

Real-Time Detection of Volcanic Unrest and Eruption at Axial Seamount Using Machine Learning

Kaiwen Wang^{*1}, Felix Waldhauser¹, David Schaff¹, Maya Tolstoy², William S. D. Wilcock², and Yen Joe Tan³

Abstract

Axial Seamount, an extensively instrumented submarine volcano, lies at the intersection of the Cobb–Eickelberg hot spot and the Juan de Fuca ridge. Since late 2014, the Ocean Observatories Initiative (OOI) has operated a seven-station cabled ocean bottom seismometer (OBS) array that captured Axial's last eruption in April 2015. This network streams data in real-time, facilitating seismic monitoring and analysis for volcanic unrest detection and eruption forecasting. In this study, we introduce a machine learning (ML)-based real-time seismic monitoring framework for Axial Seamount. Combining both supervised and unsupervised ML and double-difference techniques, we constructed a comprehensive, high-resolution earthquake catalog while effectively discriminating between various seismic and acoustic events. These events include earthquakes generated by different physical processes, acoustic signals of lava–water interaction, and oceanic sources such as whale calls. We first built a labeled ML-based earthquake catalog that extends from November 2014 to the end of 2021 and then implemented real-time monitoring and seismic analysis starting in 2022. With the rapid determination of high-resolution earthquake locations and the capability to track potential precursory signals and coeruption indicators of magma outflow, this system may improve eruption forecasting by providing short-term constraints on Axial's next eruption. Furthermore, our work demonstrates an effective application that integrates unsupervised learning for signal discrimination in real-time operation, which could be adapted to other regions for volcanic unrest detection and enhanced eruption forecasting.

Cite this article as Wang, K., F. Waldhauser, D. Schaff, M. Tolstoy, W. S. D. Wilcock, and Y. J. Tan (2024). Real-Time Detection of Volcanic Unrest and Eruption at Axial Seamount Using Machine Learning, *Seismol. Res. Lett.* **XX**, 1–12, doi: [10.1785/0220240086](https://doi.org/10.1785/0220240086).

[Supplemental Material](#)

Introduction

Earthquakes hold fundamental information about the structure and dynamics of active volcanoes and the processes that control eruptions. Submarine volcanoes, in particular, are good cases for studying volcano dynamics because they erupt in short intervals (Rubin *et al.*, 2012; Sinton *et al.*, 2002) and active source seismic profiling allows for detailed imaging of seafloor structures (Park *et al.*, 2007; Arnulf *et al.*, 2014; Chrapkiewicz *et al.*, 2022). However, seismic monitoring using ocean-bottom seismometers (OBS) is technically challenging and expensive, and the continuous waveforms include signals from various seismic sources and ocean noise. At Axial Seamount, an active submarine volcano located 1400 m below the sea surface at the intersection of the Juan de Fuca ridge and the Cobb–Eickelberg hot spot, the Ocean Observatories Initiative (OOI) operates a seven-station cabled OBS array (Kelley *et al.*, 2014). This OBS array records various signals from earthquakes, marine mammals, and airgun shots from active seismic experiments, among other sources. The

earthquakes have significantly contributed to the understanding of the structure and inner workings of the volcano, which last erupted in April 2015, about four months after the network started recording. A complex ring fault system is imaged (Wilcock *et al.*, 2016; Waldhauser *et al.*, 2020) above a shallow magma chamber (Arnulf *et al.*, 2014). The OBS recordings are complemented by other geophysical, chemical, temperature, and video camera measurements onsite (Chadwick *et al.*, 2016; Smith *et al.*, 2018; Wilcock *et al.*, 2018) operated by the OOI, making Axial Seamount one of the best-instrumented submarine volcanoes.

1. Lamont–Doherty Earth Observatory, Columbia University, Palisades, New York, U.S.A., <https://orcid.org/0000-0001-7838-6984> (KW); <https://orcid.org/0000-0002-1286-9737> (FW); 2. School of Oceanography, University of Washington, Seattle, Washington, U.S.A., <https://orcid.org/0000-0001-7224-7246> (WSDW); 3. Earth and Environmental Sciences Programme, Faculty of Science, The Chinese University of Hong Kong, Hong Kong S.A.R., China, <https://orcid.org/0000-0001-6377-7886> (YJT)

*Corresponding author: kw2988@columbia.edu

© Seismological Society of America

In the four months before and during the eruption of Axial volcano in April 2015, the OBS array recorded ~136,000 earthquakes. Identification and classification of these events is challenging due to both the high spatial density and the high rate of seismicity, and the diverse nature of brittle failure leading up to an eruption. In addition, eruption-specific non-earthquake signals, such as acoustic signals generated from hot lava reaching the sea floor that help map areas of lava flow (Tan *et al.*, 2016; Wilcock *et al.*, 2016; Le Saout *et al.*, 2020), need to be identified and separated from regular earthquakes. The detection, classification, and characterization of seismic sources, and especially precursory signals, would have to be performed rapidly, ideally in near-real time, if the information gained from the analysis is to be used in forecasting future eruptions for hazard mitigation purposes.

Here, we take advantage of recent advances in both supervised and unsupervised machine learning (ML) methods (Holtzman *et al.*, 2018; Bergen *et al.*, 2019; Beroza *et al.*, 2021; Mousavi and Beroza, 2023) to significantly improve the event detection and discrimination capabilities in a monitoring framework. This work builds on current monitoring efforts at Axial that use standard detection and location methods (Wilcock *et al.*, 2016) and correlation-based double-difference relocation (Waldhauser *et al.*, 2020). We present a suite of tools that rapidly process and analyze continuous waveform data to produce high-precision, deep-magnitude event catalogs in near-real time. Specifically, the new system is trained to detect potential precursory signals of mixed-frequency earthquakes (MFEs), which have relatively lower-frequency content compared with the volcanotectonic earthquakes, presumably caused by the movement of volatiles or magma as opposed to brittle failures triggered by stress changes on faults (Wang *et al.*, 2024). These signals were discovered using an unsupervised spectral feature extraction method (Holtzman *et al.*, 2018) (Details in [Data and Methods](#) section) before the last Axial eruption in 2015, where they emerged from background seismicity about 15 hr before the lava reached the sea floor (Wang *et al.*, 2024). The method also detects acoustic signals that are generated when hot lava reaches the cold sea floor after an eruption (Tan *et al.*, 2016; Wilcock *et al.*, 2016; Le Saout *et al.*, 2020; Tepp *et al.*, 2020), allowing for rapid tracking of lava flow in time and space.

Data and Methods

Data and current monitoring

We use the continuous seismic data from the OOI-OBS array available through the EarthScope Consortium DMC as an input to our workflow. The cabled OBS array comprises two broadband stations and five short-period stations on top of Axial seamount (Fig. 1b), all data sampled at 200 Hz. Currently, the data are automatically processed in near-real time at the University of Washington (Wilcock *et al.*, 2016) using standard routine analysis with a Kurtosis phase picker (Baillard *et al.*, 2014) and the HYPOINVERSE (HINV)

