

JGR Solid Earth

RESEARCH ARTICLE

10.1029/2024JB028735

Key Points:

- We propose a new seismic source detection and location approach based on the source-scanning algorithm and 3D image segmentation
- This approach outperforms human analysts and popular artificial intelligence (AI) and non-AI based methods in characterizing intense aftershock sequences
- The resulting catalog of the 2016 Kaikōura earthquake sequence suggests a continuous fault system surrounded by extensive fracturing

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

F. Tan,
fengzhou.tan@gmail.com

Citation:

Tan, F., Kao, H., Yi, K. M., Nissen, E., Goerzen, C., Hutchinson, J., et al. (2024). Next generation seismic source detection by computer vision: Untangling the complexity of the 2016 Kaikōura earthquake sequence. *Journal of Geophysical Research: Solid Earth*, 129, e2024JB028735. <https://doi.org/10.1029/2024JB028735>

Received 17 JAN 2024

Accepted 24 APR 2024

Author Contributions:

Conceptualization: Fengzhou Tan, Honn Kao, Kwang Moo Yi, Edwin Nissen

Funding acquisition: Fengzhou Tan, Honn Kao, Kwang Moo Yi, Edwin Nissen

Investigation: Chet Goerzen, Jesse Hutchinson, Dawei Gao, Amir M. Farahbod

Methodology: Fengzhou Tan, Honn Kao, Kwang Moo Yi

Software: Fengzhou Tan, Dawei Gao

Supervision: Honn Kao, Kwang Moo Yi, Edwin Nissen

© 2024 The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Next Generation Seismic Source Detection by Computer Vision: Untangling the Complexity of the 2016 Kaikōura Earthquake Sequence

Fengzhou Tan^{1,2} , Honn Kao^{1,2} , Kwang Moo Yi³ , Edwin Nissen¹ , Chet Goerzen^{1,2}, Jesse Hutchinson⁴, Dawei Gao⁵, and Amir M. Farahbod²

¹School of Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada, ²Geological Survey of Canada, Pacific Geoscience Centre, Sidney, BC, Canada, ³Department of Computer Science, University of British Columbia, Vancouver, BC, Canada, ⁴Ocean Networks Canada, University of Victoria, Victoria, BC, Canada, ⁵School of Geosciences and Info-physics, Central South University, Changsha, China

Abstract Seismic source locations are fundamental to many fields of Earth and planetary sciences, such as seismology, volcanology and tectonics. However, seismic source detection and location are challenging when events cluster closely in space and time with signals tangling together at observing stations, such as they often do in major aftershock sequences. Though emerging algorithms and artificial intelligence (AI) models have made processing high volumes of seismic data easier, their performance is still limited, especially for complex aftershock sequences. In this study, we propose a novel approach that utilizes three-dimensional image segmentation—a computer vision technique—to detect and locate seismic sources, and develop this into a complete workflow, Source Untangler Guided by Artificial intelligence image Recognition (SUGAR). In our synthetic and real data tests, SUGAR can handle complex, energetic earthquake sequences in near real time better than skillful analysts and other AI and non-AI based algorithms. We apply SUGAR to the 2016 Kaikōura, New Zealand sequence and obtain five times more events than the analyst-based GeoNet catalog. The improved aftershock distribution illuminates a continuous fault system with extensive fracture zones beneath the segmented, discontinuous surface ruptures. Our method has broader applicability to non-earthquake sources and other time series image data sets.

Plain Language Summary Detecting and locating earthquakes is fundamental to seismology, volcanology, and tectonics. A number of emerging algorithms, including some based upon artificial intelligence (AI), have made processing large volumes of seismic data much easier. However, their performance is still limited, especially in clustered aftershock sequences whose signals overlap at observing seismographs. We propose a new, AI computer vision-based approach to this problem, and develop it into a complete earthquake detection and location workflow, named SUGAR. Tests on synthetic and real earthquake data sets show that SUGAR characterizes complex earthquake sequences better than other AI and non-AI algorithms or professional analysts. We apply SUGAR to the complex aftershock sequence of the 2016 M_w 7.8 Kaikōura, New Zealand earthquake, detecting five times more events than the analyst-based GeoNet catalog. Whereas surface breaks of the Kaikōura earthquake are highly discontinuous, our improved aftershock distribution supports a continuous fault system surrounded by extensive fracture zones at depth. Our method has broader potential for other types of seismic sources and image series.

1. Introduction

Seismic waves are generated from a variety of sources, such as earthquakes, magmatism, landslides, wind, and even traffic (Lay & Wallace, 1995). Detecting and locating these sources provides fundamental insights into the underlying physical processes and underpins a variety of practical applications. For example, accurate locations of small earthquakes are used to map subsurface faulting (e.g., Ross, Idini, et al., 2019) and help illuminate earthquake nucleation (e.g., Ellsworth & Bulut, 2018), triggering (e.g., Kilb et al., 2000), afterslip (e.g., Perfettini & Avouac, 2007) and crustal rheology (e.g., Rolandone et al., 2004). For certain applications, such as operational earthquake forecasting (e.g., DeVries et al., 2018; Giulia & Wiemer, 2019), it may also be important to detect and locate seismicity quickly. However, achieving an earthquake catalog with high precision, completeness and consistency in a timely manner can be challenging.

Validation: Chet Goerzen,
Jesse Hutchinson, Amir M. Farahbod

Visualization: Fengzhou Tan

Writing – original draft: Fengzhou Tan

Writing – review & editing:

Fengzhou Tan, Honn Kao, Kwang Moo Yi,
Edwin Nissen

In sequences clustered densely in time and space, signals from individual events overlap at recording stations, complicating the traditional detection and location procedure of phase picking (e.g., labeling *P* and *S* waves), phase association, and hypocenter inversion (Y. Liao et al., 2012; Hainzl, 2016; Tan et al., 2020; W. Zhu et al., 2022). The overlapping signals make phase association ambiguous, in turn destabilizing the hypocenter inversion and increasing its uncertainties. In such instances, analysts often associate phases by trial and error, grouping different phase subsets—each of which produces a location that satisfies the travel time table—until a preferred solution is obtained. These subjective decisions are tedious and prone to bias and errors (see Section 3.1 “Synthetic test”). Various automatic phase association algorithms have been developed, including those using recurrent neural network (Ross, Yue, et al., 2019), graph neural network (McBrearty & Beroza, 2023; Si et al., 2024), and Bayesian Gaussian mixture model (W. Zhu et al., 2022), and those non-artificial intelligence methods (e.g., M. Zhang et al., 2019; Woollam et al., 2020). Although these algorithms facilitate faster decision-making, their performance in complex aftershock sequences remains a persistently challenging area of investigation.

Another group of automatic workflows, known as waveform-based methods, rarely run into phase association ambiguity but can instead suffer from detection and resolution issues (Kao & Shan, 2004; Tan et al., 2020; W.-Y. Liao et al., 2022). Waveform-based methods locate sources without first identifying individual seismic phases (Li et al., 2020). In waveform stacking methods, preprocessed waveforms at individual stations are summed along a theoretical travel time curve to generate a time series of brightness images over the study area, with the brightest point in this sequence designated the energy source (McBrearty et al., 2019; Tan et al., 2020; W.-Y. Liao et al., 2022; Li et al., 2020; Dokht et al., 2022). However, signals from smaller earthquakes cannot form global brightness maxima in the presence of nearby larger events, limiting the efficacy of this approach in densely clustered sequences. Another approach is to apply neural networks directly to an array of seismic records and output source information (X. Zhang et al., 2020; Shen & Shen, 2021). These methods use real seismograms and manual labels for training, and the trained models are usually suitable for realtime monitoring and early warning systems (X. Zhang et al., 2021; Münchmeyer et al., 2021). However, these methods usually assume a single event within each time window (e.g., 50 s), and resolving complex earthquake sequences is further limited by the size and accuracy of the training data set (X. Zhang et al., 2020; Shen & Shen, 2021). Another strategy is to search for small earthquakes by cross-correlating new waveforms with those from existing template events (Gibbons & Ringdal, 2006; Ross, Trugman, et al., 2019; Shelly, 2020). Template matching can increase the number of events detected, but may miss those with markedly different locations and/or focal mechanisms from the templates (Park et al., 2022; Yoon et al., 2015).

The 2016 Kaikōura, New Zealand earthquake sequence along the Pacific-Australian plate boundary (Figure 1) has proven especially challenging for earthquake detection and location methods due to its unusual complexity (Hamling, 2020; Hamling et al., 2017; Kaiser et al., 2017; Litchfield et al., 2018) and a number of controversial aspects could be clarified by a more detailed aftershock catalog. The 13 November 2016 M_w 7.8 mainshock ruptured at least 17 crustal faults within the Marlborough fault system, many of them previously unmapped or considered inactive, while the well-known, underlying Hikurangi subduction interface may or may not have slipped during the mainshock (see Discussion in Hamling (2020)). Coseismic slip models derived from seismological, geodetic, field and tsunami observations differ markedly in the number of ruptured faults (including contribution of the subduction interface), their detailed geometries and dip angles, and the rupture pathway and dynamics (Bai et al., 2017; Cesca et al., 2017; Clark et al., 2017; Diederichs et al., 2019; Hamling et al., 2017; Hollingsworth et al., 2017; Wang et al., 2018; Xu et al., 2018). Several studies have used aftershocks to help address these questions, but the resolution of the available data is very limited (Cesca et al., 2017; Chamberlain et al., 2021; Hamling et al., 2017; Lanza et al., 2019; Mouslopoulou et al., 2019; Nicol et al., 2018). These studies either directly adopt the New Zealand's routine GeoNet catalog (GNS Science, 1970), or use it as the basis for relocation or template matching. The GeoNet catalog (downloaded on 21 June 2021) has 10,861 events in the Kaikōura region (41.1–43.1°S, 172.1–175.1°E) from 13 November to 31 December 2016, with fewer than 40% of these reviewed by analysts. Moreover, we found upon inspection that even the fully reviewed hours contain numerous uncategorized seismic signals, making the catalog highly incomplete. A more comprehensive aftershock catalog could help delineate the subsurface fault structure and better characterize the mainshock source process and regional seismotectonics.

The main aim of this study is to establish a robust system to delineate extremely complex earthquake sequences where previous methods fail. We develop a novel workflow, Source Untangler Guided by Artificial intelligence

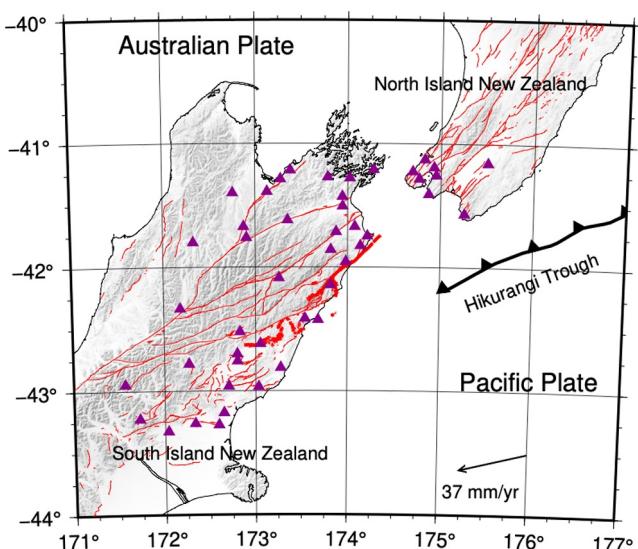


Figure 1. Tectonic setting of the 2016 Kaikōura M_w 7.8 earthquake. Purple triangles are stations used in our study, thick red lines are mainshock surface ruptures (Litchfield et al., 2018), and thin red lines are other active faults without observed surface rupture (GNS Science, 2016). The location of the Hikurangi trough is after Williams et al. (2013). The relative plate motion vector is from the EarthScope Consortium (<https://www.unavco.org/software/geodetic-utilities/plate-motion-calculator/plate-motion-calculator.html>).

represent a seismic event (Figure 2b). In the output score video, we characterize local maxima above a certain threshold as seismic sources, allowing multiple events to be recognized within a single time frame.

Semantic segmentation of 3D image is well studied in computer science and can be handled by the 3D U-Net neural network (Çiçek et al., 2016). The 3D U-Net used in this study follows the structure in Çiçek et al. (2016) but only has two steps in both the contracting and expanding paths (encoder depth = 2) (Figure 2c). The input image is a block of $60 \times 60 \times 60$ pixels (Figure 2a), while the output image is a block of $20 \times 20 \times 20$ pixels, corresponding to the central area of the input block (Figure 2b). Each pixel in east and north represents 4 km and each time frame represents 0.5 s. A margin of 20 pixels in all three dimensions effectively enables the network to use the surrounding information to better predict the output values within the central area. We have a total of 194,559 training samples and finish the training after 20 epochs, which takes ~60 hr on an 8G GeForce RTX 2080 graphic card. Full details of the model and training process can be found Text S1 in Supporting Information S1.

An advantage of this approach is that it can be trained using synthetic data, since though synthetic seismograms may be overly simplistic, brightness images generated from them are indistinguishable from those calculated using real data. This overcomes the difficulties of using real seismograms and catalogs for training, such as limited data volumes and inaccurate phase and hypocenter labels. For example, one may train this network to identify marsquakes (Clinton et al., 2021) or moonquakes (Latham et al., 1971) despite the scarcity of real, labeled waveform data. Moreover, 3D U-Net may be trained to detect diverse seismic sources including volcanic and non-volcanic tremor and landslides, or even applied to other image time series such as Interferometric Synthetic Aperture Radar (InSAR) surface deformation maps.

2.2. The SUGAR Workflow

The neural network as trained above solves the most difficult part, that is, phase association, of the whole earthquake detection and location job. Next, we insert it into a complete workflow for detecting and locating earthquakes. There are four processes in total, namely, brightness video generation, 3D U-Net prediction,

image Recognition (SUGAR), and apply it to the 2016 Kaikōura aftershock sequence. A secondary aim is to resolve some outstanding questions about this sequence. In Section 2, we introduce SUGAR, which differs from other methods by treating earthquake detection and location as a three-dimensional (3D) semantic segmentation problem in computer vision. In Section 3, we test SUGAR with both synthetic and real data and compare its performance against three human analysts, a state-of-the-art deep-learning based algorithm, and a non-AI workflow; in all cases, SUGAR outperforms them by a substantial margin. In Section 4, we apply SUGAR to the Kaikōura earthquake sequence and detail the aftershock distributions. Finally, in Section 5, we further discuss the possible rupture path, off-fault structures, and the limitation and potential of our method.

2. Methodology

We first introduce the idea of using computer vision to detect seismic sources (Section 2.1) and then further develop it into a complete earthquake detection and location workflow (Section 2.2).

