Construction of Long-Term Seismic Catalog with Deep Learning: A Workflow for Localized Self-Attention RNN (LoSAR)

Yijian Zhou[1], Hongyang Ding[2], Abhijit Ghosh[3], and Zengxi Ge[4]

[1]Department of Earth and Planetary Sciences, University of California Riverside
[2]Institute of Theoretical and Applied Geophysics, Peking University
[3]University of California, Riverside
[4]Peking University

April 26, 2024

Abstract

Characterizing fault behaviors prior to large earthquakes through long-term seismicity is crucial for seismic hazard assessment, yet constructing high-resolution catalogs over extended periods poses significant challenges. This study introduces LoSAR, a novel deep learning-driven workflow that enhances phase picking by Localizing a Self-Attention Recurrent neural network with local data, addressing the generalization problem common in data-driven approaches. We apply LoSAR to two distinct regions that are both featured by recent large earthquakes: (1) preseismic period of the Ridgecrest-Coso region (2008-2019), and (2) pre-postseismic period of the East Anatolian Fault Zone (EAFZ, 2020-2023/04). Through detailed comparisons, we demonstrate that LoSAR offers slightly higher detection completeness than the QTM matched filter catalog, while boosts an over 100 times faster processing and a superior temporal stability, avoiding low-magnitude gaps during background periods. Against PhaseNet and GaMMA, two established phase picker and associator, LoSAR proves more scalable and generalizable, achieving roughly 2.5 times more event detections in the EAFZ case, along with a ~7 times higher phase association rate. By leveraging the two enhanced catalogs and b-value analysis, we gain insights into the preseismic fault behaviors: (1) The Ridgecrest faults are characterized by sparse and distributed seismicity across a band of ~20 km, revealing multiple orthogonal preexisting faults; coupled with a low b-value that signifies this area as a persistent asperity; (2) The Erkenek-Pütürge segment of EAFZ exhibits complex fault geometry that forms a persistent rupture barrier, which consists of a hidden conjugate fault system that presents as a ~10-km wide fault zone.

Construction of Long-Term Seismic Catalog with Deep Learning: A Workflow for Localized Self-Attention RNN (LoSAR)

Yijian Zhou[1*], Hongyang Ding[2], Abhijit Ghosh[1], and Zengxi Ge[2]

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

[2]Institute of Theoretical and Applied Geophysics, Peking University, Beijing, China

Corresponding author: Yijian Zhou (yijian.zhou@email.ucr.edu)

Highlights:

- We developed a workflow to Localize a Self-Attention RNN (LoSAR) phase picking model, which realizes stable and efficient phase picking.
- The Ridgecrest region has weak but distributed seismicity before the mainshock, and the low b-value suggest a persistent asperity.
- The Erkenek-Pütürge segment of East Anatolian Fault Zone is associated with geometrical complexity, thus represents a persistent barrier.

Abstract

Characterizing fault behaviors prior to large earthquakes through long-term seismicity is crucial for seismic hazard assessment, yet constructing high-resolution catalogs over extended periods poses significant challenges. This study introduces LoSAR, a novel deep learning-driven workflow that enhances phase picking by Localizing a Self-Attention Recurrent neural network with local data, addressing the generalization problem common in data-driven approaches. We apply LoSAR to two distinct regions that are both featured by recent large earthquakes: (1) preseismic period of the Ridgecrest-Coso region (2008-2019), and (2) pre-postseismic period of the East Anatolian Fault Zone (EAFZ, 2020-2023/04). Through detailed comparisons, we demonstrate that LoSAR offers slightly higher detection completeness than the QTM matched filter catalog, while boosts an over 100 times faster processing and a superior temporal stability, avoiding low-magnitude gaps during background periods. Against PhaseNet and GaMMA, two established phase picker and associator, LoSAR proves more scalable and generalizable, achieving roughly 2.5 times more event detections in the EAFZ case, along with a ~7 times higher phase association rate. By leveraging the two enhanced catalogs and b-value analysis, we gain insights into the preseismic fault behaviors: (1) The Ridgecrest faults are characterized by sparse and distributed seismicity across a band of ~20 km, revealing multiple orthogonal preexisting faults; coupled with a low b-value that signifies this area as a persistent asperity; (2) The Erkenek-Pütürge segment of EAFZ exhibits complex fault geometry that forms a persistent rupture barrier, which consists of a hidden conjugate fault system that presents as a ~10-km wide fault zone.

Plain Language Summary

Understanding how faults behave before big earthquakes can help us better prepare for these events, but tracking these faults over a long time is tricky. In our research, we've developed a new method called LoSAR that uses advanced technology to improve how we detect earthquakes by tailoring it with specific local data. This method helps us avoid common issues found in other data-based techniques. We tested LoSAR in two areas known for their significant earthquakes: the Ridgecrest-Coso region in the USA and the East Anatolian Fault Zone (EAFZ) in Turkey, covering events before, during, and after these big earthquakes. Our findings show that LoSAR is not only faster but also more consistent in detecting earthquakes than other methods. Particularly in Turkey, LoSAR was able to identify many more seismic events accurately. From this work, we've learned that the Ridgecrest area shows signs of long-standing stress across a broad network of faults, indicating a high potential for earthquakes. Similarly, in the EAFZ, we discovered a complex network of faults that acts as a barrier to earthquake spread. These insights are crucial for better understanding earthquake risks and preparing for future seismic events.

1. Introduction

Microseismicity provides a direct indication to the fault structure and slip behavior at depth. Such strategy is especially useful when the fault slips at a high rate, e.g. during early aftershock period (e.g. Zhou et al., 2022a; Ding et al., 2023; Shelly et al., 2024), or other intense seismic sequences (e.g. Ross et al., 2020; Tan et al., 2021; Gong and Fan, 2022). However, a huge portion of faults generate low seismicity rate during the interseismic period because of a high locking ratio or a low fault slip rate (e.g. Jiang and Lapusta, 2016; Bletery et al., 2020; Chamberlain et al., 2021; Uchida and Bürgmann, 2021; Zhou et al., 2022b), while, unfortunately, these faults are also prone to large earthquakes (Sykes, 2021; Lay and Nishenko, 2022). To study such low-seismicity faults, a long-term observation is always necessary. For example, Schurr et al. (2020) built a seismic catalog for 7 years before the 2014 Iquique earthquake, and found that the pre-mainshock seismicity complements the coseismic slip; Sugan et al. (2023) observed a 8-year migration of seismicity towards the nucleation area of the 2016 central Italy seismic sequence. Technically, the construction of long-term catalogs requires a workflow that is both computationally efficient and of high detection completeness, which is still a challenging task.

Currently, two types of cataloging workflow can realize a state-of-the-art performance: (1) the PAL-style workflow that follows "phase Picking – phase Association – event Location" procedure (e.g. Zhou et al., 2021b; Zhang et al., 2022; Zhu et al., 2022b), and (2) the matched filter technique (MFT) that utilizes pre-detected events as templates to detect similar events (e.g. Ross et al., 2019a; Shelly, 2020; Neves et al., 2022). The detection completeness of PAL-style workflows is basically dependent on the phase picking algorithm. In recent years, algorithms based on artificial intelligence (AI), specifically deep learning, realize outstanding phase picking performance in terms of the detectability and picking precision (e.g. Zhu and Beroza, 2018; Zhou et al., 2019; Mousavi et al., 2020; Yu and Wang, 2022; Sun et al., 2023). Most of these models, e.g. PhaseNet (Zhu and Beroza, 2018), are trained on regional or global datasets, aiming at building a general model that works for various data not included in the training set. However, systematic tests show that the AI pickers can suffer from inconsistent performance among data in different regions (e.g. Chai et al., 2020; Jiang et al., 2021; Münchmeyer et al., 2022; Zhu et al., 2022a; Park et al., 2023; Bornstein et al., 2024), indicating a lower generalizability compared with traditional rule-based algorithms, such as short-term-average over long-term average (STA/LTA). The MFT methods can realize even higher detection ability than AI pickers (e.g. Mousavi et al., 2019; Zhou et al., 2021a; Yoon and Shelly, 2024), but its low computational efficiency makes it difficult to process big data (Ross et al., 2019b). Moreover, the detection results of MFT may be biased by incomplete templates (e.g. Herrmann and Marzocchi, 2020). In summary, the AI-based picker is the most promising method that combines both high efficiency and high detectability, whereas further improvements are needed to realize a consistent picking performance on a large spatiotemporal range of data.

In this paper, we introduce a novel cataloging workflow powered by deep learning, featuring the Localization of Self-Attention RNN (LoSAR) for phase picking, tailored with local data. This approach effectively addresses the challenge of generalization faced by deep learning models. We apply the LoSAR workflow on two cases that covers a local-to-regional scale and years-to-decade time length, in order to demonstrate its advantage in building long-term catalogs.

2. Methods

2.1 Overview of the LoSAR Workflow

As reviewed in the Introduction section, the first-generation AI pickers attempt to provide a pre-trained model suitable for all datasets, which has not been very successful so far. Instead, we designed a new workflow that generates a local training set to obtain a locally optimized neural network (Figure 1). This workflow is composed of a detection module followed by a location module, and the detection module consists of two major steps: the model training step and the model application step.

In the model training step of the earthquake detection module, we utilize the PAL method (Zhou et al., 2021b) to construct a local training set (Figure 1). PAL utilizes rule-base algorithms for phase picking and association, thus does not have generalization problem as for deep learning. Tests on the 2019 Ridgecrest aftershock sequence show that PAL realizes around 2 times the number of detections by the Southern California Seismic Network (SCSN) (Zhou et al., 2021b), and PAL has been successfully applied in multiple regions and seismic sequences (e.g. Zhou et al., 2021a; Zhou et al., 2022b; Ding et al., 2023). Since the detection of PAL is basically made by the STA/LTA algorithm, which detects both earthquakes and pulse-like noises (e.g. anthropogenic noise or data glitches), we usually set a relatively high triggering threshold to avoid high false detection ratio and to reliably detect high signal-to-noise ratio (SNR) events. The PAL detections serve as the training set for our deep learning model, which can realize stable detections for weak signals. Note that deep learning models usually require a large number of training samples to tune the hyper-parameters, but the number of training samples is dependent on the model complexity. Thus, for light-weight models (e.g. Zhou et al., 2019; Yu and Wang, 2022), as that used in this study (see the next subsection for details), a relatively small training set is required to optimize the neural network (Mousavi et al., 2020; Zhou et al., 2021a).

After the model training step, we simply substitute the PAL picker with the locally-trained SAR picker (LoSAR), and associate the LoSAR picks with PAL associator (Figure 1). The PAL associator groups pairs of P&S picks into events based on their travel time-location relationship, and it also obtains a grid-searched location in the meantime. The magnitude calculation is also completed by the PAL associator in the local magnitude scale (M_L), based on the S-wave amplitude and hypocentral distance (please refer to Zhou et al., 2021b for more details). We will show in this paper that this LoSAR workflow realizes >2.5 times more event detections compared with PAL, and is of a much higher detection stability and accuracy. It also generalizes well among very different tectonic settings, spatial scales, and network configurations.

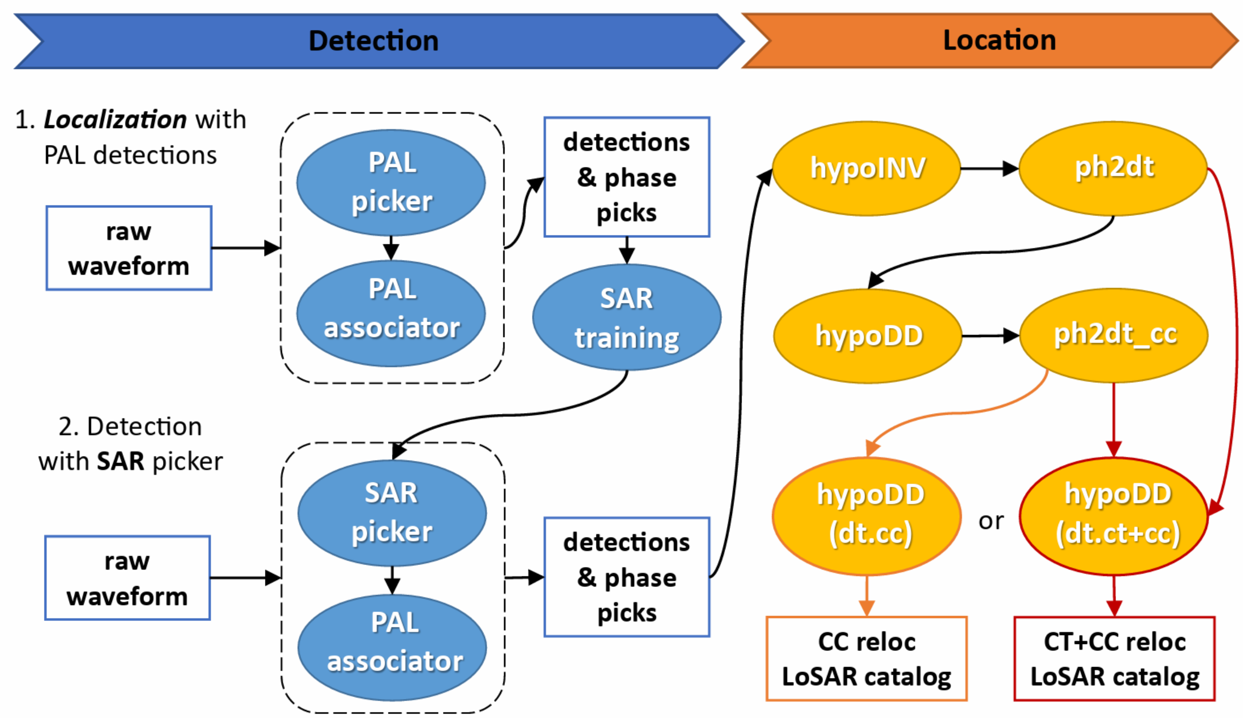


Figure 1. The Localized Self-Attention RNN (LoSAR) workflow. The blue and yellow modules denote the detection and location algorithms, respectively. PAL refers to an rule-based cataloging workflow developed by Zhou et al. (2021b); SAR refers to Self-Attention RNN phase picking model developed in

this study. CT/*dt.ct* and CC/*dt.cc* refer to catalog-based and cross-correlation-based differential travel time, following the hypoDD terminology. The *ph2dt_cc* module is a CC-based differential time calculation method developed in this study.

To locate and relocate the detections, we build interface for HypoInverse (abbreviated as HypoINV thereafter, Klein, 2002) and HypoDD (Waldhauser and Ellsworth, 2000; Waldhauser, 2001) (Figure 1). HypoINV is a widely adopted algorithm for absolute location, which minimize the travel times in an iterative manner, where the weight for the phases are adaptively determined based on their epicentral distance and residual time (Klein, 2002). HypoDD is a double-difference (DD) relocation algorithm that minimizes the differential travel times between event pairs to constrain their relative locations (Waldhauser and Ellsworth, 2000). Given that the fault structure is manifested by the relative locations in seismicity imaging, the relocation process can significantly improve the imaging resolution (e.g. Waldhauser and Ellsworth, 2000; Trugman and Shearer, 2017; Lomax and Savvaidis, 2022). Notably, the differential time (*dt*) data used in hypoDD can come from catalog picks (CT, i.e. *dt.ct*) or cross-correlation (CC, i.e. *dt.cc*), and they can be jointly inversed or used individually. The best practice of using *dt.ct* and *dt.cc* comes from their different characteristics: *dt.ct* covers a larger inter-event distance, whereas the precision is relatively low, because it is dependent on the phase picking and location accuracy; *dt.cc* realize a sub-sampling-rate measurement for the differential time and is unaffected by picking errors, but it covers a much smaller distance, because it relies on waveform similarity between event pairs (as summerized in Waldhauser, 2001). Thus, jointly using *dt.ct* and *dt.cc* in some way is usually suggested. In our LoSAR workflow, we provide two approaches of combining *dt.ct* and *dt.cc* (Figure 1): (1) sequentially relocate with only *dt.ct* and *dt.cc*, and (2) relocate with *dt.ct* in the first step and jointly inverse *dt.ct & dt.cc* in the second relocation. The first approach is suitable for a relatively dense seismic network, because the near-source stations tend to have higher CC values; the second approach can maintain more events under a relatively sparse network, while also take advantage of the high-precision *dt.cc* data.

The calculation of differential times is the fundamental step for hypoDD relocation. The hypoDD software (Waldhauser, 2001) provides a *ph2dt* module that calculate the *dt.ct* data from the input phase file. It forms a chain of *dt*-links between events by searching neighboring events within a certain radius. For each event, a maximum number of neighboring events are preset in this process, in order to lower down the computational complexity and model errors. Following a similar strategy, we develop a *ph2dt_cc* module to calculate the high-precision *dt.cc* data with waveform cross-correlation (Figure S10). It first finds all possible event pairs by comparing the location differences and the common station picks. For each event pair, only events within a certain hypocentral separation and with certain number of shared stations are selected as candidates. To control the quality of *dt* measurements, we only use stations within an epicentral distance of about 100 km, and too small events (e.g. M_L<0) cannot be linked to each other. To avoid too many measurements, we also limit the maximum number of stations for each event pair, and each event can only be linked to a maximum number of neighbors. Secondly, we calculate the CC-derived *dt* for each candidate event pair. In the CC calculation, users can set the window length, which channel to use, and the filtering frequency band. The weight for each phase used in the hypoDD is determined by the square root of CC value. After the CC calculation, we further select the *dt* measurements by discarding that with a too large *dt* or a too small CC, and the minimum number of station criteria still applies afterwards. Detailed parameters will be given for each real case in the following sections.