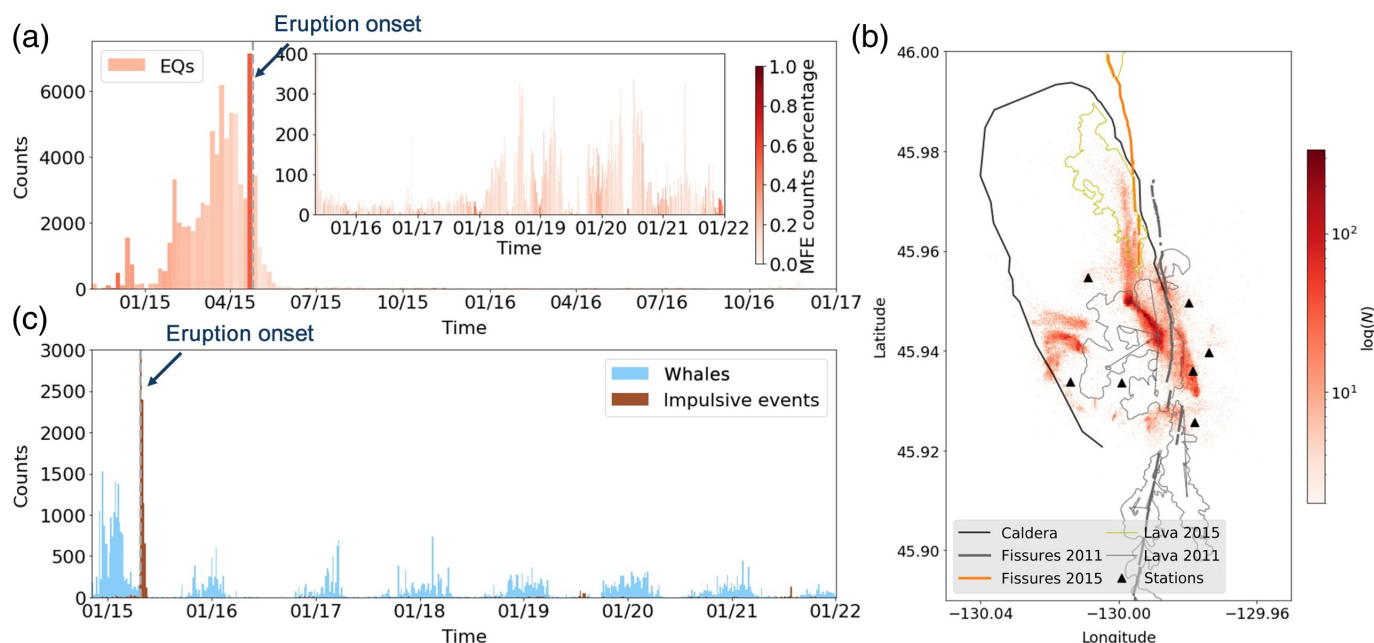
location algorithm (Klein, 2002). Each newly detected and located earthquake is then automatically relocated at the Lamont–Doherty Earth Observatory with respect to a high-resolution earthquake (base) catalog using the near-real-time double-difference relocation system (DD-RT) (Waldhauser *et al.*, 2020; Waldhauser, 2009). The base catalog was computed from a simultaneous double-difference inversion of both Kurtosis picks and precise correlation delay times (Waldhauser *et al.*, 2020), with initial locations derived from a grid-search analysis (NonLinLoc [NLL], Lomax *et al.*, 2000) of the Kurtosis picks in a 3D earthquake tomographic *P*- and *S*-wave model (Baillard *et al.*, 2019). Earthquake magnitudes are calculated by estimating a seismic moment for each phase arrivals and using the median value for all arrivals to compute a moment magnitude (Wilcock *et al.*, 2016).

New processing

Here, we replace the current routine processing step with a supervised ML method (QuakeFlow, Zhu *et al.*, 2023) to detect and characterize seismic events, and an unsupervised ML method (SpecUFEx, Holtzman *et al.*, 2018) to discriminate between various seismic sources. Our ML-based framework is illustrated in Figure 2. Wang *et al.* (2024), in a retroactive analysis, applied these tools to 4 months of continuous waveforms leading up to the April 2015 eruption to develop a catalog of diverse, labeled seismic source types, including precursory MFEs and impulsive signals generated by lava reaching the cold seafloor (Wilcock *et al.*, 2016). This analysis provides the foundation for our ML-based processing.

Base catalog

We developed a new base catalog (2014–2021) for the DD-RT system, essentially following Waldhauser *et al.* (2020), but using PhaseNet (Zhu and Beroza, 2019) for picking *P*- and *S*-phase arrivals in the continuous waveforms, and GaMMA (Zhu *et al.*, 2022) for associating them into seismic events (minimum of five picks required for each event). SpecUFEx (Holtzman *et al.*, 2018) is then used to discriminate between the various types of signals and their underlying sources (details explained in the next subsection), including separating earthquakes from all other seismic sources such as whale calls (Wang *et al.*, 2024). The ML phase picks are then used to locate events with a grid-search method, NLL (Lomax *et al.*, 2000), in a local 3D seismic velocity model (Baillard *et al.*, 2019). The ML catalog is relocated using cross-correlation (Schaff and Waldhauser, 2005) and the double-difference method (Waldhauser and Ellsworth, 2000) using parameters similar to the ones described in Waldhauser *et al.* (2020). Close to 1.4 billion correlation delay times are computed on pairs of filtered (4–50 Hz) seismograms with correlation coefficients $C_f > 0.8$ and hypocentral separation < 2 km. The correlation data together with the delay times formed from the ML picks are inverted for relative locations using the hypoDD algorithm (Waldhauser, 2001) to obtain a



high-precision earthquake catalog to be used as base catalog in the DD-RT monitoring system (Waldhauser, 2009; Waldhauser *et al.*, 2020).

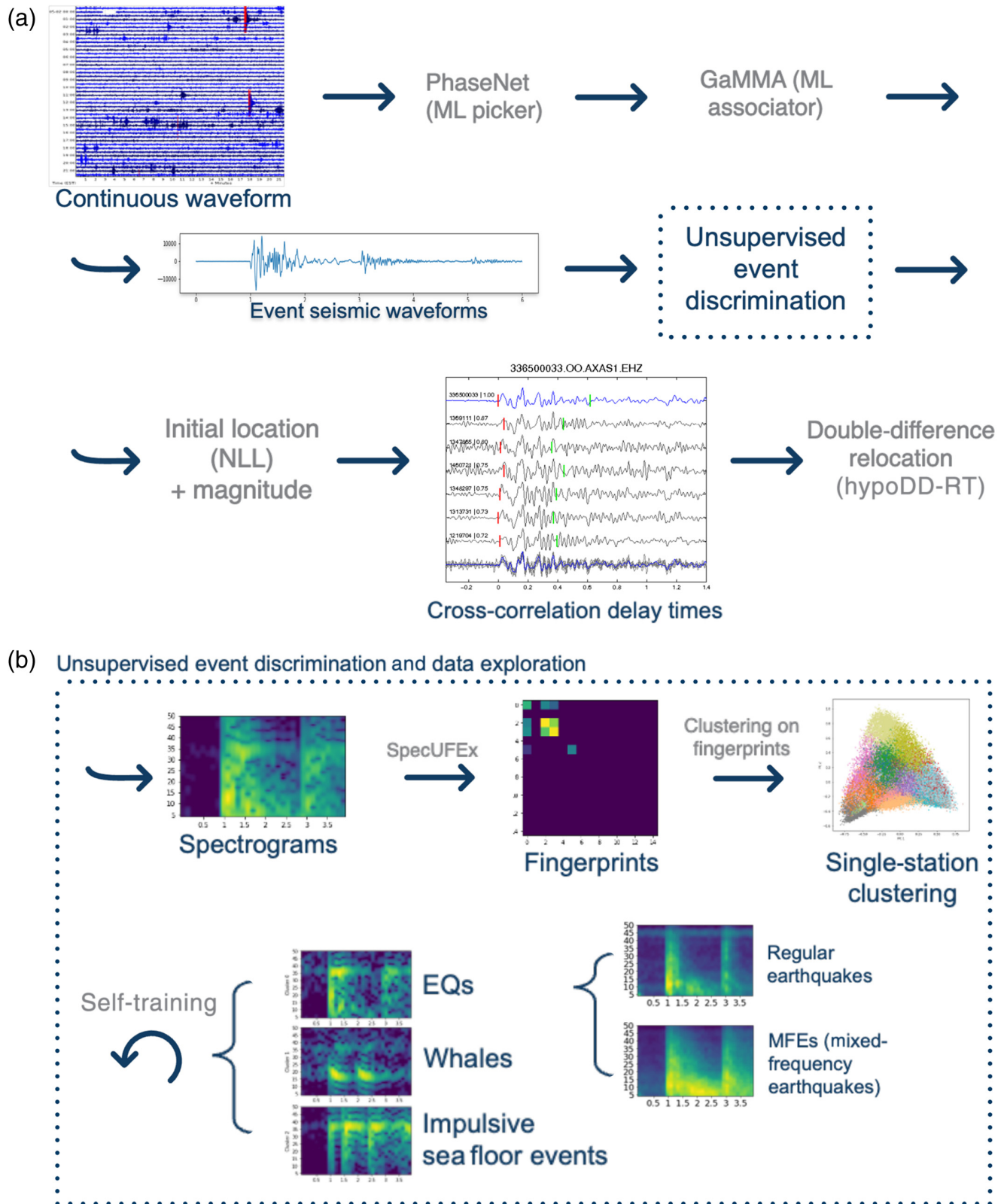
Unsupervised ML for event discrimination

We use the unsupervised ML method SpecUFEx (Holtzman *et al.*, 2018) to discriminate between various types of seismic events recorded by the Axial OBSs. This spectral feature extraction method was originally developed for audio signal recognition (Cotton and Ellis, 2011) and has been adapted to characterize seismic signals in various settings, such as earthquakes in geothermal fields and along crustal faults, acoustic emissions in lab experiments, and icequakes and seismic noises at glaciers (Holtzman *et al.*, 2018, 2021; Sawi *et al.*, 2022). SpecUFEx generates low-dimensional fingerprints for each signal that represent the spectral patterns of the original spectrograms. These fingerprints are then clustered to find groups of similar signals (Fig. 2).