2.1. Source Detection Through Computer Vision

We visualize the origin times and epicentral locations of seismic events through brightness videos generated for a study area by migrating and stacking waveforms according to a travel time moveout (Kao & Shan, 2004; Li et al., 2020) (Figure 2a). In this way, we treat the seismic source detection and location task as a 3D image segmentation problem which uses a neural network to label each pixel with a score, standing for the likelihood that its location and time

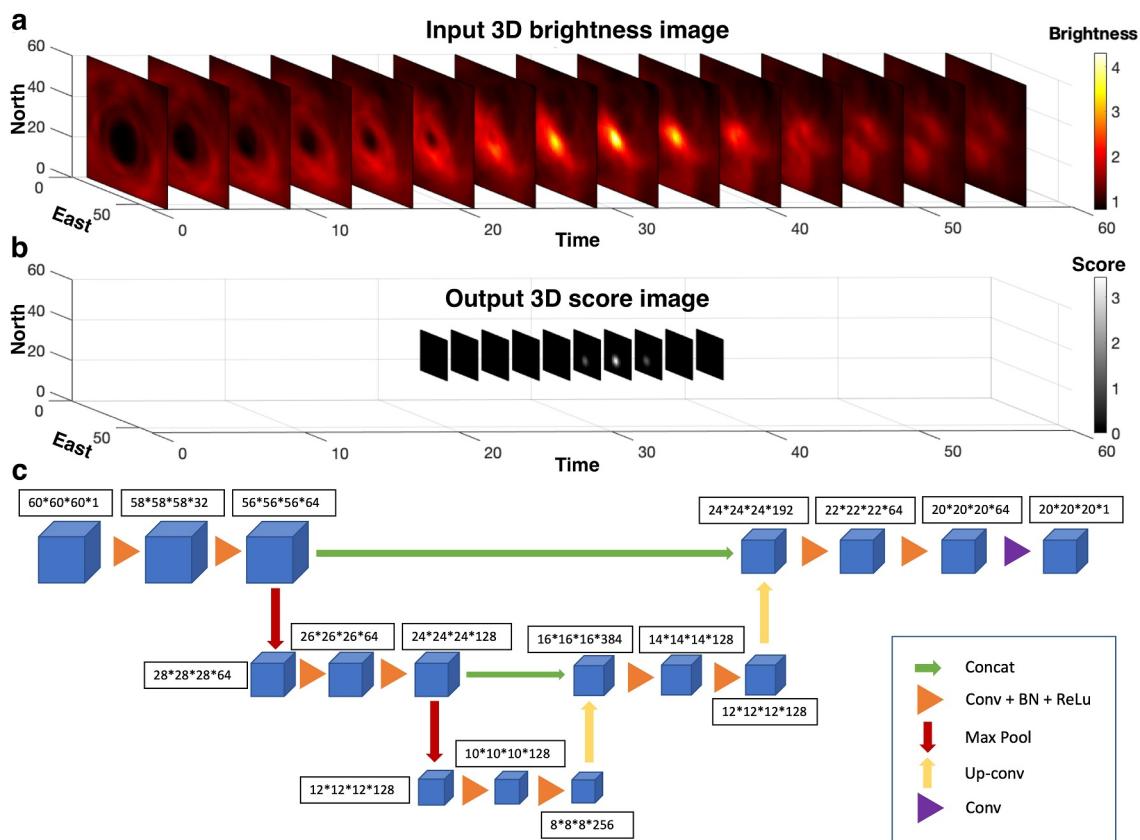


Figure 2. Illustration of the deep neural network used in SUGAR. (a) An example of the $60 \times 60 \times 60$ pixel input brightness block of the neural network, with snapshots plotted every four frames for demonstration purposes. (b) An example of the $20 \times 20 \times 20$ pixel output (label) block of the neural network, with snapshots plotted every two frames. (c) The 3D U-Net architecture used in SUGAR. Concat = 3D crop and concatenation layer; Conv = 3D convolutional layer; BN = batch normalization; ReLu = rectified linear activation function; Max Pool = 3D max pooling layer; Up-conv = 3D transposed convolutional layer.

hypocenter refinement with guided phase-picking, and magnitude determination (Figure 3). We describe each process in detail in the following four subsections.

2.2.1. Brightness Video Generation

For the first workflow process, we use a modified source-scanning algorithm (Kao & Shan, 2004; Tan et al., 2019) to generate brightness videos for each hour. We divide the study area into a grid with spacing of 4 km in latitude and longitude and a fixed depth of 10 km. Sticking to this depth is a common practice especially in rough location of crustal earthquakes and Tan et al. (2020) have demonstrated that the influence to the final hypocenter location is very limited given that the precise focal depth will be determined at a later stage. Assuming each grid node to be a potential energy source at each time increment, we calculate the theoretical P and S arrival times at each station and stack the waveform segment 1.5 s before until 4.5 s after the predicted arrivals. The waveforms are filtered to 5–20 Hz before stacking. After Tan et al. (2019), the brightness value at grid node i at time step t can be expressed as

$$B(i,t) = \left[\frac{1}{N} \sqrt{\frac{1}{w} \sum_{k=1}^w S_{it}^2(k)} \right]^3 \quad (1)$$

where N is the number of stations, w is the number of sample points in the sliding window, and the k th sample in the stacked waveform is calculated by

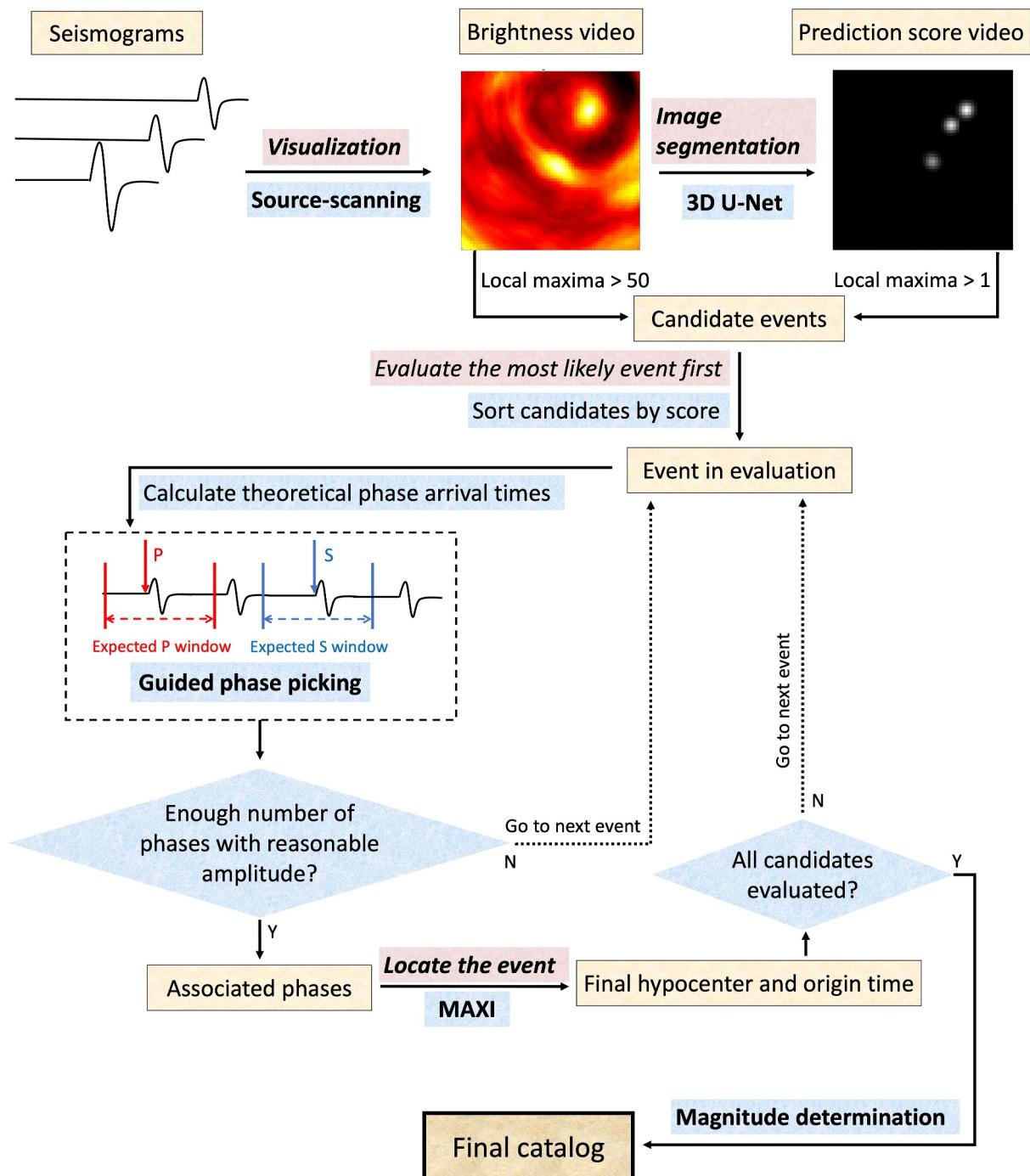


Figure 3. Schematic flowchart of SUGAR. MAXI: maximum intersection method.

$$S_{ij}(k) = \sum_{j=1}^N \left[\frac{|s_j(t + tt_{ij} + (k - \frac{1}{4}w) \times \Delta t)|}{m} \right]^{\frac{1}{n}} \quad (2)$$

where s is the seismogram at each station, Δt is the sampling time interval, and tt_{ij} is the theoretical travel time from grid node i to station j . Before stacking, seismograms are forced to be non-negative by taking the absolute

values, and then normalized by the median value (m) minute by minute without overlap. Although this normalization method would cause visible steps in the waveform when the background noise level changes dramatically in two minutes, the resulting brightness video will not show discontinuity.

The brightness values B in all space and time form the brightness video. Global maxima in each frame are a function of time ($B_{\max}(t)$). We find all peaks on $B_{\max}(t)$ and mark their times and corresponding map locations. If the brightness value of the n th peak is over a threshold (e.g., 50), the location and time of the n th peak is considered a candidate event. Meanwhile, the time period from the $(n-1)$ th peak to the $(n+1)$ th peak is considered a buffer zone that does not allow additional candidates. These candidates take their brightness values as scores when entering process 3. The buffer zone will be skipped in process 2 when searching for candidate events because the neural network does not have many training samples with large brightness values and may not work properly.

The computing time in this workflow process is not a big concern, especially with parallel processing. It takes ~ 20 min to finish this step (46 stations each with an hour-long data, 10,000 grid nodes, 0.5 s time increment) with one serial process and only ~ 1 min using parallel computing with 100 threads (AMD EPYC 7452 32-Core Processor).

2.2.2. 3D U-Net Prediction

In the second workflow process, we use the neural network to transform the brightness video into a prediction score video and then identify candidate events. We use a 3D sliding window to cut the continuous brightness video into a series of blocks of $60 \times 60 \times 60$ pixels with 20-pixel overlaps in all dimensions for use as the input to the 3D U-Net. The output blocks of $20 \times 20 \times 20$ pixels are concatenated together to cover the central part of the study area without any gap or overlap in either space or time, forming the neural network prediction score video. All local maxima in the score video over a certain threshold are mapped back to latitude, longitude and origin time as potential candidate events. For each hour, all candidate events from processes 1 and 2 are combined to create a preliminary catalog to take forward to the next workflow process. The prediction of an hour-long brightness video ($100 \times 100 \times 7200$ pixels) takes less than a minute with a 8G GeForce RTX 2080 graphic card.

2.2.3. Hypocenter Refinement With Guided Phase-Picking

In the third workflow process, we evaluate each candidate event by phase picking and trial localization, and either dismiss it or refine it to a final hypocenter and origin time. We begin by sorting the candidate events with decreasing scores. Starting from the event with the highest score, we use its preliminary location and origin time to calculate predicted P and S arrival times at each station. Then, we pick phases from seismograms around the predicted times. The picking method may have different choices, such as using the kurtosis function (e.g., Baillard et al., 2014; Tan et al., 2019) or a machine learning model (e.g., W. Zhu & Beroza, 2019; Mousavi et al., 2020). Machine learning models may have better accuracy with small to moderate size earthquake signals recorded at land based stations, whereas a simple kurtosis function may be more generalized with synthetic data, ocean bottom seismometer data, whale call signals, low frequency earthquakes and other situations that the machine learning models have not been heavily trained on. Here we apply the original model of the Earthquake Transformer (Mousavi et al., 2020) in our real data tests and the application to the 2016 Kaikōura sequence, and use a kurtosis-based method, simplified from Tan et al. (2019), in the synthetic tests. The kurtosis rate is calculated by

$$Kr(t) = K(t + dt) - K(t) \quad (3)$$

where $K(t)$ is the kurtosis function calculated by a 2-s moving window and dt equals to 5 times the sampling interval. A pick is marked at time t if $Kr(t) > 3$, and we don't allow multiple picks within 3 s in the same channel. For either picking method, the amplitude of the pick is recorded as the maximum amplitude within 5 s of the pick time or until the next pick, whichever is earlier.

We consider both arrival time and amplitude constraints to identify phases that belong to the current candidate event, after Tamaribuchi (2018). For a given candidate event, we first select the 20 closest stations around the epicenter and find P and S arrivals within ± 3 s of the predicted arrival times. The candidate is deleted if there are fewer than two P phases or two S phases. For each phase, we estimate a local magnitude (Richter, 1935) with constants after Rhoades et al. (2021):

$$M_{phase} = \log(A) - (0.51 - 0.79 \times 10^{-3} \times R - 1.67 \times \log(R)) \quad (4)$$

where A is the synthesized Wood-Anderson amplitude (Uhrhammer et al., 2011) of the phase amplitude, R is the hypocentral distance in km. To compensate the smaller P waves so that they can produce similar magnitude values as S waves, we empirically determine a magnification factor of 2 for the P wave amplitude before inputting it to Equation 4. This is only used in this process and the purpose is to identify those phases with reasonable amplitudes and eliminate those with abnormal amplitudes. The median value of all M_{phase} is defined as the reference magnitude.

Then, for all stations, we calculate a time error dt based on the theoretical arrival time and a magnitude error dm based on the reference magnitude for each phase. The criteria for a phase to be associated are:

$$dt_P < \begin{cases} 1.5s, & \text{if } d < 0.5 \\ 2.1s, & \text{if } 0.5 < d < 1.5 \\ 3.0s, & \text{if } d > 1.5 \end{cases} \quad (5)$$

$$dt_S < \begin{cases} 3.0s, & \text{if } d < 0.5 \\ 3.0s, & \text{if } 0.5 < d < 1.5 \\ 4.0s, & \text{if } d > 1.5 \end{cases} \quad (6)$$

and

$$dm < 1.2 \quad (7)$$

where d is the epicentral distance in degree. We define this distance dependent time error different from the original application (Tamaribuchi, 2018) to accommodate the sparser station distribution in the Kaikōura region and larger epicentral distances. The candidate is deleted if there are fewer than two qualified P phases or S phases.

