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

Kaiwen Wang*[1], Felix Waldhauser[1], David Schaff[1], Maya Tolstoy[2], William Wilcock[2], and Yen Joe Tan[3]

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 seismometers (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 and effectively discriminated between various seismic and acoustic events. These signals 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 rapid determination of high-resolution earthquake locations and the capability to track potential precursory signals and co-eruption indicators of magma outflow, this system may improve eruption forecasting by providing short-term constraints on Axial's next eruption. Furthermore, our work demonstrated an effective application that integrates unsupervised learning for signal discrimination in real-time operation, which could be generalized to other regions for volcanic unrest detection and enhanced eruption forecasting.

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, as they erupt in short intervals (Rubin et al., 2012; Sinton et al., 2002) and active source seismic profiling allows for detailed imaging of sub-seafloor structures (Arnulf et al., 2014; Park et al., 2007; 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, for example, an active submarine volcano located 1,400 meters below the sea surface at the intersection of the Juan de Fuca Ridge and the Cobb-Eickelberg hot spot, the 7 cabled OBSs operated by the Ocean Observatories Initiative (OOI) (Kelley et al., 2014), record signals from earthquakes, marine mammals, and airgun shots from active source experiments. 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 (Smith et al., 2018; Wilcock et al., 2018; Chadwick Jr et al., 2016) operated by the OOI, making Axial Seamount one of the best instrumented submarine volcanoes.

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 (Wilcock et al., 2016; Tan 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 (Bergen et al., 2019; Beroza et al., 2021; Mousavi and Beroza, 2023) to significantly improve 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-

1. Lamont -Doherty Earth Observatory, Columbia University, Palisades NY; 2. School of Oceanography, University of Washington, Seattle WA; 3. Earth System Science Programme, Faculty of Science, Chinese University of Hong Kong, Hong Kong
*Corresponding author: kw2988@columbia.edu

magnitude event catalogs in near real-time. Specifically, the new system is trained to detect potential precursory signals of mixed-frequency earthquakes (MFEs) presumably caused by the movement of volatiles or magma. These signals were discovered using an unsupervised spectral feature extraction method (Holtzman et al., 2018) before the last Axial eruption in 2015, where they emerged from background seismicity about 15 hours before the lava reached the seafloor (Wang et al., 2024). The method also detects acoustic signals that are generated when hot lava reaches the cold sea floor after an eruption (Wilcock et al., 2016; Tan et al., 2016; Le Saout et al., 2020), allowing for rapid tracking of the lava flow in time and space.