Following Holtzman *et al.* (2018), we converted the waveforms of each event in the base catalog into spectrograms. We use vertical-component recordings and cut the event window from 1 s before the *P* arrival to 3 s after it. This window was chosen to capture the main features of the events (initial arrival, coda, and reflections in the water column, see Fig. 3) while minimizing excessive background noises that may hamper performance. We set the spectrogram frequency range from 4 to 50 Hz to avoid high-frequency instrumental noise and low-frequency microseism noise at the OBS stations. We followed Wang *et al.* (2024) in preprocessing the spectrograms; however, here we used a catalog of all event types (earthquake and non-earthquake signals) as opposed to a subset of pre-eruption earthquakes in their study. We also integrated information from multiple stations and performed array-based analysis in this study, whereas their work focused on single-station clustering.

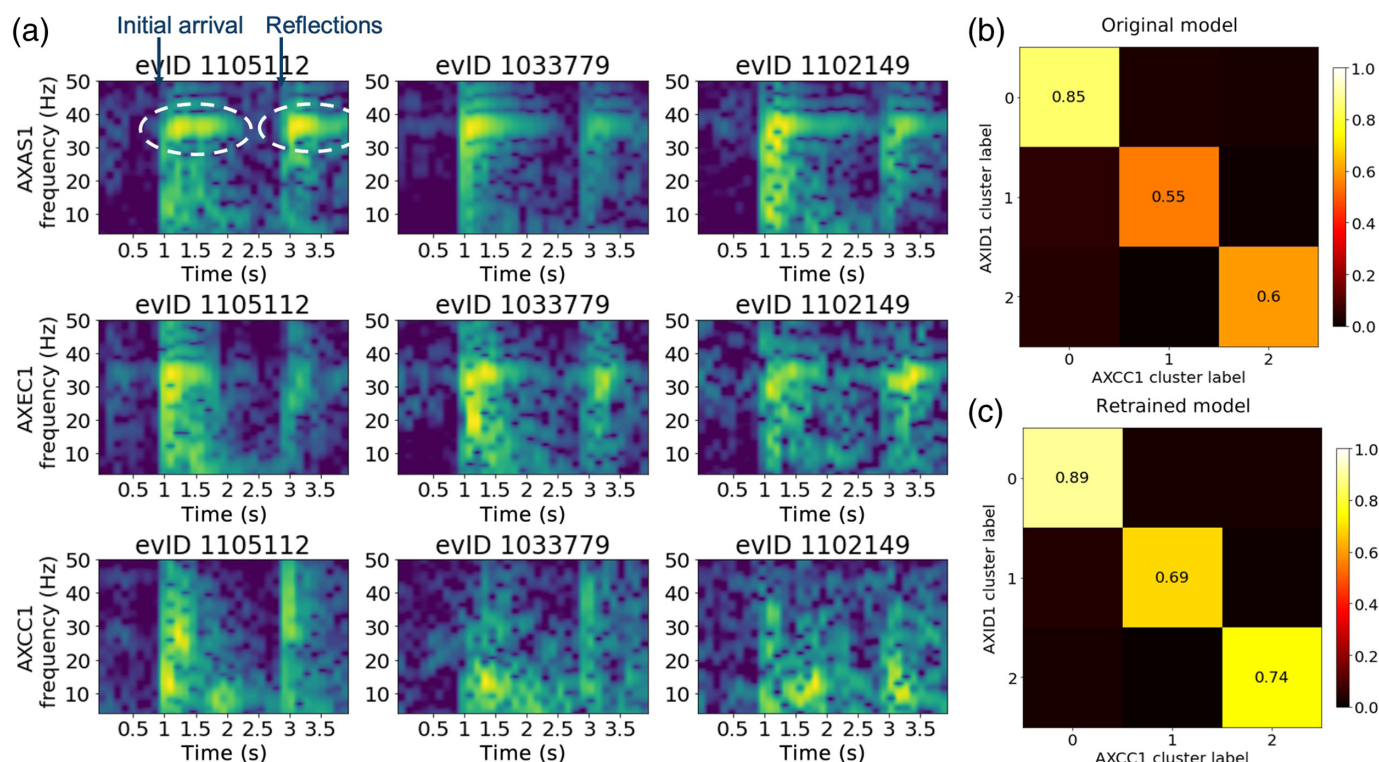
Figure 1. Histograms of different types of sources automatically characterized during seismic monitoring and their locations. (a) Earthquake rate within 5-day wide bins. Bars of earthquake rates are colored by the percentage of mixed frequency earthquakes (MFEs) in each bin. The main figure shows activity before and during the 2015 eruption. The gray dashed line marks the eruption onset. The inset shows a zoom of the post-eruption period from May 2015 to the end of 2021. (b) Map of earthquake density in log scale. Eruptive fissure and lava flow locations of the 2015 and 2011 eruptions are plotted with colored contours, showing approximate locations of the impulsive seafloor events. (c) The blue and brown bars show whale calls and impulsive seafloor events rate, respectively. The color version of this figure is available only in the electronic edition.

After generating the event spectrograms, we proceeded with a two-stage feature extraction process (Holtzman *et al.*, 2018) with nonnegative matrix factorization and a hidden Markov model (HMM). These two stages of data compression reduce data dimensionality and eliminate features that are common to all signals. From the output of HMM, we computed event fingerprints by counting state transitions. These fingerprints are low-dimensional representations of the original event spectrograms while preserving the key features of the time-variant spectral patterns. We then performed principal component analysis (PCA) and retained the top principal components that explain 80% of the total variance, followed by K-means clustering on the principal components. Here, we built feature dictionaries and fingerprints on a single-station basis to eliminate the effect of station-dependent noise on clustering. Figure 3a shows some examples of event spectrograms at a broadband station (AXCC1) and two short-period stations (AXAS1 and AXEC1). At these OBS stations, we observe station-dependent background noises that trigger especially during phase arrivals.



This noise is observed at all times for the five short-period stations while being absent at the two broadband stations. In Figure 3a, we show triggered noise at around 36.5 Hz for station AXAS1 and 33.5 Hz for station AXEC1. Based on their characteristics, we think these station-dependent noises are

Figure 2. Overview of (a) machine learning (ML)-based workflow for catalog construction and (b) unsupervised event discrimination. The color version of this figure is available only in the electronic edition.



likely due to resonance within the instrument after excitement by any incoming waves.

After obtaining the single-station-based event characterization and classification models, we combined the clustering labels across the network and employed self-training (Yarowsky, 1995) to retrain these single-station classification models. Details of the retraining process are described in the next section. The final retrained models for event characterization and spectral feature clustering are saved for use in the real-time system.