All the associated phases are input into the maximum intersection method (Font et al., 2004; Tan et al., 2019). The searching area is limited to ± 6 km in latitude and longitude from the preliminary epicenter with a depth range of 0–60 km. The event quality Q_e is calculated by the actual number of intersections divided by the total possible number of intersections. Any event with $Q_e \geq 0.6$ or the number of total phases ≥ 10 will be located and logged.

All phases used in determining one event will not be used for other events. Any other phase that meets the criteria of dt_P and dt_S from any used phase will also be eliminated. This location refining process starts with the candidate with the highest score and repeats until all candidates have been reviewed. A five-minute overlap between hours is applied to make sure events at the beginning and the end of each hour are well considered. Note that the reference magnitude here is not recorded and the magnitude will be estimated more robustly in the last workflow process (see following subsection).

The kurtosis phase picking takes only a couple minutes while the Earthquake Transformer takes ~ 10 min running with CPUs. The running time of the location step depends on what velocity model is used. It takes ~ 3 – 10 min to finish this process with a 1D velocity model depending on how many earthquakes occur in the hour. And it takes ~ 10 – 50 min with a 3D velocity model.

2.2.4. Magnitude Determination

The neural network prediction scores in workflow process 2 are not linearly related to earthquake magnitude, and the reference magnitude in workflow process 3 is not accurate. Thus, an additional process is required to determine the magnitude formally. We still use Equation 4 but the variable A is defined as the largest synthesized Wood-Anderson amplitude among all channels and phase picks at that station. We also add a station correction term after Rhoades et al. (2021).

The magnitude determination process runs in less than a minute for a typical hour-long data set. The entire workflow generally takes less than an hour to process an hour of data with a 3D velocity model and the Earthquake Transformer phase picking model, except for the first few hours after the mainshock when there are more than 200 earthquakes per hour. The recommended processing unit is “hour” and parallel computing can be achieved among different hours. With this implement, we finished processing 49 days of data (46 stations) within 10 days. The number of stations may affect the processing time mainly due to the Earthquake Transformer step, while the number of grid nodes in the study area is less important since the source-scanning step is running fairly fast. Further improvements can be made by implementing GPU to all workflow processes.

3. Performing Beyond Human and Other Workflows

In this section, we use a series of synthetic and real data tests to demonstrate SUGAR's superior performance compared with professional analysts (included as co-authors) and other state-of-the-art deep-learning based and non-AI workflows.

3.1. Synthetic Test

We test SUGAR with a range of neural network prediction score thresholds and earthquake rates. SUGAR outperforms a human analyst in precision and comprehensiveness in all cases.

For the synthetic test, we generate five hours of synthetic seismograms in the same way as for the training data set but with a different set of random origin times for the earthquakes and random noise with standard deviation of 0.0001 cm/s (a realistic value suggested by real data). Therefore, the continuous waveforms are completely new and the neural network has never encountered the resulting brightness videos in the training process. The five testing hours (No. 1 to 5) contain 511, 356, 190, 118, and 21 earthquakes, respectively. The phases in the synthetic waveforms usually look clearer, but it doesn't make the test easier. We run SUGAR with this data set and ask a professional seismic analyst who has extensive experience with Kaikōura aftershocks to work on it for comparison. SUGAR and the analyst use the same “iasp91” 1D velocity model (Kennett & Engdahl, 1991). The analyst has the prior knowledge that the data set is synthetic for the Kaikōra region and the waveforms were generated with a different velocity model. And the analyst has his own quality control and choice of software (Antelope, <https://antelopeusersgroup.org/>).

We compare the performance of SUGAR and the analyst according to their origin time and location accuracy, precision, recall and the F_1 value (Chinchor, 1992; Taha & Hanbury, 2015). We define a true positive detection as an identified earthquake with origin time difference <2 s and epicentral difference $<0.2^\circ$ (~ 22 km) compared to the known “ground truth” hypocenters. If these criteria are not met, the identified event is defined as a false positive detection. The recall is equal to the total number of true positive detections divided by the total number of ground truth events. The precision is equal to the total number of true positive detections divided by the total number of detections. The F_1 value is a combination of precision and recall (Chinchor, 1992; Taha & Hanbury, 2015), calculated as

$$F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (8)$$

The five-hour average F_1 value is defined as the mean of the five individual F_1 values for each hour.

We test SUGAR with the prediction score thresholds from 0.5 to 3.0. While there is a trade-off between the precision and the recall, SUGAR yields a five-hour average F_1 value above 0.90 in the score threshold range 0.6–2.6, with the maximum F_1 value of 0.93 at a threshold of 2.3 (Figure S2 in Supporting Information S1). The manual catalog, for comparison, only yields a five-hour average F_1 value of 0.64. The SUGAR-derived hourly results and their comparison with the manual catalog are shown in Figure 4. In hours No. 1 to 4, the manual catalog has precision of 0.7–0.9 and recall of 0.2–0.7, while the SUGAR catalog has precision of >0.9 and recall of 0.7–0.9 over a wide range of thresholds (Figure 4 inset). In hour No. 5, SUGAR can detect all 21 events correctly without false positives with the threshold ranging between 1.4 and 2.4, while the manual catalog misses one event and has one false positive detection. We conclude that SUGAR is not sensitive to the score threshold selection and consistently outperforms the human analyst by achieving much higher precision and recall in the five hours with different earthquake rates.

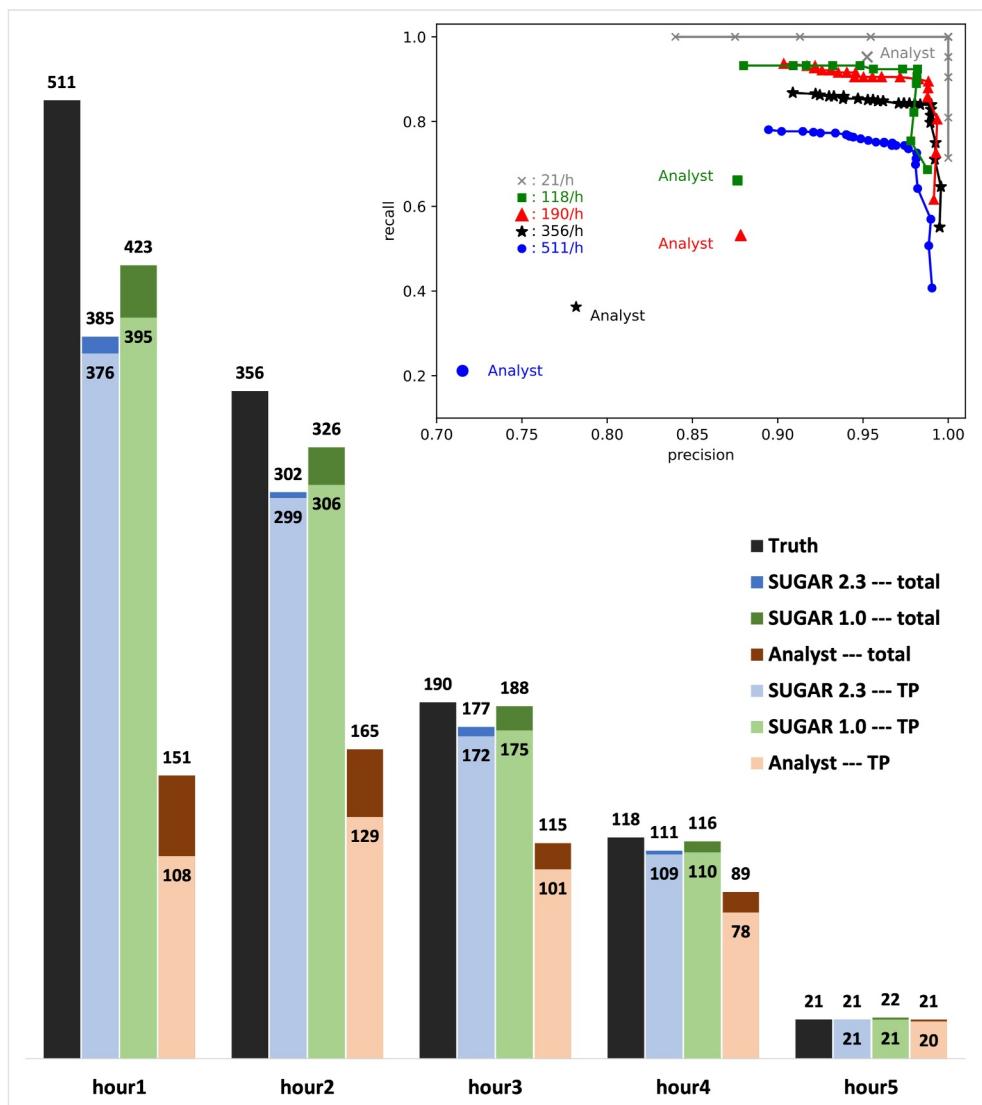


Figure 4. Summary of the synthetic test results. Testing hours No. 1 to 5 have 511, 356, 190, 118, and 21 earthquakes, respectively. Bars show the number of earthquakes in the ground truth, the SUGAR and the manual catalogs with lighter and darker colors indicating true positive and false positive detections, respectively. SUGAR has two catalogs with two different prediction score thresholds, 2.3 and 1.0. The inset shows the trade-off between SUGAR precision and recall with different thresholds, symbolized and colored by testing hour (with the number of ground truth events labeled in the key). Single points with corresponding colored symbols represent the analyst's performance in the same testing hours.

We choose two representative thresholds, 2.3 and 1.0, to further demonstrate SUGAR's superiority (Figure 4, Tables S2–S6 in Supporting Information S1). For example, in hour No. 1, the SUGAR catalog has 385 detections with the threshold of 2.3. Around 98% of those events (376) are true positives, recalling 74% of all 511 events with only 2% false detections. The manual catalog only has 151 detections and only 72% of them (108) are true positives, recalling 21% of all events with 28% false detections.

Locations in the two catalogs are shown in Figure 5, together with the known, input “ground truth” epicenters. The false detections in the manual catalog give the fictitious appearance that the regional seismicity is scattered over a wide area. In contrast, the SUGAR catalog forms denser clusters that more closely resemble the ground truth with many fewer false detections. Similarly, in hours No. 2 to 5, no matter which threshold is chosen, the performance of SUGAR is much better than the human analyst. Detailed locations for hours No. 2 to 5 are shown in Figures S3–S6 in Supporting Information S1. Quantitatively, the average origin time error, epicentral error and depth error for the SUGAR catalog (score threshold = 1.0) are 0.4 s, 2.6 and 3.9 km, respectively, whereas those values

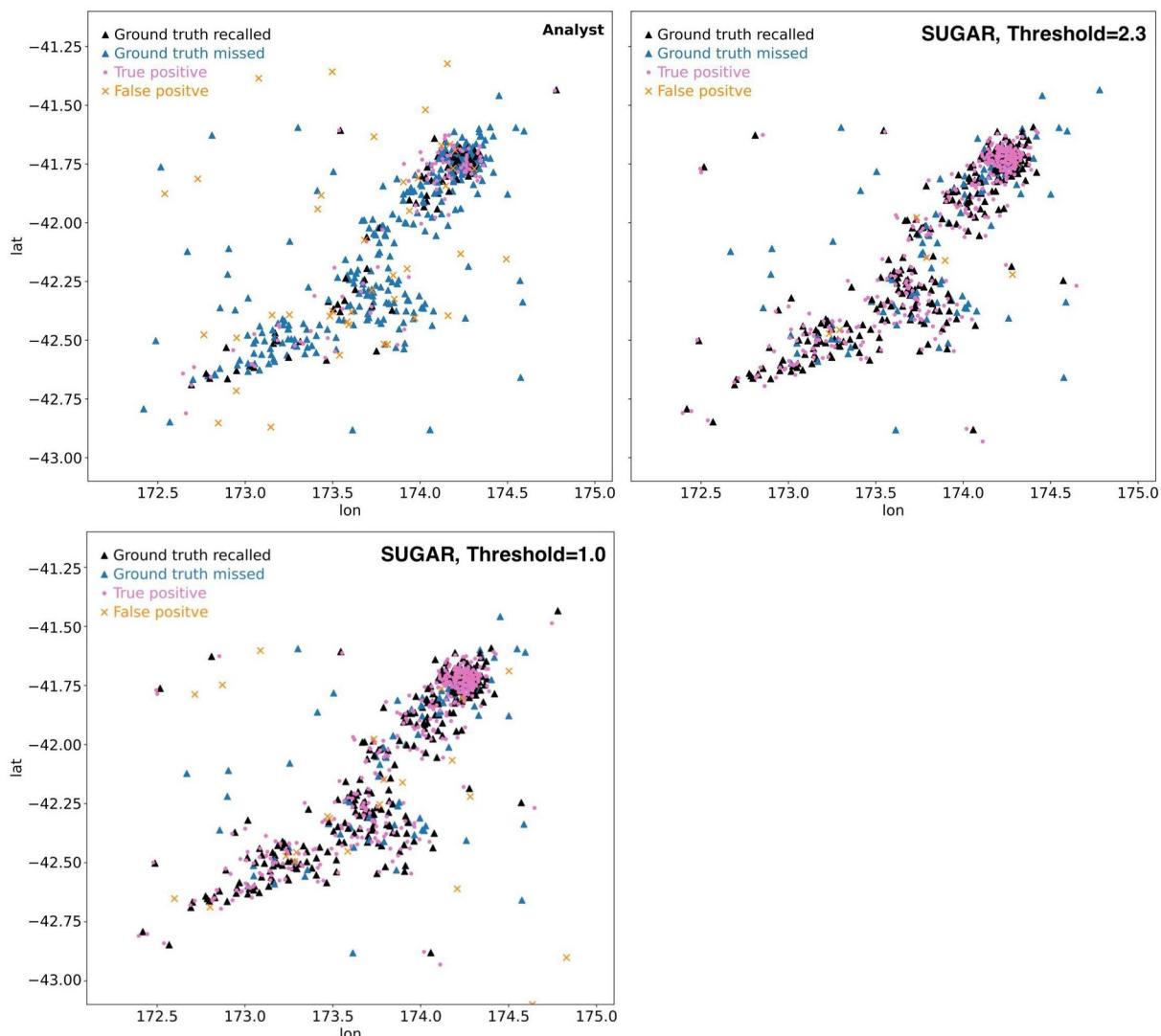


Figure 5. A comparison between the analyst's catalog and the SUGAR catalog (score threshold = 2.3, and threshold = 1.0) in testing hour No. 1. Black triangles are ground truth locations that have been recalled by the corresponding catalog. Blue triangles are ground truth locations that have been missed by the corresponding catalog. Dots represent true positive event locations while crosses represent false positive detection locations.

for the manual catalog are 0.5 s, 4.2 and 4.9 km. This shows that the origin time and location accuracy of the SUGAR catalog are also better than the manual catalog. Comparing with the ground truth, the average magnitude error of SUGAR is 0.19.