2.2 Self-Attention RNN (SAR) Model for Phase Picking

The performance of a deep learning model is decided by multiple factors, including the model structure, target function (labeling strategy), training data, and training parameters. Here, we introduce a new phase picking neural network stemming from our previous work (Zhou et al., 2019), but incorporates the recent advances in this field targeting at the above factors (Figure 2).

We adopt a Self-Attention RNN (SAR) model for phase picking. RNN is a typical deep learning model for sequence modelling, and has been widely adopted in various tasks related to time series, including speech recognition (e.g. Graves and Jaitly, 2014; Hannun et al., 2014) and natural language processing (e.g. Cho et al., 2014; Sutskever et al., 2014). Compared with Transformer (Vaswani et al., 2017), which is recently proven to be powerful in sequence-related tasks, RNNs have much fewer parameters, thus are more data-efficient and require less computational resources, making them a practical choice for small datasets (Ezen-Can, 2020; Peng et al., 2023). RNNs process sequences in a recurrent manner: generating outputs for each time step by integrating the current input with a hidden state that captures previously encountered sequence information. Zhou et al. (2019) first adopt a 2-layer bi-directional Gated Recurrent Unit (GRU) RNN in the phase picking task, where the RNN is trained to classify each 1-s time step as noise, P-wave, or S-wave. In this study, we adopt the same GRU structure with 2 bi-directional layers, and 128 hidden size, but divide the raw data into frames of a much smaller step length (0.5 s) and stride (0.1 s) to increase the theoretical phase picking precision (Figure 2a). Moreover, we add a multi-head self-attention layer (Bahdanau et al., 2014; Vaswani et al., 2017) after the GRU layers (Figure 2a).

Multi-head self-attention is a mechanism that enhances the performance by allowing the model to simultaneously focus on different parts of the input sequence from different representation subspaces, which makes a key component of the Transformer architecture (Vaswani et al., 2017). In a single self-attention head, the input sequence (i.e. the output of GRU layers in our case) is transformed into three vectors: queries (Q), keys (K), and values (V), which are then used to compute attention scores that determine how much focus each element of the input sequence should have on every other element. This process captures dependencies regardless of their distance in the sequence, making the model able to capture complex dependencies that span across long sequences. Multi-head self-attention improves upon this by dividing the Q-K-V vectors into multiple independent heads, performing the self-attention process parallelly. Thus, each attention head learns to focus on different features of the input sequence, allowing the model to capture a richer array of relationships within the data. The outputs of all attention heads are then concatenated and linearly transformed to produce the final output, which combines the diverse learned representations. Similar attention mechanisms have been proved effective in enhancing the seismic phase picking performance (e.g. Mousavi et al., 2020; Zhang et al., 2023a).

For the model target in the training stage, we label only the frames containing P & S arrivals as P and S, respectively, and all other frames as Noise (Figure 2a). This labeling strategy forms a small but finite weight for the P&S arrival times, which guides the model to focus on these features and makes more stable phase detection. Since the invention of such labeling strategy by Zhu and Beroza (2018) for PhaseNet, it has been widely adopted by most of the following deep learning models (e.g. Mousavi et al., 2020; Yu and Wang, 2022; Sun et al., 2023). In the model prediction stage, instead of treating the SAR output as classifications, we only use the output prediction probability, so that users can set the detection threshold based on their own problems. In line with PhaseNet, we also set the default triggering threshold for SAR as 0.3 to balance the detection completeness and accuracy. We regard a group of consecutive frames with a prediction probability above the threshold as a P or S pick, and take the median time of these frames as the picked phase arrival. For the cases when multiple P and S picks exist in a sliding window, we group these picks into all possible P&S pairs, only requiring that the S wave always arrives later than the P wave in a pair. Also, since we apply SAR in a sliding window manner, and that the sliding windows have about half the length overlapping, we will merge the picks from different windows if they have similar P&S picks.

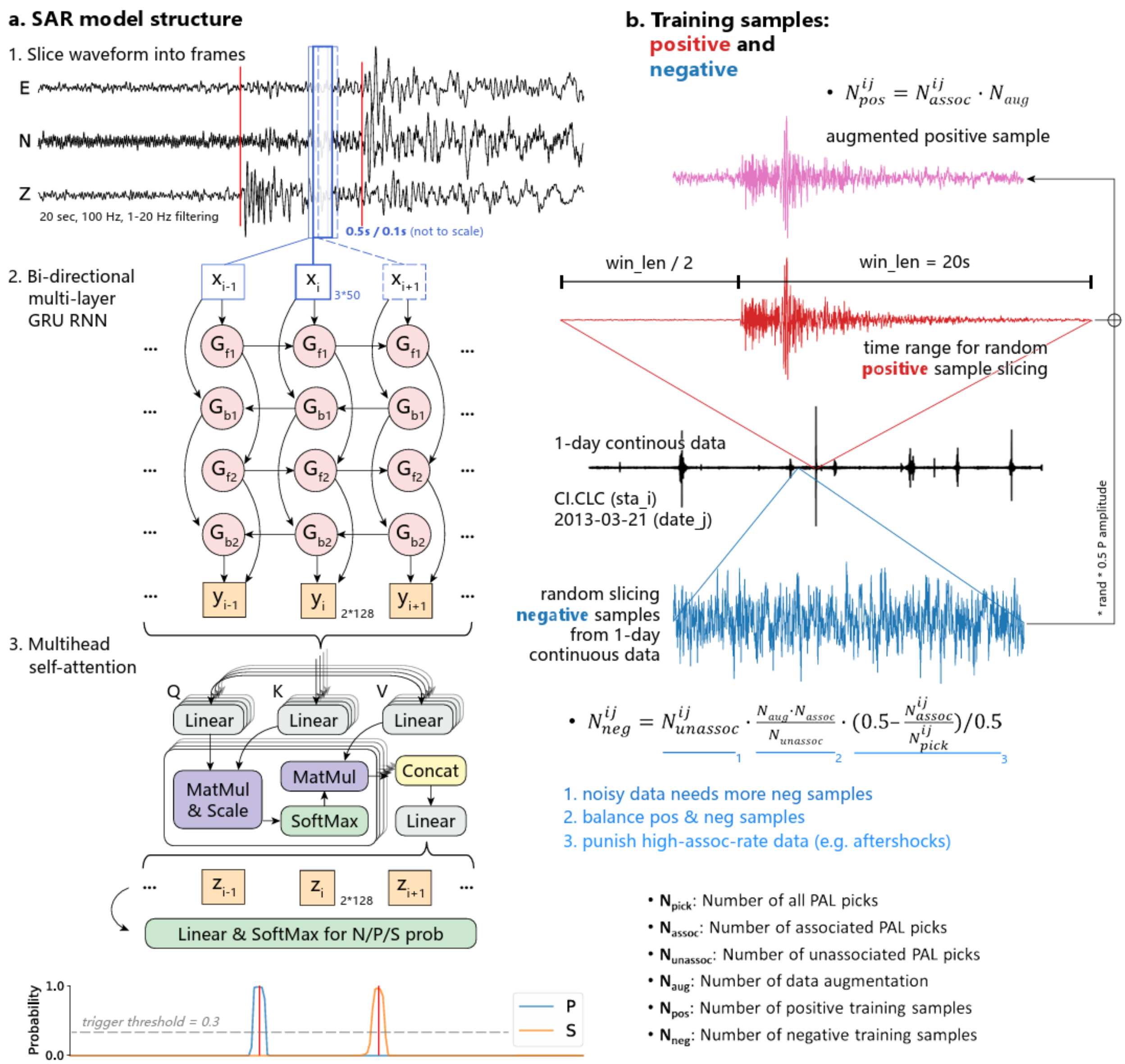


Figure 2. The SAR model structure and training sample slicing strategy. (a) SAR model structure. The filtered 3-channel seismogram is plotted in black curves, and the red vertical lines mark the P&S arrival times. The data processing units are denoted by: x_i for the i^{th} input time step; G_f and G_b for forward and backward Grated Recurrent Unit (GRU); y_i for the i^{th} output of the GRU RNN; Q/K/V for Query, Key, and

Value input in the multi-head self-attention terminology, and they are all set as the RNN output (i.e. y) in the case of self-attention; “Linear” for fully-connected layer; “MatMul” for matrix multiplication; “Concat” for concatenation; z_i for the i^{th} output of the multi-head self-attention layer; “N/P/S prob” for Noise, P wave, and S wave prediction probability (b) Training sample slice strategy. Blue, red and pink waveforms for negative, positive, and augmented positive samples, respectively.

The training samples fundamentally determine how the model behaves. To train the SAR model with the local earthquakes detected by PAL, we design a sampling strategy to properly balance the model’s ability of detecting weak signals and identifying different types of noises (Figure 2b). We slice both positive (i.e. earthquakes) and negative (i.e. noise) samples for training. The positive samples are sliced surrounding the PAL-picked P&S arrivals (Figure 2b). We randomly position the P arrival at the first half of the time window, so that the model has a consistent detection ability for random signal positions that it will encounter in the real applications. Data augmentation was employed to increase the diversity of the training set and improve the generalizability of the SAR model. This involves adding real noise, randomly sliced from the same station-date for a specific P&S pick. For each augmented sample, the noise amplitude is scaled by a random ratio between 0 and 0.5, multiplied by the maximum P-wave amplitude. In this way, the number of positive samples is:

$$N_{pos}^{ij} = N_{assoc}^{ij} \cdot N_{aug} \, , \tag{1}$$

where *i & j* is the station and date index, respectively; N_{pos} is the number of positive samples; N_{assoc} is the number of associated picks; N_{aug} is the number of augmentations set by the users. Note that the users need to set the number of augmentations, so that the number of positive samples is large enough (empirically, >100,000). The negative samples are sliced randomly in each station-date pair (Figure 2b), while excluding the time ranges that has PAL-picked P&S arrivals. The number of negative samples on a certain station-date is decided by the number of associated and unassociated PAL picks on that station-date:

$$N_{neg}^{ij} = N_{unassoc}^{ij} \cdot \frac{N_{aug} \cdot N_{assoc}}{N_{unassoc}} \cdot \frac{0.5 - \frac{N_{assoc}^{ij}}{N_{pick}^{ij}}}{0.5} \, , \tag{2}$$

where *i & j* is the station and date index, respectively; N_{neg} is the number of negative samples; N_{pick}, N_{assoc} & $N_{unassoc}$ are the number of all picks, number of associated and unassociated picks, respectively; N_{aug} is the number of augmentations set by the users. The logic for each term of this equation is that: (1) the number of unassociated PAL picks indicate the noise level, the more noise a station-date is, the more negative samples we need to slice on it; (2) the number of positive and negative samples need to be balanced in the training process; (3) the pick association ratio indicate the true positive rate of the PAL picks, and a large number of unassociated picks and a high association rate can occur together in an intense sequence (e.g. aftershock sequences), thus we should reduce the number of negative samples in this case. Note that if the association ratio is larger than 0.5, we do not slice any negative samples on that station-date. In this way, we can obtain a temporally stable number of negative samples (Figure S7). We will show in this study that the above sampling strategy realizes stable detection performance.

The training of SAR is performed in a mini-batch manner. We feed both positive and negative samples in each iteration, and train for about 15-20 epochs to make about 100,000 total training iterations. Note that with the third term of Equation 2, the total number of negative samples will be smaller than the positives, thus we feed 128 positives and a smaller number of negatives so that they experience the same number of epochs. We adopt the Adam optimizer (Kingma and Ba, 2014) with a learning rate of 10^{-4}. As will be shown in the real cases here (Figure S8) and in our previous experiments for RNN (Zhou et al., 2019), the SAR model behaves very stable in the training process, without showing any signs of overfitting.

3. Comparison of Cataloging Workflows

To test the detectability and generalizability of our LoSAR workflow, we apply it to two cases that differ in tectonic settings and spatiotemporal scales: (1) the Ridgecrest-Coso (California) region from 2008 to 2019/07, covering its long-term preseismic period and 20-days' early aftershocks; and (2) the East Anatolian Fault Zone (EAFZ, Turkey) from 2020 to 2023/04, which covers ~3 years' preseismic period and ~3 months' aftershocks. Both cases contain large earthquakes that significantly change the seismicity rate and patterns, which makes the cataloging more difficult, and is thus suitable for technical discussions. In these cases, we compare LoSAR with other popular cataloging workflows in terms of earthquake detection completeness, stability, and phase association rate.

3.1 Case 1: Ridgecrest-Coso Region (2008-2019/07)

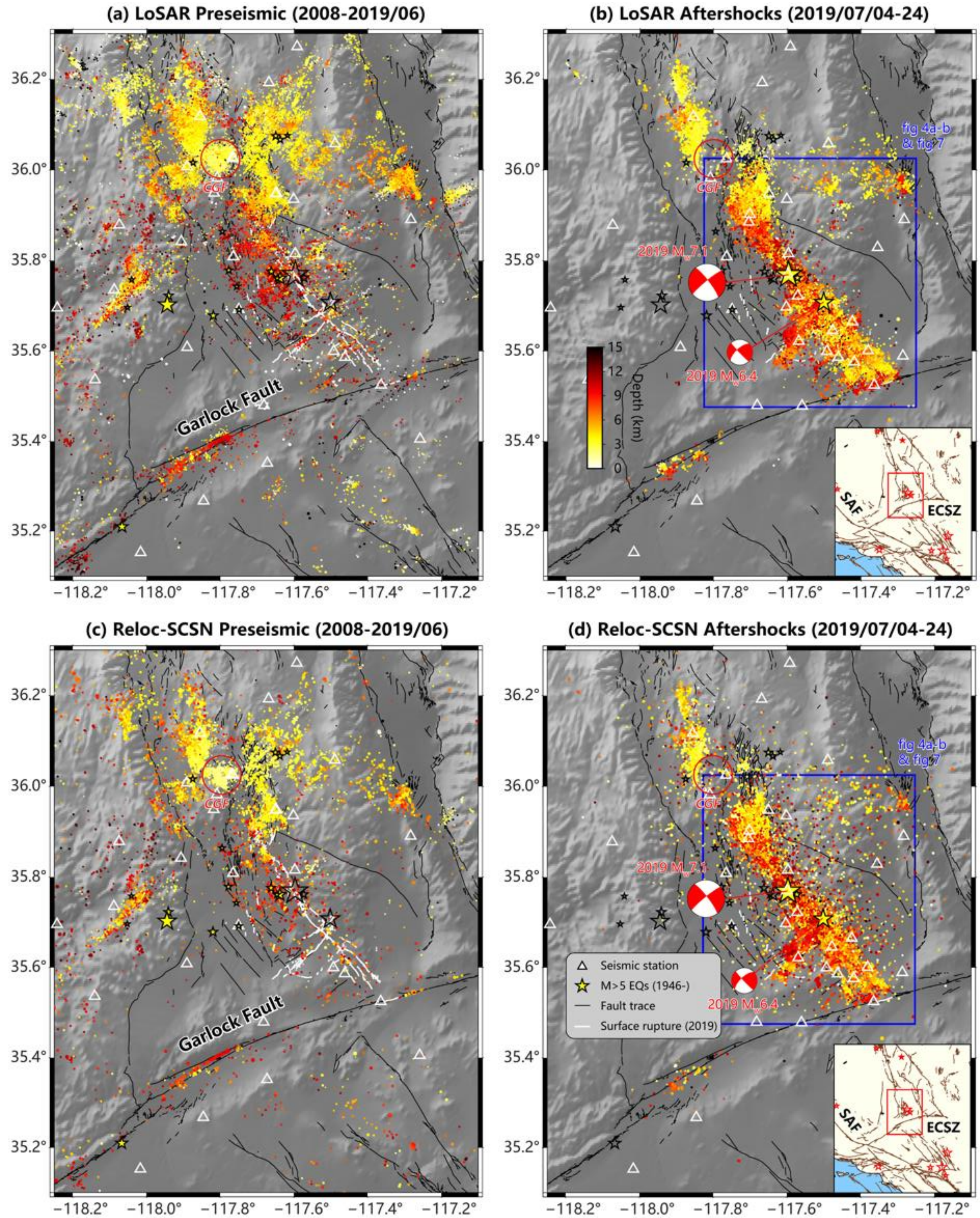


Figure 3. Study region and catalog comparison for Case 1: Ridgecrest-Coso (2008-2019/07). (a-b) plot the LoSAR catalog generated in this study, with preseismic period (2008-2019/06) and aftershock period (2019/07/04-24) shown separately. Seismicity is plotted as dots that have its depth denoted by color and size varies by magnitude. M>5 earthquakes after 1946 are marked by yellow stars. Seismic stations used are denoted by white hollow triangles; active faults are plotted as black lines; surface rupture caused by the 2019 Ridgecrest earthquakes are marked by white lines. The area of Coso Geothermal Field (CGF) is marked by a red circle. The insets show location of the study area in a larger scale, with the San Andreas Fault (SAF) and Eastern California Shear Zone (ECSZ) marked. (c-d) are the same as (a-b), but for the relocated SCSN catalog (Hauksson et al., 2012). The blue box in (b) & (d) marks the location coverage in Figure 4a-b & 7.

3.1.1 Background and Motivation

The 2019 Ridgecrest, California sequence comprised a M_w 6.4 foreshock on 2019/07/04, followed by a M_w 7.1 mainshock on 2019/07/06, and was featured by intense aftershock activities that reveal widespread orthogonal structures (Ross et al., 2019a; Shelly, 2020). It is associated with a young fault system (Goldberg et al., 2020; Hauksson and Jones, 2020; Xu et al., 2020) and the majority of the ruptured faults are not mapped before the earthquake (Thompson Jobe et al., 2020). Tectonically, the Ridgecrest sequence is situated within the East California Shear Zone (ECSZ), an approximately 100-km-wide band dominated by right-lateral strike-slip faults that accommodates the relative motion between the Pacific and North American plates (Oskin et al., 2008; Spinler et al., 2010; Tymofyeyeva and Fialko, 2015). In this study, we focus on an area not only covering the Ridgecrest ruptures, but also two close-by tectonic units (Figure 3): (1) the central Garlock fault (McGill and Sieh, 1993; Ganev et al., 2012; Hatem and Dolan, 2018), a major fault cutting off the Ridgecrest faults on its south; and (2) the Coso Geothermal Field (CGF), one of the largest three geothermal fields in California that has been inducing intense seismicity for decades (Schoenball et al., 2015; Trugman et al., 2016; Im et al., 2021).

