Identification of Repeating Earthquakes: Misconception of Waveform Similarity and a Physics-based Solution

Dawei Gao[1,2] and Honn Kao[1,2]*

[1]School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada.

[2]Pacific Geoscience Centre, Geological Survey of Canada, Sidney, British Columbia, Canada.

Corresponding author: Honn Kao (honn.kao@canada.ca)

Key Points:

- There is no simple relationship between cross-correlation coefficient (CC) and inter-event separation.
- CC is affected by many factors and thus lacks the resolution to determine two events as true repeating or just neighboring earthquakes.
- To reliably identify repeating earthquakes, we should rely on the precise estismation of both rupture radius and inter-event separation.

Abstract

Identification of repeating earthquakes (repeaters) usually depends on waveform similarity expressed as the corresponding cross-correlation coefficient (CC) above a prescribed threshold, typically ranging from 0.70 to 0.98. However, the robustness and effectiveness of such a strategy have never been thoroughly examined. In this study, we examine whether CC is a valid proxy for repeater identification through both synthetic and real earthquake experiments. We reveal that CC is controlled by not only the inter-event distance but also many other factors, including station azimuth, epicentral distance, velocity structure, etc. Consequently, CC lacks the resolution in identifying true repeaters. We propose a physics-based approach that considers both inter-event separation and rupture radius. For an event pair to be true repeaters, their inter-event separation must be smaller than the rupture radius of the larger event. Our results imply that a systematic recheck of previously identified repeaters and associated interpretations/hypotheses may be important and necessary.

Plain Language Summary

Repeating earthquakes (repeaters) are events that occur repeatedly on the same fault patch with nearly identical waveforms. They provide important insights into a variety of geophysical subjects such as fault behavior, subsurface structure change, inner core rotation, and nucleation process of earthquake and landslide. The identification of repeaters is usually solely based on waveform similarity, but the criteria can vary significantly from one case to another. With both synthetic and real data, we find that waveform similarity is controlled by many factors, in addition to inter-event distance. Therefore, a higher degree of waveform similarity does not necessarily imply a smaller hypocenter separation, and vice versa. Our results undoubtedly suggest that waveform similarity alone is insufficient to reliably identify true repeaters. We propose a physics-based approach that considers both inter-event separation and earthquake source dimension. For an event pair to be true repeaters, their inter-event separation must be smaller than the rupture radius of the larger event. Our results imply that previously identified repeaters and associated interpretations/hypotheses may be unreliable and hence need a systematic reexamination.

1. Introduction

Repeating earthquakes (repeaters) are events that recurrently rupture the same fault patch with the same focal mechanisms, often characterized by nearly identical waveforms (Uchida and Bürgmann, 2019). These events are of great importance in many aspects of geophysics, such as monitoring subtle temporal changes of crustal properties (e.g., Poupinet et al., 1984; Schaff and Beroza, 2004; Sawazaki et al., 2015; Pacheco et al., 2017) and oceanic temperature (Wu et al., 2020), estimating fault creep (e.g., Nadeau and Johnson, 1998; Uchida et al., 2003, 2006; Matsubara et al., 2005; Yu, 2013; Materna et al., 2018), investigating inner core rotation (e.g., Li and Richards, 2003; Zhang et al., 2005, 2008; Tkalčić et al., 2013), evaluating the precision of earthquake locations (e.g., Li and Richards, 2003; Meier et al., 2004; Schaff and Richards, 2011; Jiang et al., 2014), and providing insights into the nucleation process of earthquakes (Kato et al., 2012; Kato and Nakagawa, 2014; Meng et al., 2015; Huang and Meng, 2018) and landslides (Yamada et al., 2016).

Over the past few decades, repeaters are reported world-wide, in both tectonic and nontectonic settings (Uchida and Bürgmann, 2019). The most commonly used scheme for identifying repeaters is to examine waveform similarity by setting a threshold in the value of cross-correlation coefficient (CC) between a waveform pair. The employed CC threshold is somewhat arbitrary depending on the available data, ranging in 0.70–0.90 for regions with sparse network coverage (usually with one channel/station), and 0.90–0.98 for areas with denser instrumentation (usually using a minimum of two channels/stations, Table S1). With the increasing computing power, detecting repeaters through waveform similarity has become a routine process in seismology (e.g., Hotovec-Ellis and Jeffries, 2016; Tepp, 2018; Chamberlain et al., 2018). However, to our best knowledge, the robustness and effectiveness of this classical strategy have never been thoroughly examined. In early times, the lack of examination was largely due to limited data availability (especially near-field observations) and/or poor waveform quality. As the number of seismograph stations increases rapidly in recent years, such data constraints no longer exist. Yet, many recent studies simply follow the conventional appraoch without questioning the original assumption (Table S1).

The focus of this study, therefore, is to investigate whether waveform similarity is a valid proxy for repeater identification. We first examine how the CC varies with inter-event separation and uncover the overlooked factors through a large number of synthetic experiments. We then illustrate that waveform similarity indeed lacks the resolution to determine whether two events are true repeaters or not using a dense local borehole array data in Parkfield, California. To more reliably identify repeating earthquakes, we propose a physics-based approach that considers both inter-event separation and the rupture areas. We validate our approach using events occurred in the Fox Creek area, Alberta, Canada, where earthquake source parameters are well constrained by local stations.

2. Synthetic Experiments

Figure 1a illustrates the configuration of our synthetic experiments. We place one event (the template event) at the centre of an array. Then we incrementally shift the other event (the matched event) with the same focal mechanism in either north-south (Figure 1a) or vertical direction (Figure 1b). The technical details of our experiment setup and CC calculation are presented in the Supporting Information (Texts S1 and S2).

2.1 Constraining Inter-event Separation Using Single-channel Data

Single-channel CC has been used in numerous previous studies to infer the existence of repeaters (Table S1), thus we first examine how the CC varies with source separation using single-channel (i.e., E, N, or Z) data. In Figure 2a and 2b, we present the results of a representative case, namely, a strike-slip earthquake (template event) at the depth of 3 km with a station 5 km away from the epicenter.