Data and Methods

Data and current monitoring

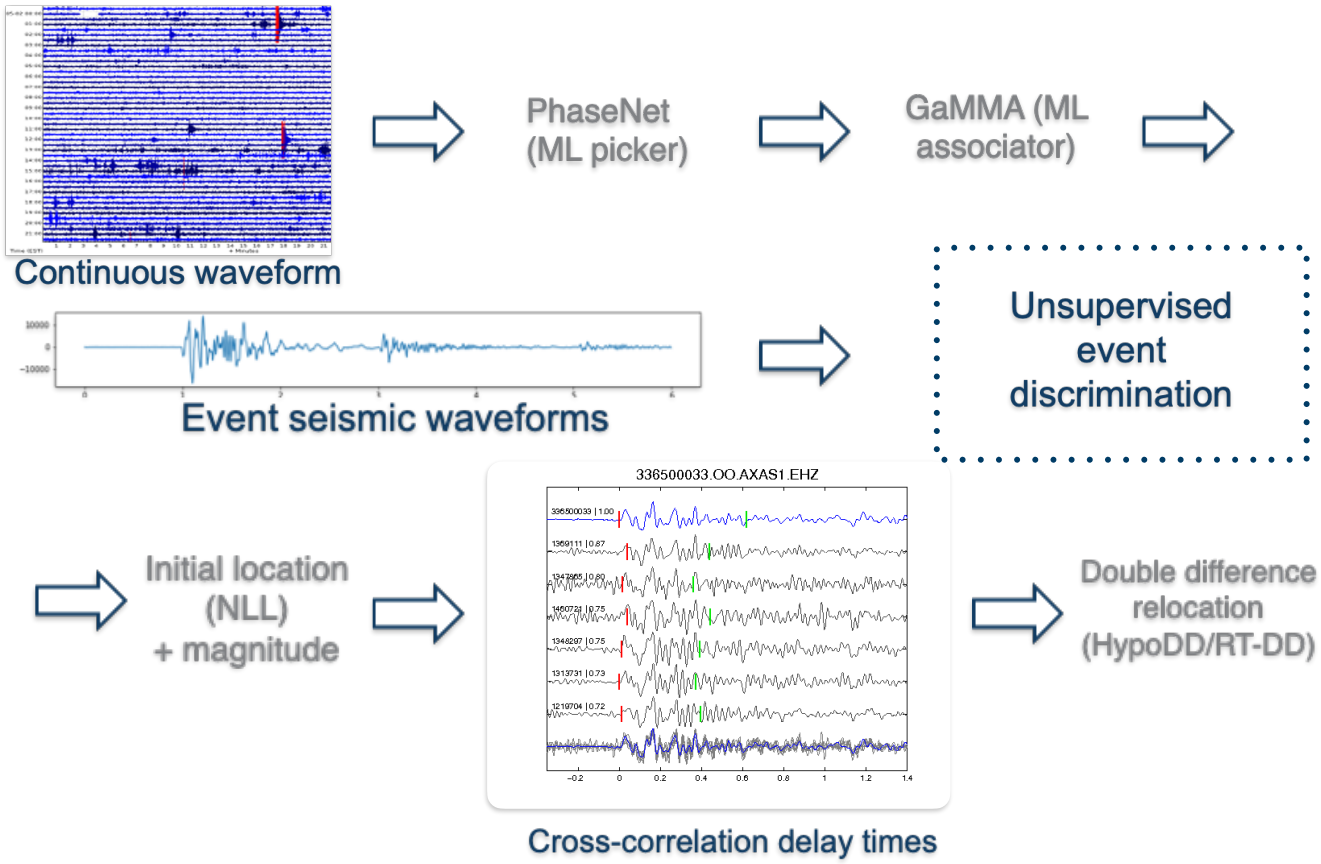
We use the continuous seismic data from the OOI OBS array available through the IRIS DMC as input to our workflow. The cabled OBS array comprises two broadband stations and five short-period stations on top of Axial seamount (Figure 4c), all data sampled at 200 Hz. Currently, the data is automatically processed in near-real-time at UW (Wilcock et al., 2016) using standard routine analysis with a Kurtosis phase picker (Baillard et al., 2014) and the Hypoinverse location algorithm (Klein, 2002). Each newly detected and located earthquake is then automatically relocated at LDEO with respect to a high-resolution earthquake (base) catalog using waveform cross-correlation and double-difference methods as implemented in the DD-RT system (Waldhauser et al., 2020) (DD-RT, 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 (NLL, Lomax et al., 2000, 2009) of the Kurtosis picks in a 3D earthquake tomographic P- and S-wave model (Baillard et al., 2019).

