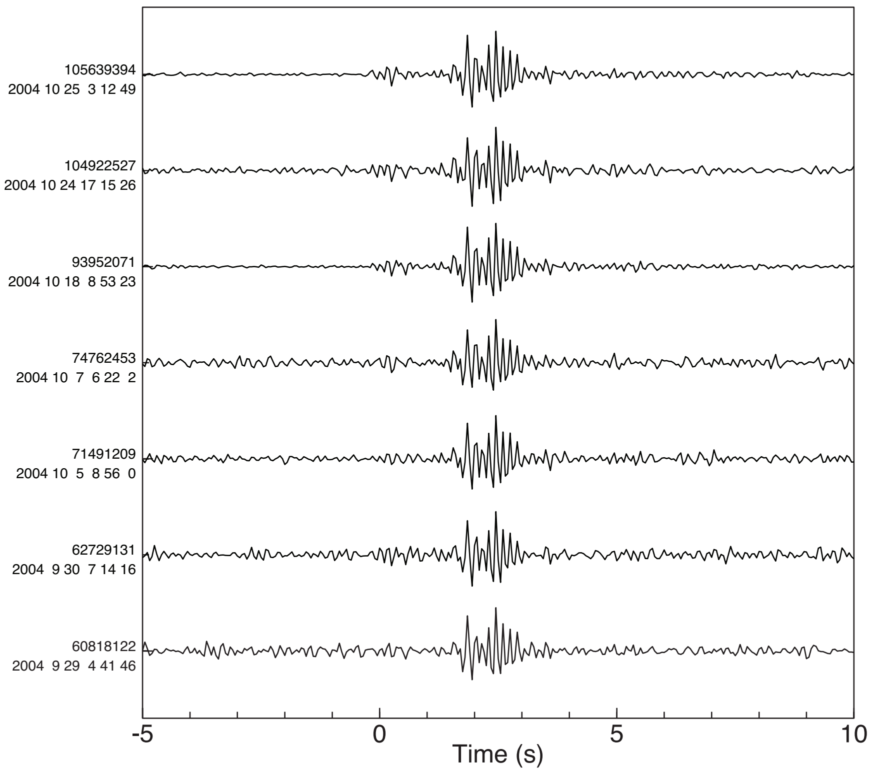

Figure 4. Seven earthquakes with similar waveforms obtained from MP results from 90 days of data from Parkfield station VARB. Seismogram sample number and data/time for the plotted start times are shown to the left. Seismograms are self-scaled to the same maximum amplitude. None of these earthquakes are included in the NCSN catalog or the template-matched catalog of Neves et al. (2023).

Our Parkfield MP results mainly yield earthquake-like signals but also can extract seismometer calibration pulses and other artificial signals. One of the more interesting signals that MP identifies is tremor, which is known to occur in the Parkfield region (e.g., Nadeau & Dolenc, 2005; Nadeau & Guilhem, 2009; Shelly, 2009; Shelly & Hardebeck, 2010a; Shelly, 2017). Figure 5 shows MP output for station VARB at the time of a tremor-like event. Notice that the MP r value rises more slowly than in the earthquake examples of Figures 1 and 2, reflecting the emergent and "ringy" nature of the tremor signal. Figure 6 shows a set of tremor events from 2004-10-21 to 2004-11-16 that our cluster extraction procedure obtained from the VARB MP results with $r_{\mathrm{min}} = 0.90$ and Figure 7 shows a different set of tremor events (mostly occurring on 2004-07-18) obtained from the VCAB MP results with $r_{\mathrm{min}} = 0.95$. Because the tremor signal is nearly monochromatic, high correlation values are obtained over a 5-s window even when the signals are not "similar" in the traditional sense. In this case, the optimal time shifts are not well-defined as can be seen in the shifting positions of the tremor envelopes in Figure 7.

3.2 Parkfield discussion

Parkfield is one of the most well-studied regions of active seismicity in the world and has been the target of many previous studies, making it a good test case for the application of the MP in seismology. We find that MP calculations can readily extract groups of similar earthquakes, many of which are not contained in existing catalogs. Fully exploiting the power of our Parkfield MP results is beyond the scope of this paper, but it seems very likely that one could obtain more complete catalogs of similar events, which would benefit analyses that are based on properties of these events, such as changes in their repeat times (e.g., K. H. Chen et al., 2010). Studies that utilize repetitive earthquakes, such as locating faults and determining their geometries at depth, studying time-dependent fault slip, analyzing changes in crustal velocity structures over time, or understanding earthquake physics in general (Vidale et al., 1994; Anooshehpoor & Brune, 2001; T. Chen & Lapusta, 2009; Khoshmanesh et al., 2015), would be affected directly or indirectly by this. A more detailed analysis of our Parkfield MP results would involve integrating the results from different stations into a single catalog of similar events and locating and assigning magnitudes to any newly identified events. If desired, the MP-identified event times could be used to perform additional cross-correlation calculations to obtain a more complete measure of the similarity of all event pairs within each clus-

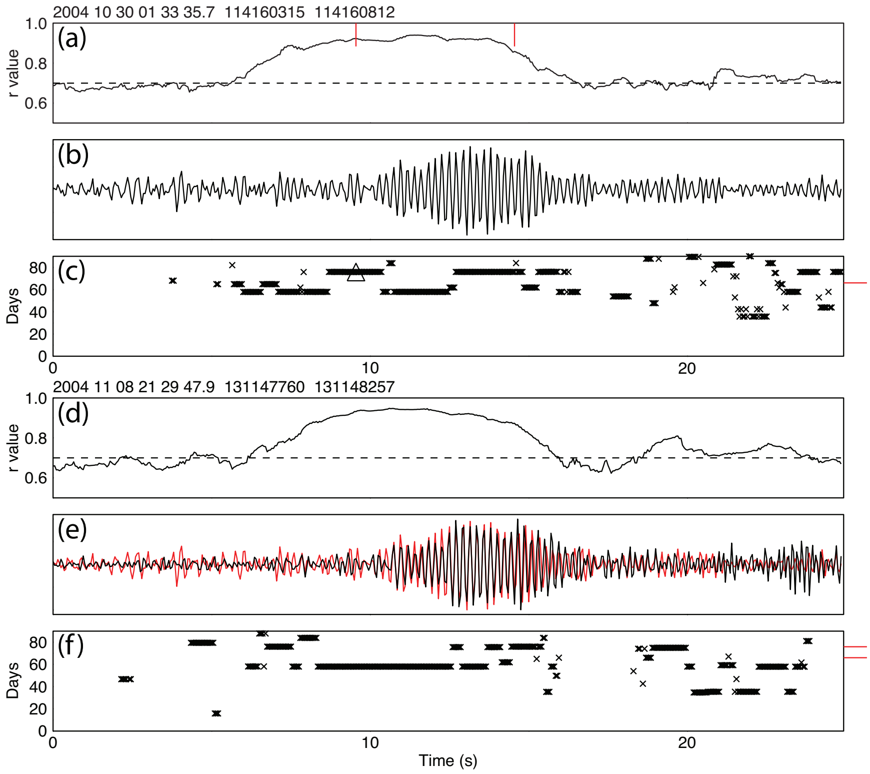

Figure 5. An example of how tremor can be extracted from the Parkfield MP output. (a) MP r values for 25 s of data from Parkfield station VARB (starting at 2004-10-30 01:33:36). The vertical red tic marks show the specific 5-s window we use to extract the index value used in panels d–f. (b) The VARB seismogram during the same time period. (c) The MP index values, showing when in the 90-day period (see y-axis) the maximum correlation with the sliding window in panel a occurred. The triangle indicates the index value at the start of the time window shown in panel a and used for panels d–f. Panels (d) – (f) are similar to panels a–c, but show the results at the time defined by the index value in panel c, which resolves another tremor-like pulse on 2004-11-08. For comparison, the seismogram from panel a is plotted in red in panel e.