Results

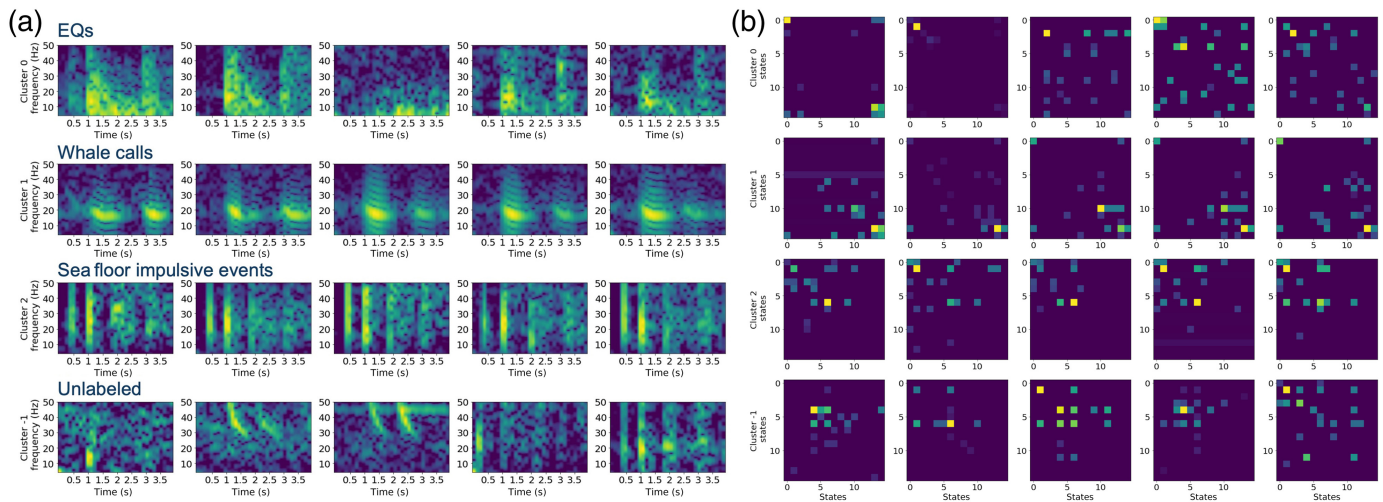
Our objective in this study is to distinguish between different event sources in the catalog so that we can perform refined seismic monitoring of different sources during real-time operation. Here, we chose the number of clusters for each station through a trial and error process. Several classic clustering evaluation metrics were tested, including summed squared distances (Nainggolan *et al.*, 2019), Silhouette coefficient (Rousseeuw, 1987), Davies–Bouldin index (Davies and Bouldin, 1979), and Calinski–Harabasz index (Caliński and Harabasz, 1974). However, these metrics did not reveal a sharp elbow (a turning point on the performance versus cluster number graph) that could clearly define the optimal cluster numbers (Nainggolan *et al.*, 2019). Therefore, we opted for evaluating clustering performance through visual inspection of the results, aiming to identify the minimum cluster number that effectively separates the signals of interest in the feature space.

Our results show different levels of clustering in the spectral feature space. On a first order, we identified three main clusters that share common spectral patterns across the

Figure 3. (a) Examples of event spectrograms at broadband station AXCC1 and two short-period stations AXAS1 and AXEC1. Arrows mark initial arrival and reflections in the water column. Station-dependent noise on the initial arrival and reflected ones marked by the white dashed ellipses. Contingency matrix of cluster labels at two OBS stations AXCC1 and AXID1 for (b) single-station models and (c) retrained models. The color version of this figure is available only in the electronic edition.

network. The three clusters are a group of earthquakes, a group of whale calls, and a group of seafloor impulsive events (Figure 4). These three main clusters define the first-order structure of the feature space, effectively separating events into earthquakes and non-earthquake sources. Further analysis within the earthquake cluster reveals higher-level subgroupings. The higher-level clusters reflect subtle differences in the spectral patterns, which help separate seismicity generated by different physical processes during the eruption. The temporal distribution of signals in different classes is shown in Figure 1a,c.

To improve on the single-station clustering approach used by Wang *et al.* (2024), we employed self-training techniques for model retraining (Yarowsky, 1995). We started by creating an initial training dataset that was composed of only events with consistent labels across the network. With this dataset, we trained an initial model and predicted pseudo labels using the initial model. High-confidence predictions from these pseudo-labels were incorporated back into the training set for iterative retraining. This iterative process allows the model to gradually improve by learning from its



predictions. By integrating clustering information from other stations in the network, we significantly improved the coherence and accuracy of the single-station cluster labels. Figure 3b,c illustrates this enhancement. The contingency matrix between station AXCC1 and station AXID1, which measures the relationship between two sets of cluster labels, shows a notable improvement in label coherency after the retraining process.

Having consistent labels across the network demonstrates that source effects dominate over path effects in event clustering. We show that we effectively discriminate signals from different sources, as shown in Figure 4.

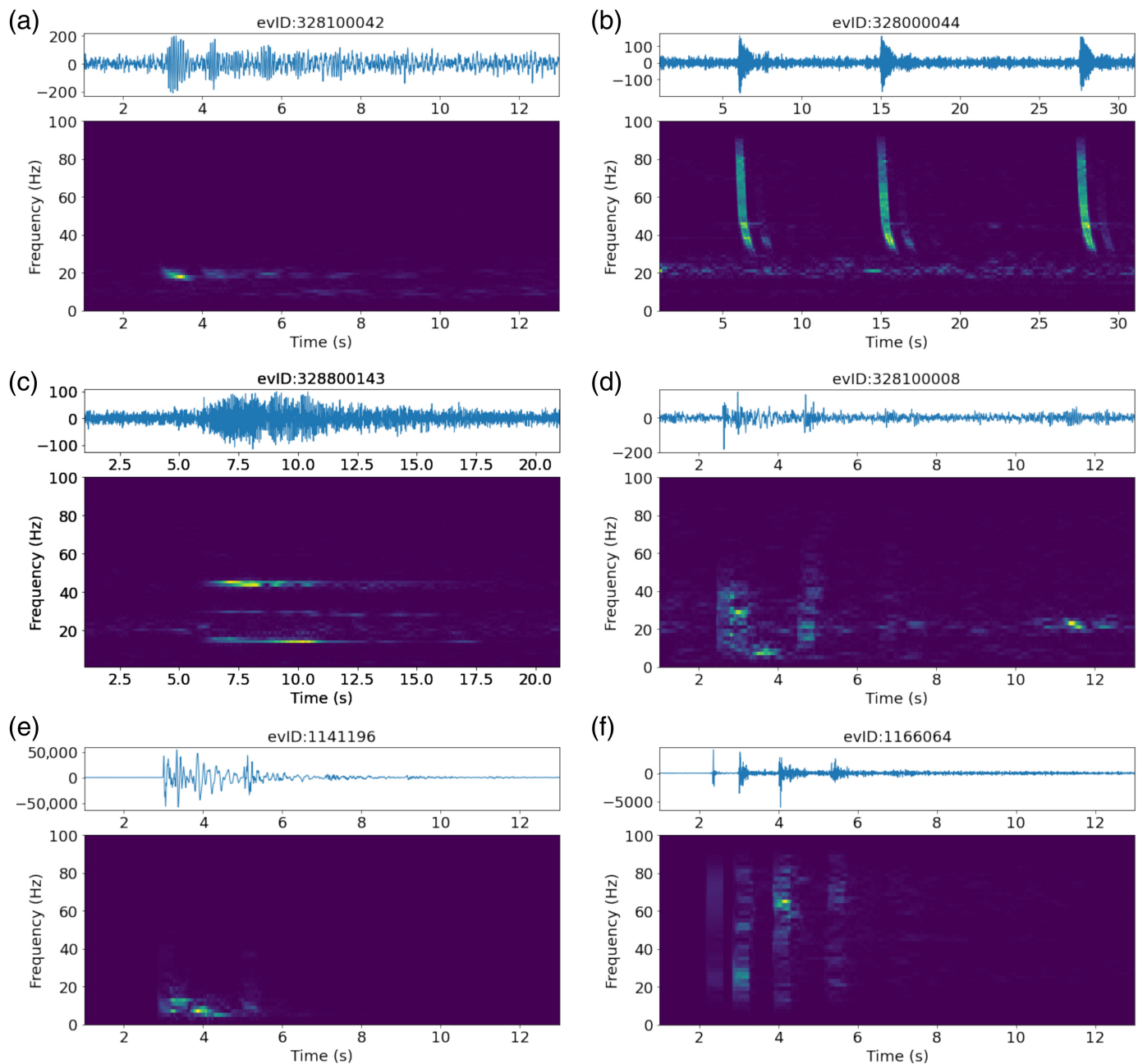
The initial ML catalog includes 200,891 earthquakes, 53,018 whale calls (predominantly fin whales), 12,406 seafloor impulsive events generated by lava–water interaction during the eruption in 2015, and 7,508 unlabeled events. We see consistent spectral patterns and fingerprints within each of the three primary event classes. It is worth noting that the earthquake class shows more variability in spectral patterns compared to the two non-earthquake classes, suggesting they are generated by complex physical processes. Further exploration of the earthquake class revealed two subgroups that have different spatiotemporal behavior and correlate with different physical processes: tidally modeled earthquakes predominantly on the caldera ring faults and precursory mixed frequency earthquakes (MFEs) associated with pre-eruption magmatic processes (Wang *et al.*, 2024).