For horizontal inter-event separation, our results indicate that single-channel waveforms can have very different sensitivities (Figure 2a). In general, the CC value decreases with increasing hypocentral separation. It quickly drops from 1 when the two sources are perfectly co-located to <0.5 when the pair is ~1 km apart. Beyond that, the CC curves appear to fluctuate between 0.2 and 0.4 without a clear monotonic trend. This implies that using the CC value to constrain the difference between two nearby hypocenters may not be ideal once the separation is on the order of kilometers.

Another important point in Figure 2a is that the CC value may be strongly affected by the combined effect of focal mechanism and relative position between the source and station. This effect is best illustrated by Station 3 as the inter-event distance increases. For all 3 channels, the CC value decreases when the matched event shifts northward from 0 to -1.3 km. Once passing the -1.3 km mark, the CC value has a sudden drop on both E and Z channels but continues to increase on the N channel. This unexpected result happens when Station 3 is located very close to one of the assumed nodal planes (Figure 1a). As the matched event shifts northward, Station 3's position moves across the nodal plane and therefore causes polarity reversal on the Z and E channels. When the inter-event separation reaches -2.6 km, Station 3 is nearly of equal distance to both the template and matched events (Figure 1a), leading to identical waveforms on the N channel but reversed shapes on the other two channels (Figure S1). Consequently, the final (maximum) CC values would be 1 for the N channel (taken when the two P phases coincide) and ~0.5–0.6 for the Z and E channels (taken when the two P phases are offset by half a cycle), even though the two events are 2.6 km apart. We have tested other types of focal mechanisms (pure normal or thrust-faulting) and the profound effect remains (Figure S2).

Unlike the cases of horizontal separation, the CC curves obtained with different channels and stations overall show similar trends when the two sources are vertically apart (Figure 2b), hinting that using the CC value to constrain the vertical inter-event separation is probably independent of data channel and station azimuth. Especially for the vertical channel, stations with different azimuths can have identical sensitivities to the inter-event separation when the focal mechanism is pure strike-slip (Figure 2b, right panel). Notice that the CC curves derived from the E and N channels of Station 1 are identical to those from the N and E channels of Station 3, respectively (Figure 2b), due to the symmetrical station location on the focal sphere (Figure 1a). Results of these tests once again suggest that a larger CC does not necessarily represent a larger separation once the vertical separation exceeds a certain threshold (~0.5 km). We also find that results from different focal mechanisms are comparable (Figure S2). Last but not the least, the CC value generally drops much faster with increasing vertical source separation (Figure 2a vs. 2b) as a result of more minor discrepancies between waveforms. In other words, the CC seems to be much more sensitive to capture the vertical source shift than the horizontal.

The simple tests above demonstrate that, in addition to inter-event distance, CC can be severely affected by the specific channel used, combined effect of focal mechanism and relative position between the source and station, and source separation direction (horizontal vs. vertical).

2.2 Constraining Inter-event Separation Using Single-station (3-channel) Data

If data from all three channels are included, we find that the CC sensitivity to source separation increases dramatically for the cases of horizontal separation (e.g., Figures 2a vs. S3a) but insignificantly for those of vertical separation (e.g., Figures 2b vs. S4a). For a given horizontal separation, Stations 1 and 3 tend to have the lowest and highest CC values, respectively (Figures 2c and S3), strongly suggesting that station azimuth is an important factor that cannot be overlooked. In contrast, the influences of focal depth, epicentral distance, and source focal mechanisms seem to be limited (Figure S3). Our results indicate that a station approximately in line with the template and matched events can be more effective in delineating the inter-event separation (e.g., Station 1 in our case, Figure 1a).

The computed CC overall is very sensitive to vertical inter-event separation with the only exception when the source is deep and the station is close (e.g., D=10 km and R=5 km, Figures 2d and S4). For a close station (R=5 km) and a shallow source (D=3 km), even a very small (0.2 km) vertical separation can lead to a dramatic drop of CC to <0.8 (Figure S4a and S4c), but the sensitivity gets worse when the source is deeper (Figure S4b and S4d). This is mainly a velocity structure effect caused by smaller seismic velocity variation at deep depths. In other words, the CC sensitivity would become higher when the corresponding velocity structure (and therefore the observed waveforms) are more complicated. An important observation to point out is that the CC is very sensitive to vertical inter-event separations when the epicentral distance is large (e.g., R = 50 or 150 km), regardless of the focal depth (Figure S4e-l). This is opposite to what is expected for earthquake depth determination as seismic phase arrival times at distant stations usually have less depth constraint. It turns out that waveforms at distant stations can have better developed depth phases (i.e., seismic phases reflected from either the free surface or Moho). Consequently, a subtle change of source depth may lead to a significant waveform difference and therefore an apparent CC drop.

Therefore, our experiments in this section further demonstrate that CC can be affected by the number of channels used, station azimuth, velocity structure, and epicentral distance.

2.3 Constraining Inter-event Separation Using Multi-station Data

For areas with excellent network coverage, it is common to use a minimum of two stations (usually only the vertical channel) for identifying repeaters (Table S1). The majority of prior work (Table S1) calculate CC separately for each station. This approach essentially uses more stations with different azimuths and/or epicentral distances but may not necessarily improve the sensitivity if all available stations happen to be the ones with lower sensitivities (Figure 2c and 2d). An alternative way is to calculate the CC simultaneously across the network (e.g., Yao et al., 2017) which includes the constraint of traveltime moveout. In such a case, the computed CC can be extremely sensitive to hypocentre difference (Gao and Kao, 2020). We refrain from investigating the multi-station scenario as the CC sensitivity is known to be strongly affected by network geometry (Chamberlain and Townend, 2018; Gao and Kao, 2020), and thus no general/common rules can be inferred objectively.

In summary, our synthetic experiments reveal that CC is a very complex function of many aforementioned factors. A higher CC value does not necessarily represent a smaller inter-event separation, and vice versa. Therefore, in contrast to the conventional wisdom, our synthetic results indicate that CC is not a robust indicator of two events being true repeaters or not.

3. Verification With Real Earthquake Examples

The High-Resolution Seismic Network (HRSN, Figure 3a) is a dense local array of borehole seismometers deployed in the Parkfield area, California, and operated by the Berkeley Seismological Laboratory. The HRSN waveform data generally have exceptionally high signal-to-noise ratio (SRN) and hence are ideal for the purpose of this study to verify whether waveform similarity is a good proxy of repeater identification. Here we take three events (No. 1-3, Figure 3a) from two well studied repeating earthquake clusters in Parkfield (Abercrombie, 2014). Among them, events No. 1 and 3 belong to the same cluster with similar source areas while event No. 2 occurred on a different fault patch.