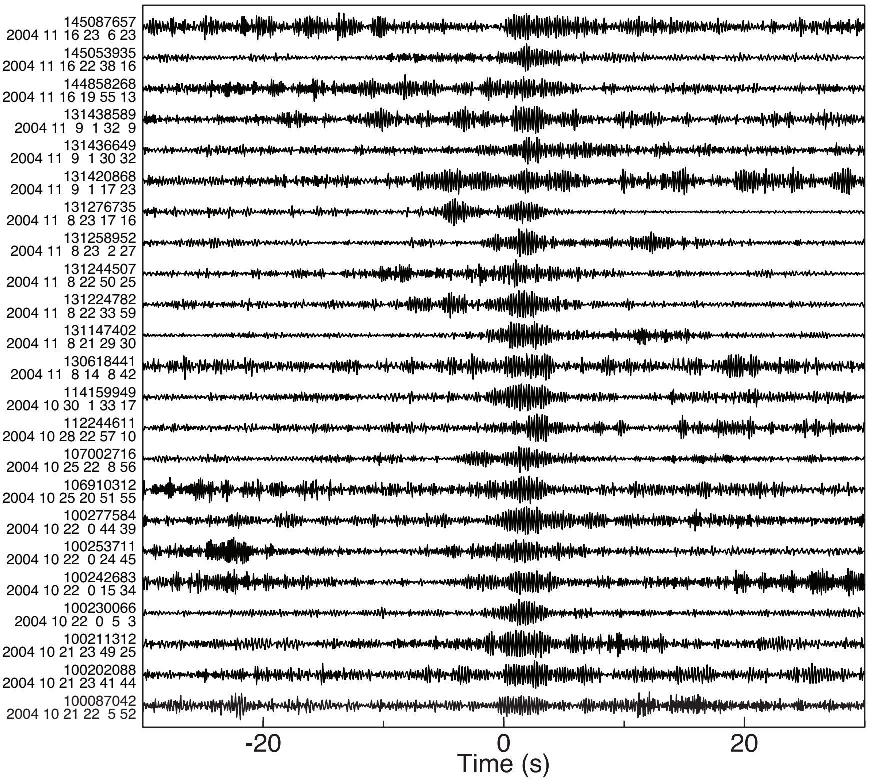

Figure 6. Examples of tremor recorded by Parkfield station VARB, as grouped into an event cluster using the MP output. Date/times and sample numbers are shown to the left for the start of each time window. Seismograms are self-scaled to the same maximum amplitude.

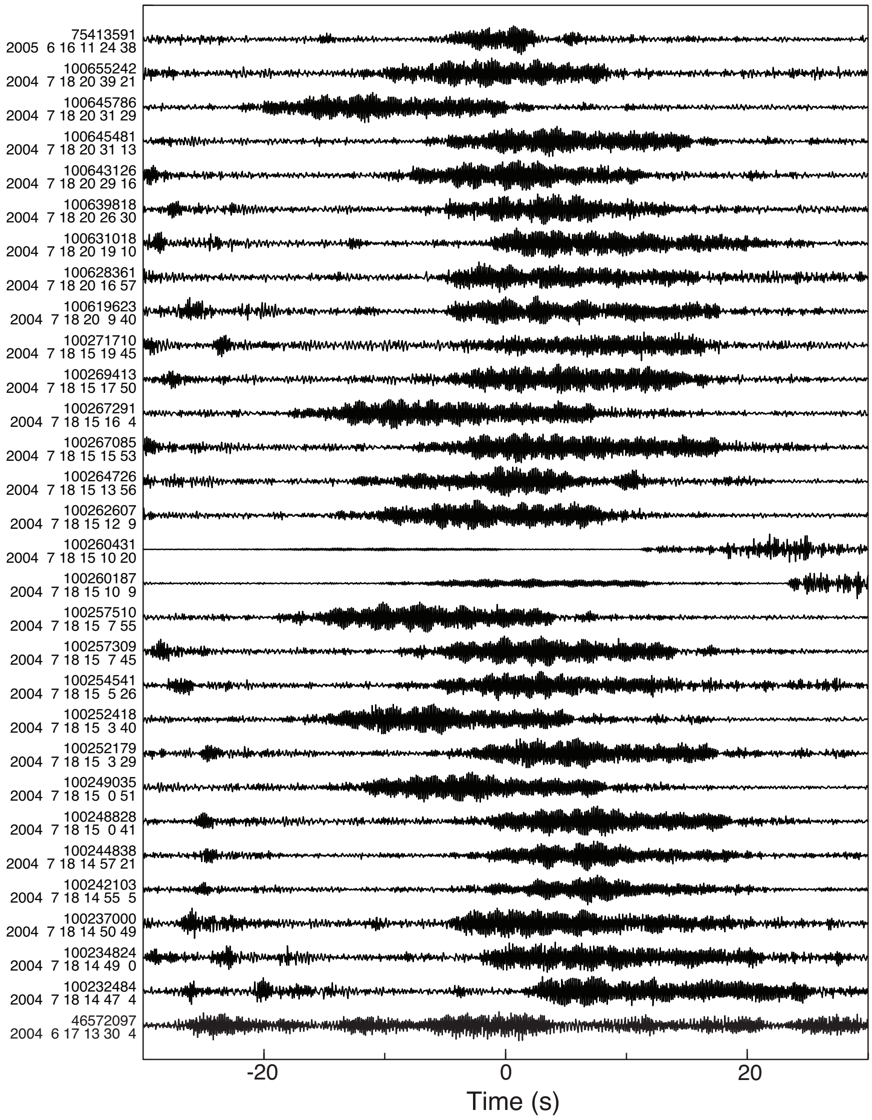

Figure 7. Examples of tremor recorded by Parkfield station VCAB, as grouped into an event cluster using the MP output. Date/times and sample numbers are shown to the left for the start of each time window. Seismograms are self-scaled to the same maximum amplitude.

ter, differential amplitude information, and more precise cross-correlation timing (i.e., subsample accuracy).

We have not yet explored the limits of how low a value of $r_{\min}$ will still yield useful results. The number of similar waveform pairs increases enormously for smaller values of $r_{\min}$, but even at $r_{\min} = 0.8$ the vast majority of the MP peaks appear to represent real geophysical signals and not simply random noise fluctuations. However, poor signal-to-noise in the associated seismograms hampers identification and timing of phase arrivals. In such cases, it may be possible to stack event waveforms from the same similar event cluster in order to reduce the noise levels and facilitate phase picking, but we defer trying this idea to future work.

4 Application to global teleseismic events

To demonstrate the capability of the MP method to detect teleseismic events, we select 20 Global Seismic Network stations (ASL/USGS, 1988), broadly distributed across the globe. For each, we download vertical component (VHZ) data at a standard sample rate of 1 Hz, and band-pass filter the data from 20 sec to 500 seconds, filling any data gaps with random noise using uniformly distributed random numbers. We resample the data to a 4-s sample interval, for which 12 years of data (2007–2018) represents ∼157 million data points. We set the MP query length at 150 samples, equivalent to 600 s (10 minutes) of data. Using two P100 GPUs, it took around 2 hours for each station to compute the MP for this data set.