Although we successfully classified most events into three main classes, a few remain unlabeled after retraining with self-training (examples shown in Fig. 4). These events either had inconsistent labels across the network or were weak signals with picks at less than two stations. Upon examining these unlabeled events, we found that they form a few event classes that are less frequently seen in the dataset, such as calls from different species of whales (as shown in Fig. 5b,c) and ship noise. Fin whales are the most common species that pass by the OBS array every year during the winter season and thus

Figure 4. Examples of (a) spectrograms and (b) corresponding fingerprints for different clusters. The color version of this figure is available only in the electronic edition.

makeup one of the three main clusters after event clustering. Their calls are sinusoidal signals in the 15–20 Hz range, as shown in Figure 5a. Some can have a higher frequency note of around 30 Hz (Weirathmueller *et al.*, 2017). In contrast, Sei whales and Blue whales are less common, less than a tenth of Fin whale calls. Thus, it is challenging to classify them with unsupervised methods. However, they have spectral content distinct from the other event classes (examples in Fig. 5b,c). Based on their spectral features, we designed a simple frequency-dependent classifier that computes energy ratio in the 30–50 Hz frequency band relative to the 4–50 Hz band to identify them during real-time operation.

The DD-RT base catalog, obtained after relocating the ML catalog of 200,891 earthquakes, includes 144,329 precisely located events constrained by 1.7 million ML *P* and *S* picks and 400 million correlation delay times. The mean root mean square of the delay times is 25 ms for picks and 4 ms for the correlation data. Relative location errors from a bootstrap analysis (Waldhauser and Ellsworth, 2000) are 39 m laterally and 53 m vertically for events within the network (see Waldhauser *et al.*, 2020, for details) The relocated earthquakes image the Axial ring fault structures and the intersection/interaction between the Juan de Fuca ridge and Axial volcano on its eastern side. The overall structures are similar to previous catalogs, although this new catalog has higher precision picks and more correlation measurements that contribute to higher resolution locations. In real-time operation, any new detection will be correlated with its neighboring events in this base catalog to allow rapid relocation with the double-difference algorithm (Waldhauser, 2009). Note that the computational cost of the DD-RT process is small and does not scale with the size of the base catalog.



Discussion

The quality of the base catalog is crucial in double-difference monitoring because it determines the accuracy of future events (Waldhauser, 2009). We have an opportunity to compare the newest catalog presented here to previous catalogs (Wilcock *et al.*, 2016; Waldhauser *et al.*, 2020) with respect to differences and robustness of arrival time measurements and hypocenter locations. Both the Wilcock *et al.* (2016) and the Waldhauser *et al.* (2020) catalog used the same set of phase picks generated by a Kurtosis phase picker, with events in the Wilcock *et al.* (2016) catalog located using HINV and those in the Waldhauser *et al.* (2020) catalog relocated using correlation-based double differences. Here, we make the comparison by matching the individual *P* and *S* wavepicks in our new ML-DD

Figure 5. Examples of spectrogram and corresponding waveform for different earthquake and non-earthquake sources: (a) Fin whale, (b) Sei whale, and (c) Blue whale; (d) volcanic earthquakes, (e) precursory MFEs, and (f) coeruption impulsive sea-floor events. The spectrograms are plotted with 100 samples (0.5 s) window length and 90 samples overlap. The color version of this figure is available only in the electronic edition.

catalog to the Kurtosis picks in the Wilcock *et al.* (2016) and Waldhauser *et al.* (2020) catalog. We match two picks if they are within 0.05 s and of the same phase type. Then we select a set of events with all picks matched to the same event ID in the other catalog. This subset of matched IDs includes a total of 104,522 events. Figure 6 plots the differences between the

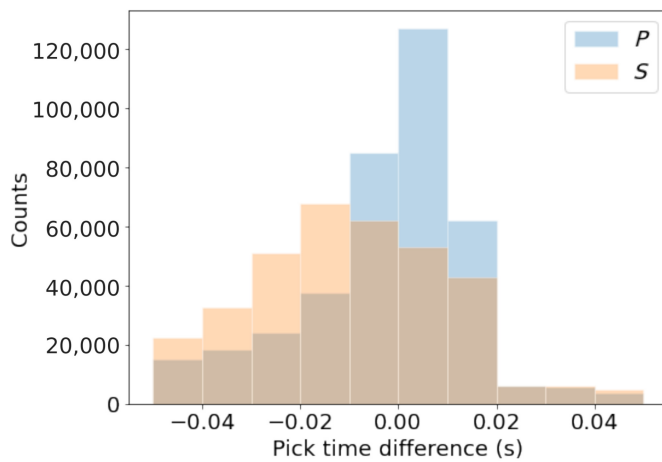


Figure 6. Histograms of the differences between matched Kurtosis and PhaseNet picks. The pick time difference is plotted as Kurtosis pick time and PhaseNet pick time. The color version of this figure is available only in the electronic edition.

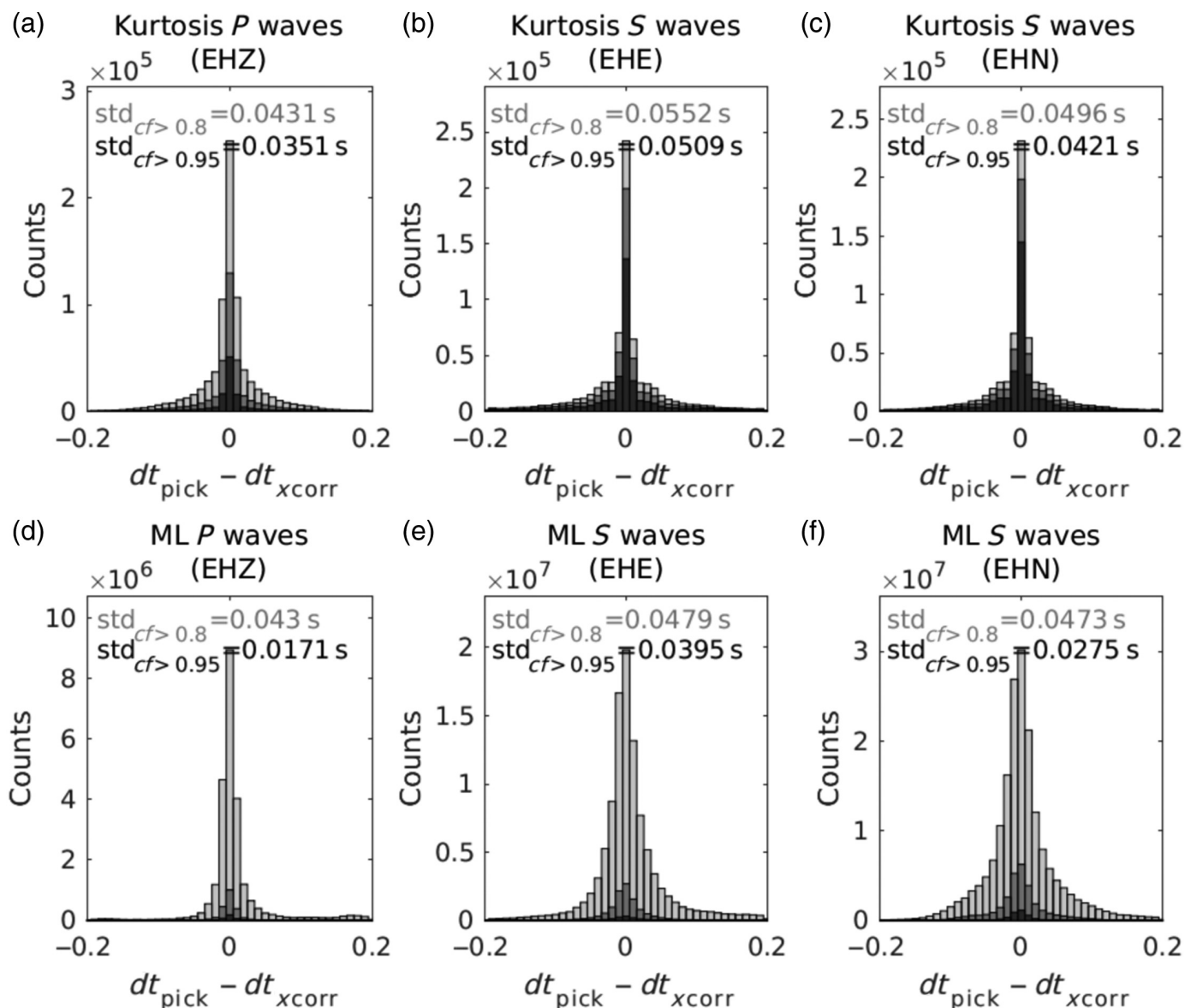
matched *P* and *S* wavepicks by the ML picker and the Kurtosis picker. The comparison shows that the differences between the matched picks are generally less than 0.02 s, while showing systematic biases in the pick times for both *P* and *S* waves. The Kurtosis picker consistently picks earlier on *S* waves compared to the PhaseNet picks, whereas it tends to pick later on *P* waves for 60% of the picks but with a distribution that is skewed to the earlier side. The delay in the *P* picks for the Kurtosis picker may be due to its nature of measuring distribution change in a window, and the differences in *S* wavepicks might be related to complications caused by the *P* coda and *S* to *P*, *P* to *S* conversions before the *S* arrival. The effect of consistently picking converted phases on hypocenter location is minimized during double-difference relocation (Waldhauser and Tolstoy, 2011).