New processing

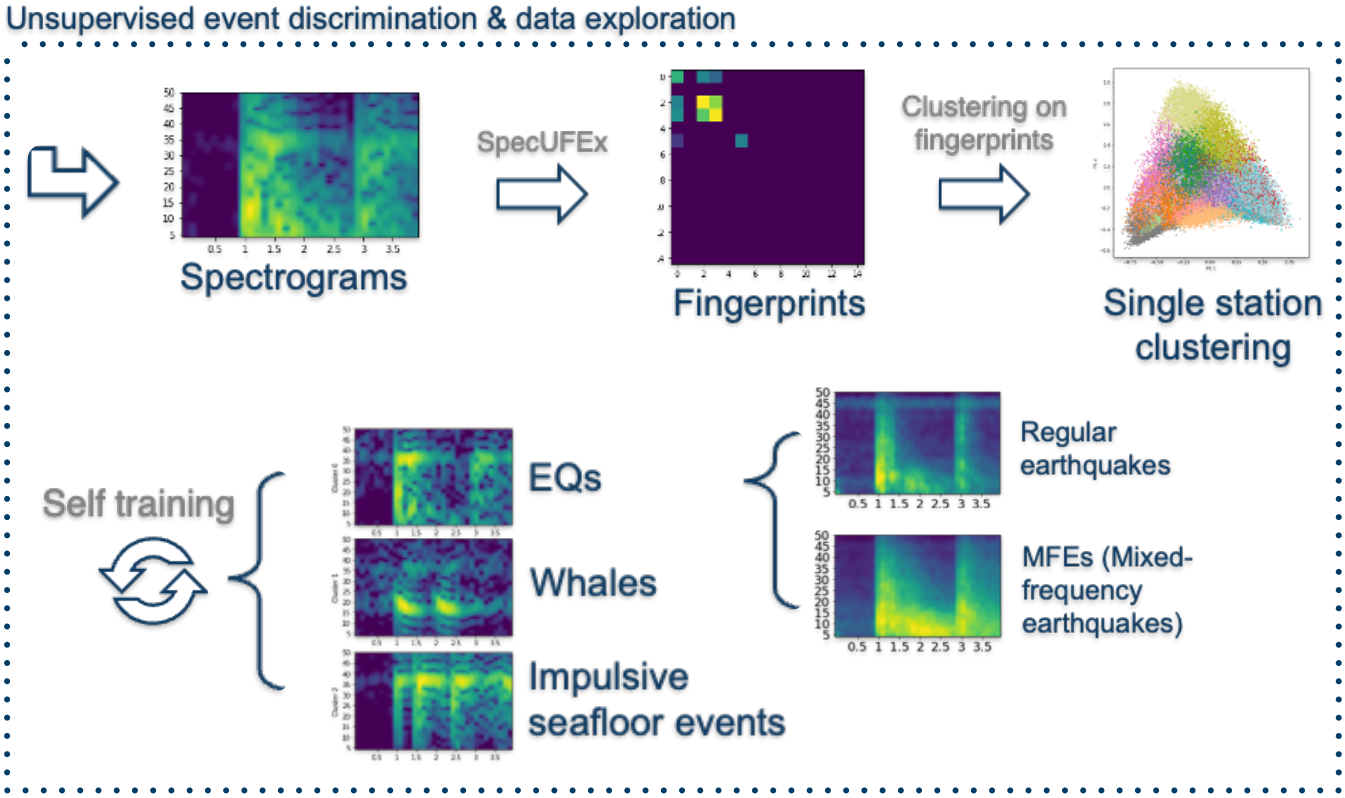
Here, we replace the current routine processing step with a supervised ML method (QuakeFlow, Zhu et al., 2023) to detect and characterize seismic events, and an unsupervised ML method (Specufex, Holtzman et al., 2018) to discriminate between various seismic sources. Our ML-based framework is illustrated in Figure 1. 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 deep-magnitude catalog of diverse, labeled seismic source types, including precursory MFEs and impulsive signals generated by lava reaching the cold seafloor (Wilcock et al., 2016). These labeled data provide the foundation for our initial ML-based processing step.

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. 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



(a) Overview of the ML workflow



(b) Unsupervised event disrimination

Figure 1: ML-based workflow for catalog construction and event discrimination.

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, NonLinLoc (Lomax et al., 2000, 2009), in a local 3-D 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 $Cf > 0.8$ and hypocentral separation < 2 km. The correlation data together with the delay times formed from the ML picks are inverted for relative location using the HypoDD algorithm to obtain a high-precision base catalog to be used in the DD-RT monitoring system (Waldhauser, 2009).

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 spectral fingerprints for each earthquake signal which are then clustered to find groups of similar signals (Figure 1).

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 second before the P-arrival to 3 seconds after it. This window was chosen to encompass the initial arrival, coda, and reflections in the water column (Figure 2). We selected this window length to capture the main features of the events 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 prepossessing the spectrograms, however, here we used a catalog of all event types (earthquakes 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.