Figure 8 shows MP results from station ESK in Scotland for a pair of $M \sim 5.2$ earthquakes occurring in central America in 2007 and 2017 at an epicentral distance of 78 degrees. Notice the very similar surface waves, as shown by the red and black lines in panel (e), with the waveform similarity including the body waves in front of the Rayleigh wave arrival and the surface-wave coda. The MP indices for the 2007 event consistently point to the 2017 event, but the MP indices for the 2017 event point to several events, not simply the 2007 event. This index information can be used to extract similar event clusters. We follow the same approach described in the Parkfield section, requiring in the global case that 50 points (200 s) of MP r values exceed $r_{\min}$ before each saved t_{peak} value and that saved t_{peak} values corresponding to nearly the same MP index be at least

30 s apart. We used a maximum allowed separation time of 250 s to join events between clusters.

Figure 9 shows the similar event cluster that contains the two events shown in Figure 8. All seven events are contained in the ANSS catalog (US Geological Survey, 2017) and range from magnitude 4.9 to 6.0. In contrast to our Parkfield results, our initial analysis of similar event clusters obtained from the ESK MP results using $r_{\min}$ does not identify any events not already in existing catalogs. It is possible that new events could be identified if we examine lower values of $r_{\min}$, but we defer this to future work. However, the MP results are nonetheless useful because they efficiently identify groups of similar events that are suited for more detailed processing. For example, waveform cross-correlation could be used to improve their location accuracy (e.g., Waldhauser & Schaff, 2007) or they could be used to examine possible temporal variations along seismic ray paths, such as the PKP paths from earthquake doublets used to examine inner-core differential rotation (e.g., Zhang et al., 2005; Tkalvcic et al., 2013; Yang & Song, 2020; Pang & Koper, 2022).

4.1 Detection of unusual teleseismic events

In our global event detection experiment (Section 4), we expect significant moveout of peaks in the MP between stations at different distances for the onsets of the same event. In order to account for these variable arrival times, we perform a alignment procedure prior to stacking the MPs for all stations (e.g., Shearer, 1994; Ekstöm, 2006). In essence, our method, which we call 'move-max' (moveout for maximum cross-correlation) uses a grid search and a local velocity model to find potential source location(s) and their corresponding time shifts that result in a maximum value of the stacked MP. A snapshot of four days of stacked MPs for 12 stations is shown in Figure 10. In this figure, it can be seen that each stacked MP peak is associated with a specific teleseismic event. Notably, among the detected events we identify the 2018 Mayotte seismovolcanic event – a very long-period seismic event that was not originally detected by seismic networks due to its emergent onset, but repeated several times in the subsequent months (Lemoine et al., 2020). As we can see in Figure 10, the stacked MP for the Mayotte event has a clear peak, showing the utility of the method for identifying unusual seismic events.

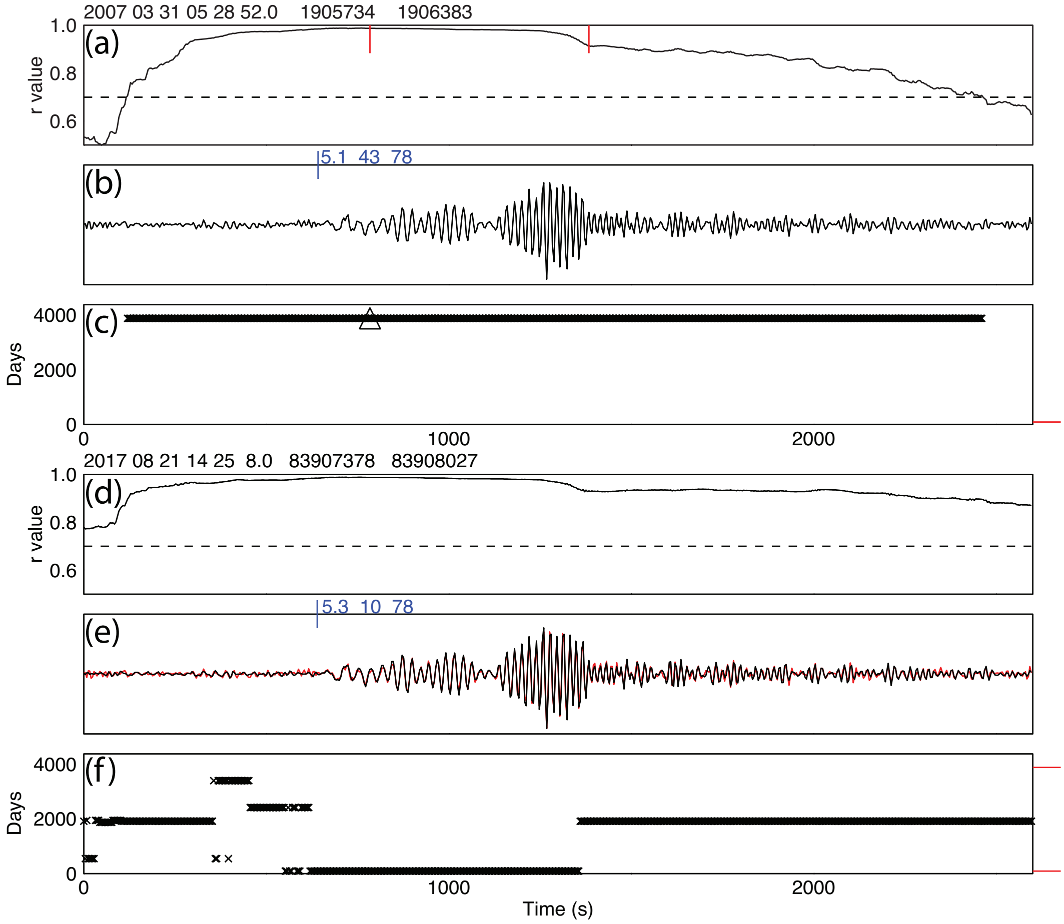

Figure 8. An example of how similar teleseismic earthquake pairs can be extracted from MP output. (a) MP r values for about 2500 s of data from station ESK in Scotland (starting at 2007-03-31 05:28:48), which includes an ANSS catalog earthquake of M 5.1. The red tic marks show the specific 600-s window we use to extract the index value used in the other panels. (b) The ESK seismogram (filtered to periods of 20 to 500 s) during the same time period. (c) The MP index values, showing when in a 12-year period (see y-axis) the maximum correlation with the sliding window in panel (a) occurred. The triangle indicates the index value at the start of the time window shown in panel a. Panels (d) – (f) are similar to (a) – (c), but show the results at the time defined by the index value in panel c, which resolves a M 5.3 earthquake on 2004-11-05 that is very similar to the 2004-10-02 event. For comparison, the seismogram from panel a is plotted in red in panel e, showing that it is nearly identical to the later record (plotted in black).