We work on a time period from 2008 to 2019, because of the significant improvement in the seismic network starting from 2008 (Hutton et al., 2010), and to overlap with the decade-long QTM matched filtering catalog that covers 2008-2017 (Ross et al., 2019b). This time frame offers an exceptional opportunity to directly compare our novel workflow with MFT. For the aftershock period, we also align our investigations with the MFT catalog by Ross et al. (2019a), which spans the first 20 days since the M_w 6.4 foreshock. Both time periods of the QTM catalog utilize the routine SCSN catalog as templates, adopting a detection threshold of 12 times the median absolute deviation of CC, and are relocated with Growclust (Trugman and Shearer, 2017) that inverse the CC-based differential times. Also involved in the comparison is the relocated SCSN catalog (Hauksson et al., 2012), which has a similar number of detections as the routine SCSN catalog, but has a much higher relocation precision that comes from the CC-based relocation. To make a fair comparison, we utilize the same set of stations maintained by SCSN, which is composed of broad-band and short-period stations of 3- and 1-component recordings, along with several near-source temporary stations deployed after the 2019 mainshock (Figure 3 & S1a-b).

3.1.2 Detection and Location

For this case, we run the LoSAR workflow for the ~11-year preseismic period and the 20-day aftershock period separately, due to the considerable differences in station distribution (Figure 3 & S1) and seismicity rate. This approach also allows for an assessment of the workflow’s performance under two end-member situations: a scenario of very intense sequence observed by a rather dense seismic network, and a pure interseismic period (no large earthquakes) with regular observational condition. We set the LoSAR parameters specifically for the preseismic period, since it is less investigated; and we keep the same set of parameters for the aftershock period to make the results more comparable. In running PAL, we set the STA/LTA triggering threshold as 12 (defined by energy), S wave searching window as 12 s after the P pick, in line with the scale of study region and the average inter-station distance. For PAL association, we require at least 4 stations to have an original time deviation <1.2 s, and a maximum P-wave travel time residual <0.8 s. The PAL obtains 61,053 / 451,694 and 49,737 / 440,942 event detections / associated picks for the preseismic and aftershock period, respectively. In running SAR, we set the window length as 20 s and the sliding stride as 10 s in the predictions for continuous data. We augment the original training samples by 2 times, making a total of 812,771 / 451,689 positive / negative samples for training in the preseismic period, and 792,784 / 123,018 positive / negative training samples for the aftershock period. Note that there are ~10% remaining samples serve as the validation set. The waveforms in a time window are band-pass filtered to 1-20 Hz and normalized in both the training and application stage. We made 15 epochs of training until the accuracy is stable (Figure S8). As described in the last section, we apply the local-trained SAR picker and the same PAL associator to enhance the PAL detections. This gives 165,393 / 1,277,974 and 122,933 / 1,078,038 event detections / associated picks for the preseismic and aftershock period, respectively. In summary, for Ridgecrest-Coso, the LoSAR workflow achieves a ~2.5-2.7-fold increase in detection number compared to PAL.

To locate and relocate the LoSAR detections, considering a rather dense network, we adopt the first approach described in Section 2.1 (Figure 1), utilizing *dt.ct* and *dt.cc* separately for a two-step relocation. Similar as the detection part, we adopt the same set of location parameters for both the preseismic and aftershock period. For the hypoINV absolute location, the distance weighting is configured as follows: 0-50 km assign full weight, >100 km zero weight, and 50-100 km is assigned a cosine tapered in between. A similar weighting scheme is employed for the time residual between 0.25 and 0.75 s. For S waves, the weights are further adjusted by a factor of 0.6, considering a larger picking error relative to P waves. The above weighting strategy results in 95.4% events well located, and an average lateral and vertical uncertainty of 1.11 km and 2.22 km, respectively. In the hypoDD relocation process with *dt.ct*, we pair events within 8 km (i.e. WDCT, in hypoDD terminology), P & S waves weighted respectively as 1 and 0.6, and inverse for 4 iterations. This results in a relative location error of about 120 m for epicenter and 150 m for depth, under the least square criteria by hypoDD. Given that the least-square location error reported by hypoDD tends to significantly underestimate the real uncertainty, we test the *dt.ct* relocation results under different velocity models, and found highly consistent distributions (Figure S16 & S32). In the final CC relocation stage, we measure the waveform-based differential time on 1-20 Hz band-pass filtered waveforms, with the P and S window length set as 2.5 s and 4 s, starting from 0.5 s and 0.2 s before the phase arrival, respectively. We do not calculate *dt* between $M_L<0$ earthquakes to lower down the computational costs. A maximum number of 200 neighboring events are preset in the candidate neighbor selection step, sorted by the separation distance. After the CC calculation, we discard the measurements with CC<0.35, or *dt_p*>0.5 s, or *dt_s*>0.8 s, and the event pairs with <4 stations fulfilling these criteria are further dropped. The above selections finally maintain 18,421,488 P and 12,506,600 S differential time measurements for the preseismic period. For the aftershocks, these numbers are 13,718,530 for P and 8,422,550 for S. In the hypoDD relocation process, we link events within 4 km (i.e. WDCC, in hypoDD terminology), set both P and S waves' weights as 1 and inverse for 4 iterations, because S waves have higher stability in waveform cross-correlation. Finally, we obtained 101,193 and 61,578 well-relocated events for the preseismic and aftershock period, respectively. For both catalogs, the average relocation error reported by hypoDD is about 40 m and 60 m along horizontal and vertical directions, respectively. Note that this location error is in the relative and average sense, and reflects more about the goodness of data fitting, instead of the real uncertainty.

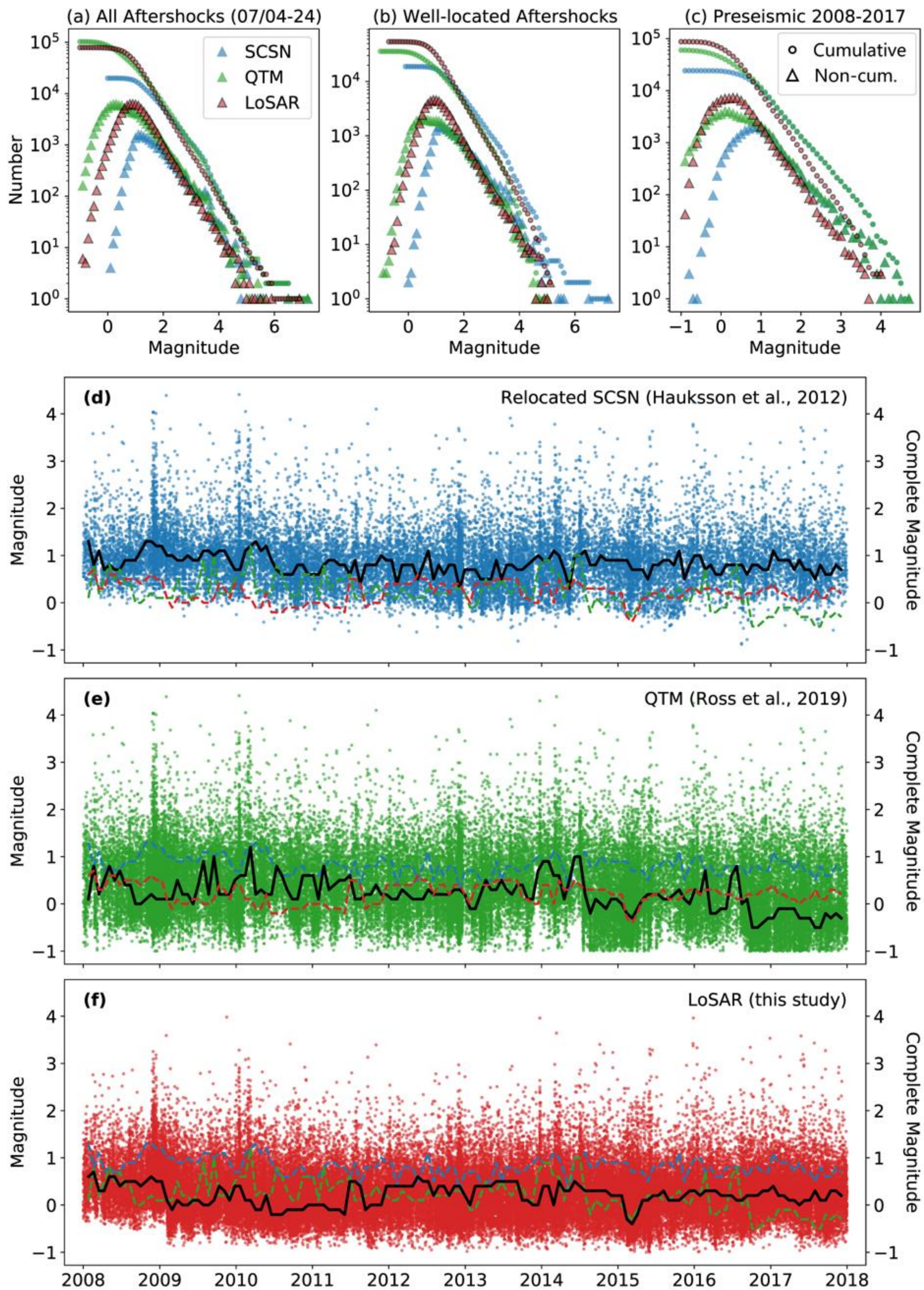
3.1.3 Catalog Comparisons

The cataloging methodology and parameters outlined in the preceding subsection yield the seismicity distribution shown in Figure 3. The LoSAR catalog and the relocated SCSN catalog exhibit an overall consistent distribution, particularly in areas of major intense clusters. Nonetheless, the LoSAR catalog demonstrates a significantly enhanced detection capability, rendering previously indistinct seismicity structures much more discernible (Figure 3). To make a more comprehensive and quantitative comparison, we plot the frequency-magnitude distribution (FMD) and magnitude-time sequences of the relocated SCSN, the QTM catalog, and the LoSAR catalog for both the aftershock and preseismic period (Figure 4).

FMDs serve to illustrate not only the total number of detections, but also the distribution of events across different magnitude bins. Ideally, a FMD should adhere to the GR law (Gutenberg and Richter, 1944), meaning that for the magnitude range above a completeness threshold (approximately the magnitude of maximum non-cumulative distribution, Wiemer and Wyss, 2000), the occurrence frequency of earthquakes is expected to follow a power-law distribution relative to their magnitudes (i.e. a linear relationship when plotted on a logarithmic scale). Furthermore, the objective of earthquake detection is also to achieve the smallest possible magnitude of completeness, or equivalently, the largest possible cumulative count. Applying the above criteria to analyze the FMDs, it becomes evident that LoSAR has a comparable or superior detectability compared with QTM, especially for the preseismic period (Figure 4a-c). While QTM shows a higher detection ability for the intense aftershock sequence (Figure 4a), it is noteworthy that ratio of well-located events determined by the Growclust algorithm falls below 35% (Figure 4b). This result also suggests that a trade-off between the quantity of detections and the quality of their locations, emphasizing the necessity of keeping comparable relocation precision for an equitable comparison of detectability across methodologies. In addition to overall event counts, the slope of these FMDs also offer critical insights. First, an apparent inconsistency in detection completeness exists in both the SCSN and QTM catalogs between the M 0-2 and M >2 events, as evidenced by the varying slopes in the FMDs (Figure 4a-c). Second, also for the SCSN & QTM catalog, there is a noticeable shift in the FMDs’ slope around M 3.5 (Figure 4a-c). This shift is partially caused by the adoption of different magnitude scales (e.g. M_l, M_w, or M_{lr}, as detailed at https://scedc.caltech.edu/eq-catalogs/change-history.html), which can affect the catalog-based b-value studies. The LoSAR catalog does not exhibit the aforementioned issues, demonstrating a consistent detection performance across the magnitude range in this study.

The magnitude-time sequence reveals the temporal evolution of seismicity, which is modulated by external sources, such as tectonic loading, as well as internal triggering between earthquakes (Ogata, 1988; Zhuang et al., 2002; Zaliapin and Ben-Zion, 2020; Hsu et al., 2024). Thus, statistical studies will require a catalog to have temporally stable detection capability for both the intense sequences and the background seismicity. In our case in the Ridgecrest-Coso preseismic period, we observe that the SCSN catalog is not temporally consistent, displaying high detectability mainly during the intense seismic sequences (Figure 4d). This pattern arises from the SCSN catalog's compilation process, which involves a semi-automated detection supplemented by manual inspections, particularly for intense seismic sequences (as detailed at https://scedc.caltech.edu/eq-catalogs/change-history.html). The issue of temporal inconsistency is not alleviated, but being exacerbated in the QTM catalog, as indicated by the highly variable lower magnitude limit (Figure 4e). It is probably due to the fact that earthquakes within an intense sequence tend to have high waveform similarity, making them easy to be detected by matched filtering; Conversely, the background seismicity comes from a wider variety of faults and asperities, which inherently diminishing the effectiveness of matched filter. In contrast, our LoSAR catalog realizes markedly improved stability in detection throughout this 11-year period (Figure 4f), without showing any notable gaps in the detection of lower magnitude events. This robustness suggests that the SAR model is adept at capturing the statistical features of seismic events over a large spatiotemporal range, thus making it a more generalized algorithm for long-term earthquake detection.

In our final comparison, we assess the computational efficiency of the LoSAR workflow against QTM. The entire LoSAR process (including running PAL) is executed within ~7 days with 1 Nvidia GeForce RTX 2080 GPU card and 1 Intel Xeon E5-2695 CPU. Contrarily, the construction of the QTM catalog requires 200 Nvidia P100 GPU cards, with a runtime exceeding 60 days. Although the study area of QTM is ~20 times larger, there is still a 100-fold difference in computational efficiency. This significant difference is attribute to the computationally intensive nature of cross-correlation and that the linear increase in runtime with the addition of templates in matched filtering.



(a) All Aftershocks (07/04-24)
(b) Well-located Aftershocks
(c) Preseismic 2008-2017
SCSN
QTM
LoSAR
Cumulative
Non-cum.
Number
Magnitude
(d)
Relocated SCSN (Hauksson et al., 2012)
(e)
QTM (Ross et al., 2019)
(f)
LoSAR (this study)
Complete Magnitude

443

Figure 4. Comparison of frequency-magnitude distribution (FMD) and magnitude-time sequence. The blue, green, and red color denote the SCSN catalog (Hutton et al., 2010; Hauksson et al., 2012), the QTM catalog (Ross et al., 2019a; Ross et al., 2019b), and the LoSAR catalog in this study, respectively. (a-c) plot the FMD comparison for the total aftershock detections, well located aftershocks, and the preseismic period (2008-2017), respectively. The dots and triangles denote cumulative and non-cumulative distribution in FMD. (d-f) plot the magnitude-time comparisons for the preseismic period (2008-2017). The thick black lines denote the magnitude of completeness in the current panel, and the colored dash lines plot that for two other catalogs. Note that only the events within the location coverage of Figure 3 are included in the statistics.

In conclusion, the LoSAR workflow not only exhibits superior detection capabilities and stability compared to the matched filter approach, but also significantly outperforms in computational efficiency. This makes it particularly advantageous for applications across extensive spatiotemporal scales.

3.2 Case 2: East Anatolian Fault Zone (2020-2023/04)

3.2.1 Background and Motivation

On February 6, 2023, Southeast Turkey's Kahramanmaraş region was struck by an M_w 7.8 earthquake followed by another M_w 7.6 within ~9 hours, marking one of the largest continental earthquake doublets ever documented (Dal Zilio and Ampuero, 2023; Hussain et al., 2023; Jia et al., 2023; Xu et al., 2023; Zhang et al., 2023b; Ren et al., 2024). The 2023 Turkey earthquake doublet occurs on the East Anatolian Fault Zone (EAFZ), a roughly ~600-km-long fault zone defining the plate boundary between the Anatolian and Arabian plates (Figure 5). Driven by the collision between the Arabian and Eurasian plates, the Anatolian plate is extruded westward at a rate of approximately 20-25 mm/yr, along with a counterclockwise rotation (McClusky et al., 2000; Bulut et al., 2012; Barbot and Weiss, 2021; Güvercin et al., 2022). This plate motion give rise to the predominantly left-lateral strike-slip characteristics observed along the EAFZ. Additionally, the fault slip rate along the EAFZ is relatively low, ranging from ~4 mm/yr on the SW segments to ~10 mm/yr on the central and NE segments (Cavalié and Jónsson, 2014; Walters et al., 2014; Aktug et al., 2016; Weiss et al., 2020). It is noteworthy that the two mainshocks of the 2023 doublet are associated with different fault systems (Figure 5): the first M_w 7.8 mainshock (denoted as *M1*) occurs on the major plate boundary faults, whereas the second M_w 7.6 mainshock (denoted as *M2*) ruptures the intraplate faults. The distinct aftershock patterns observed across this dual faults suggest contrasting fault properties (Ding et al., 2023; Güvercin, 2024).