3.1 CC between Non-repeaters

We first calculate the CC between non-repeaters, i.e., No. 1 and 2. We only use data from stations nearly free from noise contamination, as hinted by the flat waveforms before the P wave arrival (one example is shown in Figure 3b). The most striking result of our analysis is that the CC derived from unfiltered 3-channel waveforms indeed differ significantly among different stations, ranging from 0.76 to just above 0.95 (Figure 3c). Such a wide CC range is consistent with the inference from our synthetic tests that the CC can be severely affected by station azimuth and/or source-receiver position/path even under the noise-free circumstances. Additionally, the CC may be further affected by local structures of velocity discontinuities as this region is featured by complex fault zones (Figure 3a). Our study reveals that the waveforms of non-repeaters can show apparent difference (top panel in Figure 3b), slight difference (middle panel in Figure 3b) or little difference (bottom panel in Figure 3b) at different stations. For any given station, the CC values of different channels can be either similar or different (Figure S5). Together, waveform similarity indeed lacks the resolution to decide the two events to be repeaters or not.

Because nearly all prior works practically identify repeaters through filtered waveforms for the purpose of mitigating the noise impact, we then examine the effects of commonly used band-pass filters (Table S2). Our results indicate that the CC obtained from different stations all show a clear increasing trend when the passband becomes narrower (Figures 3c and S5). Especially for the very narrow but very popular 1-4 Hz band-pass filter used by many previous studies (Table S2), 9 out of 10 stations have CC >0.98 (Figure 3c), which is the highest CC threshold used in the literature in selecting repeaters (Table S1). This simple experiment highlights the overlooked fact that filtering could remove the important frequency content in the signal that distinguishes the physical separation of the two events, in addition to reducing the unwanted noise. For example, even a very wide band-pass filter (1-20 Hz) would remove the very high frequency signal with poor similarity and thus lead to very similar waveforms as shown in Figure 3d. What makes it worse is that filtering would change both the shape and width of the P wave and therefore make the subtle difference in the S-P differential traveltime (0.012s in Figure 3d) unresolvable, effectively throwing away the most critical information on the relative distance between the two sources. The results here strongly imply that filtering would lead to misidentification of repeaters if the selection criterion is solely based on waveform similarity. We also tested the effect of template window length (T_{win}) associated with different filters (Table S2) in calculating CC (Text S2) and the results are comparable (Figure S6). Two examples of how filtering increases the waveform similarity at close and distant stations, respectively, are presented in Figures S7 and S8 for reference.

3.2 CC between True Repeaters

Interestingly, we also find that, for true repeaters like events No. 1 and 3, the CC value obtained from different stations still differ significantly from each other (Figure S9). The unfiltered waveforms can be nearly identical at one station (Figure S10a) but also can be of minor difference at another station even with nearly no noise (zoom-in box in Figure S10b). Without noise contamination, the waveform difference between true repeaters may arise from the variability of the rupture process (such as the slight difference in earthquake initiation point) (Uchida, 2019) and/or seismic velocity change (e.g., Poupinet et al., 1984; Sawazaki et al., 2015; Pacheco et al., 2017). With the band-pass filters applied, the waveform discrepancy overall becomes much smaller as indicated by the increasing CC values (Figure S9). This is similar to, but less dramatic as, the case of non-repeaters.

Taken together, non-repeaters indeed can have very similar waveforms (bottom panel in Figure 3b) while the waveforms of true repeaters may display minor difference (Figure S10b). In contrast to the traditional view, our observations undoubtedly suggest that waveform similarity is not a good proxy for repeater identification, especially with band-pass filters applied.

4. A Physics-based Solution

Although waveform similarity can be useful in recognizing potential repeaters (e.g., Sáez et al., 2019), the most fundamental concern of two events being repeaters or not is whether their ruptures significantly overlap with each other. In case of two events rupturing the same fault patch but from different nucleation points, they should be deemed repeaters even though the corresponding waveform similarity may not be perfect (Uchida, 2019). Therefore, a physics-based approach considering both inter-event separation and the source rupture areas should be the most reasonable solution. Specifically, we define two events to be repeaters if their inter-event distance is no larger than the source dimension of the bigger event. In other words, the hypocenter of the smaller event must fall within the rupture area of the bigger event.

There are different ways in seismology to characterize the rupture area of an earthquake source (Stein and Wysession, 2009). One of the most common, and perhaps the easiest, method is to estimate the equivalent rupture radius (ERR), which is defined as the radius of a circle whose area is the same as the source rupture area. The ERR can be estimated from the event's stress drop ($\Delta\sigma$) via the well-established circular dislocation model (Brune, 1970):

$$ERR = \sqrt[3]{\frac{7M_o}{16\,\Delta\sigma}} \quad (1)$$

where M_o is the event's scalar seismic moment. The $\Delta\sigma$ value can either be reasonably assumed (Table S3) or directly derived (e.g., Abercrombie, 2014; Ellsworth and Bulut, 2018).

For the distance of inter-event separation, however, it is always a challenge to get a precise measurement unless a very dense local array is available (Cheng et al., 2007). In case of limited data, we propose a variant of the double difference method (HypoDD; Waldhauser and Ellsworth, 2000) that minimizes the residual between observed and predicted relative S-P differential traveltime through three-dimensional (3D) grid search to precisely estimate the inter-event distance. We explain the detail of our method, named the differential traveltime double-difference (DTDD) method, in the Supporting Information (Text S3).

Figure 4 presents an example of using the DTDD method to determine the precise relative position of three events with similar waveforms recorded at 4 nearby stations in Fox Creek, Alberta, Canada. Among them, events No. 4 and 6 are found to be ~200 m apart, consistent with the subtle difference in S-P time (left yellow zoom-in window in Figure 4b). In comparison, events No. 5 and 6 are effectively co-located (Figures 4c and S11) as indicated by the identical S-P time at all four stations (right yellow zoom-in window in Figure 4b). The corresponding ERRs of these three events are 60 (No. 4), 31 (No. 5), and 70 (No. 6) m, respectively. Therefore, we conclude that events No. 5 and No. 6 are true repeating events (i.e., ERR > inter-event distance), whereas events No. 4 and No. 6 are at most neighboring events (ERR < inter-event distance).