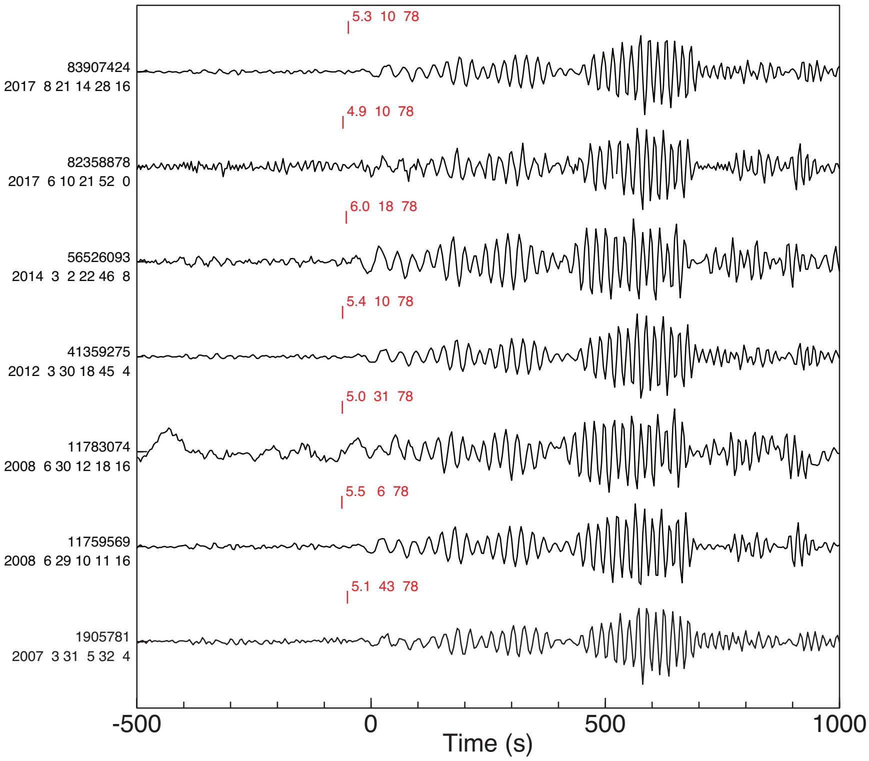

Figure 9. Seven earthquakes with similar waveforms obtained from MP results from 12 years of data from station ESK. Seismogram sample number and data/time for the plotted start times are shown to the left. Seismograms are self-scaled to the same maximum amplitude. All of these earthquakes are in the ANSS catalog and there predicted Rayleigh wave arrivals time are shown as the vertical tic marks, labeled with magnitude, catalog source depth (km), and epicentral distance (degrees).

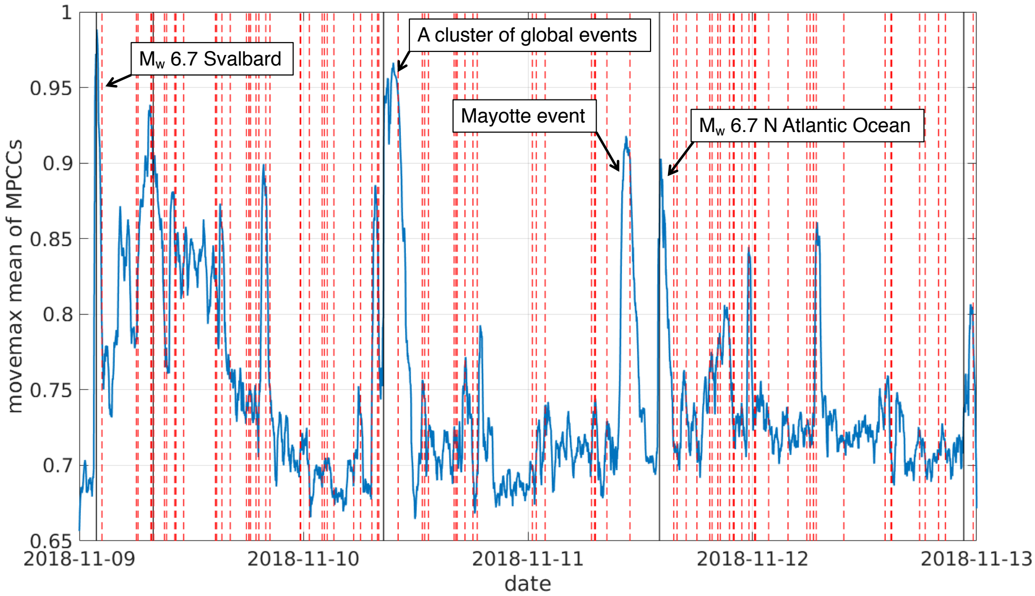

Figure 10. Teleseismic event detection using the matrix profile (MP) method. Shown are stacked MPs from 19 Global Seismic Network stations from a four day period in November 2018. Each MP is calculated from 12 years of data at 0.25 Hz sample rates and band-pass filtered between 20 and 500 second periods. We then apply a move-max filter similar to the method used by Shearer (1994) and Ekstöm (2006) to account for moveout of different seismic sources, and stack them. The MP stack increases when a large teleseismic event occurs (including catalogued and uncatalogued events). Dashed red and solid black vertical lines indicate the origin times of global catalogued events with magnitude below and above 5.5, respectively.

In practice, for novel event detection we need to deal with the large number of peaks associated with regular events (e.g., earthquakes). A user should first eliminate the section of data created by regular catalogued events in order to search for novel events among those resulting from MP thresholding. There may also be some peaks that correspond to regular earthquakes with smaller magnitudes that go undetected by traditional methods of event detection, but could be identified by template matching. As a result, the process of removing known events from MP detected events in order to search for novel events might become time-consuming. A modified version of MP was recently developed by Mercer et al. (2022) which checks for novelty when calculating MP by concatenating continuous waveforms from known templates. The output of this algorithm is a 'contrast profile', which is useful for novel seismic event detection as both the catalogued events and the template-matched events of the catalog automatically are covered (i.e., do not result in a peak in the MP; Mercer et al., 2022).

5 Practical considerations for seismology

In terms of principle and framework, the MP works in the same manner as standard template matching for seismic data. Preprocessing and experimental setup are therefore similar to those involved in template matching. Here, however, we emphasize some of the most important technical details regarding the setting up of MP experiments.

5.1 Selecting the subsequence length

Unlike other methods that have multiple tunable parameters, the MP method only has one parameter, the subsequence length. Our choice of subsequence length can be guided by some general principles – for example, maximizing the difference between the background noise MP value (BNMP) and the peak MP values for the events of interest. If the subsequence length is long then the sliding window will include any noise before or after the event and that can lower the CC between similar events substantially. Conversely if we use a very short subsequence length the BNMP value increases as the probability that two random noise sections with fewer data points matches with each other increases. Both of these issues are encountered when regular template matching is used to determine the template length, and they are usually remedied by choosing a length that includes most of the signal for smaller targeted events (e.g., Ross et al., 2019). As a result of the same underlying reason, we recommend (and use) a subsequence length between 50 and 100 percent of the shortest expected signal duration.

For our large experiment on the Parkfield data set, we set the subsequence length to 5 seconds (100 data points at our 20 Hz sample rate). In determining this length, we considered the duration of small local events below magnitude zero as well as the duration of LFEs observed at nearby stations (Shelly et al., 2009; Shelly & Hardebeck, 2010b). We observe that the BNMP range is about 0.55 to 0.7, compared with MP values above 0.9 for background seismicity and aftershocks of the 2004 Parkfield earthquake (Figure 1 and Section 5.4), providing sufficient contrast to enable event detection.