3.2. Solving the Phase Association Ambiguity

The source of false positive detection is misidentification and misassociation of seismic phases, such as marking a depth phase as an *S* phase or associating phases from different events to a single event. The more such mistakes are made, the greater the mislocation will be. On the other hand, one may decide to minimize these mistakes by dismissing the unclear phases, which leads to a low recall. In this subsection, we use a particularly closely-timed cluster of synthetic events to demonstrate the outstanding phase association ability of SUGAR.

Figure 6 shows the example of eight earthquakes within one minute (Table S7 in Supporting Information S1), taken from hour No. 1 in the synthetic test. The first three events (No. 1–3) all occur within 0.7 s, with events 1 and 2 separated by ~80 km with a 0.9-unit difference in magnitude, and events 1 and 3 separated by only ~30 km with

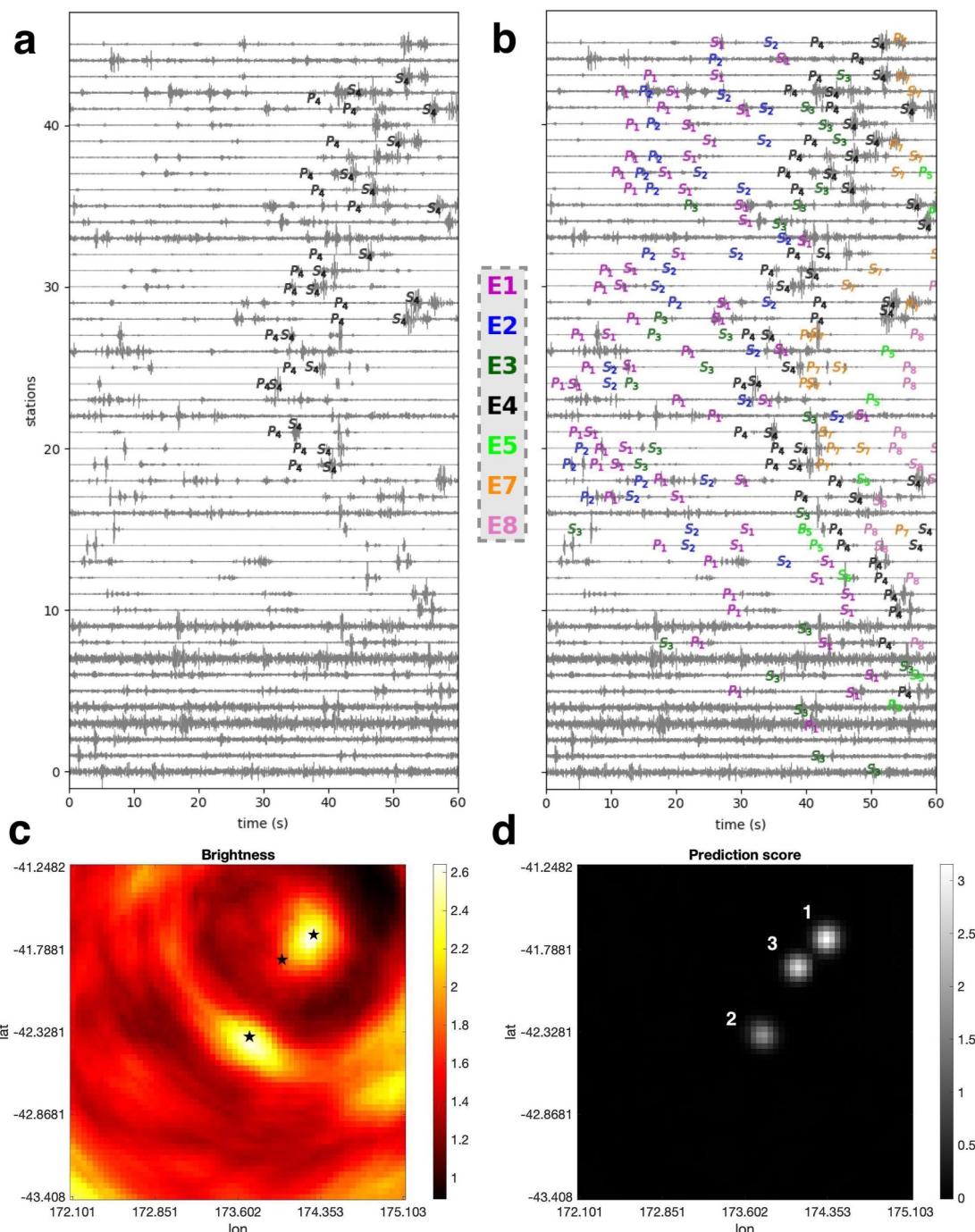


Figure 6. An example of eight synthetic earthquakes clustered within one minute. (a) Manual picks for the one-minute waveform segment (filtered to 5–20 Hz) at 46 local and regional stations. Only the biggest event No. 4 is identified by the analyst. (b) The SUGAR picks for the same segment. Seven events are identified. P and S phases for each event are marked with different colors. Numbers after letter P or S correspond to event numbers in Table S7 in Supporting Information S1. (c) The brightness image at time 0 s. Stars indicate ground truth locations for events No. 1–3. (d) The prediction score image for the brightness image in (c). Numbers correspond to events No. 1–3.

a 0.4-unit magnitude difference. Event 4 is the largest (M 4.6) among all eight events, occurring 9 s before events 5–7 which all occur within 2 s. Events 6 (M 3.1) and 7 (M 3.8) are separated by only \sim 18 km.

The waveforms are so complicated that the analyst can only identify the biggest event (Figure 6a). In contrast, SUGAR successfully locates seven of eight events without any false detection regardless of the score threshold

choice from 0.5 to 2.7 (Figure 6b, Movie S1 in Supporting Information S1). Event 6 is missed because the larger event 7 occurs at a nearby location within 2 s so that its phases overlap with those from event 6 at most stations.

We next demonstrate how SUGAR identifies the first three concurrent events (at around time 0) in more detail. The brightness image and the neural network prediction score snapshot are shown in Figure 6. In the brightness image (Figure 6c), we see one global maximum but also many local maxima, and it is difficult to determine visually or through simple mathematical rules which of them are true earthquakes. Moreover, some events may not form local maxima on the brightness image snapshot at all; for example, event 3 is impossible to verify directly from the brightness image. However, the prediction score snapshot shows three distinct local maxima (Figure 6d), corresponding to three individual events at their correct locations. The complete brightness and score videos for the whole minute can be found in Movie S1 in Supporting Information S1.

The situation described above is common in major aftershock sequences. Automatic phase association methods as well as human analysts may have trouble dealing with this problem correctly (W. Zhu et al., 2022; Ross, Yue, et al., 2019; Woollam et al., 2020; Tamaribuchi, 2018; Tan et al., 2019; M. Zhang et al., 2019). A common strategy is to uphold the overall precision of the earthquake catalog at the cost of missing many densely clustered events by providing one best-fitting solution for the most prominent event of the cluster, or simply ignoring the entire cluster if no satisfactory solution can be obtained. SUGAR, with its completely different approach, utilizes a deep neural network to identify real earthquakes from the brightness videos, and performs more satisfactorily as a result.

Nevertheless, we write a caveat here that false positive detections do exist in the SUGAR catalog depending on the threshold selected. And these detections are challenging to be identified by visual inspection. This is because that any detection made by SUGAR will align a group of phases according to expected travel times and amplitudes. Thus, false positive detections may look nicely in the waveform section plot whereas some true positive associations may look unconvincing due to fewer number of picks (Figure S7 in Supporting Information S1). Further details of the ineffectiveness of visual inspection can be found in Text S2 in Supporting Information S1. However, given the low false positive rate and the additional process of inter-event relocation after SUGAR (see Section 4), it shouldn't bias the interpretation of the final results.

3.3. Real Data Test

To validate the performance of SUGAR on real data, we choose the third hour in the 2016 Kaikōura earthquake sequence (13:00:00 to 14:00:00 UTC on Nov. 13) as a challenging test case and make a wide comparison with the GeoNet catalog, three analysts, a popular deep-learning based algorithm and a non-AI workflow. The same “iasp91” 1D velocity model is used in all tests.

We apply the score threshold of 1.0, for which SUGAR's precision is 93%–95% in the synthetic test. The GeoNet catalog has only 39 events in this hour (Figure 7a), while the SUGAR catalog has 228 (Figure 7b), about six times as many.

We ask three experienced analysts to work on the same hour independently with their own quality control criteria. Analyst B is the same person in the synthetic test. Analysts B and C use the same waveforms as SUGAR. However, the station coverage that works best for SUGAR excludes many distant stations and may not be favored by analysts. Therefore, analyst A has access to 20 additional, more distant stations (Figure S13 in Supporting Information S1) to optimize the location process. All analysts have prior knowledge that the data set corresponds to the third hour of the 2016 Kaikōura earthquake sequence. Analysts A, B and C locate 184, 127, and 43 events, respectively. Even with 20 more stations, analyst A locates about 19% fewer events than SUGAR does (Figure 7c). Analyst B only locates about half as many events as SUGAR (Figure 7d), while analyst C has performed similarly to GeoNet (Figure 7e).

We choose the “PhaseNet + Gamma” (PhG) as a representative deep learning based workflow, which initially utilizes a supervised deep learning model to pick phases, and then applies an unsupervised Bayesian Gaussian mixture model to associate those phases based on both their arrival times and amplitudes (W. Zhu & Berkoza, 2019; W. Zhu et al., 2022). While most of the default parameters work fine in this workflow, we find that the number of picks required for an earthquake can dramatically change the result. When the minimum number of picks is 10, the PhG detects 141 earthquakes (Figure 7f), 101 fewer events than SUGAR. When we lower the threshold to 7, the number of total detections rises to 212 (Figure 7g), approaching the number for SUGAR. We

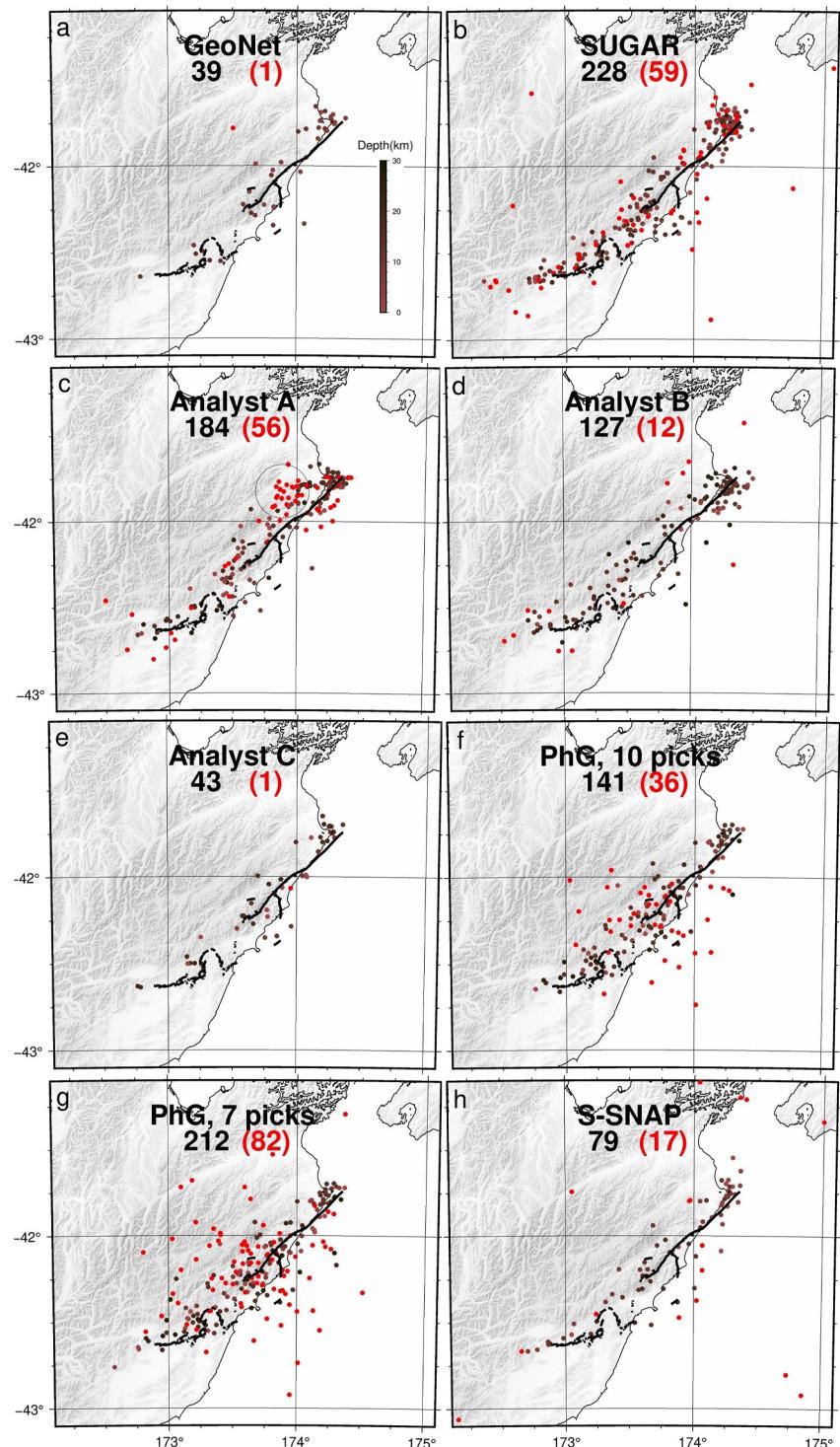


Figure 7. Results of the one-hour real data test for different workflows. Each panel title writes the method name, number of event resolved (black) and the number of event that cannot be cross validated (red). Black lines represent the Kaikoura mainshock surface ruptures (Litchfield et al., 2018). Brown and red dots represent the earthquakes that have been and cannot be cross validated, respectively. (a) GeoNet locations. (b) SUGAR-derived locations, determined using 46 stations. (c) Analyst A's locations using 66 optimal stations. (d) Analyst B's locations using the same 46 stations as for SUGAR. (e) Analyst C's locations using the same 46 stations. (f) PhG locations using the same 46 stations and the phase number threshold of 10. (g) PhG locations using the same 46 stations and the phase number threshold of 7. (h) S-SNAP locations using the same 46 stations.

also tested the Earthquake Transformer (Mousavi et al., 2020) and the MALMI phase association method (Shi et al., 2022). Though the Earthquake Transformer makes good picks, either its own associator or the MAIMI associator gives much worse results than the PhG with fewer detections and more scattered locations.

We choose the Seismicity-Scanning based Navigated Automatic Phase-picking (S-SNAP) (Tan et al., 2019, 2020) as a representative for non-AI workflow, because like SUGAR it uses the source-scanning algorithm to guide phase-picking and it performs fairly well among the non-AI algorithms. S-SNAP detects and locates 79 events in this hour and most locations are close to surface rupture traces (Figure 7h).