After generating the event spectrograms, we proceeded with a two-stage feature extraction process with Non-negative matrix factorization (NMF) and 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 2a shows some examples of event spectrograms at a broad-band station (AXCC1) and two short-period stations (AXAS1 and AXEC1). At these OBS stations, we observe station-dependent background noises that trigger especially at phase arrivals. This noise is observed at all times for the five short-period stations, while absent at the two broad-band stations. In Figure 2a, 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 likely due to resonance within the instrument and excited by any incoming waves.

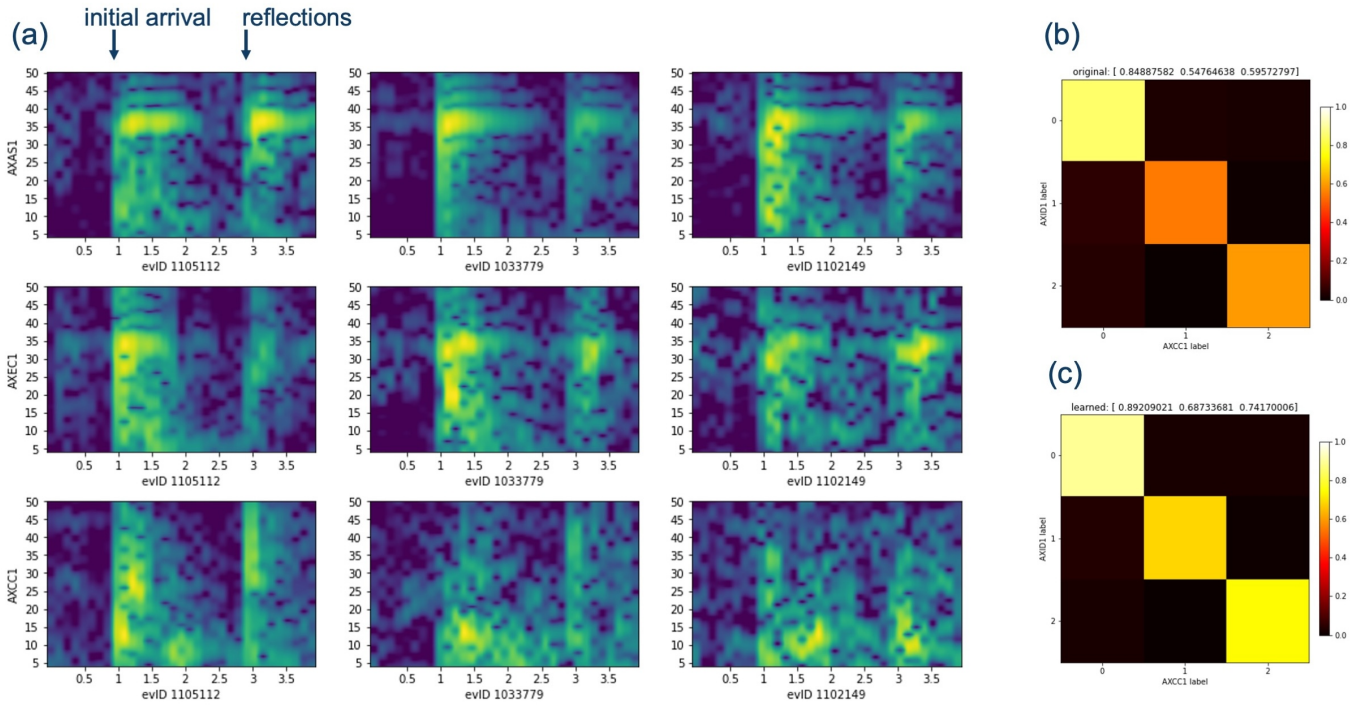


Figure 2: (a) Examples of event spectrograms at a broad band station AXCC1 and two short period stations AXAS1 and AXEC1. Initial arrival and reflections in the water column are marked by arrows. Contingency matrix of cluster labels at two OBS stations AXCC1 and AXID1 for (b)single-station models and (c)retrained models.

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.