5.2 Selecting sample rate and band-pass frequency range

As with most seismic data mining algorithms, lowering the sample rate, and therefore reducing the number of data points, can significantly reduce the run time for calculating the MP. Therefore in order to calculate the MP efficiently, we recommend (and use) downsampling of the data to the Nyquist frequency of the highest frequency feature that we are interested in. For example, if we are mining seismic data to search for events below 10 Hz (e.g., local events with $M \sim 1$ or LFEs), the sample rate can be set to 20 Hz. In the case of the Parkfield experiment, our target subsequence duration was 5 seconds and our filter band was 2–8 Hz, and therefore we downsampled the data from 40 Hz (original sample rate) to 20 Hz.

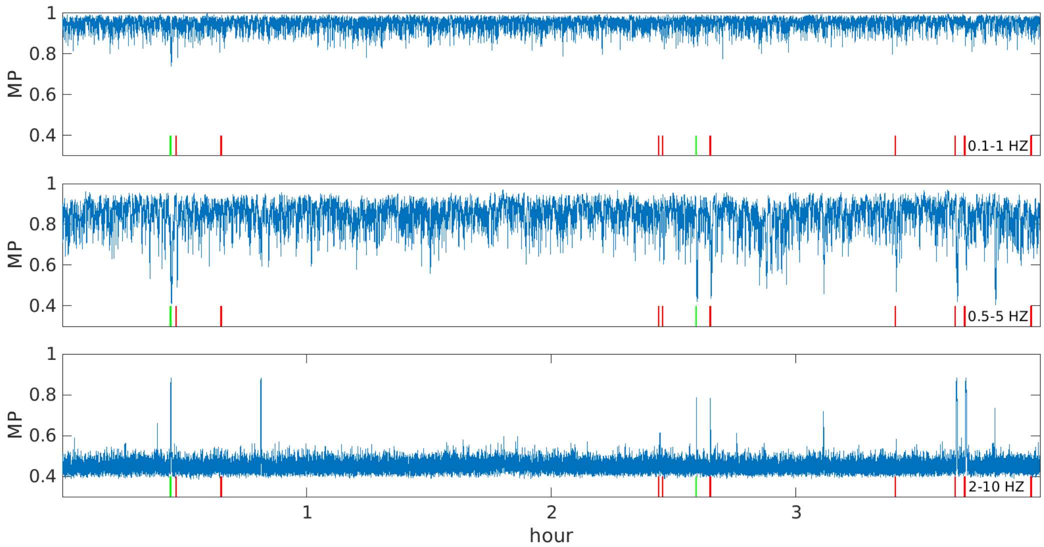

Figure 11. Effect of bandpass filters and microseisms on the matrix profile (MP). MPs shown here are calculated for 4 hours of local seismic data from station PGH, from Parkfield, CA, using frequency bands of (from top to bottom) 0.1–1 Hz, 0.5–5 Hz and 2–10 Hz, and a sample rate of 20 Hz. Red and green lines indicate the NCSN catalog events (at hypocentral distances <70 km from the station) origin time and the P arrival phase, if available, respectively.

Similar to most other seismic data mining applications, selecting an appropriate frequency band-pass filter is a routine and important preprocessing step before applying the MP method. It is trivial to select a frequency range in which the frequency content of the target events is higher than the background noise. For example, 1–15 Hz is commonly used for local earthquake detection (e.g., Schaff & Waldhauser, 2005) and for LFE detection 2–8 HZ has been used in previous studies (e.g., Shelly et al., 2009). In terms of template matching of local events, the choice of bandpass appears to vary between groups based on expert opinion or preference. (Peng & Zhao, 2009), for example, used 2-4 HZ whereas (Ross et al., 2019) used 2-15 HZ. To the best of our knowledge, the effects of the bandpass filter on template matching results have not been extensively studied. As a result of the MP, we are able to discover the dynamic correlation between noise and noise, and between noise and event signals, and therefore can provide guidance in choosing the appropriate bandpass for cross-correlation analysis.

This is illustrated in Figure 11. In the lower frequency bands (top and middle panels), which overlap with the microseismic band, background noise MP (BNMP) values are consistently high (0.80–0.95), indicating that the background noise is repeating and dominating the MP – indicating that both waveforms are affected by microseisms. Interestingly, in these cases, the MP values at the arrival times of seismic phases are lower, particularly in the 0.5–5 Hz case. This indicates that the presence of coherent seismic radiation from earthquakes in that frequency range disrupts the microseisms, and leads to temporarily less similar waveform subsequences. The 2–10 Hz frequency range (bottom panel), is optimal for capturing local earthquakes and excluding microseisms; the BNMP is low (~ 0.5) and significantly higher MP peaks (≥ 0.75) coincide with most of the catalog events, as well as highlighting a few possible uncatalogued events – showing that the MP is sensitive to local events even when a short waveform is examined. As a result of this, many true events can be misidentified or noise sections can be viewed as true events if the data are not bandpassed carefully for template matching purposes.

5.3 Selecting the time series length

The main reason for the success of the MP and other similarity search-based methods in seismic event detection is that events with similar locations have similar waveforms—the source-receiver path having the strongest influence on the recorded signal. The likelihood of the MP method identifying a good match, even to an event in a location where seismicity is infrequent, increases as the length of the time series, and thus the number of recorded events, increases. Aside from the cost of computation, the only disadvantage of including more data in the MP calculation is the possibility that individual sections of background noise may find a better match, thereby increasing the level of the BNMP as a whole (Figure 12). The experiment for station VARB at Parkfield demonstrates, however, that the MPs for background noise and for seismic events have a high enough contrast for event detection, showing that increasing the length of the data time series does does not create a practical issue for event detection.

Figure 12 illustrates the detection capability of the MP with respect to data length. In this figure, we find that the background noise MP (BNMP) increases with the time

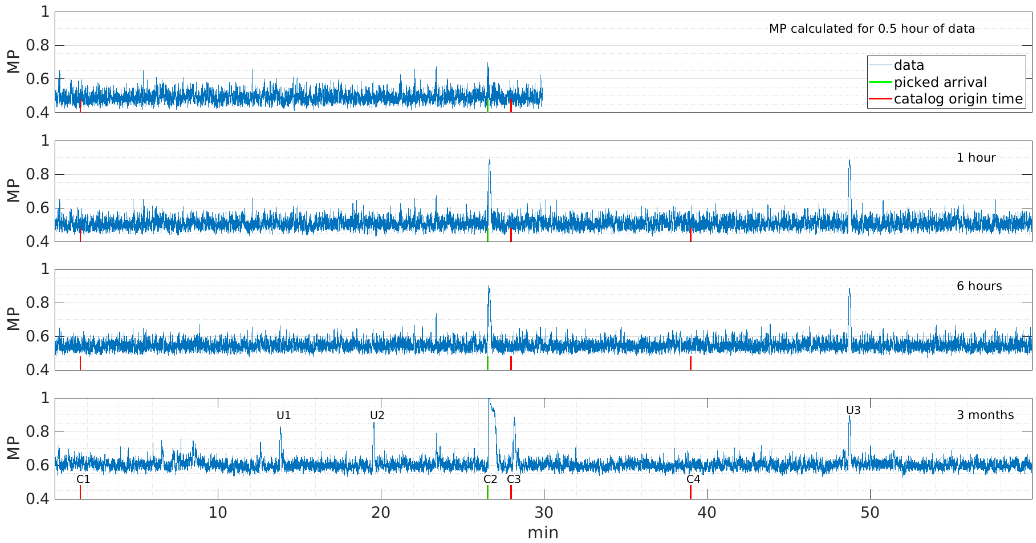

Figure 12. Effect of waveform time series length on the matrix profile (MP). Shown are examples of the MP calculated using (from top to bottom) 0.5 hours, 1 hour, 6 hours and 3 months of 20 Hz waveform data (station PGH, Parkfield, CA, bandpass filtered between 2 and 8 Hz). Up to one hour of the MP is shown, starting at 2004/10/14, 08:00:00 UTC. In this one hour period, four events (C1–C4) were reported in the NCSN catalog within 300 km of the station, but only one of them had a P arrival phase reported at this station. These events were small, and at a range of hypocentral distances (C1: M_L1.17, 70 km, C2: M_L1.30, 6 km C3: M_L0.93, 24 km, C4: M_L1.21, 61 km).