We finally note that both event pairs have high CC values (0.83 between events No. 4 and 6, 0.88 between events No. 5 and 6; Figure 4b). Consequently, the determination of these event pairs being true repeaters or not can be arbitrary depending on the choice of the CC threshold (e.g., the choice of 0.80, 0.85, or 0.90 will lead to completely different outcomes). Our results clearly indicate that the CC can drop below 0.9 even for true repeating events when the waveforms of the smaller event are contaminated with noise (event No. 5; Figure 4b). It further underscores the challenge in identifying repeaters based solely on waveform similarity with the presence of noise.

5. Discussion

There are two key paramaters, i.e., ERR and inter-event distance, in our proposed framework of identifying repeaters. To uniquely estimate the ERR of a small event, it is necessary to specify the value of $\Delta\sigma$ in the popular Brune model (Equation 1). While most previous studies consider $\Delta\sigma$ to be no more than 10 MPa (Table S3), detailed analysis of dense local borehole array data in the Parkfield area suggests that $\Delta\sigma$ of small (M_L ~2) events can be as high as tens of MPa (Abercrombie, 2014). Since a smaller $\Delta\sigma$ value will yield a larger ERR, underestimation of $\Delta\sigma$ is likely to misclassify neighboring events as repeaters, and vice versa. Therefore, the uncertainty due to a poorly constrained (or wrongly assumed) $\Delta\sigma$ value should be treated with caution.

The DTDD method that we develop to estimate the inter-event distance with limited waveform data relies on precise measurement of the S-P differential traveltime (Text S3). For a typical crustal model (i.e., Vp = 6.5 km/s and Vp/Vs = 1.73), a 0.01s difference in the S-P time corresponds to a hypocentral difference of ~90 m (Hayward and Bostock, 2017). Consequently we need to pay attention to two potential issues. First, the application of digital filtering, such as the band-pass filters used in previous studies, may change the width of P wave, leading to a slight shift between the P and S phases (one example is shown in Figure 3d). The small bias (0.012 s in the case shown in Figure 3d) is equivalent to a mislocation of ~100 m that is sufficient to cause misinterpretation for events with small source dimension. Therefore, we prefer to use unfiltered broadband waveforms in the measurement of S-P differential traveltime to avoid any possible bias from waveform filtering.

The second issue is the resolution limit defined by the sampling rate of waveform data. It can be particularly problematic if the original sampling rate is less than 100 Hz (i.e., ≥0.01s between samples) so the hypocentral uncertainty becomes comparable to the source dimension of small events. A straightforward solution is to pre-process waveforms with interpolation to increase its apparent sampling rate before measuring the S-P times (Li et al., 2007, 2011). Similarly, the grid size used in the DTDD source-searching process should be much smaller than the source dimension of the targeted events to achieve optimal spatial resolution.

We note that, in the extreme case of limited data from only one or two stations, the DTDD solution can be highly non-unique. Consequently, a priori constraints must be introduced to quantitatively estimate the inter-event distance. One commonly adopted remedy is to require the two events to occur on a given fault plane (e.g., Li et al., 2007). Another commonly adopted constraint comes from the ambient tectonic loading rate, i.e., the recurrence interval between two repeaters should be proportional to the size of the second event (e.g., Li et al., 2007, 2011; Bohnhoff et al., 2017). If the two events occur very closely to each other in time, they are more likely to be neighboring events because the fault patch ruptured during the first event has not healed yet.

6. Conclusion

In this study we reveal that CC can be severely affected by many factors, including the choice of one specific channel or all three channels, combined effect of focal mechanism and relative position between the source and station, station azimuth, epicentral distance, velocity structure, orientation of the source separation (horizontal vs. vertical), network geometry, and the filter's frequency bandwidth. In reality, noise, heterogeneity in the crust, and variation in the Moho depth may further contribute to the complication of CC sensitivity. Therefore it is almost impossible to reliably identify repeaters solely based on a given CC value, implying that a systematic recheck of previously identified repeaters and associated interpretations/hypotheses may be important and necessary.

To more reliably identify true repeaters, we propose a physics-based approach that considers both ERR and inter-event separation. For an event pair to be true repeaters, their inter-event separation must be no larger than the ERR of the larger event. For the precise estimation of inter-event distance in case of limited data, we develop the DTDD method which relies on the relative S-P differential traveltime. Finally we illustrate the effectiveness of the DTDD method and validate the physics-based approach in selecting repeaters using earthquakes occurred in the Fox Creek area, Alberta, Canada. The findings of this work has far-reaching impact on not only repeating earthquake reseach but also other waveform-similarity-based studies.

Acknowledgements

We are grateful to Lupei Zhu for providing the FK code that is used in generating synthetic seismograms. We specially thank Bei Wang for helping calculate the stress drops of the Fox Creek events, and thank Didem Cambaz and Edwin Nissen for the help offered in this study. Insightful discussions with Toshihiro Igarashi, Naoki Uchida, David P. Schaff, Makoto Naoi, Jianlong Yuan, Stan Dasso, Kelin Wang, Ryan Visser, Rachel E. Abercrombie, Fengzhou Tan, Ramin Mohammad Hosseini Dokht, Wen-che Yu, Lingsen Meng, Hui Huang, Jean Schmittbuhl, Marco Bohnhoff, Christopher Wollin, Tomoaki Nishikawa, Emily Warren-Smith, Luis A. Dominguez, Tianhaozhe Sun, and Haipeng Luo are much appreciated. The HRSN and Fox Creek event waveform data used in this study were downloaded from the Northern California Earthquake Data Center (NCEDC) (http://ncedc.org/hrsn/) and Incorporated Research Institutions for Seismology (IRIS) (http://ds.iris.edu/ds/nodes/dmc/), respectively. Seismic data are processed with Obspy (*Beyreuther et al*., 2010; https://github.com/obspy/obspy/). Figures are made with Matplotlib (*Hunter*, 2007; https://matplotlib.org) and Inkscape (https://inkscape.org). This study is partially supported by a University of Victoria Fellowship (DG), the Induced Seismicity Research Project of NRCan (HK), Geoscience BC (HK), and a NSERC Discovery Grant (HK). This paper is NRCan contribution 2021xxxx.