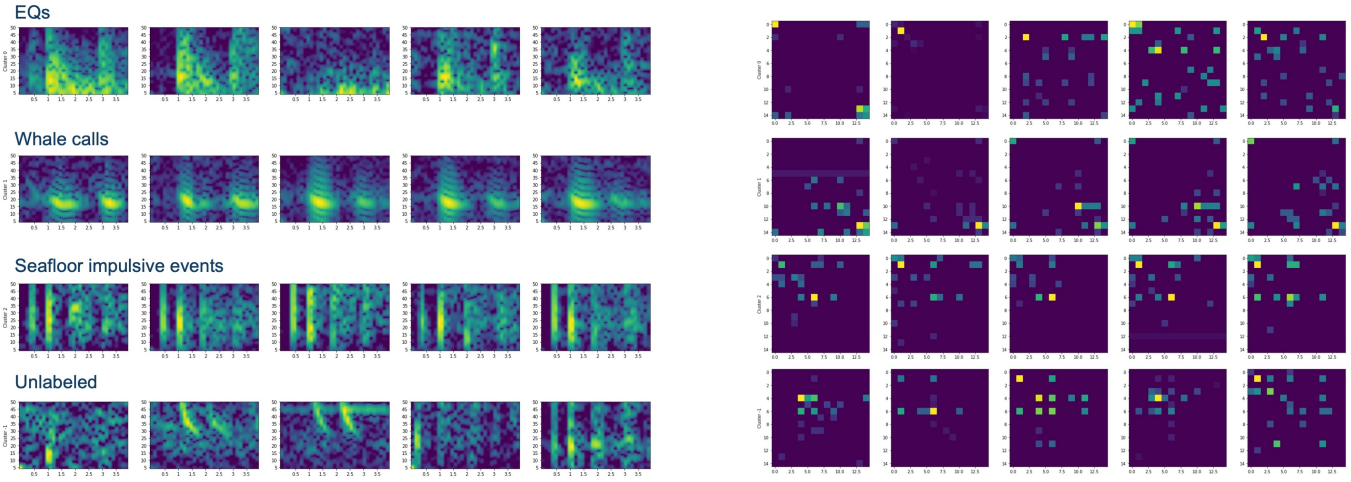
Results

Our objective in this study is to distinguish between different event sources in the catalog so that we can perform refined seismic monitoring on 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), Calinski-Harabasz index (Caliński and Harabasz, 1974). However, these metrics did not reveal a sharp elbow that could

clearly define the optimal cluster numbers (Nainggolan et al., 2019). Therefore, we resorted to 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 result shows different levels of clustering in the spectral feature space. On the first order, we identified three main clusters that share common spectral patterns across the network. The three clusters are a group of earthquakes, a group of whale calls, and a group of seafloor impulsive events (Figure 3). 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 sub-groupings. The higher-level clusters reflect subtle differences in the spectral patterns, which help to separate seismicity generated by different physical processes during the eruption. The temporal distributions of signals in different classes are shown in Figure 4a and 4b.

To improve on the single-station clustering approach used in Wang et al. (2024), we employed self-training techniques for model retraining (Yarowsky, 1995). We started by creating an initial training dataset that was comprised 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 own 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 2b and 2c illustrates this enhancement. The contingency matrix between station AXCC1 and station AXID1 shows a notable improvement in label coherency after the retraining process.



(a) Spectrograms

(b) Fingerprints

Figure 3: Examples of spectrogram and corresponding fingerprints for different clusters. (a) spectrogram examples. (b) fingerprint examples.

Having consistent labels across the network ensures that source effect is dominant over path effect in event clustering. We show that we effectively discriminate signals from different sources, as shown in Figure 3.

The initial ML catalog includes 200,891 earthquakes, 53,018 whale calls (predominantly from 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 sub-groups 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).