series length. An MP peak can be seen coinciding with event C2 in all cases, although the strength of the peak is only significantly above the BNMP level (i.e., > 0.8) for MP durations 1 hour or longer. Event C3 is only detected in the 3 month-long MP, suggesting that an event with a similar location and waveform to C3 only occurred in that longer period. Events C1 and C4 are very distant; neither visual inspection nor MP values indicate a phase arrival at the station (see Figure S1). Three additional peaks (U1–U3) above 0.8 are visible in the MP for the 3 month case; U3 is also present in the one-hour and six-hour cases, and its MP peak rises in concert with the C2 peak, suggesting that it may represent an uncatalogued event located near the C2 event. Overall we see that even though the BNMP increases when more data is included, there is a higher probability of detecting more events with longer time series.

Thus, finding the optimal time series length for the MP method depends on the event density in time and space, background noise and available computational resources. As mentioned in Section 3 above, using two NVIDIA P100 GPU cards it is possible to calculate MP for several months of local network data (at 20 Hz sampling) in a feasible run time (i.e., hours; Zimmerman et al., 2019).

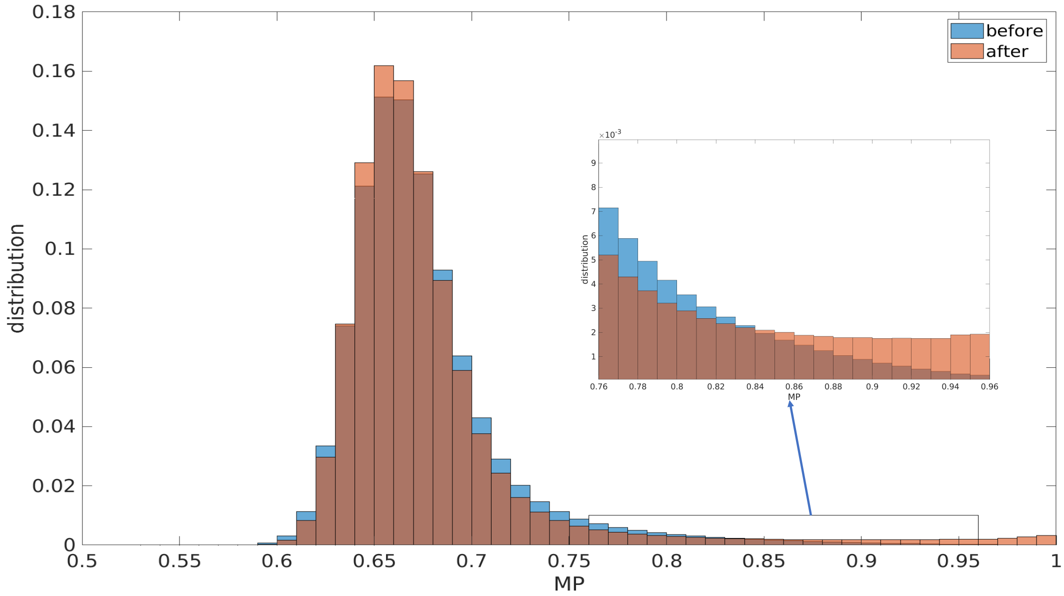

Figure 13. The distribution of continuous matrix profile (MP) values for 20 days before (2004/09/03–2004/09/23; blue) and 20 days after (2004/10/01–2004/10/21; orange) the Parkfield mainshock. Areas of overlap between the two distributions show as brown.

5.4 Selecting a matrix profile threshold for event detection

Choosing an appropriate MP threshold ($r_{\min}$)for seismic event detection depends on multiple factors – for example, the type of data being used (e.g., global, regional, local, lab experiments), the quality of the data, its noise characteristics, and the desired sensitivity. As with the other factors affecting the design of the MP computations, the value and variability of the BNMP is a central consideration. Here we suggest some strategies to choose a threshold.

1. If some *a priori* knowledge exists for the target seismic events such as an earthquake catalog, one can observe the behavior of the MP before, during and after the events of interest and adjust the threshold accordingly. For example for the Parkfield earthquake aftershocks, the MP between events is in the range 0.55 to 0.70 (i.e., the BNMP level), but during the onset of events rises to above 0.9, returning gradually to the background level during the coda. In this case a relaxed threshold could be set around 0.8 and a more conservative threshold could be set around 0.9.
2. A more formal way of defining the detection threshold might be to assume that the majority of the seismic data consists of ambient noise. The assumption might

fail in the cases of particularly energetic early aftershock sequences or seismic swarms, but is valid in most cases (Yanovskaya & Koroleva, 2011). If we plot the histogram of the MP values, the great majority of the population should be around BNMP values, although we typically observe a slight skew towards higher values, perhaps reflecting some degree of correlation in the noise. However we also observe a second mode in the distribution corresponding to even higher MP values for repeated or near-repeated events (Figure 13). Thus, we can define a threshold based on the boundary between these two clusters. In Figure 13 we can see that for the Parkfield experiment this boundary approximately falls around 0.85.

5.5 Clustering and associating matrix profiles from multiple stations

Combining results from multiple stations will improve the reliability of MP data products. The observation of a rise in the MP at several stations with reasonable lag time and/or moveout with distance can be used to eliminate false detections from station glitches or calibration pulses. It is straightforward to stack the MP values for several stations, however the variable lag time to the earthquake onset causes misalignment of the MP peaks. As long as the event durations at the stations are not longer than possible lag times, this does not present a serious problem. However where we expect a long lag time, such as for teleseismic events at GSN stations, or a very short event duration, such as for events close to the lower detection limit in a local network, we need to align the MPs from stations to account for moveout. For instance, MPs from multiple stations were stacked to obtain the results in Section 4.1.

6 Future work

Here, we aim to present the concept of MP to the seismology community and illustrate its capabilities with a few key examples. In this study, we demonstrated that MP can be used to detect events or clusters of similar events that do not appear in template matching catalogs. This concept is novel and new, and by taking advantage of the sensitivity of the method, a number of existing studies can be revisited using the MP approach. In this regard, studies related to foreshocks, aftershocks, swarms, volcanic eruptions, among others, may be included. According to our preliminary investigation, MP can provide accurate detection of seismic events, phase picking, and magnitude estimation of seismic events. As MP is able to detect signals that are below the noise level, it is challenging to compare the MP results with those of other methods or to those of human experts, since MP's accuracy is essentially greater than those of other methods. As a result, investigating the superiority of MP in comparison with other methods requires the careful development of an experimental setup and benchmarks, as well as the possible use of synthetic tests, which is within the scope of our future research.