We estimate the pick uncertainty by comparing the Kurtosis and ML phase pick delay times between an event pair with the corresponding cross-correlation delay time measurements. The delay time differences for the *P* picks with respect to the correlation measurements on the vertical component and the *S* picks compared to those on the two horizontal components are shown in Figure 7. We see that the ML catalog has more picks and more correlation measurements (1–2 orders of magnitude more compared with the Kurtosis catalog). The standard deviations are comparable for the $C_f > 0.8$ group, which is the threshold we use to select correlation measurements for double-difference relocation. However, the ML picks have much lower standard deviations with respect to data from highly correlated waveforms ($C_f > 0.95$): 17 ms for *P* waves, and 40 ms/28 ms for *S* waves on east (E)/north (N) components, compared with 35 ms (*P*) and 51 ms/42 ms (*S* on E/N components) for the Kurtosis picks. This suggests that ML picks are of higher accuracy than Kurtosis picks, and PhaseNet performs better when

picking low signal-to-noise ratio data (more picks with similar uncertainty to the Kurtosis catalog).

We also compared the locations of earthquakes in the three catalogs. To make a straightforward comparison, we plot the locations of the same subset of events with matched IDs. Figure 8a shows the comparison between the final double-difference located ML catalog (this study) and the Waldhauser *et al.* (2020) catalog. We see that the high-resolution earthquake locations in this study delineate sharp-focused structures, including the varying geometry on the western wall of the caldera ring faults and the complex structures on the eastern wall at the intersection with the ridge. The Waldhauser *et al.* (2020) locations have similar patterns overall, but, are more diffuse. To understand which processing step contributed to this improvement, we also made comparisons of the initial locations before the double-difference relocation. Figure 8b,c shows the differences between the initial NLL locations in the ML catalog (this study) and the initial NLL locations in the Waldhauser *et al.* (2020) catalog (Fig. 8b), and HINV locations in Wilcock *et al.* (2016) catalog (Fig. 8c). The differences between Figure 8b,c shows the contribution from using a 3D velocity model. We can see that in Figure 8b, the initial NLL locations of the two DD catalogs match well, whereas, in Figure 8c, the NLL locations and the HINV locations show significant differences. This suggests that the use of a 3D velocity model NLL, as opposed to a 1D velocity model, contributed considerably to the improvement of the final earthquake location results. Figure 8b shows the location differences caused by the absolute picks, whereas Figure 8a shows the differences after double-difference relocation. In Figure 8b, the overall pattern of the two catalogs agrees, and the differences are sharpened in Figure 8a. This suggests that improvements in pick accuracy and especially the addition of cross-correlation measurements helped in refining the complex seismicity structure at Axial Volcano.

Since the beginning of 2022, we have been operating the ML-based seismic monitoring workflow in real-time. Seismic activity following the last eruption has remained at a relatively low rate (Fig. 1a). The overall seismicity pattern is generally consistent with the active faults and structures activated during the 2015 eruption, with recent seismicity bursts more concentrated in the southern part of the caldera near the hydrothermal field International District (Kelley *et al.*, 2014) (Fig. S2, available in the supplemental material to this article). As of April 2024, the automatic near-real-time analysis of the continuous waveform data has identified 37,974 earthquakes since the end of the base catalog, 30,152 whale calls (93% of them from Fin whales), 701 misidentified impulsive events, and 1,700 unlabeled events. We set a relatively tight threshold during real-time operation to minimize the number of misclassified events. Visual inspection of the unlabeled events shows that they are either weak events recorded by only a few stations or have inconsistent labels across the network.



Since the 2015 eruption, Axial Seamount has been continuously inflating and has recovered 90%–95% of its previous eruption level (Chadwick *et al.*, 2023). However, its inflation rate has gradually decreased over the past few years and leveled out in the last year, deviating from the steady inflation pattern observed in previous cycles (Chadwick *et al.*, 2016, 2022) (the inflation rate seems to start picking up again in early 2024). This irregular behavior makes it challenging for eruption forecasting solely from the deformation data. In this study, we demonstrate an ML workflow to track different types of seismic events, including short-term precursory events, in real-time. Our current real-time process operates on an hourly basis during periods of low activity. The real-time ML-DD catalog can be accessed at [Data and Resources](#). In routine ML-based processing, continuous seismic data are fed into the workflow in 15 s overlapping windows, and the window size will decrease to 6 s during high-seismicity-rate periods. This allows us to achieve near-real-time operation, such as every 5 min, when

Figure 7. Histogram of differences between pick and correlation delay times, shown for three correlation coefficient thresholds (light gray: $cf > 0.8$; gray: $cf > 0.9$; black: $cf > 0.95$). Differences are shown for Kurtosis picks (a–c) and PhaseNet picks (d–f) for the three components. The standard deviations of different cross-correlation coefficient groups are labeled on each subplot.

we approach the next eruption. It should be noted that in real-time operation, precursory events and eruption signals will be classified from the initial catalog before running the relocation step, further saving computational time for seismic monitoring. This system, with its high-resolution earthquake catalog and real-time analysis capability, complements the current deformation-based long-term forecasting methods by providing valuable short-term constraints. It may enhance eruption forecasting at Axial Seamount and potentially other volcanoes in submarine or terrestrial environments.