The eight catalogs have very different number of events ranging from 39 to 228. SUGAR, Analyst A and PhG (7 picks) all have ~200 detections while others may have missed a huge amount of events. While there is no ground truth for these real data, it is very difficult to fully verify all the detections in different catalogs. A practical way is to try using all the catalogs to cross validate each other. If an event appears in at least two catalogs, we assume that it is a true detection. Given the large travel time errors result from the “iasp91” velocity model, we allow 4 s difference in origin time and 0.2° in epicentral distance. Those cross validated events are marked as browns, whereas the remaining events are marked as red in Figure 7. Generally, catalogs with smaller number of detections would have fewer remaining (questionable) events. To have a more detailed comparison, we focus on the SUGAR, analyst A, analyst B and PhG (7 picks) catalogs, which have relatively more detections.

Most remaining events in the SUGAR catalog fall around surface rupture traces and ~20 are further away (Figure 7b). We cannot be sure that these events are false detections, but we should expect ~20 artifacts given the total of 228 detections and the false detection rate of ~7% or slightly higher switching to the real data. Analyst A's catalog has fewer scattered events, but shows a unique cluster of ~20 events (Figure 7c, black circle) in the northern part which is not visible in any other catalog. This is suspicious because there is no known fault rupture in that region and none of other studies have shown those seismicity even in the entire aftershock sequence. In addition, analyst A has done a synthetic test with the same 66 stations and shown a tendency to create false detections in that region. These all suggest that analyst A's catalog not only has false detections, but also creates an artifact cluster that can lead to significant bias. This is even worse than those scattered artifacts which may be easily removed by additional processes such as inter-event relocation. Therefore, we conclude that SUGAR (228 detections) performs much better than analyst A (184 detections).

Analyst B's catalog has much fewer detections and the remaining 12 events scatter away from the faults (Figure 7d), which is well expected given the precision of analyst B in the synthetic tests. Analyst B's catalog generally forms a wider distribution compared with the SUGAR catalog. Therefore, we conclude that SUGAR performs better than analyst B with much more detections and a slightly better location accuracy. We note that the different performance between analysts are due to subjective preference even though they use the same quality control criteria: a minimum of four picks for any event to be locatable. Although analyst A has 20 more distant stations, most picks are not made at those stations. Therefore, the effect of accessing more data cannot explain the major differences between results of analysts A and B. Subjective preferences are very common among seismologists, and can easily cause inconsistency when multiple analysts are working together on the same sequence.

The PhG catalog (7 picks) has the most similar number of detections compared with SUGAR. The remaining events, however, form a circle pattern around the center of the study region and many of them (>30) are far away from the faults (Figure 7g). Also, the seismicity generally forms a much wider distribution than the SUGAR catalog, even with the minimum of 10 picks requirement, indicating much larger location errors. We conclude that the PhG workflow is not as accurate as SUGAR and may create many more artifacts when lowering the phase pick threshold and pursuing a high number of detections.

To conclude, this real data test shows that SUGAR could be preferable to human analysts and other AI and non-AI workflows for processing busy aftershock sequences, achieving desirable comprehensiveness, precision, efficiency, and consistency. Waveform examples of various qualities and additional descriptions can be found in Text S3 and Figure S14–S19 in Supporting Information S1.

4. Application to the 2016 Kaikōura Earthquake Sequence

We now apply SUGAR to the complex 2016 Kaikōura earthquake sequence. We consider a 48.5 day period from the November 13 mainshock until the end of 2016, and use the 46 stations in Figure 1 and Table S9. Following discussion of known timing issues in some regional strong motion stations (Jerome Salichon, pers. comm.), we

find small ($\sim \pm 1$ s) timing errors at just 9 of our selected stations; testing shows these errors influence our results only minimally. The New Zealand-wide 3D seismic velocity model 2.3 (Eberhart-Phillips et al., 2022) is used in both process 1 and 3. The score threshold in process 2 is 1.0, giving a precision above 0.93 in the synthetic test, and other parameter settings are listed in Table S8 in Supporting Information S1. SUGAR may not work well for large earthquakes ($M > 6.5$) since it assumes point sources and lacks many training samples at these magnitudes. SUGAR therefore misses the mainshock, and starts yielding reliable results after its wave train passes, ~ 2 min after its origin time. Similarly, the Earthquake Transformer may not pick phases for large earthquakes, probably also due to a lack of training samples, a problem encountered in other studies (e.g., Gong & Fan, 2022). This may explain why SUGAR also misses the largest aftershock reported by GeoNet (2016-11-14 00:34:23, M6.7); this event was detected properly by the 3D U-Net and would have been cataloged had we used the kurtosis-based phase picker. At smaller magnitudes, the Earthquake Transformer generally performs better than the kurtosis-based phase picker. For completeness, we manually add GeoNet solutions for the mainshock and largest aftershock to our catalog.

During the 48.5-day study period, SUGAR detects and locates 67,660 events (Figure S20 in Supporting Information S1 and Table S10), more than six times the equivalent of 10,861 in the GeoNet catalog. Magnitudes determined by GeoNet are known to be inconsistent with moment magnitude, and Hutchinson et al. (2024) re-calculated trace amplitudes, corrected for Wood-Anderson response, and proceeded to determine station magnitudes according to the equations of Rhoades et al. (2021). We also use the Rhoades et al. (2021) magnitude scale, such that we can compare SUGAR magnitudes with those of Hutchinson et al. (2024). The magnitudes are generally in agreement but those of Hutchinson et al. (2024) are larger than those of SUGAR by 0.15 on average (Figure S21 in Supporting Information S1). A few events uncorrected by Hutchinson et al. (2024) exhibit larger magnitude differences. After removing the average 0.15 difference, the magnitude distributions of the two catalogs are shown in Figure S22 in Supporting Information S1. We calculate magnitudes of completeness (MC) for different time periods using the maximum curvature method with a +0.2 correction (Woessner & Wiemer, 2005). In the SUGAR catalog, the MC is 2.4 in the first 1.5 days after the mainshock (Figure S22a in Supporting Information S1), decreases to 1.8 over the following two days (Figure S22b in Supporting Information S1), and then stabilizes at 1.5 in December (Figures S22g–S22i in Supporting Information S1). The GeoNet catalog has a MC of 2.6 in the first 1.5 days (Figure S22a in Supporting Information S1), decreases to 2.2 in the next two days (Figure S22b in Supporting Information S1), and also stabilizes at ~ 1.5 in the end (Figures S22g–S22i in Supporting Information S1). Though MC values are similar in each catalog, they decrease faster in the SUGAR catalog, which therefore detects more smaller events in the early stage of the sequence. However, in the first few days small events are naturally buried in signals from larger events, an issue that even SUGAR cannot fully resolve.

The b value (Gutenberg & Richter, 1944) of the SUGAR catalog is consistently larger than that of the GeoNet catalog, suggesting that the latter misses many events above the magnitude of completeness (Figure S22 in Supporting Information S1). Moreover, in the early aftershock sequence the SUGAR catalog not only has many more small events of magnitude < 4 , but also many more moderate events of magnitude 4–5 (Figure S22a in Supporting Information S1). We try to visually inspect all 235 $M \geq 4$ events in the SUGAR catalog. While many of them cannot be corroborated in this way, we present two examples that were missed not only by the GeoNet catalog, but also by Lanza et al. (2019) and Chamberlain et al. (2021). The first is a M 4.8 event ~ 4 min after the mainshock (Figure S23 in Supporting Information S1), and the second is a M 4.6 event ~ 23 min after the mainshock (Figure S24 in Supporting Information S1). These events concentrate early in the aftershock sequence when signals tangle together (like in Figure 6) and where previous studies would naturally fail. Among all 235 events, 27 are classified “unlikely event” by the first author. Assuming these are false detections, the precision would be $\sim 89\%$, slightly lower than the synthetic test. This is probably because these events concentrate in the first days of the sequence. We should keep in mind that the assumption may not hold true due to the ineffectiveness of visual inspection (see Text S2 in Supporting Information S1). Therefore, instead of this subjective classification, we recommend a more objective approach of setting a phase pick threshold (e.g., 15) in order to filter out high quality events. In our application, however, we keep all detections at this stage and rely on the following inter-event relocation process to remove many of the low-quality and false detections.

We perform double-difference relocation on the SUGAR catalog using the Growclust package (Trugman & Shearer, 2017). For cross correlation calculation, we filter the waveforms to 1–10 Hz and cut 2 s segments around P and S picks (or predicted arrival times if no pick has been made). Event pairs with at least eight > 0.65 cross

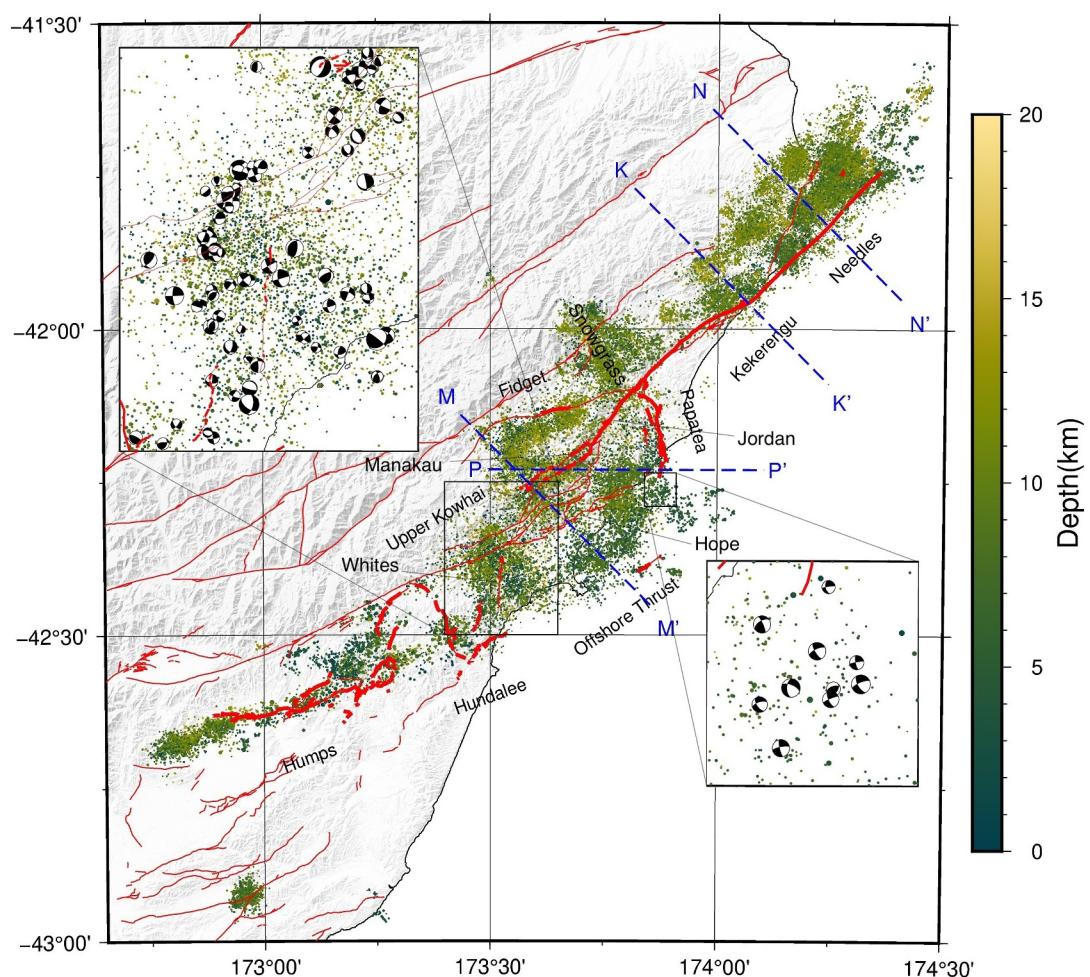


Figure 8. SUGAR-derived and Growclust-relocated hypocenters of the 2016 Kaikōura sequence, colored by focal depth. Thick red lines are mainshock surface ruptures (Litchfield et al., 2018) and thin red lines are other active faults without observed surface rupture in the 2016 earthquake (GNS Science, 2016). Insets show focal mechanisms from Chamberlain et al. (2021) and this study for two regions of particular interest.

correlation coefficient values are included in the relocation. This strict requirement results in only 46,440 events being successfully relocated (Table S11). After relocation, 23 of those 27 “unlikely” events are removed. This is therefore an effective way of retaining only high-quality events, and ensuring that low-quality or false detections will not bias interpretations of the relocated catalog. To enhance the quality further, we retain only those clusters with 10 or more events, trimming the number further to 41,392 (Figure 8). While relative uncertainties of the relocated events from bootstrapping are below 1 km, we lack accurate estimates of absolute uncertainties, but expect these to be a few kilometers based on most travel time residuals being less than 1 s. Chamberlain et al. (2021) deploy a match filtering based workflow with the catalog of Lanza et al. (2019) as templates, and use Growclust to refine their locations. With their quality control, 14,276 events are successfully relocated for the same time period and study region, far fewer than the 41,392 reported here. We also augment the focal mechanisms of Chamberlain et al. (2021) by calculating focal mechanisms for 12 additional $M > 4.0$ earthquakes and 43 additional smaller events using the “FPFIT” program (Reasenberg & Oppenheimer, 1985). The 55 new focal mechanism solutions are listed in Table S12.

4.1. Aftershock Patterns and Fault Structure

Almost all aftershocks are shallow (<15 km), implying that they occurred ~10 km above the presumed subduction interface in this area (Williams et al., 2013). They generally follow the surface ruptures but extend beyond them both along and across strike (Figure 8). In many places they form diffuse clouds rather than discrete planes,

suggesting complex secondary faulting distinct from the mapped surface ruptures. We describe these features and compare them with previous studies from the south to the north.

At the southwest end of the aftershock distribution, near where the mainshock initiated, seismicity forms a narrow trend subparallel to the Humps fault trace, but also extends further southwest beyond the observed surface rupture (Figure 8). Earthquakes along the eastern Humps fault fall mostly northwest of its surface trace, consistent with the northwestward dip angle of $\sim 60^\circ$ (Litchfield et al., 2018). However, clusters of shallow events away from the fault trace cannot be explained by this dip angle, but rather suggest a complex fault zone structure, consistent with the markedly discontinuous surface ruptures and small scale coseismic block rotation (Wang et al., 2020). Further northeast, seismicity follows the curvilinear trend of the Humps and Hundalee faults but without the sharp, northward deviation observed in the surface ruptures, suggesting a straighter linkage between these faults below the surface.