References

Abercrombie, R. E. (2014). Stress drops of repeating earthquakes on the San Andreas fault at Parkfield. *Geophysical Research Letters*, *41*(24), 8784-8791.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 530-533.

Bohnhoff, M., Wollin, C., Domigall, D., Küperkoch, L., Martínez-Garzón, P., Kwiatek, G., ... & Malin, P. E. (2017). Repeating Marmara Sea earthquakes: indication for fault creep. *Geophysical Journal International*, *210*(1), 332-339.

Bouchon, M., Karabulut, H., Aktar, M., Özalaybey, S., Schmittbuhl, J., & Bouin, M. P. (2011). Extended nucleation of the 1999 Mw 7.6 Izmit earthquake. *science*, *331*(6019), 877-880.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of geophysical research*, *75*(26), 4997-5009.

Chamberlain, C. J., & Townend, J. (2018). Detecting Real Earthquakes Using Artificial Earthquakes: On the Use of Synthetic Waveforms in Matched-Filter Earthquake Detection. *Geophysical Research Letters*, *45*(21), 11-641.

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. *Seismological Research Letters*, *89*(1), 173-181.

Cheng, X., Niu, F., Silver, P. G., Horiuchi, S., Takai, K., Iio, Y., & Ito, H. (2007). Similar microearthquakes observed in western Nagano, Japan, and implications for rupture mechanics. *Journal of Geophysical Research: Solid Earth*, *112*(B4).

Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. *Nature Geoscience*, *11*(7), 531-535.

Gao, D., & Kao, H. (2020). Optimization of the Match-Filtering Method for Robust Repeating Earthquake Detection: The Multisegment Cross-Correlation Approach. *Journal of Geophysical Research: Solid Earth*, *125*(7), e2020JB019714.

Hayward, T. W., & Bostock, M. G. (2017). Slip behavior of the queen Charlotte plate boundary before and after the 2012, MW 7.8 Haida Gwaii earthquake: evidence from repeating earthquakes. *Journal of Geophysical Research: Solid Earth*, *122*(11), 8990-9011.

Hotovec-Ellis, A. J., & Jeffries, C. (2016). Near real-time detection, clustering, and analysis of repeating earthquakes: Application to Mount St. Helens and Redoubt volcanoes. In *Seismological Society of America Annual Meeting*.

Huang, H., & Meng, L. (2018). Slow unlocking processes preceding the 2015 Mw 8.4 Illapel, Chile, earthquake. *Geophysical Research Letters*, *45*(9), 3914-3922.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *IEEE Annals of the History of Computing*, *9*(03), 90-95.

Jiang, C., Wu, Z., Li, Y., & Ma, T. (2014). “Repeating Events” as Estimator of Location Precision: The China National Seismograph Network. *Pure and Applied Geophysics*, *171*(3-5), 413-423.

Kato, A., & Nakagawa, S. (2014). Multiple slow-slip events during a foreshock sequence of the 2014 Iquique, Chile Mw 8.1 earthquake. *Geophysical Research Letters*, *41*(15), 5420-5427.

Kato, A., Obara, K., Igarashi, T., Tsuruoka, H., Nakagawa, S., & Hirata, N. (2012). Propagation of slow slip leading up to the 2011 Mw 9.0 Tohoku-Oki earthquake. *Science*, *335*(6069), 705-708.

Li, A., & Richards, P. G. (2003). Using earthquake doublets to study inner core rotation and seismicity catalog precision. *Geochemistry, Geophysics, Geosystems*, *4*(9).

Li, L., Chen, Q. F., Cheng, X., & Niu, F. (2007). Spatial clustering and repeating of seismic events observed along the 1976 Tangshan fault, north China. *Geophysical research letters*, *34*(23).

Li, L., Chen, Q. F., Niu, F., & Su, J. (2011). Deep slip rates along the Longmen Shan fault zone estimated from repeating microearthquakes. *Journal of Geophysical Research: Solid Earth*, *116*(B9).

Materna, K., Taira, T. A., & Bürgmann, R. (2018). Aseismic transform fault slip at the Mendocino triple junction from characteristically repeating earthquakes. *Geophysical Research Letters, 45*(2), 699-707.

Matsubara, M., Yagi, Y., & Obara, K. (2005). Plate boundary slip associated with the 2003 Off-Tokachi earthquake based on small repeating earthquake data. *Geophysical Research Letters*, *32*(8).

Meier, T., Rische, M., Endrun, B., Vafidis, A., & Harjes, H. P. (2004). Seismicity of the Hellenic subduction zone in the area of western and central Crete observed by temporary local seismic networks. *Tectonophysics*, *383*(3-4), 149-169.

Meng, L., Huang, H., Bürgmann, R., Ampuero, J. P., & Strader, A. (2015). Dual megathrust slip behaviors of the 2014 Iquique earthquake sequence. *Earth and Planetary Science Letters*, *411*, 177-187.

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.

Pacheco, K., Nishimura, T., & Nakahara, H. (2017). Seismic velocity changes of P and S waves associated with the 2011 Tohoku-Oki earthquake (M w 9.0) as inferred from analyses of repeating earthquakes. *Geophysical Journal International*, *209*(1), 517-533.

Poupinet, G., Ellsworth, W. L., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5719-5731.

Sáez, M., Ruiz, S., Ide, S., & Sugioka, H. (2019). Shallow nonvolcanic tremor activity and potential repeating earthquakes in the Chile Triple Junction: seismic evidence of the subduction of the active Nazca–Antarctic spreading center. *Seismological Research Letters*, *90*(5), 1740-1747.

Sawazaki, K., Kimura, H., Shiomi, K., Uchida, N., Takagi, R., & Snieder, R. (2015). Depth-dependence of seismic velocity change associated with the 2011 Tohoku earthquake, Japan, revealed from repeating earthquake analysis and finite-difference wave propagation simulation. *Geophysical Journal International*, *201*(2), 741-763.

Schaff, D. P., & Beroza, G. C. (2004). Coseismic and postseismic velocity changes measured by repeating earthquakes. *Journal of Geophysical Research: Solid Earth*, *109*(B10).