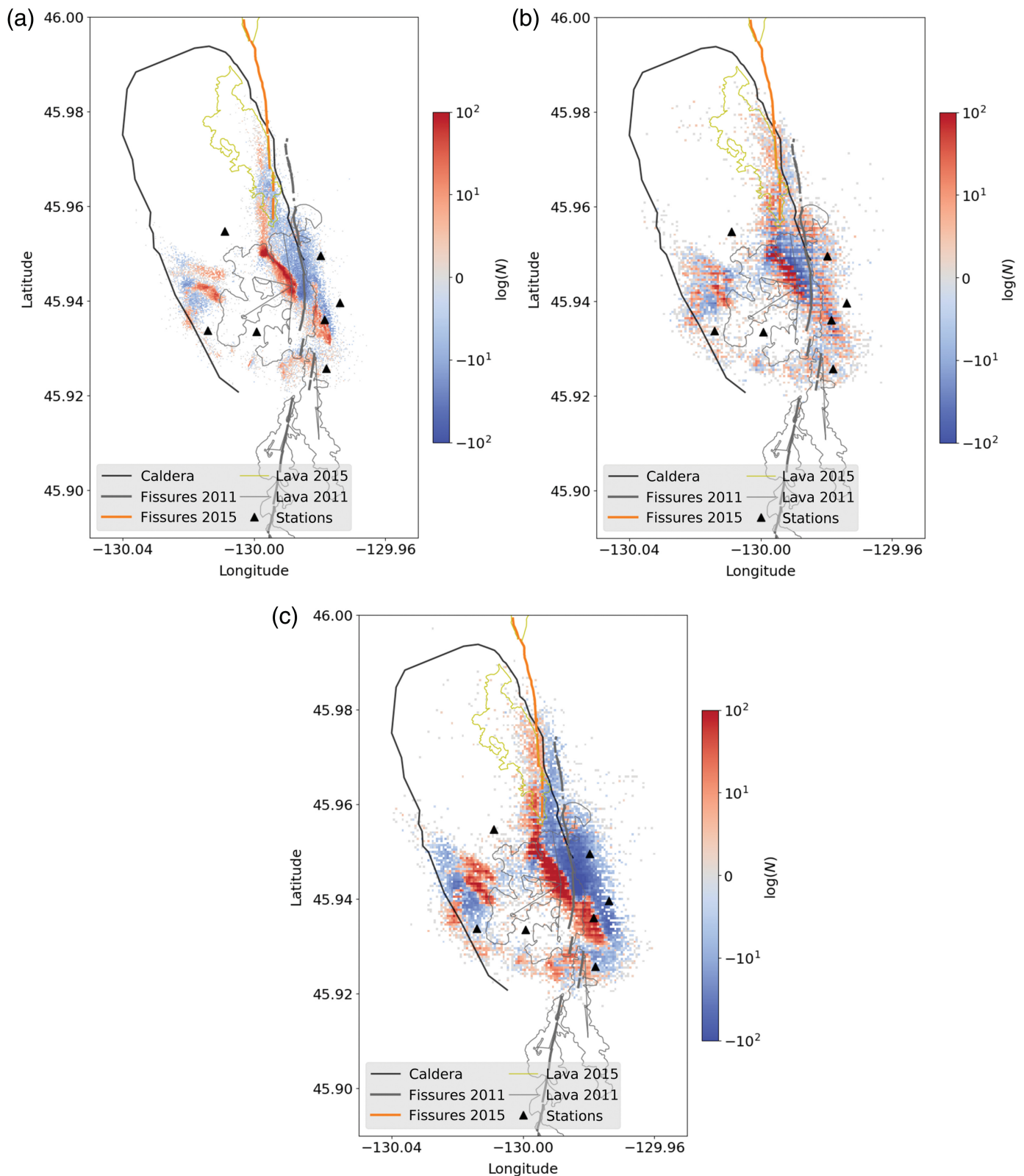


Figure 8. Comparison of earthquake locations derived in this study and those in the catalogs of [Wilcock et al. \(2016\)](#) and [Waldhauser et al. \(2020\)](#). Comparison between (a) the ML-DD catalog (this study) with the [Waldhauser et al. \(2020\)](#) catalog, (b) ML-NLL initial locations with the Kurtosis-NLL initial locations in the [Waldhauser et al. \(2020\)](#) catalog. (c) ML-NLL initial locations with the Kurtosis-HINV initial

locations in the [Wilcock et al. \(2016\)](#) catalog. The three plots are earthquake density differences between the catalogs in log-scale counts. The red and blue colors show where the ML catalog has more and fewer events compared to the other catalog. The bin size is 25 m \times 25 m in (a) and 50 m \times 50 m in (b) and (c). The same comparisons with overlaid density plots are shown in Figure S1.

Furthermore, new data from a temporary OBS array may improve our ability to constrain event locations outside the current array's coverage. Future plans include extending our ML workflow using the new OBS and terrestrial array stations and migrating the system to the cloud for more accessible operation.

Conclusions

Our real-time seismic monitoring framework effectively integrates automated ML- and double-difference analysis for high-precision, deep-magnitude catalog production. This framework not only handles routine seismic processing and earthquake location that are available in existing monitoring systems but is also equipped with new modules that classify various event types in real-time without human assistance, using our pretrained semisupervised models. These unique modules are tailored for complex submarine volcanic environments such as Axial Seamount with characteristic sources that may be indicative of eruption-related processes. Beyond routine seismic monitoring, we are now able to discriminate and track different types of seismic events as they occur (Fig. 1), including precursory MFEs (Fig. 5e) that potentially indicate the preparation of an eruption (Wang *et al.*, 2024) and seafloor impulsive events (Fig. 5f) that can be used to track magma outflows during an eruption (Tan *et al.*, 2016; Wilcock *et al.*, 2016; Caplan-Auerbach *et al.*, 2017). This may help improve the short-term eruption forecasting for Axial Seamount's next eruption and eruptions at other submarine and possibly terrestrial volcanoes.

Data and Resources

The seismic data used in this study are downloaded from the EarthScope Consortium Data Management Center (DMC) with network code "OO". The continuously updated earthquake catalog is available on our Axial Seamount real-time monitoring website <https://axialdd.ideo.columbia.edu> (last accessed June 2024). The supplemental material for this article includes earthquake catalog comparisons and real-time earthquake locations until April 2024.

Declaration of Competing Interests

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

Acknowledgments

The authors would like to thank two anonymous reviewers and the Editor Alicia Hotovec-Ellis for providing helpful comments and suggestions. This study benefited from fruitful discussions with Eric Beauce, Ben Holtzman, and Theresa Sawi. This work is funded by National Science Foundation Award Numbers NSF OAC-2103741 (project SCOPED) and NSF OCE-1951448.

References

Arnulf, A., A. Harding, G. Kent, S. Carbotte, J. Canales, and M. Nedimović (2014). Anatomy of an active submarine volcano, *Geology* **42**, no. 8, 655–658.

Baillard, C., W. C. Crawford, V. Ballu, C. Hibert, and A. Mangeney (2014). An automatic kurtosis-based P- and S-phase picker

designed for local seismic networks, *Bull. Seismol. Soc. Am.* **104**, no. 1, 394–409.

Baillard, C., W. S. Wilcock, A. F. Arnulf, M. Tolstoy, and F. Waldhauser (2019). A joint inversion for three-dimensional P and S wave velocity structure and earthquake locations beneath axial seamount, *J. Geophys. Res.* **124**, no. 12, 12,997–13,020.

Bergen, K. J., P. A. Johnson, M. V. de Hoop, and G. C. Beroza (2019). Machine learning for data-driven discovery in solid earth geoscience, *Science* **363**, no. 6433, eaau0323, doi: [10.1126/science.aau0323](https://doi.org/10.1126/science.aau0323).

Beroza, G. C., M. Segou, and S. Mostafa Mousavi (2021). Machine learning and earthquake forecasting—Next steps, *Nat. Commun.* **12**, no. 1, 4761, doi: [10.1038/s41467-021-24952-6](https://doi.org/10.1038/s41467-021-24952-6).

Caliński, T., and J. Harabasz (1974). A dendrite method for cluster analysis, *Commun. Stat. Theory Meth.* **3**, no. 1, 1–27.

Caplan-Auerbach, J., R. Dziak, J. Haxel, D. Bohnenstiehl, and C. Garcia (2017). Explosive processes during the 2015 eruption of axial seamount, as recorded by seafloor hydrophones, *Geochem. Geophys. Geosyst.* **18**, no. 4, 1761–1774.

Chadwick, W., Jr, J. Paduan, D. Clague, B. Dreyer, S. Merle, A. Bobbitt, D. Caress, B. Philip, D. Kelley, and S. Nooner (2016). Voluminous eruption from a zoned magma body after an increase in supply rate at axial seamount, *Geophys. Res. Lett.* **43**, no. 23, 12,063–12,070.

Chadwick, W. W., Jr, W. S. Wilcock, S. L. Nooner, and J. W. Beeson (2023). Is axial seamount just napping? An update on the latest inflation and seismic data, available at <https://agu.confex.com/agu/fm23/meetingapp.cgi/Paper/1326844> (last accessed June 2024).

Chadwick, W. W., Jr, W. S. Wilcock, S. L. Nooner, J. W. Beeson, A. M. Sawyer, and T.-K. Lau (2022). Geodetic monitoring at axial seamount since its 2015 eruption reveals a waning magma supply and tightly linked rates of deformation and seismicity, *Geochem. Geophys. Geosyst.* **23**, no. 1, e2021GC010153, doi: [10.1029/2021GC010153](https://doi.org/10.1029/2021GC010153).

Chrapkiewicz, K., M. Paulatto, B. Heath, E. Hooft, P. Nomikou, C. Papazachos, F. Schmid, D. Toomey, M. Warner, and J. Morgan (2022). Magma chamber detected beneath an arc volcano with full-waveform inversion of active-source seismic data, *Geochem. Geophys. Geosyst.* **23**, no. 11, e2022GC010475, doi: [10.1029/2022GC010475](https://doi.org/10.1029/2022GC010475).

Cotton, C. V., and D. P. Ellis (2011). Spectral vs. spectro-temporal features for acoustic event detection, *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, 16–19 October 2011, pp. 69–72.

Davies, D. L., and D. W. Bouldin (1979). A cluster separation measure, *IEEE Trans. Pattern Analysis Machine Intellig.* **1**, no. 2, 224–227, doi: [10.1109/TPAMI.1979.4766909](https://doi.org/10.1109/TPAMI.1979.4766909).