At middle latitudes of the Kaikōura rupture area (42°S – 42.5°S), we observe many clusters of aftershocks between the major active faults. One broad cluster falls around the reverse-sinistral Whites fault, which connects the Hundalee and Upper Kowhai faults (Figure 8). Although the Whites fault trends north–south, most strike-slip focal mechanisms in this region are sinistral northwest–southeast or dextral northeast–southwest, consistent with the bulk stress regime of the Kaikōura region, and perhaps indicative of structures oblique to the Whites fault. The wide variety of focal mechanisms here (Figure 8 inset) also suggests a complex fault zone structure. Further northeast, seismicity appears to fill the entire 3D volume from the inland Fidget fault all the way to offshore, without clear concentrations along surface traces of the Fidget, Manakau, Upper Kowhai or Hope faults, and with a wide variety of focal mechanisms (Figures 8 and 9). Our hypocentral locations, together with available focal mechanisms from Chamberlain et al. (2021) and this study, support an offshore thrust fault dipping northwestwards. Its dip angle may change with depth and it may split into multiple branches as it approaches the surface (Figure 9), possibly including the newly found Armers Beach fault (Nicol et al., 2022). However, the average dip of $\sim 35^\circ$ is consistent with the value suggested by Mouslopoulou et al. (2019). This structure may extend inland beneath Kaikōura peninsula to intersect with steeper, onshore crustal faults at depth, possibly accounting for localized uplift of the peninsula observed in differential lidar data (Clark et al., 2017).

Seismicity and focal mechanisms suggest that this offshore thrust may connect the Hundalee fault in the south with the Papatea fault in the north. The cross section P–P' shows a similar, west-dipping thrust at the southern end of the Papatea fault (Figures 8 and 9), with a steep dip above ~ 10 km depth and a shallower dip below that, consistent with geodetic observations (Diederichs et al., 2019). The multiple surface strands of the Papatea fault (Langridge et al., 2018) may also splay off this gently-dipping structure at depth. In addition, we observe a cluster of events south of the Papatea fault trace with consistent strike-slip focal mechanisms (Figure 8 inset). This may represent a southward extension of the west-dipping, sinistral-reverse Papatea fault, or alternatively an eastward extension of the Hope fault with a subvertical fault plane and dextral motion.

In the northeastern part of the Kaikōura rupture area (41.5°S – 42°S), aftershocks roughly follow the NE-trending, dextral Kekerengu and Needles faults. However, our catalog also appears to delineate a northwestward extension of the Papatea fault across the Kekerengu fault, named the Snowgrass fault zone by Chamberlain et al. (2021). Further northeastward, we observe some clusters deviating from the main Kekerengu–Needles fault trace by 10–20 km (Figures 8 and 10), too large to be explained by the fault dip of 50 – 70° (Litchfield et al., 2018). Though uncertainties for individual events may be large, their collective trend and range of focal mechanisms cannot be explained simply as noise. We therefore suspect that distinct structures northwest of and subparallel to the Kekerengu–Needles faults were activated during or after the Kaikōura mainshock. In profile K–K', this northwestward structure appears to dip steeply southeastwards, whereas in profile N–N', it may dip moderately ($\sim 45^\circ$) northwestwards, as indicated by local focal mechanisms. Around the main Kekerengu–Needles fault, a wide variety of focal mechanisms are observed, suggesting complex kinematics within a broad zone around the primary right-lateral faulting.

Finally, SUGAR successfully detects a southern cluster of events captured by the GeoNet catalog near $\sim 42.9^\circ\text{S}$, 172.9°E , ~ 30 km south of the mainshock epicenter and completely disconnected from the 2016 surface ruptures. There is no training sample generated for the neural network in this area. This further verifies that SUGAR has learned how to recognize earthquakes rather than simply how to match new images with training samples as for template matching.

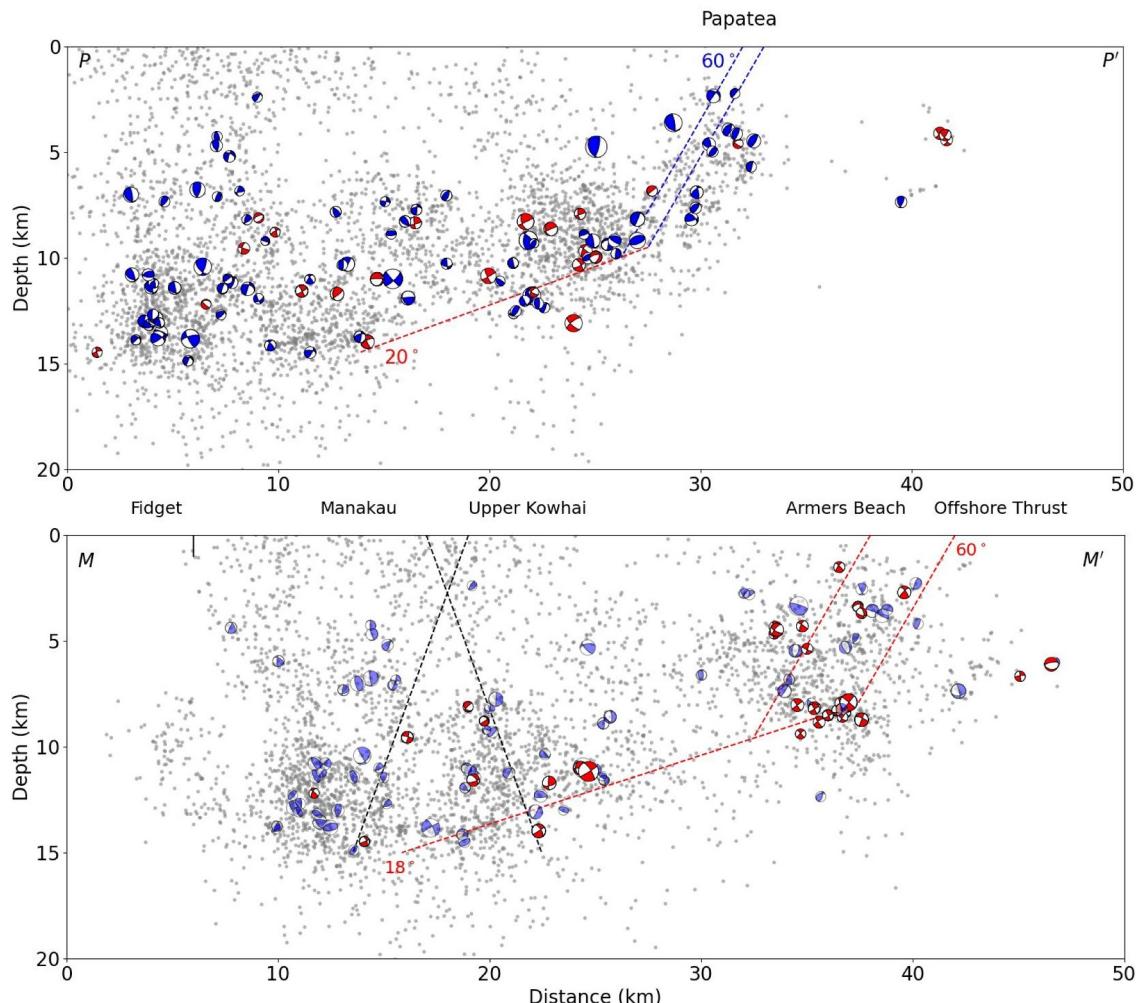


Figure 9. Cross sections P–P' and M–M' (Figure 8) showing earthquakes within 5 km horizontally from the line. Focal mechanisms from Chamberlain et al. (2021) and this study are viewed from the side and scaled by magnitude. Red and blue beachballs represent thrust and other events, respectively. Dashed lines in red and blue show inferred fault segments discussed in the text, whereas black lines mark structures after Litchfield et al. (2018).

5. Discussion

The most conspicuous feature of the SUGAR-derived Kaikōura catalog is that the spatial distribution of aftershocks exhibits no obvious gaps, contrasting with the discontinuous surface ruptures (Figure 8). Although this characteristic of the sequence has been observed in previous studies (e.g., Chamberlain et al., 2021; Lanza et al., 2019), our SUGAR catalog provides more events and in turn, more focal mechanisms solutions, resolving new details of the faulting involved. For example, while the Papatea fault is conspicuously lacking in aftershocks in earlier studies, we not only locate clusters of aftershocks around its southern end, but may also delineate a curved structure at depth, similar to the offshore thrust further south. Although aftershocks are not a direct indicator of co-seismic slip, this structural similarity and the connectivity of seismicity suggest a likely mainshock rupture path through the Humps, Hundalee, offshore thrust and Papatea faults, in agreement both with dynamic modeling (Ulrich et al., 2019) and another aftershock study (Chamberlain et al., 2021). However, the mainshock may have ruptured the alternative path through the inland crustal faults (Humps, Hundalee, Upper Kowhai and Jordan Thrust), given the connectivity of the aftershocks along this trend and with support from other dynamic models (Ando & Kaneko, 2018). One possibility is that reverse slip on the offshore thrust unclamped the inland crustal faults, allowing both pathways to rupture at the same time or in very quick succession. This simultaneous rupture scenario was first proposed by Herman et al. (2023) on the basis of Coulomb stress change calculations and finite element models of fault slip, but with a linking role for the underlying Hikurangi subduction interface.

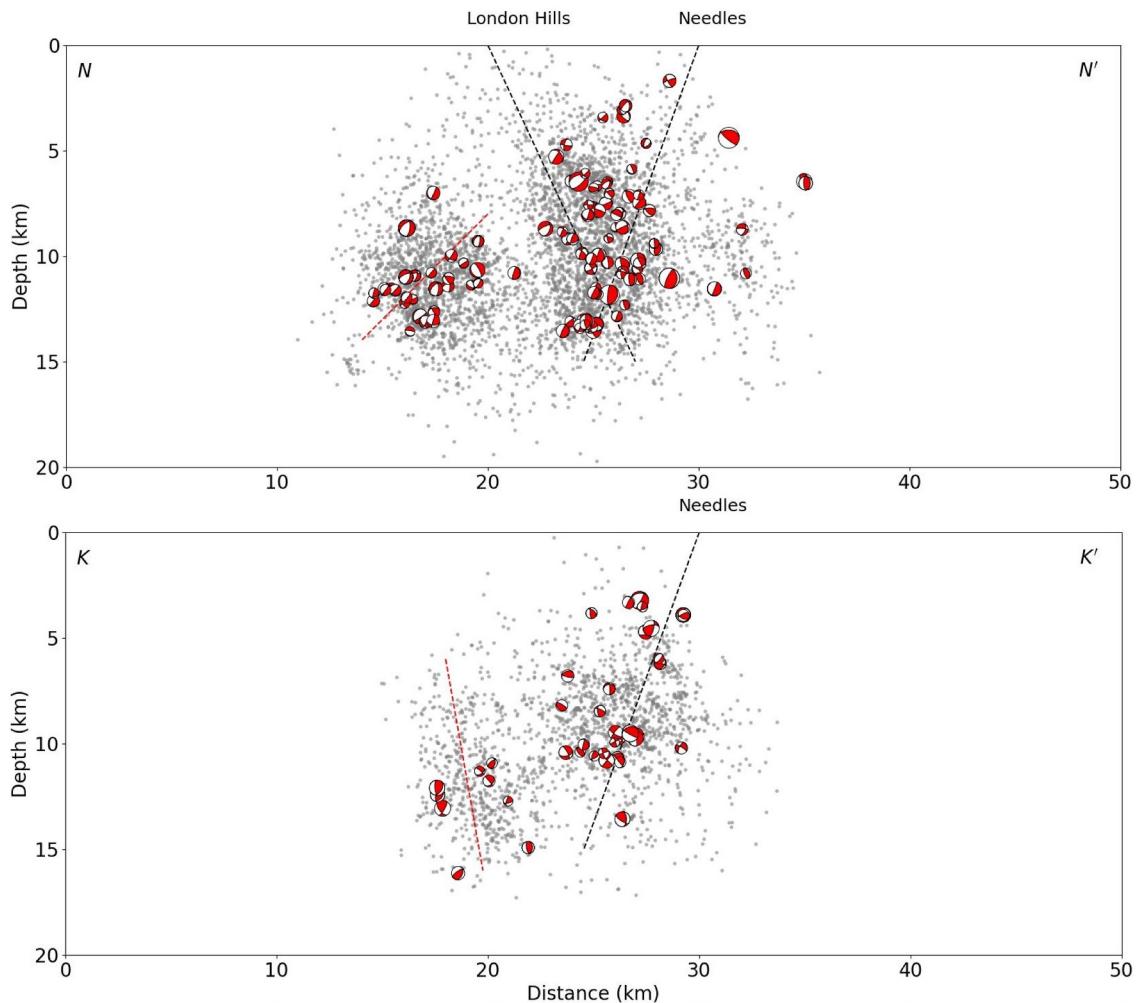


Figure 10. Cross sections N–N' and K–K' (Figure 8) showing earthquakes within 5 km horizontally from the line. Focal mechanisms from Chamberlain et al. (2021) and this study are viewed from the side and scaled by magnitude. Dashed lines in red show inferred faults discussed in the text, whereas black lines mark structures after Townsend and Little (1998) and Litchfield et al. (2018).

However, this seems unlikely because the “slab2” geometry (Hayes, 2018) used in their study is ~10 km too shallow and misaligned with regional intraslab seismicity (Aziz Zanjani et al., 2021).

Another noticeable feature of the SUGAR catalog is the broad distribution of aftershocks away from surface ruptures of the Kaikōura mainshock. This was observed by previous aftershock studies (e.g., Chamberlain et al., 2021; Lanza et al., 2019), but has not been discussed in detail, perhaps due to insufficient numbers of events and/or a lack of spatial resolution. Our SUGAR catalog, with its many times more detections, reveals more robust off-fault clusters with distinct separation from the main faults (e.g., on the northwest side of the Needles fault, Figure 10). Expected location errors of a few kilometers cannot explain the abundance of off-fault events, nor their consistent trends, dip directions, and focal mechanisms. We interpret that the wide distribution of aftershocks is not simply the result of off-fault deformation associated with damage zones surrounding the major, surface-rupturing crustal faults. The diffuse seismicity with various focal mechanisms is generally much wider than typical off-fault deformation zones of a few hundred meters to a few kilometers (Perrin et al., 2021). Moreover, the seismicity does not decay away from mapped surface ruptures in the way that damage zone aftershocks are expected to (Perrin et al., 2021). We speculate that much of the seismicity occurs on unmapped, secondary crustal faults of a variety of orientations, which may connect to the main faults at depths, forming a wide and complex shear zone. However, since this is the first earthquake sequence detected and located by

SUGAR, we are cautious of overinterpreting the seismicity distribution; more sequences should be analyzed with SUGAR before like-for-like comparisons can be made.

The biggest limitation of SUGAR is that the neural network is currently trained with the seismic network configuration around the Kaikōura area as a site-specific model. When applied to other regions with their unique station distributions, seismic events will be represented by different shapes in the brightness videos, such that the neural network trained in the Kaikōura region will be ineffective at identifying them. This situation is akin to applying a neural network trained to identify grizzly bears to a data set of polar bears. To solve this problem, one may train a generalized model with a variety of station distributions or apply transfer learning for 3D image segmentation (Chen et al., 2019) based on the “Kaikōura model.” However, since SUGAR can use synthetic brightness videos for training, and the training process only takes a few days, we recommend retraining the network from scratch in each new study region. Furthermore, in applying SUGAR to other areas, some parameters in the training process and workflow may need to be adjusted empirically. A few weeks might therefore be needed to train and tune the program for optimal results. Once trained, however, SUGAR is near real-time.