Schaff, D. P., & Richards, P. G. (2011). On finding and using repeating seismic events in and near China. *Journal of Geophysical Research: Solid Earth*, *116*(B3).

Stein, S., & Wysession, M. (2009). *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons.

Tepp, G. (2018). A repeating event sequence alarm for monitoring volcanoes. *Seismological Research Letters*, *89*(5), 1863-1876.

Tkalčić, 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.

Uchida, N. (2019). Detection of repeating earthquakes and their application in characterizing slow fault slip. *Progress in Earth and Planetary Science*, *6*(1), 40.

Uchida, N., & Bürgmann, R. (2019). Repeating earthquakes. *Annual Review of Earth and Planetary Sciences*, *47*, 305-332.

Uchida, N., Matsuzawa, T., Hasegawa, A., & Igarashi, T. (2003). Interplate quasi-static slip off Sanriku, NE Japan, estimated from repeating earthquakes. *Geophysical Research Letters*, *30*(15).

Uchida, N., Matsuzawa, T., Hirahara, S., & Hasegawa, A. (2006). Small repeating earthquakes and interplate creep around the 2005 Miyagi-oki earthquake (M= 7.2). *Earth, planets and space*, *58*(12), 1577-1580.

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, *90*(6), 1353-1368.

Wang, B., Harrington, R. M., Liu, Y., Kao, H., & Yu, H. (2020). A study on the largest hydraulic-fracturing-induced earthquake in Canada: Observations and static stress-drop estimation. *Bulletin of the Seismological Society of America*, *110*(5), 2283-2294.

Wu, W., Zhan, Z., Peng, S., Ni, S., & Callies, J. (2020). Seismic ocean thermometry. *Science*, *369*(6510), 1510-1515.

Yamada, M., Mori, J., & Matsushi, Y. (2016). Possible stick-slip behavior before the Rausu landslide inferred from repeating seismic events. *Geophysical Research Letters*, *43*(17), 9038-9044.

Yao, D., Walter, J. I., Meng, X., Hobbs, T. E., Peng, Z., Newman, A. V., ... & Protti, M. (2017). Detailed spatiotemporal evolution of microseismicity and repeating earthquakes following the 2012 Mw 7.6 Nicoya earthquake. *Journal of Geophysical Research: Solid Earth*, *122*(1), 524-542.

Yu, W. C. (2013). Shallow-focus repeating earthquakes in the Tonga–Kermadec–Vanuatu subduction zones. *Bulletin of the Seismological Society of America*, *103*(1), 463-486.

Zhang, J., Richards, P. G., & Schaff, D. P. (2008). Wide-scale detection of earthquake waveform doublets and further evidence for inner core super-rotation. *Geophysical Journal International*, *174*(3), 993-1006.

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.

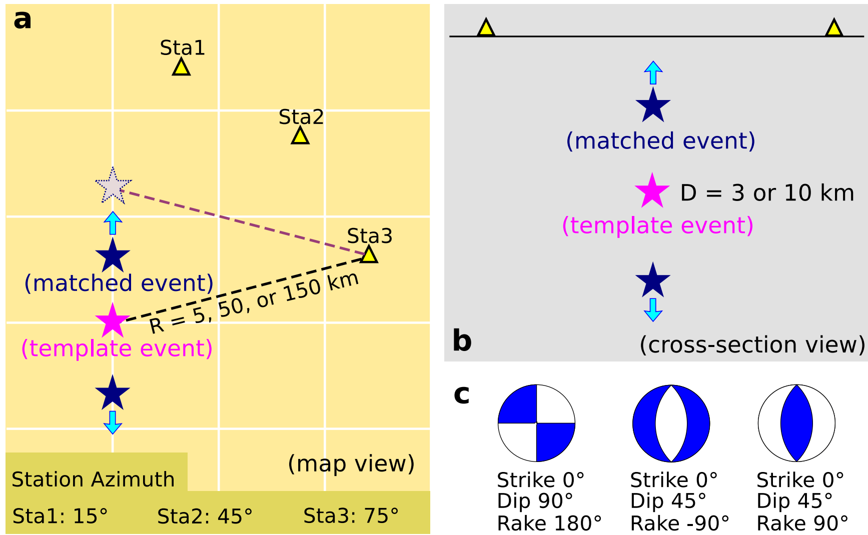

Figure 1. Schematic diagram showing the setup of the synthetic experiments. (a) For horizontal inter-event separation, two sources are placed along a line trending N–S separated by a short distance. The template event (fuchsia star) is fixed in the middle while the matched event (navy blue star) moves away from the template event in both directions. Stations (triangles) are placed at three different epicentral distances (R=5, 50, or 150 km). The gray star marks the location of the matched event such that one of the stations (Sta3 in this case) is of equal distance to both the template and matched events. (b) For vertical inter-event separation, the template event is placed at two different depths (D=3 or 10 km) with the matched event moving up or down. (c) Three different types of focal mechanisms are used in the calculation of synthetic seismograms.