7 Conclusions

We demonstrate that a recently developed method called the matrix profile (MP) is capable of detecting seismic events with a high degree of sensitivity. In terms of sensitivity, our method is comparable to autocorrelation, but is more computationally efficient. As well as utilizing the MP cross-correlation coefficient values to detect seismic events, the MP's index can also be used to retrieve similar events at the same time, thereby eliminating the need for template matching of new detected events. We can use the MP method to detect not only small regular earthquakes that are hidden in noise, but also novel events such as clusters of repeated events, tremor, and global slow earthquakes. The code needed to compute the MP efficiently is available for download on Github (`https://github.com/zpzim/SCAMP`) and guidelines for pre- and post-processing and interpretation of the MP are provided in this paper. This article is accompanied by a repository of calculated MPs for our experiments in the Parkfield region and a global teleseismic data set (`https://github.com/Naderss/MP4Seismo`). We also plan to maintain and update it with newly calculated MPs in the future.

8 Data Availability statement

The seismic data for the Parkfield and the Global study is downloaded from the Incorporated Research Institutions for Seismology Data Management Center (IRIS-DMC) using the IRISFETCH MATLAB software that can be downloaded from `http://ds.iris.edu/ds/nodes/dmc/software/downloads/irisFetch.m`. Seismic data have been preprocessed with MATLAB signal processing software. Matrix Profiles are generated by SCAMP software (`https://github.com/zpzim/SCAMP`). The computed MPs are stored and publicly available at `http://github.com/Naderss/MP4Seismo`.

Acknowledgments

We gratefully acknowledge funding from NSF project 2103976. NSS and GJF additionally acknowledge support from USGS project G20AP00092, and NASA project NNX15AM66H.

References

Allen, R. (1982). Automatic phase pickers: Their present use and future prospects. *Bull. Seismol. Soc. Am.*, *72*(6B), S225–S242.

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... Farhan, L. (2021). Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, *8*(1), 1–74.

Anooshehpoor, A., & Brune, J. N. (2001). Quasi-static slip-rate shielding by locked and creeping zones as an explanation for small repeating earthquakes at Parkfield. *Bull. Seismol. Soc. Am.*, *91*, 401–403.

ASL/USGS, A. S. L. (1988). *Global seismograph network (gsn–iris/usgs).*

Bakun, W. H., & Lindh, A. G. (1985). The Parkfield, California, Earthquake Prediction Experiment. *Science*, *229*(4714), 619–624. doi: 10.1126/science.229.4714.619

Beaucé, E., Frank, W. B., & Romanenko, A. (2018). Fast Matched Filter (FMF): An efficient seismic matched-filter search for both CPU and GPU architectures. *Seismol. Res. Lett.*, *89*(1), 165–172. doi: 10.1785/0220170181

Bergen, K. J., & Beroza, G. C. (2018). Detecting earthquakes over a seismic network using single-station similarity measures. *Geophys. J. Int.*, *213*(3), 1984–1998. doi: 10.1093/gji/ggy100

Brown, J. R., Beroza, G. C., Ide, S., Ohta, K., Shelly, D. R., Schwartz, S. Y., ... Kao, H. (2009). Deep low-frequency earthquakes in tremor localize to the plate interface in multiple subduction zones. *Geophys. Res. Lett.*, *36*. doi: 10.1029/2009GL040027.

Brown, J. R., Beroza, G. C., & Shelly, D. R. (2008). An autocorrelation method to detect low frequency earthquakes within tremor. *Geophysical Research Letters*, *35*(16).

Chakraborty, S., Tomsett, R., Raghavendra, R., Harborne, D., Alzantot, M., Cerutti, F., ... others (2017). Interpretability of deep learning models: A survey of results. In *2017 ieee smartworld, ubiquitous intelligence & computing,* *advanced & trusted computed, scalable computing & communications, cloud & big data computing, internet of people and smart city innovation (smart-world/scalcom/uic/atc/cbdcom/iop/sci)* (pp. 1–6).

Chamberlain, C. J., Hopp, C. J., Boese, C. M., Warren-Smith, E., Chambers, D., Chu, S. X., ... Townend, J. (2018). EQcorrscan: Repeating and near-repeating earthquake detection and analysis in Python. *Seismol. Res. Lett.*, *89*(1), 173–181. doi: 10.1785/0220170151

Chen, K. H., Bürgmann, R., Nadeau, R. M., Chen, T., & Lapusta, N. (2010). Postseismic variations in seismic moment and recurrence interval of repeating earthquakes. *Earth and Planetary Science Letters*, *299*(1-2), 118–125.

Chen, T., & Lapusta, N. (2009). Scaling of small repeating earthquakes explained by interaction of seismic and aseismic slip in a rate and state fault model. *J. Geophys. Res.*, *114*. doi: 10.1029/2008JB005749

Dokht, R. M. H., Kao, H., Visser, R., & Smith, B. (2019). Seismic event and phase detection using time–frequency representation and convolutional neural networks. *Seismol. Res. Lett.*, *90*(2A), 481–490. doi: 10.1785/0220180308

Earle, P. S., & Shearer, P. M. (1994). Characterization of global seismograms using an automatic-picking algorithm. *Bulletin of the Seismological Society of America*, *84*(2), 366–376.

Ekstöm, G. (2006). Global detection and location of seismic sources by using surface waves. *Bulletin of the Seismological Society of America*, *96*(4A), 1201–1212.

Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophys. J. Int.*, *165*(1), 149–166. doi: 10.1111/j.1365-246X.2006.02865.x

Gionis, A., Indyk, P., & Motwani, R. (1999). Similarity search in high dimensions via hashing. In *Proceedings of the 25th international conference on very large data bases* (p. 518–529). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Hutko, A. R., Bahavar, M., Trabant, C., Weekly, R. T., Fossen, M. V., & Ahern, T. (2017). Data products at the iris-dmc: Growth and usage. *Seismological Research Letters*, *88*(3), 892–903.

Indyk, P., & Motwani, R. (1998). Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proceedings of the thirtieth annual acm sym-*

posium on theory of computing (p. 604–613). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/276698.276876 doi: 10.1145/276698.276876

Khoshmanesh, M., Shirzaei, M., & Nadeau, R. M. (2015). Time-dependent model of aseismic slip on the central San Andreas fault from InSAR time series and repeating earthquakes. *J. Geophys. Res.*, *120*, 6658–6679.

Lemoine, A., Briole, P., Bertil, D., Roullé, A., Foumelis, M., Thinon, I., ... Hoste Colomer, R. (2020, 06). The 2018–2019 seismo-volcanic crisis east of Mayotte, Comoros islands: seismicity and ground deformation markers of an exceptional submarine eruption. *Geophys. J. Int.*, *223*(1), 22-44. doi: 10.1093/gji/ggaa273

Meng, X., Peng, Z., & Hardebeck, J. L. (2013). Seismicity around parkfield correlates with static shear stress changes following the 2003 mw6. 5 san simeon earthquake. *Journal of Geophysical Research: Solid Earth*, *118*(7), 3576–3591.

Mercer, R., Alaee, S., Abdoli, A., Senobari, N. S., Singh, S., Murillo, A., & Keogh, E. (2022). Introducing the contrast profile: a novel time series primitive that allows real world classification. *Data Mining and Knowledge Discovery*, *36*(2), 877–915.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Comms.*, *11*. doi: 10.1038/s41467-020-17591-w

Nadeau, R. M., & Dolenc, D. (2005). Nonvolcanic tremors deep beneath the san andreas fault. *Science*, *307*(5708), 389–389.