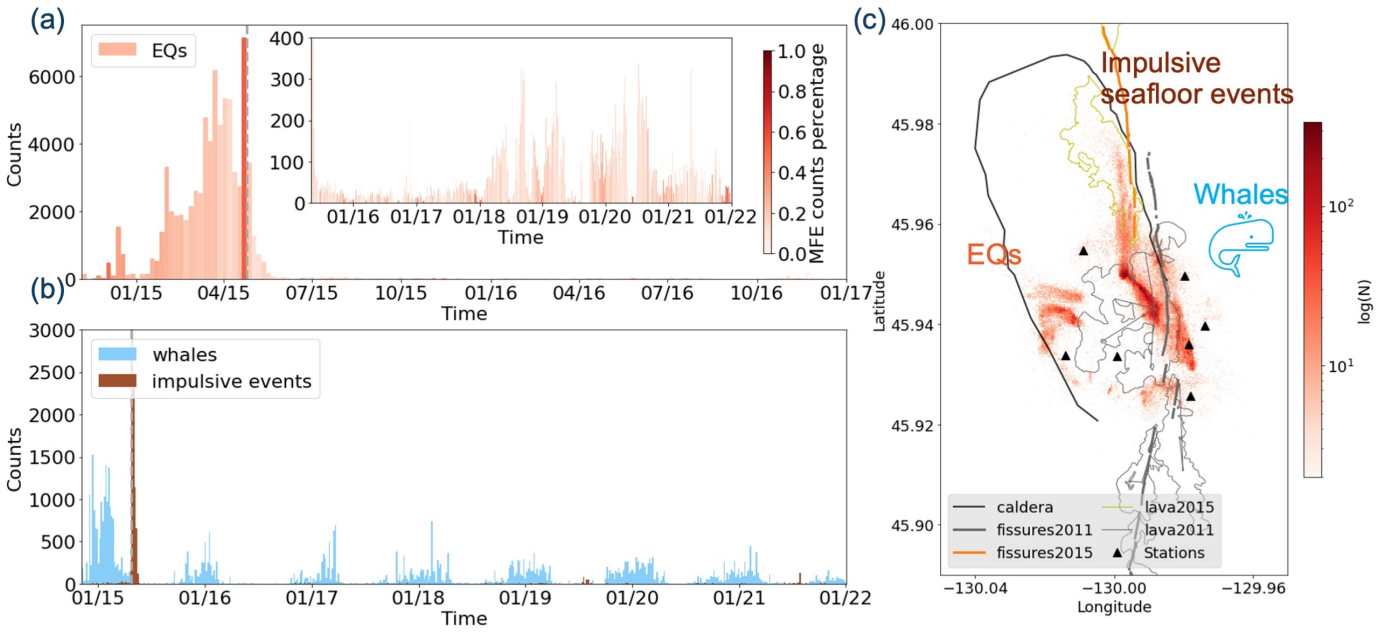
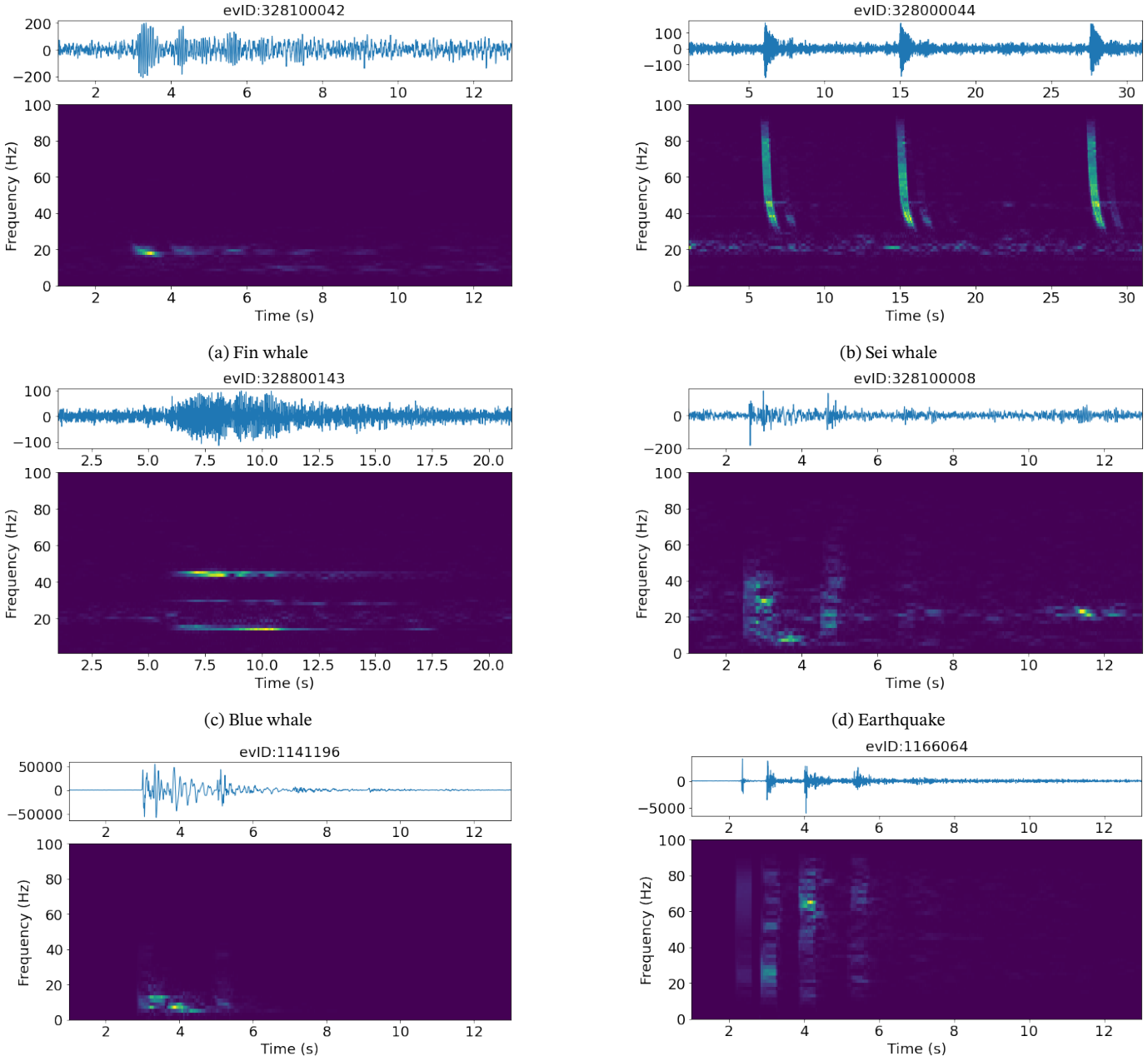