In this study, we build a seismic catalog covering the entire EAFZ, a task that presents technical challenges due to the large spatial extent, and the relatively sparse seismic network, with its inter-station distances ranging from about 30 to 60 km (Figure 5 & S1c). This network density is representative of most regions outside of the well-monitored areas like California and Japan. We start our analysis from 2020, because of the significant enhancement of Turkey seismic network in that year (Figure S1c & S2b), spurred by a M_w 6.8 earthquake on January 24, 2020, that ruptured the NE side of the 2023 rupture zones (Gallovič et al., 2020; Lin et al., 2020; Melgar et al., 2020; Pousse-Beltran et al., 2020). Our study extends over a 3-month aftershock period to 2023/04, which expands our previous rapid work that only covers 1 months' early aftershocks (Ding et al., 2023) and the 2-month aftershock period examined by Güvercin (2024). We combine 3-channel broad-band stations from multiple networks, including the TU network from the Disaster and Emergency Management Presidency (AFAD) and the KO network from the Kandilli Observatory and Earthquake Research Institute (KOERI). Instead of discussing this case in the EAFZ alone, we will combine the previous case in the Ridgecrest-Coso to compare the generalization ability of different phase picking and association algorithms in earthquake detection. This is motivated by the fact that the two cases are associated with contrasting tectonic settings, spatiotemporal scales, and network configurations.

3.2.2 Detection and Location

Contrary to our first case in Ridgecrest-Coso, for this case in EAFZ, we execute the LoSAR workflow over a combined ~3.5-year pre-postseismic period. This approach offers additional validation of LoSAR's capability to handle highly variable seismicity rates. The detection and location parameters for this EAFZ case largely align with those used in the Ridgecrest-Coso case, with the exception of several parameters adjusted to accommodate the expanded scope and density of the network.

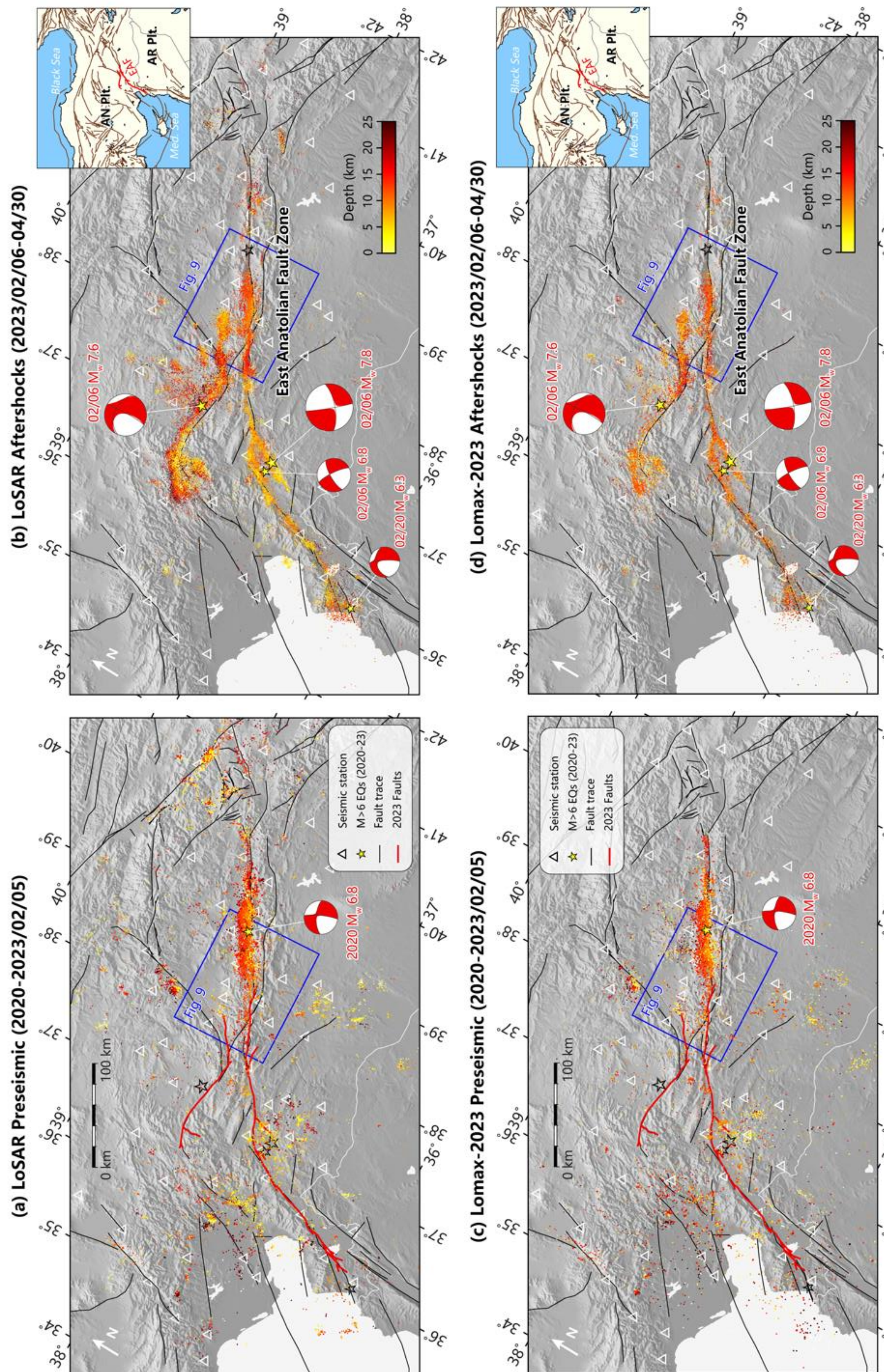

(a) LoSAR Preseismic (2020-2023/02/05)
(b) LoSAR Aftershocks (2023/02/06-04/30)
(c) Lomax-2023 Preseismic (2020-2023/02/05)
(d) Lomax-2023 Aftershocks (2023/02/06-04/30)
East Anatolian Fault Zone
Fig. 9
02/06 Mw 7.6
02/06 Mw 7.8
02/06 Mw 6.8
02/20 Mw 6.3
2020 Mw 6.8
Depth (km)
Seismic station
M>6 EQs (2020-23)
Fault trace
2023 Faults
0 km
100 km
Black Sea
Med. Sea
AN Plt.
AR Plt.
EAF

Figure 5. Study region and catalog comparison for Case 2: East Anatolian Fault Zone (2020-2023/04). (a-b) plot the LoSAR catalog generated in this study, with preseismic period (2020-2023/02/05) and aftershock period (2023/02/06-04/30) shown separately. Seismicity is plotted as dots that have its depth denoted by color and size varies by magnitude. M>6 earthquakes after 2020 are marked by yellow stars. Seismic stations used are denoted by white hollow triangles; active faults are plotted as black lines; responsible faults for the 2023 earthquake doublet are marked by red lines. The insets show location of the study area in a larger scale, with the Arabian Plate (AR Plt.) and Anatolian Plate (AN Plt.) marked. (c-d) are the same as (a-b), but for the relocated AFAD catalog by Lomax (2023). The blue boxes mark the location coverage in Figure 9.

In running PAL, the phase picker utilizes the same triggering threshold, but an 18-s S wave search window, which corresponds to a ~150-km epicentral distance. The PAL associator maintains a 4-station requirement, but different in the limit for original time deviation as <1.6 s and the P-wave travel time as <1.2 s. These specifications resulted in 47,416 / 295,494 event detections / associated picks. In running SAR, we set a larger window length of 25 s and maintain the same 10-s sliding stride. We augment the original training samples by 2 times, making a total of 516168 / 322,894 positive / negative training samples. Similarly, ~10% remaining samples are used for validation. We applied the same data preprocessing, and made 20 epochs of training until the accuracy is stabilized (Figure S8). After applying the locally-trained SAR picker and the same PAL associator, we got 140,119 / 929,944 event detections / associated picks, which represents a ~3-fold increase in the number of detections compared to PAL.

To locate and relocate these detections, we adopt the second approach described in Section 2.1 (Figure 1), utilizing both *dt.ct* and *dt.cc* in the second round of relocation. For the hypoINV absolute location, we also weight the phases by its epicentral distance and residual times, but set a less restrictive set of parameters to account for the sparser network: the distance weighting is configured as 0-80 km full weight, >160 km zero weight, and 80-160 km cosine tapered; the time residual weighting is configured in the same way between 0.4 s and 1.2 s. Similar as for Ridgecrest-Coso, the S-wave weights are further adjusted by a factor of 0.6. The hypoINV outputs ~93.1% well located events, with an average lateral and vertical uncertainty of 2.48 km and 3.73 km, respectively. Note that the location error is highly variable across our study region, due to the heterogenous station distribution. In the first round of hypoDD relocation with *dt.ct*, we pair events within 20 km (i.e. WDCT, in hypoDD terminology), and P & S waves weighted as 1 & 0.6, inversed for 4 iterations. The relocation uncertainty is also correlated with the station density, ranging from about 180-280 m for epicenter and 250-550 m for depth, under the least square criteria by hypoDD. For CC-based differential time measurements, we filter the waveforms to 1-12 Hz, considering the sparser network, while keeping the same P and S windows as in Ridgecrest-Coso. We only calculate *dt* between M_L>0.5 earthquakes and other events. We limit a maximum number of 50 neighboring events in the candidate selection step, and we require CC>0.3, *dt_p*<1 s, and *dt_s*<1.75 s, in addition to a minimum number of 4 stations fulfilling these criteria. The above processes finally obtain 3,719,995 P and 2,961,716 S differential times. For the final hypoDD relocation process that combines *dt.ct & dt.cc*, we made 2 sets of inversions. In the first round, we set WDCC and WDCT as 4 km and 10 km, and weighted the *dt.ct* data strongly; in the second round of inversion, the WDCC and WDCT are set to 2 km and 5 km, with the *dt.cc* being dominantly weighted. Using this weighting scheme, we benefit from both the high-precision *dt.cc* data and the wide coverage of *dt.ct* that helps maintain more events in relocation. Finally, we obtained 93,680 well-relocated events, with an average relocation error ranging from about 40-80 m and 70-120 m along horizontal and vertical directions, respectively. Furthermore, aiming for a higher-resolution catalog (at the cost of reducing the number of detections), and for a more realistic estimation of location uncertainty, we also applied the first approach outlined in Section 2.1 (Figure 1) to relocate the LoSAR detections. We exclusively used dt.cc for relocation, setting WDCC as 2 km and performing the inverse over 4 iterations. This yielded 56,656 events with a much lower relocation uncertainty and an overall consistent distribution of seismicity (Figure S22-S24). However, it is important to note that this relocation approach results in a considerable reduction of events above the completeness magnitude, rendering it less suitable for b-value analysis.

The detection and location strategies elaborated above result in a significantly enhanced catalog compared with the relocated AFAD catalog produced by Lomax (2023), as show in Figure 5. Such advancement is made possible by a much higher detectability and relocation precision. Similar as in Ridgecrest-Coso, the LoSAR catalog exhibit ideal FMDs and magnitude-time series in EAFZ (Figure S25), effectively capturing the seismicity features for both the preseismic and postseismic period.

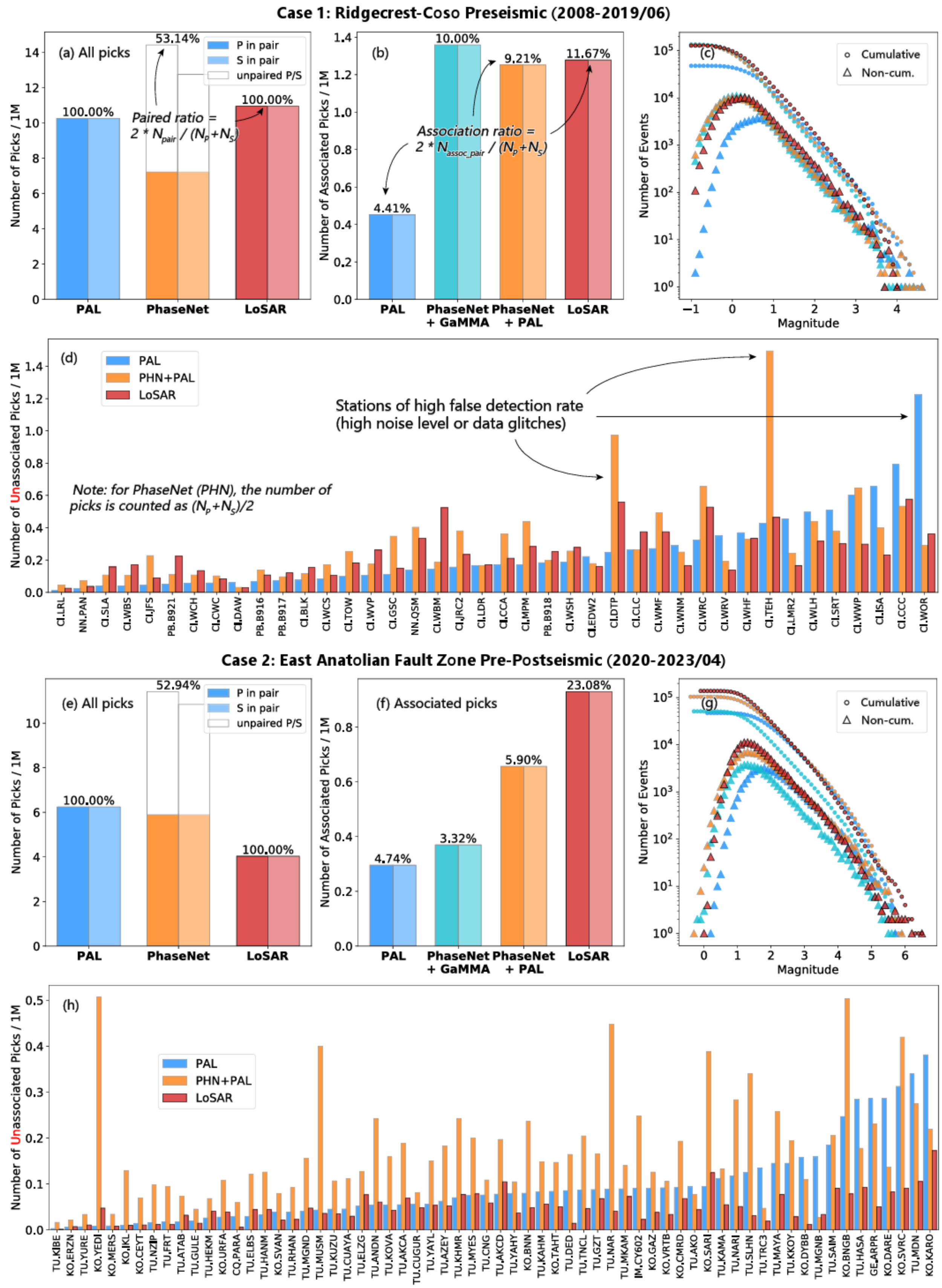

Case 1: Ridgecrest-Coso Preseismic (2008-2019/06)
(a) All picks
Number of Picks / 1M
53.14%
100.00%
100.00%
P in pair
S in pair
unpaired P/S
Paired ratio = 2 * N_pair / (N_P+N_S)
PAL
PhaseNet
LoSAR
(b)
Number of Associated Picks / 1M
10.00%
9.21%
11.67%
4.41%
Association ratio = 2 * N_assoc_pair / (N_P+N_S)
PAL
PhaseNet + GaMMA
PhaseNet + PAL
LoSAR
(c)
Cumulative
Non-cum.
Number of Events
Magnitude
(d)
PAL
PHN+PAL
LoSAR
Stations of high false detection rate (high noise level or data glitches)
Note: for PhaseNet (PHN), the number of picks is counted as (N_P+N_S)/2
Number of Unassociated Picks / 1M
CI.LRL NN.PAN CI.SLA CI.WBS CI.JFS PB.B921 CI.WCH CI.CWC CI.DAW PB.B916 PB.B917 CI.BLK CI.WCS CI.TOW CI.WVP CI.GSC NN.QSM CI.WBM CI.JRC2 CI.LDR CI.CCA CI.MPM PB.B918 CI.WSH CI.EDW2 CI.DTP CI.CLC CI.WMF CI.WNM CI.WRC CI.WRV CI.WHF CI.TEH CI.MR2 CI.WLH CI.SRT CI.WWP CI.ISA CI.CCC CI.WOR
Case 2: East Anatolian Fault Zone Pre-Postseismic (2020-2023/04)
(e) All picks
Number of Picks / 1M
52.94%
100.00%
100.00%
P in pair
S in pair
unpaired P/S
PAL
PhaseNet
LoSAR
(f) Associated picks
Number of Associated Picks / 1M
23.08%
5.90%
3.32%
4.74%
PAL
PhaseNet + GaMMA
PhaseNet + PAL
LoSAR
(g)
Cumulative
Non-cum.
Number of Events
Magnitude
(h)
PAL
PHN+PAL
LoSAR
Number of Unassociated Picks / 1M
TU.KIBE KO.ERZN TU.YURE KO.YEDI KO.MERS KO.IKL KO.CEYT TU.NZIP TU.FRT TU.ATAB TU.GULE TU.HEKM KO.URFA CQ.PARA TU.ELBS TU.HANM KO.SVAN TU.RHAN TU.MGND TU.MUSM TU.KUZU TU.CUAYA TU.ELZG TU.ANDN TU.KOVA TU.AKCA TU.CUGUR TU.YAYL TU.AZEY TU.KHMR TU.MYES TU.CNG TU.AKCD TU.YAHY KO.BNN TU.KAHM KO.TAHT TU.DED TU.TNCL TU.GZT TU.NAR TU.MKAM IM.CY602 KO.GAZ KO.VRTB KO.CMRD TU.AKO KO.SARI TU.KAMA TU.NARI TU.SLHN TU.TRC3 TU.MAYA TU.KKOY KO.DYBB TU.MGNB TU.SAIM KO.BNGB TU.HASA GE.ARPR KO.DARE KO.SVRC TU.MDN KO.KARO

Figure 6. Comparison of phase picking and association performance. The blue, cyan, orange, and red markers denote the results by PAL, PhaseNet picker + GaMMA associator, PhaseNet picker + PAL associator, and LoSAR workflow developed in this study, respectively. (a & e) plot the number of picks. The P & S picks are marked by darker and lighter colors. The paired P&S picks are filled by color, with the paired ratio annotated on top of the bar; (b & f) plot the number of associated picks. The pick association ratio is annotated on top the bar; (c & g) plot the FMD comparison, where circles and triangles denote cumulative and non-cumulative numbers, respectively; (d & h) plot the number of unassociated picks on each station. The stations are sorted by the number by PAL. (a-d) and (e-h) plot the Case 1 (Ridgecrest-Coso) and Case 2 (EAFZ), respectively. Note that for (c) & (g), only the events within the location coverage of Figure 5 are included in the statistics.