Nadeau, R. M., Foxall, W., & McEvilly, T. (1995). Clustering and periodic recurrence of microearthquakes on the san andreas fault at parkfield, california. *Science*, *267*(5197), 503–507.

Nadeau, R. M., & Guilhem, A. (2009). Nonvolcanic tremor evolution and the san simeon and parkfield, california, earthquakes. *science*, *325*(5937), 191–193.

Nadeau, R. M., & Johnson, L. R. (1998). Seismological studies at parkfield vi: Moment release rates and estimates of source parameters for small repeating earthquakes. *Bulletin of the Seismological Society of America*, *88*(3), 790–814.

Neves, M., Peng, Z., & Lin, G. (2023). A high-resolution earthquake catalog for the 2004 Mw 6 Parkfield earthquake sequence using a matched filter technique. *Seismological Society of America*, *94*(1), 507–521.

Pang, G., & Koper, K. D. (2022). Excitation of earth's inner core rotational oscillation during 2001–2003 captured by earthquake doublets. *Earth and Planetary Science Letters*, *584*, 117504.

Peng, Z., & Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield earthquake. *Nature Geoscience*, *2*(12), 877–881.

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Sci. Adv.*, *4*(2). doi: 10.1126/sciadv.1700578

Ross, Z. E., Meier, M.-A., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. *Bull. Seismol. Soc. Am.*, *108*(5A), 2894–2901. doi: 10.1785/0120180080

Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in Southern California. *Science*, *364*(6442), 767–771. doi: 10.1126/science.aaw6888

Royer, A., & Bostock, M. (2014). A comparative study of low frequency earthquake templates in northern cascadia. *Earth and Planetary Science Letters*, *402*, 247–256.

Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, *2*(6), 1–20.

Schaff, D. P., & Waldhauser, F. (2005). Waveform cross-correlation based differential travel-time measurements at the Northern California Seismic Network. *Bull. Seismol. Soc. Am.*, *95*, 2446–2461. doi: 10.1785/012004022

Shakibay Senobari, N., Funning, G. J., Keogh, E., Zhu, Y., Yeh, C. M., Zimmerman, Z., & Mueen, A. (2019). Super-Efficient Cross-Correlation (SEC-C): A fast matched filtering code suitable for desktop computers. *Seismol. Res. Lett.*, *90*, 322–334. doi: 10.1785/0220180122

Shearer, P. M. (1994). Global seismic event detection using a matched filter on long-period seismograms. *Journal of Geophysical Research: Solid Earth*, *99*(B7), 13713–13725.

Shelly, D. R. (2009). Possible deep fault slip preceding the 2004 parkfield earthquake, inferred from detailed observations of tectonic tremor. *Geophysical Re-*

search Letters, *36*(17).

Shelly, D. R. (2017). A 15 year catalog of more than 1 million low-frequency earthquakes: Tracking tremor and slip along the deep san andreas fault. *Journal of Geophysical Research: Solid Earth*, *122*(5), 3739–3753.

Shelly, D. R., Beroza, G. C., & Ide, S. (2007). Non-volcanic tremor and low-frequency earthquake swarms. *Nature*, *446*, 305–307. doi: 10.1038/nature05666

Shelly, D. R., Beroza, G. C., Ide, S., & Nakamula, S. (2006). Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature*, *442*, 188–191. doi: 10.1038/nature04931

Shelly, D. R., Ellsworth, W. L., Ryberg, T., Haberland, C., Fuis, G. S., Murphy, J., ... Bürgmann, R. (2009). Precise location of San Andreas Fault tremors near Cholame, California using seismometer clusters: Slip on the deep extension of the fault? *Geophys. Res. Lett.*, *36*(1). doi: 10.1029/2008GL036367

Shelly, D. R., & Hardebeck, J. L. (2010a). Precise tremor source locations and amplitude variations along the lower-crustal central san andreas fault. *Geophysical Research Letters*, *37*(14).

Shelly, D. R., & Hardebeck, J. L. (2010b). Precise tremor source locations and amplitude variations along the lower-crustal central San Andreas Fault. *Geophys. Res. Lett.*, *37*. doi: 10.1029/2010GL043672

Tkalvcic, H., Young, M., Bodin, T., Ngo, S., & Sambridge, M. (2013). The shuffling rotation of the earth's inner core revealed by earthquake doublets. *Nature Geoscience*, *6*(6), 497–502.

US Geological Survey, E. H. P. (2017). Advanced national seismic system (anss) comprehensive catalog of earthquake events and products: Various.

Vassallo, M., Satriano, C., & Lomax, A. (2012). Automatic picker developments and optimization: A strategy for improving the performances of automatic phase pickers. *Seismol. Res. Lett.*, *83*(3), 541.

Vidale, J. E., Ellsworth, W. L., Cole, A., & Marone, C. (1994). Variations in rupture process with recurrence interval in a repeated small earthquake. *Nature*, *368*, 624–629.

Waldhauser, F., & Schaff, D. (2007). Regional and teleseismic double-difference earthquake relocation using waveform cross-correlation and global bulletin data. *Journal of Geophysical Research: Solid Earth*, *112*(B12).

Yang, Y., & Song, X. (2020). Temporal changes of the inner core from globally distributed repeating earthquakes. *Journal of Geophysical Research: Solid Earth*, *125*(3), e2019JB018652.

Yanovskaya, T., & Koroleva, T. Y. (2011). Effect of earthquakes on ambient noise cross-correlation function. *Izvestiya, Physics of the Solid Earth*, *47*(9), 747–756.

Yeh, C.-C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., ... Keogh, E. (2016). Matrix Profile I: All pairs similarity joins for time series: A unifying view that includes motifs, discords and shapelets. In *2016 IEEE 16th International Conference on Data Mining (ICDM)* (pp. 1317–1322). doi: 10.1109/ICDM.2016.0179

Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. *Sci. Adv.*, *1*(11). doi: 10.1126/sciadv.1501057

Zhang, J., Song, X., Li, Y., Richards, P. G., Sun, X., & Waldhauser, F. (2005). Inner core differential motion confirmed by earthquake waveform doublets. *Science*, *309*(5739), 1357–1360.

Zhu, L., Peng, Z., McClellan, J., Li, C., Yao, D., Li, Z., & Fang, L. (2019). Deep learning for seismic phase detection and picking in the aftershock zone of 2008 m_w7.9 Wenchuan earthquake. *Phys. Earth Planet. Inter.*, *2093*. doi: 10.1016/j.pepi.2019.05.004

Zhu, Y., Zimmerman, Z., Shakibay Senobari, N., Yeh, C.-C. M., Funning, G., Mueen, A., ... Keogh, E. (2016). Matrix Profile II: Exploiting a novel algorithm and GPUs to break the one hundred million barrier for time series motifs and joins. In *2016 IEEE 16th International Conference on Data Mining (ICDM)* (pp. 739–748). doi: 10.1109/ICDM.2016.0085

Zimmerman, Z., Kamgar, K., Shakibay Senobari, N., Crites, B., Funning, G., Brisk, P., & Keogh, E. (2019). Matrix Profile XIV: Scaling time series motif discovery with GPUs to break a quintillion pairwise comparisons a day and beyond. In *Proceedings of the acm symposium on cloud computing* (p. 74–86). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3357223.3362721