Figure 4: Histograms of different types of sources in seismic monitoring and their locations. (a) Histogram shows earthquake total rate. The bin size is 5 days. Bars of earthquake rate is colored by MFE count percentage in each bin. Main figure shows activity before and during the 2015 eruption. Dashed gray line marks the eruption onset. Inset figure shows a zoom in view of post eruption period from May 2015 to the end of 2021. (b) Blue and brown bars show whale call and impulsive seafloor events rate. (c) Map of earthquake density in log scale. Eruptive fissure and lava flow locations of 2015 and 2011 eruption is plotted with colored contours, showing approximate locations of the impulsive seafloor events.

Although we successfully classified most events into three main classes, a few remain unlabeled after retraining, examples shown in (Figure 3). These events either had inconsistent labels across the network or were weak signals with clear picks at less than two stations. Upon examining these unlabeled events, we found that they comprised a few event classes that are less frequently seen in the dataset, such as calls from different species of whales (as shown in Figure 5b and 5c) and ship noise. Fin whales are the most common species that pass by the OBS array every year during winter seasons and thus contribute to one of the three main clusters in 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 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 distinct spectral content from the other event classes (examples in Figure 5b and 5c). Based on their spectral features, We designed a simple frequency-dependant classifier that computes energy ratio in the 30-50 Hz frequency band relative to the 4-50 Hz band to distinguish them during real-time operation.



(a) Fin whale

(b) Sei whale

(c) Blue whale

(d) Earthquake

(e) MFE

(f) Impulsive seafloor event

Figure 5: Examples of spectrogram and corresponding waveform for different species of whale calls ((a)Fin whale, (b)Sei whale, and (c)Blue whale), volcanic earthquakes, precursory MFEs, and co-eruption impulsive seafloor events. The spectrograms are plotted with 100 samples (0.5 s) window length and 90 samples overlap.