3.2.3 Comparisons with Both Cases

Utilizing this case in the EAFZ pre-postseismic period, along with our previous case in the Ridgecrest-Coso preseismic period, we conduct detailed comparisons against PhaseNet (Zhu and Beroza, 2018) and GaMMA (Zhu et al., 2022c), both of which are considered as the state-of-the-art phase picking and association algorithms, and are being widely adopted in cataloging workflows (e.g. Zhang et al., 2022; Zhu et al., 2022b). Through these comparisons, we can evaluate the scalability of these algorithms across different tectonic environments, spatiotemporal ranges, and network densities.

In executing PhaseNet and GaMMA, we adhere primarily to their default parameters and the behaviors anticipated for typical users, which are intended to be broadly applicable. For PhaseNet, we use 30-s time windows that slide in 15-s steps. They are larger than what we set for SAR picker in the two cases, and should have enough coverage. The waveforms are high-pass filtered above 1 Hz. We keep the default triggering threshold at 0.3. Picks from overlapping sliding windows are merged using a 0.5-s window, where the picks with higher prediction probabilities are selected. Note that PhaseNet pick P & S arrivals independently, and we pair them up to combine with the PAL associator. The phase pairing strategy is consistent with that used for the SAR picker; however, it necessitates manually setting an S-wave search window, which we keep the same as for PAL picker in the two cases. For GaMMA, we set the criterion at a minimum of 8 phases to declare an event detection, aligning this requirement with the 4-station criterion of the PAL associator that we applied in both the EAFZ and Ridgecrest-Coso cases. It is important to highlight that the association results from GaMMA do not necessarily ensure paired P and S phases for each station. However, we observed that directly utilizing the GaMMA outputs for hypoINV lead to a significant number of events unconstrained (Figures S11 & S12). Consequently, we opted to refine the GaMMA outputs by excluding the stations with only single P or S picks and imposed a requirement for a minimum of 4 stations.

We first compare the number of associated picks and the association ratio (Figure 6a-b & e-f). We assume that the association process filters out most of the false picks, thus the number of associated picks represents the phase detection completeness, and the association ratio indicate the detection accuracy. In the Ridgecrest-Coso case, all three combinations of AI-driven picker-associator setups (i.e. PhaseNet+GaMMA, PhaseNet+PAL, and LoSAR) achieved a similar number of associated picks, with an association ratio around 10% (Figure 6a-b). This performance markedly outperforms that of the STA/LTA method, as represented by PAL, demonstrating the advanced picking accuracy and completeness afforded by deep learning methodologies that learn from waveform features. Note that PhaseNet is originally trained on Northern California data, thus is intrinsically localized for the Ridgecrest-Coso case. Moreover, all three methods recalled ~97% events in the SCSN catalog, showing great consistency in event detection. In contrast, in the EAFZ case, the number of associated picks by LoSAR surpasses that of PhaseNet+PAL by ~1.5 times, and ~2.5 times more than PhaseNet+GaMMA (Figure 6e-f). Furthermore, both PhaseNet+PAL and PhaseNet+GaMMA demonstrate significantly lower association ratios (5.9% and 3.3%, respectively) in comparison to LoSAR's 23.1%, as well as to their performances in the Ridgecrest-Coso case (~10%). This result indicates that PhaseNet is less tailored for the EAFZ data, and that GaMMA suffers from a much larger-scale network with sparse station distributions. By comparing the FMDs (Figure 6c & g), it is evident that the three AI-driven picker-associator combinations show high consistency in Ridgecrest-Coso, as previously noted. However, PhaseNet+GaMMA displays noticeable discrepancies for the EAFZ case, missing a significant portion of events above the complete magnitude. Excluding this particular instance, the remaining seven catalogs across both cases display consistent FMDs for the portion above completeness magnitude, aligning with the expectations set by the GR law.

Lastly, we examine the number of unassociated picks across all stations in both cases (Figure 6d & h). Given the presumption that a large portion of the unassociated picks represent false detections, the spread of unassociated picks across stations serves as an indicator of a phase picker's detection stability. In our analysis, we compare three phase pickers: PAL, PhaseNet, and

LoSAR, employing PAL (based on STA/LTA) as the benchmark for evaluating PhaseNet and LoSAR. The rationale behind this comparison stems from the high sensitivity of STA/LTA algorithms to pulse-like noises (e.g. data glitches), which often results in a higher volume of unassociated picks. Consequently, when PhaseNet or LoSAR generate significantly more unassociated picks than PAL, it suggests a high false detection rate for low-SNR phases. In both cases, there is no correlation in the number of unassociated picks between PAL and the two AI-based pickers (Figure 6d & h), demonstrating the AI pickers' effective ability to discriminate against pulse-like noises. This performance aligns with expectations due to their design to analyze whole-waveform characteristics. However, even in the case of Ridgecrest-Coso (Figure 6d), a few stations were identified where PhaseNet produced a significantly higher number of unassociated picks compared to PAL. Such instability of PhaseNet is notably severer in EAFZ, where multiple stations recorded more than twice the unassociated picks than PAL (Figure 6h). These observations suggest that PhaseNet may be sensitive to certain types of low-SNR noises. A similar behavior is also observed with LoSAR's picking results, albeit to a lesser extent (Figure 6d). The above observations highlight a potential area for improvement in AI-picker's noise discrimination capabilities, and a multi-station-based picking algorithm could offer a solution (e.g. Feng et al., 2022; Sun et al., 2023).

In summary, our comparative analysis across cases with divergent characteristics demonstrate the LoSAR's superiority in detection completeness and stability over PhaseNet, as well as PAL's greater scalability in phase association compared to GaMMA. These attributes position LoSAR as a notably robust workflow for compiling long-term, large-scale seismic catalogs.

4. Characterizing Preseismic Fault Behaviors

Leveraging the high-resolution catalogs developed in the preceding section, we conduct practical analysis to demonstrate the advantages of an enhanced catalog in characterizing fault behaviors before a large earthquake. Our investigation spans both the Ridgecrest-Coso and EAFZ region, exploring whether the preseismic catalog shed light on coseismic fault behaviors, which is potentially helpful in seismic hazard assessment for other regions prior to large earthquakes. We confine the analysis within the spatial dimension, utilizing seismicity distribution to infer the fault geometry and using b-value as an indicator of stress level. The examinations of temporal seismicity evolution and b-value variation demand more rigorous analysis and statistical techniques, thus fall beyond the scope of this paper.

The seismicity imaging technique relies on the understanding that earthquakes, regardless of their size, occur on faults, including both major and subsidiary ones. Consequently, the clustered seismicity serves as a tool to detect geologically hidden faults, and to direct delineate fault structures at depth with a high resolution (e.g. Waldhauser and Schaff, 2008; Hauksson et al., 2012; Ross et al., 2019b; Shelly, 2020). However, the occurrence of seismicity is intricately linked to the stress and frictional state of faults (as reviewed in Bürgmann, 2018), which introduces several caveats in the interpretation of seismicity distribution. For example, certain fault segments can be quiescent during the interseismic period, because of the stress shadow drawn by a deeply penetrated rupture of previous large earthquakes (Jiang and Lapusta, 2016). Therefore, a lack of continuous seismicity does not necessarily imply a discontinuity in fault structure. Another commonly observed phenomenon is the complementary distribution between aftershocks and coseismic slip (Wetzler et al., 2018; Toda and Stein, 2022). In such cases, aftershocks tend to concentrate on the periphery of ruptured asperities, thereby obscure the fault dip in cross-section views. Thus, to accurately infer fault geometry, it is always recommended to jointly consider the focal mechanism solutions and the surface traces of faults or ruptures (e.g. Lu et al., 2021; Zhou et al., 2022a; Ding et al., 2023).

The b-value in the GR-law, which quantifies the relative number of large versus small earthquakes, is inversely related to the stress level, as demonstrated by numerous experimental and statistical studies (Scholz, 2015, and references therein). In this study, we obtain the spatial distribution of b-values by calculations performed on each grid cell, incorporating earthquakes within a specified radius. Given a set of earthquakes, the b-value can be estimated by the maximum-likelihood method (Aki, 1965):

$$b = \frac{\log_{10} e}{\bar{M} - M_C + \frac{\Delta M}{2}}\ , \tag{3}$$

where $\bar{M}$, M_C, and ΔM is the mean magnitude, lower cut-off magnitude, and the magnitude bin, respectively. The b-value uncertainty is estimated following Shi and Bolt (1982):

$$\delta b = 2.3 \times b^2 \times \sqrt{\sum\nolimits_{i=1}^{n} \frac{(M_i - \bar{M})^2}{n(n-1)}}\ , \tag{4}$$

where *n* is the number of events. For reliable b-value estimations, we require a minimum number of events above the cut-off magnitude, typically >80. It is important to recognize that the b-value uncertainty determined by Equation 4 tend to underestimate its true variability (Woessner and Wiemer, 2005; Amorèse et al., 2010; Marzocchi et al., 2019). Therefore, we conduct additional tests to prevent over-interpretation in our applications, which includes: (1) Adjusting the slicing radius for each grid, which affect the number of events in b-value estimations; (2) Modifying the cut-off magnitude (M_C), given that b-value estimation is positively correlated with M_C (Cao and Gao, 2002; Zhou et al., 2018); and (3) whether to adopt a uniform M_C or calculate M_C individually on each grid. Note that we opt to use a uniform M_C in the main text, based on the considerations mentioned above.

4.1 2019 Ridgecrest-ruptured fault

As reviewed in Section 3.1.1, the 2019 Ridgecrest sequence activated a largely unmapped orthogonal fault system. However, subsequent detailed geological investigation by Thompson Jobe et al. (2020) suggest that up to 50-70% of the fault traces could have been mapped before the earthquake. It is practically important to know whether those faults can be imaged by long-term seismicity, and whether this region represent a hazardous seismic gap, considering that the Ridgecrest sequence breaks a ~20-year quiescence in the ECSZ (Ross et al., 2019a; Chen et al., 2020; Hauksson and Jones, 2020).

From the event distribution prior to the earthquake (Figure 7a-b), it is evident that an over 20-km-wide zone along the Ridgecrest faults is characterized by spreading microseismicity. This pattern is more distinctly observed in our LoSAR catalog compared to both the relocated SCSN catalog and the QTM catalog (Figure S17). Though the southern segments have much lower seismicity rate, the seismicity in the LoSAR catalog exhibit multiple clusters trending orthogonally to the main faults ruptured during the 2019 M_w 7.1 earthquake, aligning with the focal mechanisms of M>2 events within those clusters (Figure S18). Additionally, these orthogonal clusters appear to be further activated following the mainshock, since they collocate with the aftershocks and mapped surface traces (Figure 7c). The above observations suggest that multiple subparallel faults extend to the SE of the 2019 M_w 6.4 foreshock, an area where previously few fault traces had been mapped. The net effect of such a fault system is a distributed shear deformation and a low slip rate on each individual fault, which agree with the weak seismic activity observed in this region.

By examining the seismicity depth distribution in the cross sections (Figure 7), we find a similar pattern of depth contour before and after the 2019 mainshocks, despite contrasting intensities of seismic activity, both complementing the coseismic slip. Furthermore, the b-value mapping results reveal that the areas of weak seismicity in the southern segments correspond to notably low b-values (Figure 8), indicating a high level of differential stress. Collectively, these findings suggest that the southern segments were strongly locked prior to the earthquake, and can be considered as a persistent asperity primed for rupture. It is noteworthy that the northern segments, responsible for the M_w 7.1 mainshock ruptures, are of relatively high b-value (Figure 8). This can be interpreted by the stress shadow effect caused by several M>5 events during the 1990s on its western subparallel faults. Nanjo (2020) identified a similar contrasting pattern in preseismic b-value distribution using an extensive dataset from over 40-yr SCSN catalog, a pattern not discernible with merely ~11 years of data, as shown in Figure 8a. This contrast in preseismic b-value near the foreshock and mainshock hypocenters provides insight into the sequential occurrence of the M_w 6.4 foreshock before the larger M_w 7.1 mainshock.

In summary, the distributed microseismicity prior to the mainshock unveils a preexisting fault system composed of multiple subparallel branches. This system, characterized by a weak seismicity and an overall low b-value, represents a persistent asperity of potential hazard.

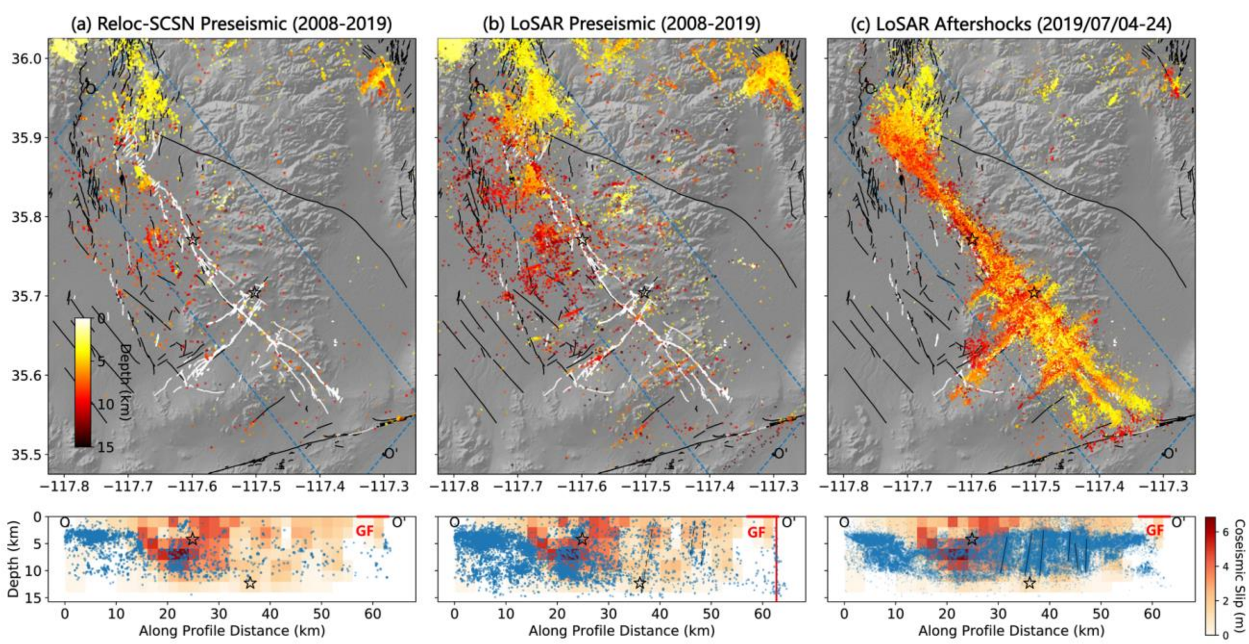


Figure 7. Seismicity on the Ridgecrest faulting area before and after the 2019 earthquakes. (a-b) plot the preseismic period of relocated SCSN catalog (Hauksson et al., 2012) and LoSAR catalog in this study, respectively, and (c) for the LoSAR catalog aftershock period. Note that the dot size and transparency for the preseismic (a-b) and aftershock (c) period are set differently, because of the very different seismicity rate. In the map view plots (upper panels), the seismicity is plotted as dots that have its color denoting the depth and size varies with the magnitude. The active faults are plotted as black lines; the surface ruptures are marked by white lines. The reference points and spatial coverage of the along-fault cross-section is marked by blue dashed rectangle. The hollow black stars denote the largest foreshock and the mainshock of the 2019 Ridgecrest sequence. In the cross-sections (lower panels), the rupture model by Yue et al. (2021) is plotted as different shades of red. The Garlock Fault (GF) on the surface and seismicity-interpreted depth extension is marked by red lines. The black lines are the imaged cross faults by seismicity.