Another limitation is that the current 3D U-Net may favor seismic sources occurring within a narrow depth range (e.g., 0–30 km) because it only searches for seismic sources at an assumed, fixed depth and relies on subsequent processes to determine focal depths and refine hypocenter locations. In our study, almost all the synthetic training and testing events are shallower than 30 km, and the average depth error for SUGAR is 4 km, smaller than the 5 km error of the manual catalog. The shallowest (<3 km) and deepest (>17 km) events have larger average depth errors of 7 km, which indicates that SUGAR is most reliable for events close to the searching depth (10 km) and yields larger uncertainties for much shallower or deeper earthquakes. The 7 km depth error is still acceptable given the complexity of the testing data set and the sparse station coverage in the Kaikōura area. As a result, even though SUGAR is less sensitive to subduction interface events beneath the 2016 crustal fault ruptures (expected at the 25–30 km depth), it is capable of identifying them in any significant abundance. We did not train or test the model with deeper events but we would expect that when earthquakes occur in a wider depth range, for example, 0–60 km, the current SUGAR workflow might not be as effective. A practical solution may be to train and apply different neural networks for different depth ranges (e.g., one for 0–30 km, the other for 30–60 km). We would need four dimensional convolution to fully consider depth in the initial search, which is not currently realistic given the high demands on GPU memory.

The core of SUGAR is a deep neural network, specifically the 3D U-Net, trained using synthetic brightness videos. The same approach may be adopted at different spatial scales, from smaller clusters of induced seismicity to regional or nationwide earthquake monitoring. The same neural network structure might also be suitable for studies of other time-series data, such as tracking the rupture process of large earthquakes with back projection (Ishii et al., 2005), locating tremors and low frequency earthquakes (Kao et al., 2009; Shelly et al., 2006), landslides (Lee et al., 2019) and volcano-tectonic events (Yoma et al., 2022), and evaluating ground deformation or infrastructure damage through InSAR or remotely sensed images (Milillo et al., 2018). The novel idea of visualizing data in a time series of images (or 3D images) and applying computer vision methods may have even wider applications to challenging problems in other disciplines beyond Earth science.

6. Conclusions

We propose a novel approach for seismic source detection by computer vision and combine it with traditional earthquake location techniques to provide a complete workflow for earthquake detection and location (SUGAR). With its ability to solve complex phase association problems, it performs better at locating aftershock sequences than human analysts and other deep-learning based and non-AI workflows. We apply SUGAR to the complex 2016 Kaikōura earthquake sequence and detect 67,660 earthquakes from 13 November to 31 December 2016, providing the most complete catalog of the immediate aftershock sequence to date, independent from all existing catalogs. The SUGAR catalog shows continuous clusters of seismicity under the highly segmented surface ruptures and reveals a connection between well-studied onshore faults and a hitherto poorly characterized offshore thrust. The diffuse nature of the seismicity indicates widespread secondary faulting extending well beyond the normal kilometric-width damage zones associated with the major faults.

Data Availability Statement

Waveform data of the Kaikōura sequence are available in the GeoNet database (GNS Science, 2021). The ObsPy package (Krischer et al., 2015) and Matplotlib (Hunter, 2007) were used in the study. Topography data (for plots) were downloaded from GMTSAR (Sandwell et al., 2011a, 2011b). The SUGAR code, sample data, and phases associated in the SUGAR catalog are available on Zenodo (Tan, 2024) and GitHub “https://github.com/tan-fengzhou/SUGAR_kaioura”.

Acknowledgments

We thank editor Rachel Abercrombie, the associate editor, reviewer Calum Chamberlain and another anonymous reviewer for their constructive suggestions. We acknowledge the New Zealand GeoNet programme and its sponsors EQC, GNS Science, LINZ, NEMA and MBIE for providing data used in this study. This study benefited from constructive discussions with Ramin M.H. Dokht, Shenyang Ge, Jerome Salichon, Andrew Schaeffer, Kelin Wang, and Hongyu Yu. This research is funded by: University of Victoria Graduate Award (FT); Mitacs Research Training Award (FT); The Charles S. Humphrey Graduate Student Award (FT); Alexander and Helen Stafford MacCarthy Muir Graduate Scholarship (FT); Melva J. Hanson Graduate Scholarship (FT); M.A. and D.E. Breckenridge Graduate Awards (FT); David McGillivray Scholarship in Science (FT); Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2019-04148 (HK); Natural Resources Canada (NRCan) Environmental Geoscience Program (HK); NRCan Public Safety Geoscience Program (HK); NSERC Discovery Grant RGPIN-2018-03788 (KMY); Canada Research Chair Program (EN); NSERC Discovery Grant RGPIN-2017-04029 (EN).

References

- Ando, R., & Kaneko, Y. (2018). Dynamic rupture simulation reproduces spontaneous multifault rupture and arrest during the 2016 Mw 7.9 Kaikōura earthquake. *Geophysical Research Letters*, 45(23), 12875–12883. <https://doi.org/10.1029/2018GL080550>
- Aziz Zanjani, F., Lin, G., & Thurber, C. H. (2021). Nested regional-global seismic tomography and precise earthquake relocation along the Hikurangi subduction zone, New Zealand. *Geophysical Journal International*, 227(3), 1567–1590. <https://doi.org/10.1093/gji/ggab294>
- Bai, Y., Lay, T., Cheung, K. F., & Ye, L. (2017). Two regions of seafloor deformation generated the tsunami for the 13 November 2016, Kaikōura, New Zealand earthquake. *Geophysical Research Letters*, 44(13), 6597–6606. <https://doi.org/10.1002/2017gl073717>
- Baillard, C., Crawford, W. C., Ballu, V., Hibert, C., & Mangeney, A. (2014). An automatic Kurtosis-based P- and S-phase picker designed for local seismic networks. *Bulletin of the Seismological Society of America*, 104(1), 394–409. <https://doi.org/10.1785/0120120347>
- Cesca, S., Zhang, Y., Mouslopoulou, V., Wang, R., Saul, J., Savage, M., et al. (2017). Complex rupture process of the Mw 7.8, 2016, Kaikōura earthquake, New Zealand, and its aftershock sequence. *Earth and Planetary Science Letters*, 478, 110–120. <https://doi.org/10.1016/j.epsl.2017.08.024>
- Chamberlain, C. J., Frank, W. B., Lanza, F., Townend, J., & Warren-Smith, E. (2021). Illuminating the pre-co-and post-seismic phases of the 2016 Mw 7.8 Kaikōura earthquake with 10 years of seismicity. *Journal of Geophysical Research: Solid Earth*, 126(8), e2021JB022304. <https://doi.org/10.1029/2021jb022304>
- Chen, S., Ma, K., & Zheng, Y. (2019). MED3D: Transfer learning for 3D medical image analysis.
- Chinchor, N. (1992). Muc-4 evaluation metrics. In *Proceedings of the 4th conference on message understanding* (pp. 22–29). Association for Computational Linguistics. <https://doi.org/10.3115/1072064.1072067>
- Çiçek, O., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. arXiv:1606.06650 [cs.CV].
- Clark, K. J., Nissen, E. K., Howarth, J. D., Hamling, I. J., Mountjoy, J. J., Ries, W. F., et al. (2017). Highly variable coastal deformation in the 2016 MW 7.8 Kaikōura earthquake reflects rupture complexity along a transpressional plate boundary. *Earth and Planetary Science Letters*, 474, 334–344. <https://doi.org/10.1016/j.epsl.2017.06.048>
- Clinton, J. F., Ceylan, S., van Driel, M., Giardini, D., Stähler, S. C., Böse, M., et al. (2021). The Marsquake catalogue from InSight, sols 0–478. *Physics of the Earth and Planetary Interiors*, 310, 106595. <https://doi.org/10.1016/j.pepi.2020.106595>
- DeVries, P. M. R., Viégas, F., Wattenberg, M., & Meade, B. J. (2018). Deep learning of aftershock patterns following large earthquakes. *Nature*, 560(7720), 632–634. <https://doi.org/10.1038/s41586-018-0438-y>
- Diederichs, A., Nissen, E. K., Lajoie, L. J., Langridge, R. M., Malireddi, S. R., Clark, K. J., et al. (2019). Unusual kinematics of the Papatea fault (2016 Kaikōura earthquake) suggest anelastic rupture. *Science Advances*, 5(10), eaax5703. <https://doi.org/10.1126/sciadv.aax5703>
- Dokht, R. M. H., Kao, H., Ghofrani, H., & Visser, R. (2022). Combining deep learning and the Source-Scanning Algorithm for improved seismic monitoring. *Bulletin of the Seismological Society of America*, 112(5), 2312–2326. <https://doi.org/10.1785/0120220007>
- Eberhart-Phillips, D., Bannister, S., Reyners, M., & Bourguignon, S. (2022). New Zealand Wide model 2.3 seismic velocity model for New Zealand [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.6568301>
- Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. *Nature Geoscience*, 11(7), 531–535. <https://doi.org/10.1038/s41561-018-0145-1>
- Font, Y., Kao, H., Lallemand, S., Liu, C.-S., & Chiao, L.-Y. (2004). Hypocentre determination offshore of eastern Taiwan using the Maximum Intersection method. *Geophysical Journal International*, 158(2), 655–675. <https://doi.org/10.1111/j.1365-246X.2004.02317.x>
- Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, 165(1), 149–166. <https://doi.org/10.1111/j.1365-246X.2006.02865.x>
- GNS Science. (1970). GeoNet Aotearoa New Zealand earthquake catalogue [Dataset]. GNS Science. <https://doi.org/10.21420/0S8P-TZ38>
- GNS Science. (2016). New Zealand active faults database 1:250,000 scale [Dataset]. GNS Science. Retrieved from <https://www.gns.cri.nz/data-and-resources/new-zealand-active-faults-database/>
- GNS Science. (2021). GeoNet Aotearoa New Zealand seismic digital waveform dataset [Dataset]. GNS Science. <https://doi.org/10.21420/G19Y-9D40>
- Gong, J., & Fan, W. (2022). Seismicity, fault architecture, and slip mode of the westernmost Gofar transform fault. *Journal of Geophysical Research: Solid Earth*, 127(11), e2022JB024918. <https://doi.org/10.1029/2022JB024918>
- Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, 574(7777), 193–199. <https://doi.org/10.1038/s41586-019-1606-4>
- Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4), 185–188. <https://doi.org/10.1785/bssa0340040185>
- Hainzl, S. (2016). Rate-dependent incompleteness of earthquake catalogs. *Seismological Research Letters*, 87(2A), 337–344. <https://doi.org/10.1785/0220150211>
- Hamling, I. J. (2020). A review of the 2016 Kaikōura earthquake: Insights from the first 3 years. *Journal of the Royal Society of New Zealand*, 50(2), 226–244. <https://doi.org/10.1080/03036758.2019.1701048>
- Hamling, I. J., Hreinsdóttir, S., Clark, K., Elliott, J., Liang, C., Fielding, E., et al. (2017). Complex multifault rupture during the 2016 Mw 7.8 Kaikōura earthquake, New Zealand. *Science*, 356(6334), eaam7194. <https://doi.org/10.1126/science.aam7194>
- Hayes, G., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2 - A comprehensive subduction zone geometry model. *U.S. Geological Survey*, 362(6410), 58–61. <https://doi.org/10.5066/F7PV6JNV>
- Herman, M. W., Furlong, K. P., & Benz, H. M. (2023). Substantial upper plate faulting above a shallow subduction megathrust earthquake: Mechanics and implications of the surface faulting during the 2016 Kaikōura, New Zealand, earthquake. *Tectonics*, 42(5), e2022TC007645. <https://doi.org/10.1029/2022TC007645>