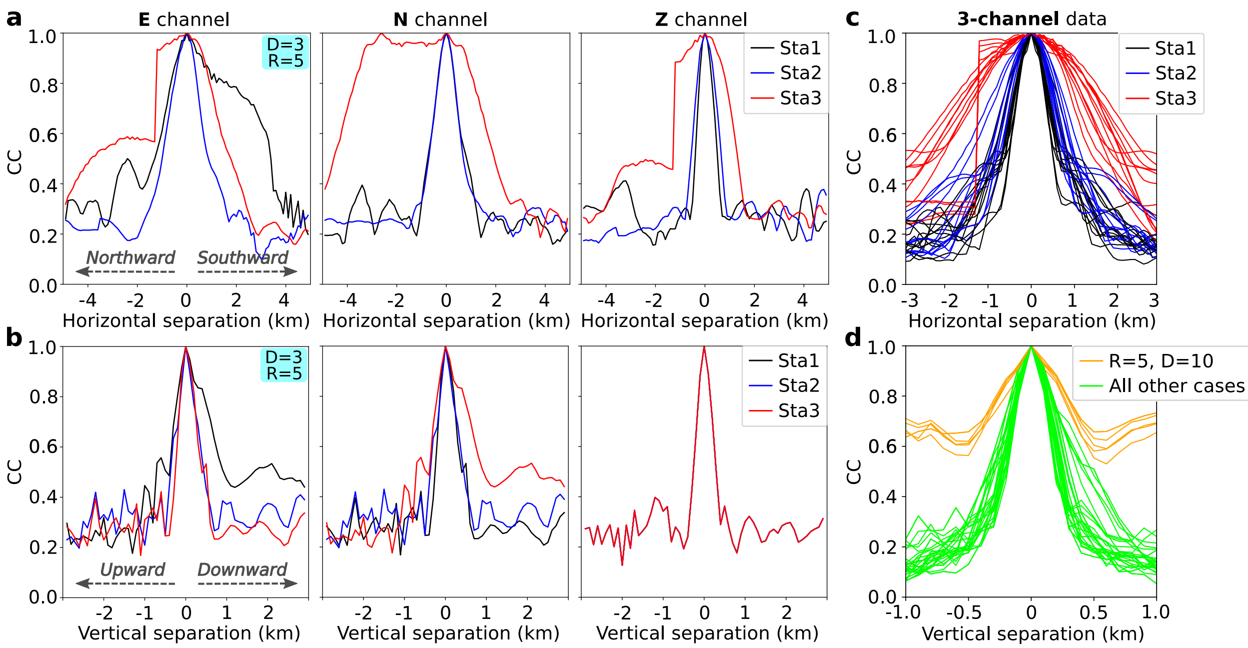

Figure 2. Results of our synthetic experiment showing CC variation as a function of horizontal (a and c) and vertical (b and d) inter-event separation. The setup of sources and receivers is depicted in Figure 1. (a) and (b) correspond to a representative case with single-channel data, whereas (c) and (d) compile all test results with single-station (3-channel) data. Individual test results are presented in Figures S3 and S4. For (a) and (c), positive and negative values along the X axis mean that the matched event is shifted to the south and north, respectively; for (b) and (d), positive and negative X axis mean that the matched event moves downward and upward, respectively.

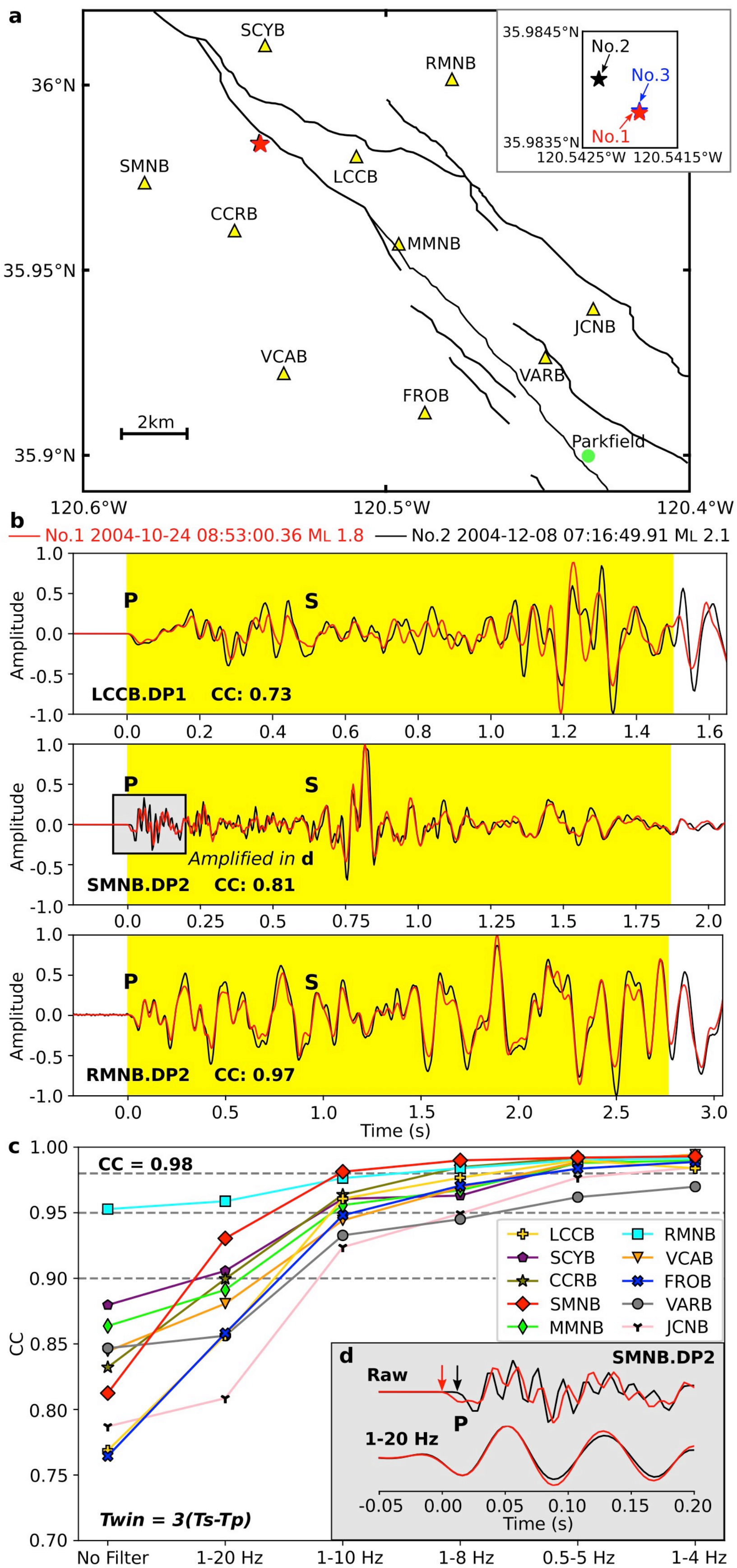
a
SCYB
RMNB
36°N
SMNB
LCCB
CCRB
MMNB
35.95°N
JCNB
VCAB
VARB
FROB
2km
Parkfield
35.9°N
120.6°W
120.5°W
120.4°W
35.9845°N
No.2
No.3
35.9835°N
No.1
120.5425°W 120.5415°W
b
No.1 2004-10-24 08:53:00.36 ML 1.8
No.2 2004-12-08 07:16:49.91 ML 2.1
Amplitude
P
S
LCCB.DP1 CC: 0.73
Amplified in d
SMNB.DP2 CC: 0.81
RMNB.DP2 CC: 0.97
Time (s)
c
CC = 0.98
CC
LCCB
SCYB
CCRB
SMNB
MMNB
RMNB
VCAB
FROB
VARB
JCNB
Twin = 3(Ts-Tp)
No Filter
1-20 Hz
1-10 Hz
1-8 Hz
0.5-5 Hz
1-4 Hz
d
SMNB.DP2
Raw
1-20 Hz
Time (s)