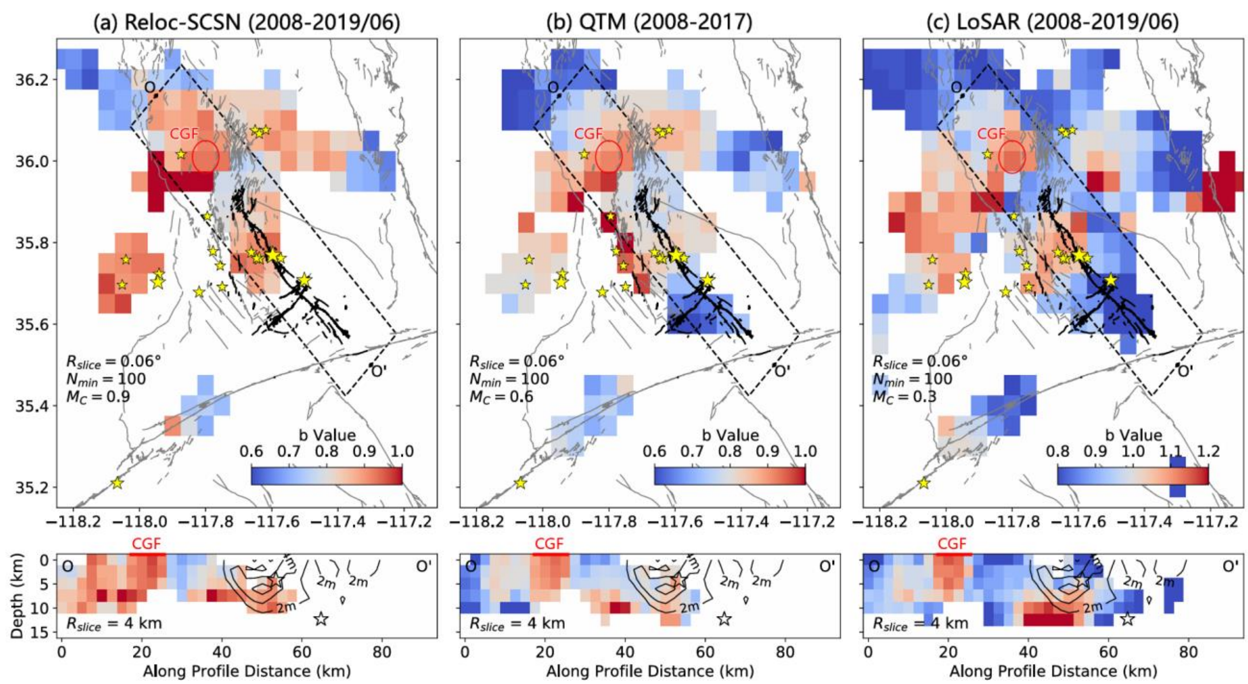


Figure 8. Comparison of b-value mapping results for the RC-Coso region during preseismic period with different seismic catalogs. (a-c) plot the b-value mapping results with the relocated SCSN catalog (Hauksson et al., 2012), the QTM catalog (Ross et al., 2019b), and the LoSAR catalog in this study, respectively. The active faults and surface ruptures are plotted by gray and black lines, respectively. The M>5 earthquakes after 1946 are marked by yellow stars. The area of Coso Geothermal Field (CGF) is marked by red circle in the map view and by horizontal line in the cross section. The coseismic slip by Yue et al. (2021) is plotted as black contours in the cross section. Annotations in the map view include: R_{slice},

the slicing radius for each grid in b-value calculation; N_{min}, the minimum number of events above the complete magnitude to calculate b-value; M_C, the magnitude of completeness.

4.2 Erkenek-Pütürge fault segment (EPF) of EAFZ

As reviewed in Section 3.2.1, the 2023 Turkey doublet is extraordinarily large for continental environments, raising questions about the mechanisms that allow an earthquake to reach such size. Instead of investigating how the rupture grows, we are curious about how it terminates. Specifically, we explore the termination of the *M1* rupture at its NE end, named as Erkenek-Pütürge fault segment (EPF), which acted as a barrier not only for *M1*, but also for the 2020 M_w 6.8 earthquake and multiple historical and paleoseismic events (Hubert-Ferrari et al., 2020; Güvercin et al., 2022; Karabacak et al., 2023). Fortunately, the observational condition for this segment of the EAFZ were optimal during our study period from 2020 to 2023, which set a solid foundation to examine the formation of this persistent barrier.

Firstly, we utilize all seismicity from 2020-2023 to image the fault structure (Figure 9). The map view (Figure 9a-b) and cross-sections (Figure 9c) reveal an along-strike variation in fault structure: from a single major fault in profile 01-04, to a major-secondary fault style in profile 05-10, and then to a wide fault zone from profile 11-15, and reverts to the major-secondary pattern in profile 16-18. A zoom-in plot with CC-relocated LoSAR catalog (Figure 10c) show that the ~10-km wide fault zone (profile 11-15) is composed of two major subparallel branches, along with several subsidiary conjugate faults in between. Intriguingly, the surface fault traces in this area also exhibit two subparallel branches with shapes similar to the seismicity observed at depth, but of a much narrower width (Figure 10c). This may imply that the fault zone broadens as it extends deeper into the crust, a scenario not commonly observed (Scholz, 1988; Ben-Zion and Sammis, 2003; Scholz, 2019). In addition to the aforementioned structural variation is a shift in the dip angle of the major fault, which gradually change from an SSE-dipping in profile 01-04 to an NNW-dipping in profile 17-18 (Figure 9c). This contrast in fault dip direction of the EPF is also reflected in the moment tensors of the 2023 *M1* (subevent inversion by Jia et al., 2023) and the 2020 M_w 6.8 & M_w 5.6 (Figure 9b). These structural variations appear to correlate with the coseismic behavior, with the unruptured segments (profile 08-14) show greater geometrical complexity, in contrast to the relatively simpler structures that were ruptured in 2023 (profile 01-07) and 2020 (profile 15-18).

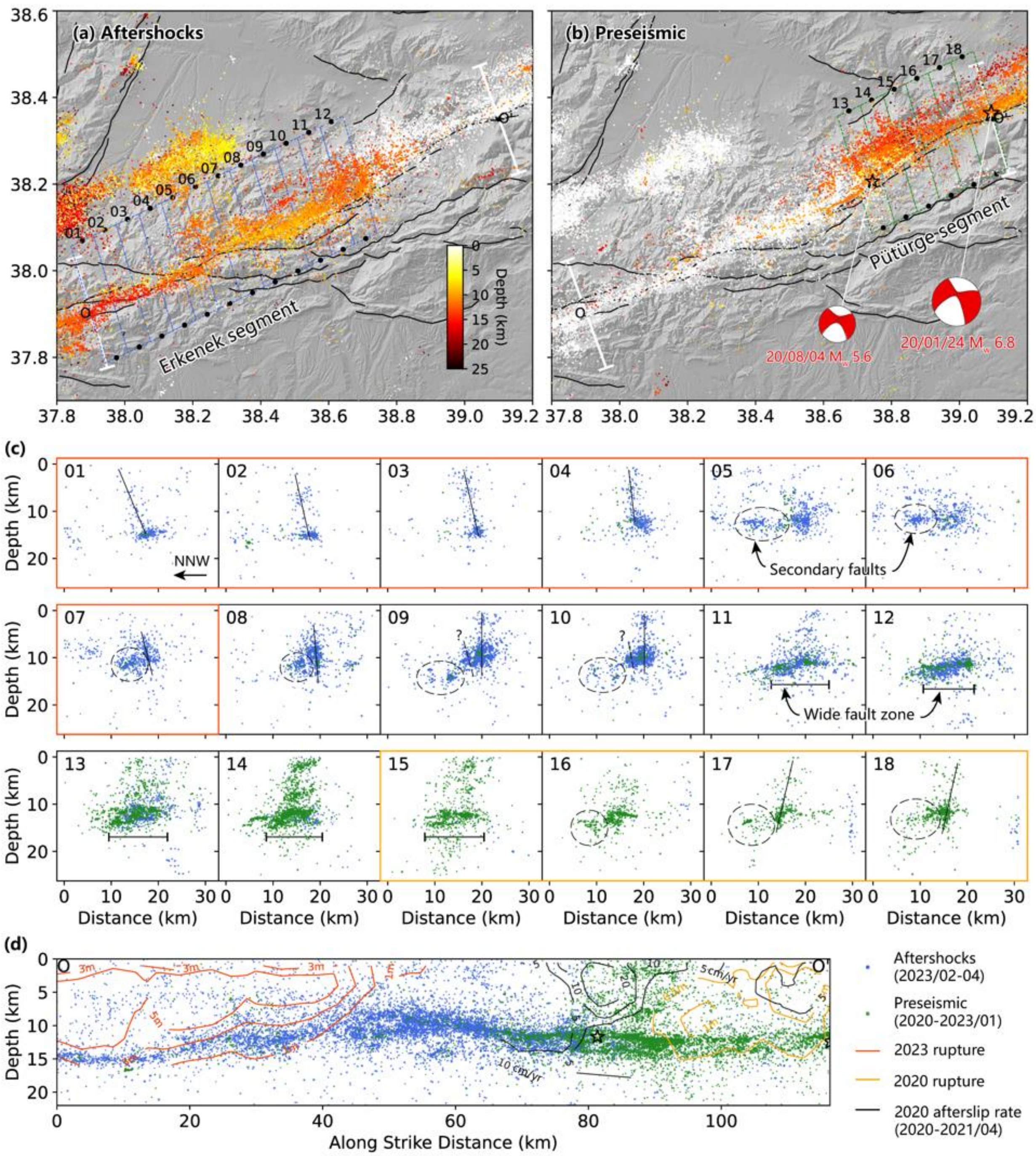
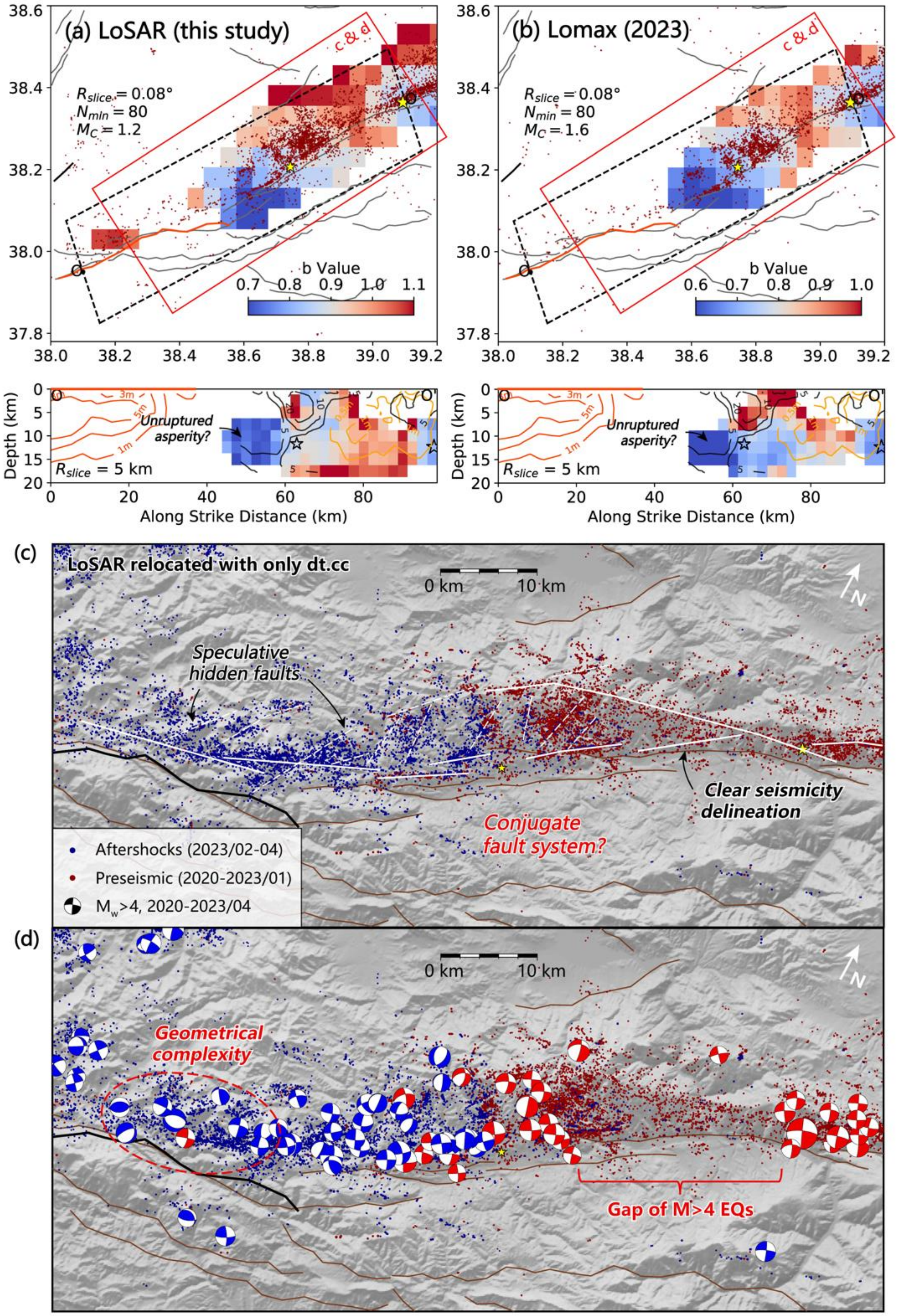


Figure 9. Seismicity and fault geometry interpretation for the Erkenek-Pütürge fault segment (EPF) of EAFZ. (a-b) plot the map-view distribution of postseismic (2023/02-04) and preseismic period (2020-2023/01) seismicity. The events in current panel / time period are color coded by the depth, and the seismicity in the other time period is plotted as white dots in the bottom. (c) plot the fault-normal cross-sections. The 2023 aftershocks and preseismic events are denoted in blue and green, respectively. The interpreted fault dip is marked as black lines. (d) plot the along-strike cross-section. The coseismic slip of

the 2023 M_w 7.8 obtained by Ren et al. (2024) is plotted in orange-red contours, the coseismic slip and afterslip slip of the 2020 M_w 6.8 obtained by Cakir et al. (2023) are plotted as orange and black contours, respectively.

We further examine the b-value distribution before the 2023 mainshocks (Figure 10a-b). Note that we exclude events occurring within 24-hr following the 2020 M_w 6.8, since it causes a transient incompleteness in the catalog that can bias the b-value estimation (Figure S27). However, we tested that using different time ranges do not alter the relative values in the distribution (Figure S30). The b-value distributions, derived from both the LoSAR catalog and that by Lomax (2023), reveal that the 2020-ruptured area is characterized by a relatively high b-value, which agree with the significant stress drop after the large earthquake. Notably, this segment also exhibits a gap of M>4 earthquakes in our study period (Figure 10d), further supporting a significantly reduced stress level. Another area of high b-value coincide with the area experiencing large afterslip following the 2020 mainshock (Cakir et al., 2023). In the map-view distribution (Figure 10a-b), the secondary faults show relatively higher b-value compared to those observed along the major fault, which suggest a contrast in the strength between major and secondary faults. The area near the NE end of the 2023 ruptures also appear to have a high b-value before the earthquake (Figure 10a). This area also displays highly variable focal mechanisms, indicating complex local structures. However, the reliability of this feature may be compromised due to the significantly fewer events that occurred before the 2023 mainshock. In addition to the high-b areas mentioned above, a markedly low-b area extends over 15 km length along strike, located southwest of the 2020 M_w 5.6 aftershock (Figure 10a-b). This area probably represents an unruptured asperity that has not experienced any M>5.5 earthquakes so far. Observations of postseismic deformation following the 2023 sequence may help to rule out the alternative hypothesis that the stress in this segment has been relieved through aseismic slips.

Back to the question posed in the beginning of this section, the high b-value observed in the northeastern part of the 2023 rupture could potentially decelerate the rupture process; whereas the presence of a significantly low-b area adjacent to it on the northeast starkly contrasts with the observation that the rupture terminated before reaching this region. Furthermore, given that fault stress is more variable than structural features over extended time scales, we posit that the geometrical complexity is the primary factor rendering the EPF a persistent rupture barrier.



(a) LoSAR (this study)
R_{slice} = 0.08°
N_{min} = 80
M_C = 1.2
c & d
b Value
0.7 0.8 0.9 1.0 1.1
(b) Lomax (2023)
R_{slice} = 0.08°
N_{min} = 80
M_C = 1.6
b Value
0.6 0.7 0.8 0.9 1.0
Unruptured asperity?
R_{slice} = 5 km
Depth (km)
Along Strike Distance (km)
(c)
LoSAR relocated with only dt.cc
0 km 10 km
N
Speculative hidden faults
Clear seismicity delineation
Conjugate fault system?
Aftershocks (2023/02-04)
Preseismic (2020-2023/01)
M_w>4, 2020-2023/04
(d)
Geometrical complexity
Gap of M>4 EQs

Figure 10. B-value mapping and detailed fault structure interpretation of the EPF area. (a-b) B-value mapping results with the LoSAR catalog by this study and the relocated AFAD catalog by Lomax (2023) during the preseismic period. The active faults and 2023-ruptured faults are plotted by gray and orange-red lines, respectively. The M>6 earthquakes during 2020-2023 are marked by yellow stars. The contours share the same meaning with Figure 9d. The red box marks the spatial coverage of (c) & (d). (c-d) provide a zoom-in plot for part of the EPF. (c) plot the distribution of CC-relocated LoSAR catalog. The 2023 aftershocks and preseismic events are denoted by blue and red dots, respectively. The white solid and dashed lines show interpreted faults with clear seismicity delineation (high confidential) and those with rather scattered seismicity (speculative), respectively. (d) plot the focal mechanism solutions of M_w>4 events during 2020-2023/04 from AFAD.

5. Conclusions

In this study, we introduce LoSAR, an innovative deep learning-driven workflow for constructing long-term seismic catalogs. It is designed based on the idea of training a deep learning model with local data. By applying LoSAR to two distinct cases in the Ridgecrest-Coso and EAFZ region, we demonstrated that LoSAR realizes a detection ability comparable to MFT, while offering markedly improved temporal stability and computational efficiency. Through direct comparisons with other state-of-the-art phase picking and association algorithms, we conclude that LoSAR has superior scalability and generalizability. These advantages position LoSAR as an ideal tool for building catalogs of a large spatiotemporal scale.