- Hollingsworth, J., Ye, L., & Avouac, J.-P. (2017). Dynamically triggered slip on a splay fault in the Mw 7.8, 2016 Kaikoura (New Zealand) earthquake. *Geophysical Research Letters*, 44(8), 3517–3525. <https://doi.org/10.1002/2016gl072228>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment [Software]. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Hutchinson, J. A., Zhu, C., Bradley, B. A., Lee, R. L., Wotherspoon, L. M., Dupuis, M., et al. (2024). The 2023 New Zealand ground-motion database. *Bulletin of the Seismological Society of America*, 114(1), 291–310. <https://doi.org/10.1785/0120230184>
- Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature*, 435(7044), 933–936. <https://doi.org/10.1038/nature03675>
- Kaiser, A., Balfour, N., Fry, B., Holden, C., Litchfield, N., Gerstenberger, M., et al. (2017). The 2016 Kaikōura, New Zealand, earthquake: Preliminary seismological report. *Seismological Research Letters*, 88(3), 727–739. <https://doi.org/10.1785/0220170018>
- Kao, H., & Shan, S.-J. (2004). The source-scanning algorithm: Mapping the distribution of seismic sources in time and space. *Geophysical Journal International*, 157(2), 589–594. <https://doi.org/10.1111/j.1365-246x.2004.02276.x>
- Kao, H., Shan, S.-J., Dragert, H., & Rogers, G. (2009). Northern Cascadia episodic tremor and slip: A decade of tremor observations from 1997 to 2007. *Journal of Geophysical Research*, 114(B11), B00A12. <https://doi.org/10.1029/2008jb006046>
- Kennett, B. L. N., & Engdahl, E. R. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2), 429–465. <https://doi.org/10.1111/j.1365-246X.1991.tb06724.x>
- Kilb, D., Gomberg, J., & Bodin, P. (2000). Triggering of earthquake aftershocks by dynamic stresses. *Nature*, 408(6812), 570–574. <https://doi.org/10.1038/35046046>
- Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., & Wassermann, J. (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem [Software]. *Computational Science & Discovery*, 8(1), 014003. <https://doi.org/10.1088/1749-4699/8/1/014003>
- Langridge, R. M., Rowland, J., Villamor, P., Mountjoy, J., Townsend, D. B., Nissen, E., et al. (2018). Coseismic rupture and preliminary slip estimates for the Papatea Fault and its role in the 2016 Mw 7.8 Kaikōura, New Zealand, earthquake. *Bulletin of the Seismological Society of America*, 108(3B), 1596–1622. <https://doi.org/10.1785/0120170336>
- Lanza, F., Chamberlain, C. J., Jacobs, K., Warren-Smith, E., Godfrey, H. J., Kortink, M., et al. (2019). Crustal fault connectivity of the Mw 7.8 2016 Kaikōura Earthquake constrained by aftershock relocations. *Geophysical Research Letters*, 46(12), 6487–6496. <https://doi.org/10.1029/2019gl082780>
- Latham, G., Ewing, M., Dorman, J., Lammlein, D., Press, F., Toksoz, N., et al. (1971). Moonquakes. *Science*, 174(4010), 687–692. <https://doi.org/10.1126/science.174.4010.687>
- Lay, T., & Wallace, T. C. (1995). Seismic sources. In *Modern global seismology* (pp. 9–10). Academic Press.
- Lee, E.-J., Liao, W.-Y., Lin, G.-W., Chen, P., Mu, D., & Lin, C.-W. (2019). Towards automated real-time detection and location of large-scale landslides through seismic waveform back projection. *Geofluids*, 2019, 1–14. <https://doi.org/10.1155/2019/1426019>
- Li, L., Tan, J., Schwarz, B., Stanek, F., Poiata, N., Shi, P., et al. (2020). Recent advances and challenges of waveform-based seismic location methods at multiple scales. *Reviews of Geophysics*, 58(1), e2019RG000667. <https://doi.org/10.1029/2019rg000667>
- Liao, W.-Y., Lee, E.-J., Mu, D., & Chen, P. (2022). Toward fully autonomous seismic networks: Backprojecting deep learning-based phase time functions for earthquake monitoring on continuous recordings. *Seismological Research Letters*, 93(3), 1880–1894. <https://doi.org/10.1785/0220210274>
- Liao, Y., Honn, K., Andreas, R., Shu-Kun, H., & Huang, B.-S. (2012). Delineating complex spatiotemporal distribution of earthquake aftershocks: An improved source-scanning algorithm. *Geophysical Journal International*, 189(3), 1753–1770. <https://doi.org/10.1111/j.1365-246X.2012.05457.x>
- Litchfield, N. J., Villamor, P., Dissen, R. J. V., Nicol, A., Barnes, P. M., A. Barrell, D. J., et al. (2018). Surface rupture of multiple crustal faults in the 2016 Mw 7.8 Kaikōura, New Zealand, earthquake. *Bulletin of the Seismological Society of America*, 108(3B), 1496–1520. <https://doi.org/10.1785/0120170300>
- McBrearty, I. W., & Beroza, G. C. (2023). Earthquake phase association with graph neural networks. *Bulletin of the Seismological Society of America*, 113(2), 524–547. <https://doi.org/10.1785/0120220182>
- McBrearty, I. W., Gomberg, J., Delorey, A. A., & Johnson, P. A. (2019). Earthquake arrival association with backprojection and graph theory. *Bulletin of the Seismological Society of America*, 109(6), 2510–2531. <https://doi.org/10.1785/0120190081>
- Milillo, P., Giardina, G., DeJong, M. J., Perissin, D., & Milillo, G. (2018). Multi-temporal InSAR structural damage assessment: The London Crossrail case study. *Remote Sensing*, 10(2), 287. <https://doi.org/10.3390/rs10020287>
- Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, 11(1), 3952. <https://doi.org/10.1038/s41467-020-17591-w>
- Mouslopoulou, V., Saltogianni, V., Nicol, A., Oncken, O., Begg, J., Babeyko, A., et al. (2019). Breaking a subduction-termination from top to bottom: The large 2016 Kaikōura earthquake, New Zealand. *Earth and Planetary Science Letters*, 506, 221–230. <https://doi.org/10.1016/j.epsl.2018.10.020>
- Münchmeyer, J., Bindi, D., Leser, U., & Tilmann, F. (2021). The transformer earthquake alerting model: A new versatile approach to earthquake early warning. *Geophysical Journal International*, 225(1), 646–656. <https://doi.org/10.1093/gji/ggaa609>
- Nicol, A., Begg, J., Saltogianni, V., Mouslopoulou, V., Oncken, O., & Howell, A. (2022). Uplift and fault slip during the 2016 Kaikōura earthquake and late quaternary, Kaikōura Peninsula, New Zealand. *New Zealand Journal of Geology and Geophysics*, 66(2), 1–16. <https://doi.org/10.1080/00288306.2021.2021955>
- Nicol, A., Khajavi, N., Pettinga, J. R., Fenton, C., Stahl, T., Bannister, S., et al. (2018). Preliminary geometry, displacement, and kinematics of fault ruptures in the epicentral region of the 2016 Mw 7.8 Kaikōura, New Zealand, earthquake. *Bulletin of the Seismological Society of America*, 108(3B), 1521–1539. <https://doi.org/10.1785/0120170329>
- Park, Y., Beroza, G. C., & Ellsworth, W. L. (2022). Basement fault activation before larger earthquakes in Oklahoma and Kansas. *The Seismic Record*, 2(3), 197–206. <https://doi.org/10.1785/0320220020>
- Perfettini, H., & Avouac, J.-P. (2007). Modeling afterslip and aftershocks following the 1992 Landers earthquake. *Journal of Geophysical Research*, 112(B7), B07409. <https://doi.org/10.1029/2006JB004399>
- Perrin, C., Waldhauser, F., & Scholz, C. H. (2021). The shear deformation zone and the smoothing of faults with displacement. *Journal of Geophysical Research: Solid Earth*, 126(5), e2020JB020447. <https://doi.org/10.1029/2020jb020447>
- Reasenberg, P., & Oppenheimer, D. H. (1985). *FPPFIT, FPPLT and FPPLAGE: Fortran computer programs for calculating and displaying earthquake fault-plane solutions* (Tech. Rep. No. 85-739). U.S. Geological Survey.
- Rhoades, D. A., Christoffersen, A., Bourguignon, S., Ristau, J., & Salichon, J. (2021). A depth-dependent local magnitude scale for New Zealand earthquakes consistent with moment magnitude. *Bulletin of the Seismological Society of America*, 111(2), 1056–1066. <https://doi.org/10.1785/0120200252>

- Richter, C. F. (1935). An instrumental earthquake magnitude scale. *Bulletin of the Seismological Society of America*, 25(1), 1–32. <https://doi.org/10.1785/bssa0250010001>
- Rolandone, F., Bürgmann, R., & Nadeau, R. M. (2004). The evolution of the seismic-aseismic transition during the earthquake cycle: Constraints from the time-dependent depth distribution of aftershocks. *Geophysical Research Letters*, 31(23), L23610. <https://doi.org/10.1029/2004GL021379>
- Ross, Z. E., Idini, B., Jia, Z., Stephenson, O. L., Zhong, M., Wang, X., et al. (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. *Science*, 366(6463), 346–351. <https://doi.org/10.1126/science.aaz0109>
- Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in Southern California. *Science*, 364(6442), 767–771. <https://doi.org/10.1126/science.aaw6888>
- Ross, Z. E., Yue, Y., Meier, M.-A., Hauksson, E., & Heaton, T. H. (2019). PhaseLink: A deep learning approach to seismic phase association. *Journal of Geophysical Research: Solid Earth*, 124(1), 856–869. <https://doi.org/10.1029/2018JB016674>
- Sandwell, D., Mellors, R., Tong, X., Wei, M., & Wessel, P. (2011a). GMTSAR: An InSAR processing system based on generic mapping tools (Tech. Rep.) [Dataset]. Scripps Institution of Oceanography. Retrieved from <https://topex.ucsd.edu/gmtsar/demgen/>
- Sandwell, D., Mellors, R., Tong, X., Wei, M., & Wessel, P. (2011b). Open radar interferometry software for mapping surface deformation. *Eos, Transactions American Geophysical Union*, 92(28), 234. <https://doi.org/10.1029/2011EO280002>
- 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. <https://doi.org/10.1785/0220190309>
- Shelly, D. R., Beroza, G. C., Ide, S., & Nakamura, S. (2006). Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature*, 442(7099), 188–191. <https://doi.org/10.1038/nature04931>
- Shen, H., & Shen, Y. (2021). Array-based convolutional neural networks for automatic detection and 4D localization of earthquakes in Hawai'i. *Seismological Research Letters*, 92(5), 2961–2971. <https://doi.org/10.1785/0220200419>
- Shi, P., Grigoli, F., Lanza, F., Beroza, G. C., Scarabello, L., & Wiemer, S. (2022). MALMI: An automated earthquake detection and location workflow based on machine learning and waveform migration. *Seismological Research Letters*, 93(5), 2467–2483. <https://doi.org/10.1785/0220220071>
- Si, X., Wu, X., Li, Z., Wang, S., & Zhu, J. (2024). An all-in-one seismic phase picking, location, and association network for multi-task multi-station earthquake monitoring. *Communications Earth & Environment*, 5(1), 1–13. <https://doi.org/10.1038/s43247-023-01188-4>
- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, 15(1), 29. <https://doi.org/10.1186/s12880-015-0068-x>
- Tamaribuchi, K. (2018). Evaluation of automatic hypocenter determination in the JMA unified catalog. *Earth Planets and Space*, 70(1), 141. <https://doi.org/10.1186/s40623-018-0915-4>
- Tan, F. (2024). Source untangle guided by artificial intelligence image recognition (sugar) [Software]. Zenodo. <https://doi.org/10.5281/zenodo.10937462>
- Tan, F., Kao, H., Nissen, E., & Eaton, D. (2019). Seismicity-scanning based on navigated automatic phase-picking. *Journal of Geophysical Research: Solid Earth*, 124(4), 3802–3818. <https://doi.org/10.1029/2018jb017050>
- Tan, F., Kao, H., Nissen, E., & Visser, R. (2020). Tracking earthquake sequences in real time: Application of seismicity-scanning based on navigated automatic phase-picking (S-SNAP) to the 2019 Ridgecrest, California sequence. *Geophysical Journal International*, 223(3), 1511–1524. <https://doi.org/10.1093/gji/ggaa387>
- Townsend, D. B., & Little, T. A. (1998). Pliocene-Quaternary deformation and mechanisms of near-surface strain close to the eastern tip of the Clarence Fault, northeast Marlborough, New Zealand. *New Zealand Journal of Geology and Geophysics*, 41(4), 401–417. <https://doi.org/10.1080/00288306.1998.9514819>
- Trugman, D. T., & Shearer, P. M. (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. <https://doi.org/10.1785/0220160188>
- Uhrhammer, R. A., Hellweg, M., Hutton, K., Lombard, P., Walters, A. W., Hauksson, E., & Oppenheimer, D. (2011). California integrated seismic network (CISN) local magnitude determination in California and Vicinity California integrated seismic network local magnitude determination in California and vicinity. *Bulletin of the Seismological Society of America*, 101(6), 2685–2693. <https://doi.org/10.1785/0120100106>
- Ulrich, T., Gabriel, A.-A., Ampuero, J.-P., & Xu, W. (2019). Dynamic viability of the 2016 Mw 7.8 Kaikōura earthquake cascade on weak crustal faults. *Nature Communications*, 10(1), 1213. <https://doi.org/10.1038/s41467-019-10912-w>
- Wang, T., Jiao, L., Tappouner, P., Shi, X., & Wei, S. (2020). Space imaging Geodesy reveals near circular, coseismic block rotation during the 2016 Mw 7.8 Kaikōura earthquake, New Zealand. *Geophysical Research Letters*, 47(22), e2020GL090206. <https://doi.org/10.1029/2020GL090206>
- Wang, T., Wei, S., Shi, X., Qiu, Q., Li, L., Peng, D., et al. (2018). The 2016 Kaikōura earthquake: Simultaneous rupture of the subduction interface and overlying faults. *Earth and Planetary Science Letters*, 482, 44–51. <https://doi.org/10.1016/j.epsl.2017.10.056>
- Williams, C. A., Eberhart-Phillips, D., Bannister, S., Barker, D. H. N., Henrys, S., Reyners, M., & Sutherland, R. (2013). Revised interface geometry for the Hikurangi subduction zone, New Zealand. *Seismological Research Letters*, 84(6), 1066–1073. <https://doi.org/10.1785/0220130035>
- Woessner, J., & Wiemer, S. (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. <https://doi.org/10.1785/0120040007>
- Woollam, J., Rietbrock, A., Leitloff, J., & Hinz, S. (2020). HEX: Hyperbolic Event eXtractor, a seismic phase associator for highly active seismic regions. *Seismological Research Letters*, 91(5), 2769–2778. <https://doi.org/10.1785/0220200037>
- Xu, W., Feng, G., Meng, L., Zhang, A., Ampuero, J. P., Bürgmann, R., & Fang, L. (2018). Transpressional rupture cascade of the 2016 Mw 7.8 Kaikoura earthquake, New Zealand. *Journal of Geophysical Research: Solid Earth*, 123(3), 2396–2409. <https://doi.org/10.1002/2017jb015168>
- Yoma, N. B., Wuth, J., Pinto, A., de Celis, N., Celis, J., Huemupan, F., & Fustos-Toribio, I. J. (2022). End-to-end LSTM based estimation of volcano event epicenter localization. *Journal of Volcanology and Geothermal Research*, 429, 107615. <https://doi.org/10.1016/j.jvolgeores.2022.107615>
- Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. *Science Advances*, 1(11), e1501057. <https://doi.org/10.1126/sciadv.1501057>
- Zhang, M., Ellsworth, W. L., & Beroza, G. C. (2019). Rapid earthquake association and location. *Seismological Research Letters*, 90(6), 2276–2284. <https://doi.org/10.1785/0220190052>
- Zhang, X., Zhang, J., Yuan, C., Liu, S., Chen, Z., & Li, W. (2020). Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method. *Scientific Reports*, 10(1), 1941. <https://doi.org/10.1038/s41598-020-58908-5>

- Zhang, X., Zhang, M., & Tian, X. (2021). Real-time earthquake early warning with deep learning: Application to the 2016 M 6.0 central Apennines, Italy earthquake. *Geophysical Research Letters*, 48(5), 2020GL089394. <https://doi.org/10.1029/2020GL089394>
- Zhu, W., & Beroza, G. C. (2019). PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1), 261–273. <https://doi.org/10.1093/gji/ggy423>
- Zhu, W., McBrearty, I. W., Mousavi, S. M., Ellsworth, W. L., & Beroza, G. C. (2022). Earthquake phase association using a Bayesian Gaussian mixture model. *Journal of Geophysical Research: Solid Earth*, 127(5), e2021JB023249. <https://doi.org/10.1029/2021jb023249>

References From the Supporting Information

- Eberhart-Phillips, D., Bannister, S., Reyners, M., & Henrys, S. (2020). New Zealand wide model 2.2 seismic velocity and Qs and Qp models for New Zealand [Dataset]. *Zenodo*. Retrieved from <https://zenodo.org/records/3779523>
- Harrington, R. M., & Brodsky, E. E. (2009). Source duration scales with magnitude differently for earthquakes on the San Andreas Fault and on secondary faults in Parkfield, CaliforniaSource duration scales with magnitude differently on the San Andreas Fault and on secondary faults. *Bulletin of the Seismological Society of America*, 99(4), 2323–2334. <https://doi.org/10.1785/0120080216>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv:1412.6980 [cs.LG].
- Zhu, L., & Rivera, L. A. (2002). A note on the dynamic and static displacements from a point source in multilayered media. *Geophysical Journal International*, 148(3), 619–627. <https://doi.org/10.1046/j.1365-246x.2002.01610.x>