Holtzman, B., N. Groebner, T. Sawi, T. Xing, M. Pec, H. Ghaffari, U. Mok, R. Skarbek, J. Paisley, and T. Mittal, *et al.* (2021). Unsupervised spectral feature extraction applied to acoustic emissions during brittle creep of basalt under dry and wet conditions, *AGU Fall Meet. Abstracts*, Vol. 2021, New Orleans, Los Angeles, 13–17 December 2021, pp. H12E-04.

Holtzman, B. K., A. Paté, J. Paisley, F. Waldhauser, and D. Repetto (2018). Machine learning reveals cyclic changes in seismic source spectra in geysers geothermal field, *Sci. Adv.* **4**, no. 5, eaao2929, doi: [10.1126/sciadv.aao2929](https://doi.org/10.1126/sciadv.aao2929).

Kelley, D. S., J. R. Delaney, and S. K. Juniper (2014). Establishing a new era of submarine volcanic observatories: Cabling Axial

- Seamount and the endeavour segment of the Juan de Fuca ridge, *Mar. Geol.* **352**, 426–450.
- Klein, F. W. (2002). User's guide to hypoinverse-2000, a Fortran program to solve for earthquake locations and magnitudes, *U.S. Geol. Surv. Open-File Rept.* 02-171.
- Le Saout, M., D. Bohnenstiehl, J. Paduan, and D. Clague (2020). Quantification of eruption dynamics on the north rift at Axial Seamount, Juan de Fuca ridge, *Geochem. Geophys. Geosyst.* **21**, no. 9, e2020GC009136, doi: [10.1029/2020GC009136](https://doi.org/10.1029/2020GC009136).
- Lomax, A., J. Virieux, P. Volant, and C. Berge-Thierry (2000). Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, in *Advances in Seismic Event Location*, C. H. Thurber and N. Rabinowitz (Editors), Kluwer, Amsterdam, the Netherlands, 101–134.
- Mousavi, S. M., and G. C. Beroza (2023). Machine learning in earthquake seismology, *Annu. Rev. Earth Planet. Sci.* **51**, 105–129.
- Nainggolan, R., R. Perangin-angin, E. Simarmata, and A. F. Tarigan (2019). Improved the performance of the k-means cluster using the sum of squared error (sse) optimized by using the elbow method, *Journal of Physics: Conf. Series*, Vol. 1361, IOP Publishing, Medan, Indonesia, 23–24 November 2019, pp. 012015.
- Park, J., J. Morgan, C. Zelt, P. Okubo, L. Peters, and N. Benesh (2007). Comparative velocity structure of active Hawaiian volcanoes from 3-d onshore-offshore seismic tomography, *Earth Planet. Sci. Lett.* **259**, nos. 3–4, 500–516.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *J. Comput. Appl. Math.* **20**, 53–65.
- Rubin, K. H., S. A. Soule, W. W. Chadwick, D. J. Fornari, D. A. Clague, R. W. Embley, E. T. Baker, M. R. Perfit, D. W. Caress, and R. P. Dziak (2012). Volcanic eruptions in the Deep Sea, *Oceanogr.* **25**, no. 1, 142–157.
- Sawi, T., B. Holtzman, F. Walter, and J. Paisley (2022). An unsupervised machine-learning approach to understanding seismicity at an alpine glacier, *J. Geophys. Res.* **127**, no. 12, e2022JF006909, doi: [10.1029/2022JF006909](https://doi.org/10.1029/2022JF006909).
- Schaff, D. P., and F. Waldhauser (2005). Waveform cross-correlation-based differential travel-time measurements at the northern California seismic network, *Bull. Seismol. Soc. Am.* **95**, no. 6, 2446–2461.
- Sinton, J., E. Bergmanis, K. Rubin, R. Batiza, T. Gregg, K. Macdonald, and S. White (2002). Volcanic eruptions on mid-ocean ridges: New evidence from the superfast spreading east pacific rise, *J. Geophys. Res.* **107**, no. b6, 17–19.
- Smith, L. M., J. A. Barth, D. S. Kelley, A. Plueddemann, I. Rodero, G. A. Ulses, M. F. Vardaro, and R. Weller (2018). The ocean observatories initiative, *Oceanogr.* **31**, no. 1, 16–35.
- Tan, Y. J., M. Tolstoy, F. Waldhauser, and W. S. Wilcock (2016). Dynamics of a seafloor-spreading episode at the east pacific rise, *Nature* **540**, no. 7632, 261–265.
- Tepp, G., R. P. Dziak, M. M. Haney, J. J. Lyons, C. Searcy, H. Matsumoto, and J. Haxel (2020). Seismic and hydroacoustic observations of the 2016–17 Bogoslof eruption, *Bull. Volcanol.* **82**, 1–21.
- Waldhauser, F. (2001). HypoDD: A computer program to compute double-difference hypocenter locations, *U.S. Geol. Surv. Open-file Rep.* 01-113, doi: [10.3133/ofr01113](https://doi.org/10.3133/ofr01113).
- Waldhauser, F. (2009). Near-real-time double-difference event location using long-term seismic archives, with application to northern California, *Bull. Seismol. Soc. Am.* **99**, no. 5, 2736–2748.
- Waldhauser, F., and W. L. Ellsworth (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.* **90**, no. 6, 1353–1368.
- Waldhauser, F., and M. Tolstoy (2011). Seismogenic structure and processes associated with magma inflation and hydrothermal circulation beneath the east pacific rise at 9°50'N, *Geochem. Geophys. Geosyst.* **12**, no. 8, doi: [10.1029/2011GC003568](https://doi.org/10.1029/2011GC003568).
- Waldhauser, F., W. Wilcock, M. Tolstoy, C. Baillard, Y. Tan, and D. Schaff (2020). Precision seismic monitoring and analysis at axial seamount using a real-time double-difference system, *J. Geophys. Res.* **125**, no. 5, e2019JB018796, doi: [10.1029/2019GC003568](https://doi.org/10.1029/2019GC003568).
- Wang, K., F. Waldhauser, M. Tolstoy, D. P. Schaff, T. Sawi, W. S. D. Wilcock, and Y. J. Tan (2024). Volcanic precursor revealed by machine learning offers new eruption forecasting capability, ESS Open Archive.
- Weirathmueller, M. J., K. M. Stafford, W. S. Wilcock, R. S. Hilmo, R. P. Dziak, and A. M. Tréhu (2017). Spatial and temporal trends in fin whale vocalizations recorded in the NE Pacific Ocean between 2003 and 2013, *PLoS one* **12**, no. 10, e0186127, doi: [10.1371/journal.pone.0186127](https://doi.org/10.1371/journal.pone.0186127).
- Wilcock, W. S., R. P. Dziak, M. Tolstoy, W. W. Chadwick, S. L. Nooner, D. R. Bohnenstiehl, J. Caplan-Auerbach, F. Waldhauser, A. F. Arnulf, and C. Baillard, *et al.* (2018). The recent volcanic history of axial seamount: Geophysical insights into past eruption dynamics with an eye toward enhanced observations of future eruptions, *Oceanogr.* **31**, no. 1, 114–123.
- Wilcock, W. S., M. Tolstoy, F. Waldhauser, C. Garcia, Y. J. Tan, D. R. Bohnenstiehl, J. Caplan-Auerbach, R. P. Dziak, A. F. Arnulf, and M. E. Mann (2016). Seismic constraints on caldera dynamics from the 2015 axial seamount eruption, *Science* **354**, no. 6318, 1395–1399.
- Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods, in *33rd Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Cambridge, Massachusetts, pp. 189–196, doi: [10.3115/981658.981684](https://doi.org/10.3115/981658.981684).
- Zhu, W., and G. C. Beroza (2019). Phasenet: A deep-neural-network-based seismic arrival-time picking method, *Geophys. J. Int.* **216**, no. 1, 261–273.
- Zhu, W., A. B. Hou, R. Yang, A. Datta, S. M. Mousavi, W. L. Ellsworth, and G. C. Beroza (2023). Quakeflow: A scalable machine-learning-based earthquake monitoring workflow with cloud computing, *Geophys. J. Int.* **232**, no. 1, 684–693.
- Zhu, W., I. W. McBrearty, S. M. Mousavi, W. L. Ellsworth, and G. C. Beroza (2022). Earthquake phase association using a Bayesian Gaussian mixture model, *J. Geophys. Res.* **127**, no. 5, e2021JB023249, doi: [10.1029/2021JB023249](https://doi.org/10.1029/2021JB023249).

Manuscript received 4 March 2024

Published online 11 July 2024