Utilizing the high-resolution long-term catalogs built for the Ridgecrest-Coso and EAFZ region, we delve into their preseismic fault behaviors by jointly analyzing the seismicity distribution and the b-value mapping. We specifically investigate one smaller fault segment of the two regions, and reached the following conclusions:

(1) The Ridgecrest faults exhibited distributed and weak microseismicity prior to the earthquake, revealing multiple preexisting orthogonal faults. The seismicity depth contour mimics that of the aftershocks, and is associated with a low b-value prior to the earthquake. These observations constitute the Ridgecrest faults a persistent asperity of high stress level.

(2) The EPF segment of EAFZ is characterized by complex fault structures, including multiple secondary fault of high b-value, and a conjugate fault system that makes a ~10-km wide fault zone. Despite the overall high b-value, a >15-km-long segment with a notably low b-value

adjacent to the 2023 rupture points to geometrical complexity as the reason EPF acts as a persistent barrier for rupture propagations.

Acknowledgments

We thank Prof. Ziyadin Cakir for sharing the rupture model and afterslip model for the 2020 Elaziğ (Turkey) earthquake. This work is jointly supported by the National Key R&D Program (grant no. 2022YFF0800601) and the University of California, Riverside.

Open Research

Figures in this paper are plotted with GMT and Matplotlib. The continuous seismic data for Ridgecrest-Coso (2008-2019), the waveform-relocated SCSN catalog, the QTM catalog, and the refined focal mechanism solutions by Cheng et al. (2023) are downloaded from the Southern California Seismic Network (SCSN, https://dx.doi.org/10.7909/C3WD3xH1). The continuous data for the East Anatolian Fault Zone (EAFZ, 2020-2023/04) is collected through multiple sources: (1) TU network, from the Disaster and Emergency Management Presidency (AFAD, https://tdvms.afad.gov.tr/continuous_data), downloaded manually, (2) KO network and Kandilli Observatory And Earthquake Research Institute (KOERI), downloaded with Obspy, and (3) GE, CQ, and IM network, all available through Obspy. The relocated AFAD catalog by Lomax (2023) is available at https://doi.org/10.5281/zenodo.8089273. The focal mechanism solutions for $M_w>4$ events (2020-2023) are also available at AFAD. The active faults data of the Ridgecrest-Coso region is available at United States Geological Survey (USGS, https://www.usgs.gov/programs/earthquake-hazards/faults); that for East Anatolian Fault Zone comes from the GEM Global Active Faults Database (Styron and Pagani, 2020), and we adopt a more detailed fault data for the EPF area from the Active Faults of Eurasia Database (AFEAD, Zelenin et al., 2022). The surface rupture data of the 2019 Ridgecrest earthquake comes from Ponti et al. (2020); the fault traces of the 2023 Turkey sequence is available by USGS (https://doi.org/10.5066/P985I7U2, last accessed 16th Mar). The coseismic slip model of the 2019 Ridgecrest earthquake comes from Yue et al. (2021); that for the 2023 Turkey earthquake comes from Ren et al. (2024); the rupture model and afterslip model for the 2020 Elaziğ (Turkey) earthquake comes from Cakir et al. (2023) through personal communication. The referred velocity models for the RC-Coso region include: CVM-S4 velocity model (available at https://strike.scec.org/scecpedia/CVM-S4, last accessed 2024/03), Hutton et al. (2010) for the whole Southern California, and Shelly (2020) for the Ridgecrest source region. The referred velocity models for the EAFZ include that from Güvercin et al. (2022), Acarel et al. (2019), and Ding et al. (2023). The LoSAR workflow is open sourced at Github: https://github.com/YijianZhou/LoSAR/releases/tag/v4.3 (https://doi.org/10.5281/zenodo.10895585).

References

Acarel, D., M. D. Cambaz, F. Turhan, A. K. Mutlu, and R. Polat (2019). Seismotectonics of Malatya Fault, Eastern Turkey, *Open Geosciences*, *11*(1), 1098-1111, doi:doi:10.1515/geo-2019-0085.

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

Aktug, B., H. Ozener, A. Dogru, A. Sabuncu, B. Turgut, K. Halicioglu, O. Yilmaz, and E. Havazli (2016). Slip rates and seismic potential on the East Anatolian Fault System using an improved GPS velocity field, *Journal of Geodynamics*, *94-95*, 1-12, doi:https://doi.org/10.1016/j.jog.2016.01.001.

Amorèse, D., J.-R. Grasso, and P. A. Rydelek (2010). On varying b-values with depth: results from computer-intensive tests for Southern California, *Geophysical Journal International*, *180*(1), 347-360, doi:10.1111/j.1365-246X.2009.04414.x.

Bahdanau, D., K. Cho, and Y. Bengio (2014). Neural machine translation by jointly learning to align and translate, *arXiv preprint arXiv:1409.0473*.

Barbot, S., and J. R. Weiss (2021). Connecting subduction, extension and shear localization across the Aegean Sea and Anatolia, *Geophysical Journal International*, *226*(1), 422-445, doi:10.1093/gji/ggab078.

Ben-Zion, Y., and C. G. Sammis (2003). Characterization of Fault Zones, *pure and applied geophysics*, *160*(3), 677-715, doi:10.1007/PL00012554.

Bletery, Q., O. Cavalié, J.-M. Nocquet, and T. Ragon (2020). Distribution of Interseismic Coupling Along the North and East Anatolian Faults Inferred From InSAR and GPS Data, *Geophysical Research Letters*, *47*(16), e2020GL087775, doi:https://doi.org/10.1029/2020GL087775.

Bornstein, T., D. Lange, J. Münchmeyer, J. Woollam, A. Rietbrock, G. Barcheck, I. Grevemeyer, and F. Tilmann (2024). PickBlue: Seismic Phase Picking for Ocean Bottom Seismometers With Deep Learning, *Earth and Space Science*, *11*(1), e2023EA003332, doi:https://doi.org/10.1029/2023EA003332.

Bulut, F., M. Bohnhoff, T. Eken, C. Janssen, T. Kılıç, and G. Dresen (2012). The East Anatolian Fault Zone: Seismotectonic setting and spatiotemporal characteristics of seismicity based on precise earthquake locations, *Journal of Geophysical Research: Solid Earth*, *117*(B7), doi:https://doi.org/10.1029/2011JB008966.

Bürgmann, R. (2018). The geophysics, geology and mechanics of slow fault slip, *Earth and Planetary Science Letters*, *495*, 112-134, doi:https://doi.org/10.1016/j.epsl.2018.04.062.

Cakir, Z., et al. (2023). Arrest of the Mw 6.8 January 24, 2020 Elaziğ (Turkey) earthquake by shallow fault creep, *Earth and Planetary Science Letters*, *608*, 118085, doi:https://doi.org/10.1016/j.epsl.2023.118085.

Cao, A., and S. S. Gao (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc, *Geophysical Research Letters*, *29*(9), 48-41-48-43, doi:10.1029/2001gl013775.

Cavalié, O., and S. Jónsson (2014). Block-like plate movements in eastern Anatolia observed by InSAR, *Geophysical Research Letters*, *41*(1), 26-31, doi:https://doi.org/10.1002/2013GL058170.

Chai, C., M. Maceira, H. J. Santos-Villalobos, S. V. Venkatakrishnan, M. Schoenball, W. Zhu, G. C. Beroza, C. Thurber, and E. C. Team (2020). Using a Deep Neural Network and Transfer

Learning to Bridge Scales for Seismic Phase Picking, *Geophysical Research Letters*, *47*(16), e2020GL088651, doi:https://doi.org/10.1029/2020GL088651.

Chamberlain, C. J., W. B. Frank, F. Lanza, J. Townend, and E. Warren-Smith (2021). Illuminating the Pre-, Co-, and Post-Seismic Phases of the 2016 M7.8 Kaikōura Earthquake With 10 Years of Seismicity, *Journal of Geophysical Research: Solid Earth*, *126*(8), e2021JB022304, doi:https://doi.org/10.1029/2021JB022304.

Chen, K., J.-P. Avouac, S. Aati, C. Milliner, F. Zheng, and C. Shi (2020). Cascading and pulse-like ruptures during the 2019 Ridgecrest earthquakes in the Eastern California Shear Zone, *Nature Communications*, *11*(1), 22, doi:10.1038/s41467-019-13750-w.

Cheng, Y., E. Hauksson, and Y. Ben-Zion (2023). Refined Earthquake Focal Mechanism Catalog for Southern California Derived With Deep Learning Algorithms, *Journal of Geophysical Research: Solid Earth*, *128*(2), e2022JB025975, doi:https://doi.org/10.1029/2022JB025975.

Cho, K., B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, *arXiv e-prints*, arXiv:1406.1078.

Dal Zilio, L., and J.-P. Ampuero (2023). Earthquake doublet in Turkey and Syria, *Communications Earth & Environment*, *4*(1), 71, doi:10.1038/s43247-023-00747-z.

Ding, H., Y. Zhou, Z. Ge, T. Taymaz, A. Ghosh, H. Xu, T. S. Irmak, and X. Song (2023). High-resolution seismicity imaging and early aftershock migration of the 2023 Kahramanmaraş (SE Türkiye) <i>M</i>_W7.9 & 7.8 earthquake doublet, *Earthquake Science*, *36*, 1-16, doi:10.1016/j.eqs.2023.06.002.

Ezen-Can, A. (2020). A Comparison of LSTM and BERT for Small Corpus, *arXiv preprint arXiv:2009.05451*.

Feng, T., S. Mohanna, and L. Meng (2022). EdgePhase: A Deep Learning Model for Multi-Station Seismic Phase Picking, *Geochemistry, Geophysics, Geosystems*, *23*(11), e2022GC010453, doi:https://doi.org/10.1029/2022GC010453.

Gallovič, F., J. Zahradník, V. Plicka, E. Sokos, C. Evangelidis, I. Fountoulakis, and F. Turhan (2020). Complex rupture dynamics on an immature fault during the 2020 Mw 6.8 Elazığ earthquake, Turkey, *Communications Earth & Environment*, *1*(1), 40, doi:10.1038/s43247-020-00038-x.

Ganev, P. N., J. F. Dolan, S. F. McGill, and K. L. Frankel (2012). Constancy of geologic slip rate along the central Garlock fault: implications for strain accumulation and release in southern California, *Geophysical Journal International*, *190*(2), 745-760, doi:10.1111/j.1365-246X.2012.05494.x.

Goldberg, D. E., D. Melgar, V. J. Sahakian, A. M. Thomas, X. Xu, B. W. Crowell, and J. Geng (2020). Complex Rupture of an Immature Fault Zone: A Simultaneous Kinematic Model of the 2019 Ridgecrest, CA Earthquakes, *Geophysical Research Letters*, *47*(3), e2019GL086382, doi:https://doi.org/10.1029/2019GL086382.

Gong, J., and W. Fan (2022). Seismicity, Fault Architecture, and Slip Mode of the Westernmost Gofar Transform Fault, *Journal of Geophysical Research: Solid Earth*, *127*(11), e2022JB024918, doi:https://doi.org/10.1029/2022JB024918.

Graves, A., and N. Jaitly (2014). Towards End-To-End Speech Recognition with Recurrent Neural Networks, in *Proceedings of the 31st International Conference on Machine Learning*, edited by P. X. Eric and J. Tony, pp. 1764--1772, PMLR, Proceedings of Machine Learning Research.

Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California, *Bulletin of the Seismological Society of America*, *34*(4), 185-188.

Güvercin, S. E., H. Karabulut, A. Ö. Konca, U. Doğan, and S. Ergintav (2022). Active seismotectonics of the East Anatolian Fault, *Geophysical Journal International*, *230*(1), 50-69, doi:10.1093/gji/ggac045.

Güvercin, S. E. (2024). 2023 Earthquake Doublet in Türkiye Reveals the Complexities of the East Anatolian Fault Zone: Insights from Aftershock Patterns and Moment Tensor Solutions, *Seismological Research Letters*, doi:10.1785/0220230317.

Hannun, A., C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, and A. Coates (2014). Deep speech: Scaling up end-to-end speech recognition, *arXiv preprint arXiv:1412.5567*.

Hatem, A. E., and J. F. Dolan (2018). A Model for the Initiation, Evolution, and Controls on Seismic Behavior of the Garlock Fault, California, *Geochemistry, Geophysics, Geosystems*, *19*(7), 2166-2178, doi:https://doi.org/10.1029/2017GC007349.

Hauksson, E., W. Yang, and P. M. Shearer (2012). Waveform Relocated Earthquake Catalog for Southern California (1981 to June 2011), *Bulletin of the Seismological Society of America*, *102*(5), 2239-2244, doi:10.1785/0120120010.

Hauksson, E., and L. M. Jones (2020). Seismicity, Stress State, and Style of Faulting of the Ridgecrest ‐ Coso Region from the 1930s to 2019: Seismotectonics of an Evolving Plate Boundary Segment, *Bulletin of the Seismological Society of America*, *110*(4), 1457-1473, doi:10.1785/0120200051.

Herrmann, M., and W. Marzocchi (2020). Inconsistencies and Lurking Pitfalls in the Magnitude–Frequency Distribution of High ‐ Resolution Earthquake Catalogs, *Seismological Research Letters*, *92*(2A), 909-922, doi:10.1785/0220200337.

Hsu, Y.-F., I. Zaliapin, and Y. Ben-Zion (2024). Informative Modes of Seismicity in Nearest-Neighbor Earthquake Proximities, *Journal of Geophysical Research: Solid Earth*, *129*(3), e2023JB027826, doi:https://doi.org/10.1029/2023JB027826.

Hubert-Ferrari, A., L. Lamair, S. Hage, S. Schmidt, M. N. Çağatay, and U. Avşar (2020). A 3800 yr paleoseismic record (Lake Hazar sediments, eastern Turkey): Implications for the East Anatolian Fault seismic cycle, *Earth and Planetary Science Letters*, *538*, 116152, doi:https://doi.org/10.1016/j.epsl.2020.116152.

Hussain, E., S. Kalaycıoğlu, C. W. D. Milliner, and Z. Çakir (2023). Preconditioning the 2023 Kahramanmaraş (Türkiye) earthquake disaster, *Nature Reviews Earth & Environment*, doi:10.1038/s43017-023-00411-2.

Hutton, K., J. Woessner, and E. Hauksson (2010). Earthquake Monitoring in Southern California for Seventy-Seven Years (1932–2008), *Bulletin of the Seismological Society of America*, *100*(2), 423-446, doi:10.1785/0120090130.

Im, K., J.-P. Avouac, E. R. Heimisson, and D. Elsworth (2021). Ridgecrest aftershocks at Coso suppressed by thermal destressing, *Nature*, *595*(7865), 70-74, doi:10.1038/s41586-021-03601-4.

Jia, Z., et al. (2023). The complex dynamics of the 2023 Kahramanmaraş, Turkey, <i>M</i>_w 7.8-7.7 earthquake doublet, *Science*, *381*(6661), 985-990, doi:doi:10.1126/science.adi0685.

Jiang, C., L. Fang, L. Fan, and B. Li (2021). Comparison of the earthquake detection abilities of PhaseNet and EQTransformer with the Yangbi and Maduo earthquakes, *Earthquake Science*, *34*(5), 425-435, doi:10.29382/eqs-2021-0038.

Jiang, J., and N. Lapusta (2016). Deeper penetration of large earthquakes on seismically quiescent faults, *Science*, *352*(6291), 1293-1297, doi:doi:10.1126/science.aaf1496.

Karabacak, V., et al. (2023). The 2023 Pazarcık (Kahramanmaraş, Türkiye) earthquake (Mw 7.7): implications for surface rupture dynamics along the East Anatolian Fault Zone, *Journal of the Geological Society*, *180*(3), doi:10.1144/jgs2023-020.

Kingma, D. P., and J. Ba (2014). Adam: A Method for Stochastic Optimization, *arXiv e-prints*, arXiv:1412.6980.

Klein, F. W. (2002). User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes*Rep. 2331-1258*, US Geological Survey.

Lay, T., and S. P. Nishenko (2022). Updated concepts of seismic gaps and asperities to assess great earthquake hazard along South America, *Proceedings of the National Academy of Sciences*, *119*(51), e2216843119, doi:doi:10.1073/pnas.2216843119.

Lin, X., J. Hao, D. Wang, R. Chu, X. Zeng, J. Xie, B. Zhang, and Q. Bai (2020). Coseismic Slip Distribution of the 24 January 2020 Mw 6.7 Doganyol Earthquake and in Relation to the Foreshock and Aftershock Activities, *Seismological Research Letters*, *92*(1), 127-139, doi:10.1785/0220200152.

Lomax, A., and A. Savvaidis (2022). High-Precision Earthquake Location Using Source-Specific Station Terms and Inter-Event Waveform Similarity, *Journal of Geophysical Research: Solid Earth*, *127*(1), e2021JB023190, doi:https://doi.org/10.1029/2021JB023190.

Lu, W., Y. Zhou, Z. Zhao, H. Yue, and S. Zhou (2021). Aftershock sequence of the 2017 Mw 6.5 Jiuzhaigou, China earthquake monitored by an AsA network and its implication to fault structures and strength, *Geophysical Journal International*, *228*(3), 1763-1779, doi:10.1093/gji/ggab443.

Marzocchi, W., I. Spassiani, A. Stallone, and M. Taroni (2019). How to be fooled searching for significant variations of the b-value, *Geophysical Journal International*, *220*(3), 1845-1856, doi:10.1093/gji/ggz541.

McClusky, S., et al. (2000). Global Positioning System constraints on plate kinematics and dynamics in the eastern Mediterranean and Caucasus, *Journal of Geophysical Research: Solid Earth*, *105*(B3), 5695-5719, doi:https://doi.org/10.1029/1999JB900351.

McGill, S., and K. Sieh (1993). Holocene slip rate of the Central Garlock Fault in southeastern Searles Valley, California, *Journal of Geophysical Research: Solid Earth*, *98*(B8), 14217-14231, doi:https://doi.org/10.1029/93JB00442.