Figure 3. CC test results with real earthquake data. (a) Map showing the distribution of earthquake epicenters (colored stars) and HRSN seismograph stations (yellow triangles). Black lines denote the surface traces of the San Andreas Fault system. The town of Parkfield is shown as a green dot. Insert shows the zoom-in locations of events No. 1-3. (b) Examples of normalized unfiltered waveforms of two events that have been verified to be non-repeaters (No. 1 and No. 2), aligned at the S wave arrival. The highlighted segment indicates the window of dynamic length (see Text S2) used for CC calculation. The gray box in the middle panel outlines the waveform segment amplified in (d). (c) Effects of filtering on the CC values between events No. 1 and No. 2 determined with individual single-station (3-channel) data and dynamic window lengths. (d) An example of waveform change due to filtering. Red and black arrows mark the P wave onset of events No.1 and No.2, respectively. The slight time difference (0.012 s) between the two arrows is overlooked after band-pass filtering between 1 and 20 Hz.

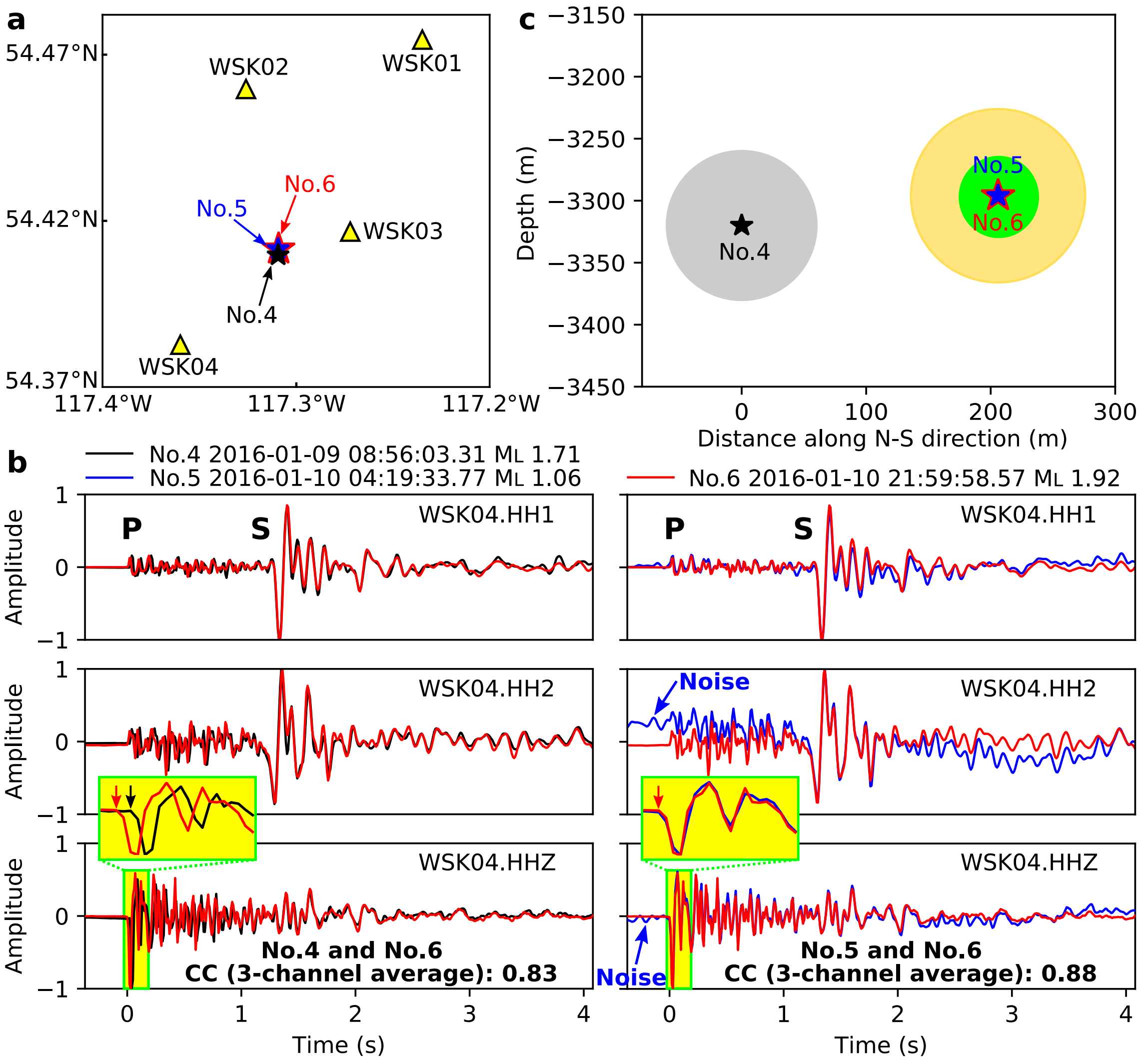

Figure 4. A physics-based approach to distinguish repeating and neighboring events. (a) Map showing the distribution of earthquake epicenters (colored stars) and seismograph stations (yellow triangles). A zoom-in map of the source area is presented in Figure S11. (b) Normalized unfiltered waveforms, aligned at the S wave arrival. Notice the apparent time difference in the S-P differential traveltime between the two cases. (c) North-south cross section showing the relative event locations. Gray, lime, and gold circles are the ERRs of events No. 4, 5 and 6, respectively. The stress drops of events No. 4 and 6 are 35 and 30 MPa, respectively, based on the spectral ratio analysis of the broadband waveforms (*Wang et al*., 2020). Since the waveforms of event No. 5 is contaminated by a high level of noise, its stress drop is assumed to be the average of events No. 4 and 6 (i.e., 32.5 MPa).