Melgar, D., A. Ganas, T. Taymaz, S. Valkaniotis, B. W. Crowell, V. Kapetanidis, V. Tsironi, S. Yolsal-Çevikbilen, and T. Öcalan (2020). Rupture kinematics of 2020 January 24 Mw 6.7 Doğanyol-Sivrice, Turkey earthquake on the East Anatolian Fault Zone imaged by space geodesy, *Geophysical Journal International*, *223*(2), 862-874, doi:10.1093/gji/ggaa345.

Mousavi, S. M., W. Zhu, Y. Sheng, and G. C. Beroza (2019). CRED: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection, *Scientific Reports*, *9*(1), 10267, doi:10.1038/s41598-019-45748-1.

Mousavi, S. M., W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking, *Nature Communications*, *11*(1), 3952, doi:10.1038/s41467-020-17591-w.

Münchmeyer, J., et al. (2022). Which Picker Fits My Data? A Quantitative Evaluation of Deep Learning Based Seismic Pickers, *Journal of Geophysical Research: Solid Earth*, *127*(1), e2021JB023499, doi:https://doi.org/10.1029/2021JB023499.

Nanjo, K. Z. (2020). Were changes in stress state responsible for the 2019 Ridgecrest, California, earthquakes?, *Nature Communications*, *11*(1), 3082, doi:10.1038/s41467-020-16867-5.

Neves, M., Z. Peng, and G. Lin (2022). A High‐Resolution Earthquake Catalog for the 2004 Mw 6 Parkfield Earthquake Sequence Using a Matched Filter Technique, *Seismological Research Letters*, *94*(1), 507-521, doi:10.1785/0220220206.

Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes, *Journal of the American Statistical Association*, *83*(401), 9-27, doi:10.1080/01621459.1988.10478560.

Oskin, M., L. Perg, E. Shelef, M. Strane, E. Gurney, B. Singer, and X. Zhang (2008). Elevated shear zone loading rate during an earthquake cluster in eastern California, *Geology*, *36*(6), 507-510, doi:10.1130/g24814a.1.

Park, Y., G. C. Beroza, and W. L. Ellsworth (2023). A Mitigation Strategy for the Prediction Inconsistency of Neural Phase Pickers, *Seismological Research Letters*, doi:10.1785/0220230003.

Peng, B., E. Alcaide, Q. Anthony, A. Albalak, S. Arcadinho, H. Cao, X. Cheng, M. Chung, M. Grella, and K. K. GV (2023). Rwkv: Reinventing rnns for the transformer era, *arXiv preprint arXiv:2305.13048*.

Ponti, D. J., et al. (2020). Documentation of Surface Fault Rupture and Ground‐Deformation Features Produced by the 4 and 5 July 2019 Mw 6.4 and Mw 7.1 Ridgecrest Earthquake Sequence, *Seismological Research Letters*, *91*(5), 2942-2959, doi:10.1785/0220190322.

Pousse-Beltran, L., E. Nissen, E. A. Bergman, M. D. Cambaz, É. Gaudreau, E. Karasözen, and F. Tan (2020). The 2020 Mw 6.8 Elazığ (Turkey) Earthquake Reveals Rupture Behavior of the East Anatolian Fault, *Geophysical Research Letters*, *47*(13), e2020GL088136, doi:https://doi.org/10.1029/2020GL088136.

Ren, C., et al. (2024). Supershear triggering and cascading fault ruptures of the 2023 Kahramanmaraş, Türkiye, earthquake doublet, *Science*, *383*(6680), 305-311, doi:doi:10.1126/science.adi1519.

Ross, Z. E., et al. (2019a). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence, *Science*, *366*(6463), 346-351, doi:10.1126/science.aaz0109.

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

Ross, Z. E., E. S. Cochran, D. T. Trugman, and J. D. Smith (2020). 3D fault architecture controls the dynamism of earthquake swarms, *Science*, *368*(6497), 1357-1361, doi:10.1126/science.abb0779.

Schoenball, M., N. C. Davatzes, and J. M. G. Glen (2015). Differentiating induced and natural seismicity using space-time-magnitude statistics applied to the Coso Geothermal field, *Geophysical Research Letters*, *42*(15), 6221-6228, doi:https://doi.org/10.1002/2015GL064772.

Scholz, C. H. (1988). The brittle-plastic transition and the depth of seismic faulting, *Geologische Rundschau*, *77*(1), 319-328, doi:10.1007/BF01848693.

Scholz, C. H. (2015). On the stress dependence of the earthquake b value, *Geophysical Research Letters*, *42*(5), 1399-1402, doi:10.1002/2014gl062863.

Scholz, C. H. (2019). *The mechanics of earthquakes and faulting*, Cambridge university press, doi:https://doi.org/10.1017/9781316681473.

Schurr, B., M. Moreno, A. M. Tréhu, J. Bedford, J. Kummerow, S. Li, and O. Oncken (2020). Forming a Mogi Doughnut in the Years Prior to and Immediately Before the 2014 M8.1 Iquique, Northern Chile, Earthquake, *Geophysical Research Letters*, *47*(16), e2020GL088351, doi:https://doi.org/10.1029/2020GL088351.

Shelly, D. R. (2020). A High‐Resolution Seismic Catalog for the Initial 2019 Ridgecrest Earthquake Sequence: Foreshocks, Aftershocks, and Faulting Complexity, *Seismological Research Letters*, *91*(4), 1971-1978, doi:10.1785/0220190309.

Shelly, D. R., D. E. Goldberg, K. Z. Materna, R. J. Skoumal, J. L. Hardebeck, C. E. Yoon, W. L. Yeck, and P. S. Earle (2024). Subduction intraslab-interface fault interactions in the 2022 <i>M</i>_w 6.4 Ferndale, California, earthquake sequence, *Science Advances*, *10*(10), eadl1226, doi:doi:10.1126/sciadv.adl1226.

Shi, Y., and B. A. Bolt (1982). The standard error of the magnitude-frequency b value, *Bulletin of the Seismological Society of America*, *72*(5), 1677-1687.

Spinler, J. C., R. A. Bennett, M. L. Anderson, S. F. McGill, S. Hreinsdóttir, and A. McCallister (2010). Present-day strain accumulation and slip rates associated with southern San Andreas and eastern California shear zone faults, *Journal of Geophysical Research: Solid Earth*, *115*(B11), doi:https://doi.org/10.1029/2010JB007424.

Styron, R., and M. Pagani (2020). The GEM Global Active Faults Database, *Earthquake Spectra*, *36*(1_suppl), 160-180, doi:10.1177/8755293020944182.

Sugan, M., S. Campanella, L. Chiaraluce, M. Michele, and A. Vuan (2023). The unlocking process leading to the 2016 Central Italy seismic sequence, *Geophysical Research Letters*, *n/a*(n/a), e2022GL101838, doi:https://doi.org/10.1029/2022GL101838.

Sun, H., Z. E. Ross, W. Zhu, and K. Azizzadenesheli (2023). Phase Neural Operator for Multi-Station Picking of Seismic Arrivals, *Geophysical Research Letters*, *50*(24), e2023GL106434, doi:https://doi.org/10.1029/2023GL106434.

Sutskever, I., O. Vinyals, and Q. V. Le (2014). Sequence to sequence learning with neural networks, *Advances in neural information processing systems*, *27*.

Sykes, L. R. (2021). Decadal seismicity prior to great earthquakes at subduction zones: Roles of major asperities and low-coupling zones, *International Journal of Geosciences*, *12*(9), 845-926, doi:https://doi.org/10.4236/ijg.2021.129046.

Tan, Y. J., F. Waldhauser, W. L. Ellsworth, M. Zhang, W. Zhu, M. Michele, L. Chiaraluce, G. C. Beroza, and M. Segou (2021). Machine‐Learning‐Based High‐Resolution Earthquake Catalog Reveals How Complex Fault Structures Were Activated during the 2016–2017 Central Italy Sequence, *The Seismic Record*, *1*(1), 11-19, doi:10.1785/0320210001.

Thompson Jobe, J. A., et al. (2020). Evidence of Previous Faulting along the 2019 Ridgecrest, California, Earthquake Ruptures, *Bulletin of the Seismological Society of America*, *110*(4), 1427-1456, doi:10.1785/0120200041.

Toda, S., and R. S. Stein (2022). Central shutdown and surrounding activation of aftershocks from megathrust earthquake stress transfer, *Nature Geoscience*, *15*(6), 494-500, doi:10.1038/s41561-022-00954-x.

Trugman, D. T., P. M. Shearer, A. A. Borsa, and Y. Fialko (2016). A comparison of long-term changes in seismicity at The Geysers, Salton Sea, and Coso geothermal fields, *Journal of Geophysical Research: Solid Earth*, *121*(1), 225-247, doi:https://doi.org/10.1002/2015JB012510.

Trugman, D. T., and P. M. Shearer (2017). GrowClust: A Hierarchical Clustering Algorithm for Relative Earthquake Relocation, with Application to the Spanish Springs and Sheldon, Nevada, Earthquake Sequences, *Seismological Research Letters*, *88*(2A), 379-391, doi:10.1785/0220160188.

Tymofyeyeva, E., and Y. Fialko (2015). Mitigation of atmospheric phase delays in InSAR data, with application to the eastern California shear zone, *Journal of Geophysical Research: Solid Earth*, *120*(8), 5952-5963, doi:https://doi.org/10.1002/2015JB011886.

Uchida, N., and R. Bürgmann (2021). A Decade of Lessons Learned from the 2011 Tohoku-Oki Earthquake, *Reviews of Geophysics*, *59*(2), e2020RG000713, doi:https://doi.org/10.1029/2020RG000713.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin (2017). Attention is all you need, *Advances in neural information processing systems*, *30*.

Waldhauser, F., and W. L. Ellsworth (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, doi:10.1785/0120000006.

Waldhauser, F. (2001). hypoDD--A program to compute double-difference hypocenter locations, *U.S. Geological Survey Open-File Report 01-113*, 25 pp.

Waldhauser, F., and D. P. Schaff (2008). Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, *Journal of Geophysical Research: Solid Earth*, *113*(B8), doi:10.1029/2007jb005479.

Walters, R. J., B. Parsons, and T. J. Wright (2014). Constraining crustal velocity fields with InSAR for Eastern Turkey: Limits to the block-like behavior of Eastern Anatolia, *Journal of Geophysical Research: Solid Earth*, *119*(6), 5215-5234, doi:https://doi.org/10.1002/2013JB010909.

Weiss, J. R., et al. (2020). High-Resolution Surface Velocities and Strain for Anatolia From Sentinel-1 InSAR and GNSS Data, *Geophysical Research Letters*, *47*(17), e2020GL087376, doi:https://doi.org/10.1029/2020GL087376.

Wetzler, N., T. Lay, E. E. Brodsky, and H. Kanamori (2018). Systematic deficiency of aftershocks in areas of high coseismic slip for large subduction zone earthquakes, *Science Advances*, *4*(2), eaao3225, doi:10.1126/sciadv.aao3225.

Wiemer, S., and M. Wyss (2000). Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan, *Bulletin of the Seismological Society of America*, *90*(4), 859-869, doi:10.1785/0119990114.

Woessner, J., and S. Wiemer (2005). Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty, *Bulletin of the Seismological Society of America*, *95*(2), 684-698, doi:10.1785/0120040007.

Xu, L., S. Mohanna, L. Meng, C. Ji, J.-P. Ampuero, Z. Yunjun, M. Hasnain, R. Chu, and C. Liang (2023). The overall-subshear and multi-segment rupture of the 2023 Mw7.8 Kahramanmaraş, Turkey earthquake in millennia supercycle, *Communications Earth & Environment*, *4*(1), 379, doi:10.1038/s43247-023-01030-x.

Xu, X., D. T. Sandwell, L. A. Ward, C. W. D. Milliner, B. R. Smith-Konter, P. Fang, and Y. Bock (2020). Surface deformation associated with fractures near the 2019 Ridgecrest earthquake sequence, *Science*, *370*(6516), 605-608, doi:10.1126/science.abd1690.

Yoon, C. E., and D. R. Shelly (2024). Distinct Yet Adjacent Earthquake Sequences near the Mendocino Triple Junction: 20 December 2021 Mw 6.1 and 6.0 Petrolia, and 20 December 2022 Mw 6.4 Ferndale, *The Seismic Record*, *4*(1), 81-92, doi:10.1785/0320230053.

Yu, Z., and W. Wang (2022). LPPN: A Lightweight Network for Fast Phase Picking, *Seismological Research Letters*, *93*(5), 2834-2846, doi:10.1785/0220210309.

Yue, H., J. Sun, M. Wang, Z. Shen, M. Li, L. Xue, W. Lu, Y. Zhou, C. Ren, and T. Lay (2021). The 2019 Ridgecrest, California earthquake sequence: Evolution of seismic and aseismic slip on an orthogonal fault system, *Earth and Planetary Science Letters*, *570*, 117066, doi:https://doi.org/10.1016/j.epsl.2021.117066.

Zaliapin, I., and Y. Ben-Zion (2020). Earthquake Declustering Using the Nearest-Neighbor Approach in Space-Time-Magnitude Domain, *Journal of Geophysical Research: Solid Earth*, *125*(4), e2018JB017120, doi:https://doi.org/10.1029/2018JB017120.

Zelenin, E., D. Bachmanov, S. Garipova, V. Trifonov, and A. Kozhurin (2022). The Active Faults of Eurasia Database (AFEAD): the ontology and design behind the continental-scale dataset, *Earth Syst. Sci. Data*, *14*(10), 4489-4503, doi:10.5194/essd-14-4489-2022.

Zhang, J., Z. Li, and J. Zhang (2023a). Simultaneous Seismic Phase Picking and Polarity Determination with an Attention‐Based Neural Network, *Seismological Research Letters*, *94*(2A), 813-828, doi:10.1785/0220220247.

Zhang, M., M. Liu, T. Feng, R. Wang, and W. Zhu (2022). LOC‐FLOW: An End‐to‐End Machine Learning‐Based High‐Precision Earthquake Location Workflow, *Seismological Research Letters*, *93*(5), 2426-2438, doi:10.1785/0220220019.

Zhang, Y., X. Tang, D. Liu, T. Taymaz, T. Eken, R. Guo, Y. Zheng, J. Wang, and H. Sun (2023b). Geometric controls on cascading rupture of the 2023 Kahramanmaraş earthquake doublet, *Nature Geoscience*, *16*(11), 1054-1060, doi:10.1038/s41561-023-01283-3.

Zhou, Y., S. Zhou, and J. Zhuang (2018). A test on methods for MC estimation based on earthquake catalog, *Earth and Planetary Physics*, *2*(2), 150-162, doi:10.26464/epp2018015.

Zhou, Y., H. Yue, Q. Kong, and S. Zhou (2019). Hybrid Event Detection and Phase‐Picking Algorithm Using Convolutional and Recurrent Neural Networks, *Seismological Research Letters*, *90*(3), 1079-1087, doi:10.1785/0220180319.

Zhou, Y., A. Ghosh, L. Fang, H. Yue, S. Zhou, and Y. Su (2021a). A high-resolution seismic catalog for the 2021 MS6.4/MW6.1 Yangbi earthquake sequence, Yunnan, China: Application of AI picker and matched filter, *Earthquake Science*, *34*(5), 390-398, doi:https://doi.org/10.29382/eqs-2021-0031.

Zhou, Y., H. Yue, L. Fang, S. Zhou, L. Zhao, and A. Ghosh (2021b). An Earthquake Detection and Location Architecture for Continuous Seismograms: Phase Picking, Association, Location, and Matched Filter (PALM), *Seismological Research Letters*, *93*(1), 413-425, doi:10.1785/0220210111.

Zhou, Y., C. Ren, A. Ghosh, H. Meng, L. Fang, H. Yue, S. Zhou, and Y. Su (2022a). Seismological Characterization of the 2021 Yangbi Foreshock-Mainshock Sequence, Yunnan, China: More than a Triggered Cascade, *Journal of Geophysical Research: Solid Earth*, *127*(8), e2022JB024534, doi:https://doi.org/10.1029/2022JB024534.

Zhou, Y., H. Yue, S. Zhou, L. Fang, Y. Zhou, L. Xu, Z. Liu, T. Wang, L. Zhao, and A. Ghosh (2022b). Microseismicity along Xiaojiang Fault Zone (Southeastern Tibetan Plateau) and the characterization of interseismic fault behavior, *Tectonophysics*, *833*, 229364, doi:https://doi.org/10.1016/j.tecto.2022.229364.

Zhu, J., Z. Li, and L. Fang (2022a). USTC-Pickers: a Unified Set of seismic phase pickers Transfer learned for China, *Earthquake Science*, *36*, 1-11.

Zhu, W., and G. C. Beroza (2018). PhaseNet: a deep-neural-network-based seismic arrival-time picking method, *Geophysical Journal International*, *216*(1), 261-273, doi:10.1093/gji/ggy423.

Zhu, W., A. B. Hou, R. Yang, A. Datta, S. M. Mousavi, W. L. Ellsworth, and G. C. Beroza (2022b). QuakeFlow: a scalable machine-learning-based earthquake monitoring workflow with cloud computing, *Geophysical Journal International*, *232*(1), 684-693, doi:10.1093/gji/ggac355.

Zhu, W., I. W. McBrearty, S. M. Mousavi, W. L. Ellsworth, and G. C. Beroza (2022c). Earthquake Phase Association Using a Bayesian Gaussian Mixture Model, *Journal of Geophysical Research: Solid Earth*, *127*(5), e2021JB023249, doi:https://doi.org/10.1029/2021JB023249.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002). Stochastic Declustering of Space-Time Earthquake Occurrences, *Journal of the American Statistical Association*, *97*(458), 369-380.