The DD-RT base catalog, obtained after relocating the ML catalog of 200,891 earthquakes, includes 144,329 precisely located events constrained by 400 million correlation delay times and 1,702,714 ML P- and S-picks. Mean RMS of the delay times are 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. Differences between ML pick and correlation times have standard deviations of 17 ms, 40 ms, and 28 ms for the P wave on vertical component and the S wave on two horizontal components (for $cf > 0.95$ pairs), which can be considered the upper bound average pick uncertainty (Figure 7). The relocated catalog of seismicity images 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.

Discussions

The quality of the base catalog is crucial in double-difference monitoring as 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) and 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 Hypoinverse 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 wave picks in our new ML-DD 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 matched P and S wave picks 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 having systematic biases on the pick times for both P and S waves. The Kurtosis picker consistently picks earlier on S waves compared to the PhaseNet picks, while 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 due to its windowing nature and the differences in S wave picks might be related to the complications caused by P coda and S to P/P to S conversions before S arrival.

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 $cf > 0.8$ group, which is the threshold we use to select correlation measurements for double-difference relocation. However, the ML picks have much lower standard deviations for the error distribution for the $cf > 0.95$ group, 17 ms for P waves, and 40 ms/28 ms for S waves on E / N components, compared with 35 ms (P) and 51 ms/42 ms (S on E/N components) for the Kurtosis picks. This

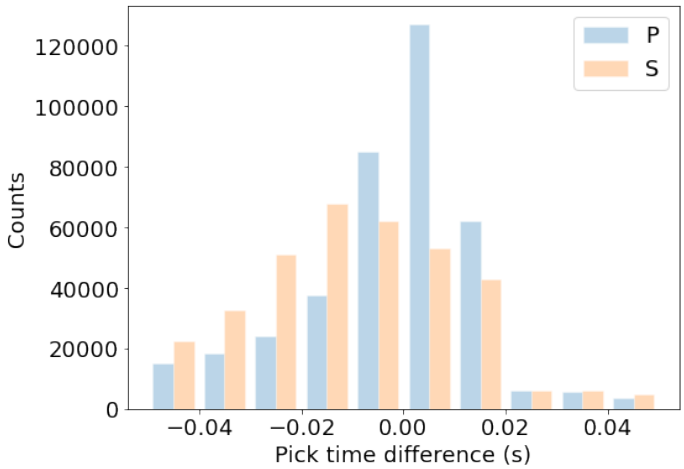


Figure 6: Histograms of the differences between matched phase picks. Pick time difference is plotted as Kurtosis pick time - PhaseNet pick time.

suggests the ML catalog has higher pick accuracy for the best constrained events, and has better performance when picking low signal-to-noise ratio data (more picks with similar uncertainty to the Kurtosis catalog).

We also compared the final earthquake location results of the three catalogs. To make a straightforward comparison of location quality, we plot the locations of the same subset of events with matched IDs. Figure 8a shows the comparison of the final double-difference earthquake locations of the ML catalog (this study) and the Waldhauser et al. (2020) catalog. We see that the high resolution earthquake locations in this study delineates 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 however more diffused. To understand which processing step contributed to this improvement, we also made comparisons of the initial locations before the double difference relocation. Figure 8b and 8c show the differences between the initial NonLinLoc (NLL) locations in the ML catalog (this study) and the initial NLL locations in the Waldhauser et al. (2020) catalog (Figure 8b), and Hypoinverse (HINV) locations in Wilcock et al. (2016) catalog (Figure 8c). The differences between Figure 8b and 8c shows the contribution of using a 3-D velocity model. We can see that in Figure 8b the initial locations of the two NLL catalogs match well for the overall patterns, whereas in Figure 8c the NLL locations and the HINV locations show significant differences. This suggests that the differences in initial locations, i.e. the use of a 3-D velocity model with NLL as opposed to a 1-D velocity model by HINV, contributed considerably to the improvement of the final earthquake location results. Figure 8b shows the location differences caused by the absolute picks, while Figure 8a shows the differences after double-difference relocation. In Figure 8b the overall pattern of two catalogs agrees and the differences are sharpened in Figure 8a. This suggests that improvements in pick accuracy and especially cross-correlation measurements helped in refining the complex structures 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 (Figure 4a). The overall seismicity pattern is generally consistent with the active faults and structures activated during the 2015 eruption, with recent seismicity bursts more

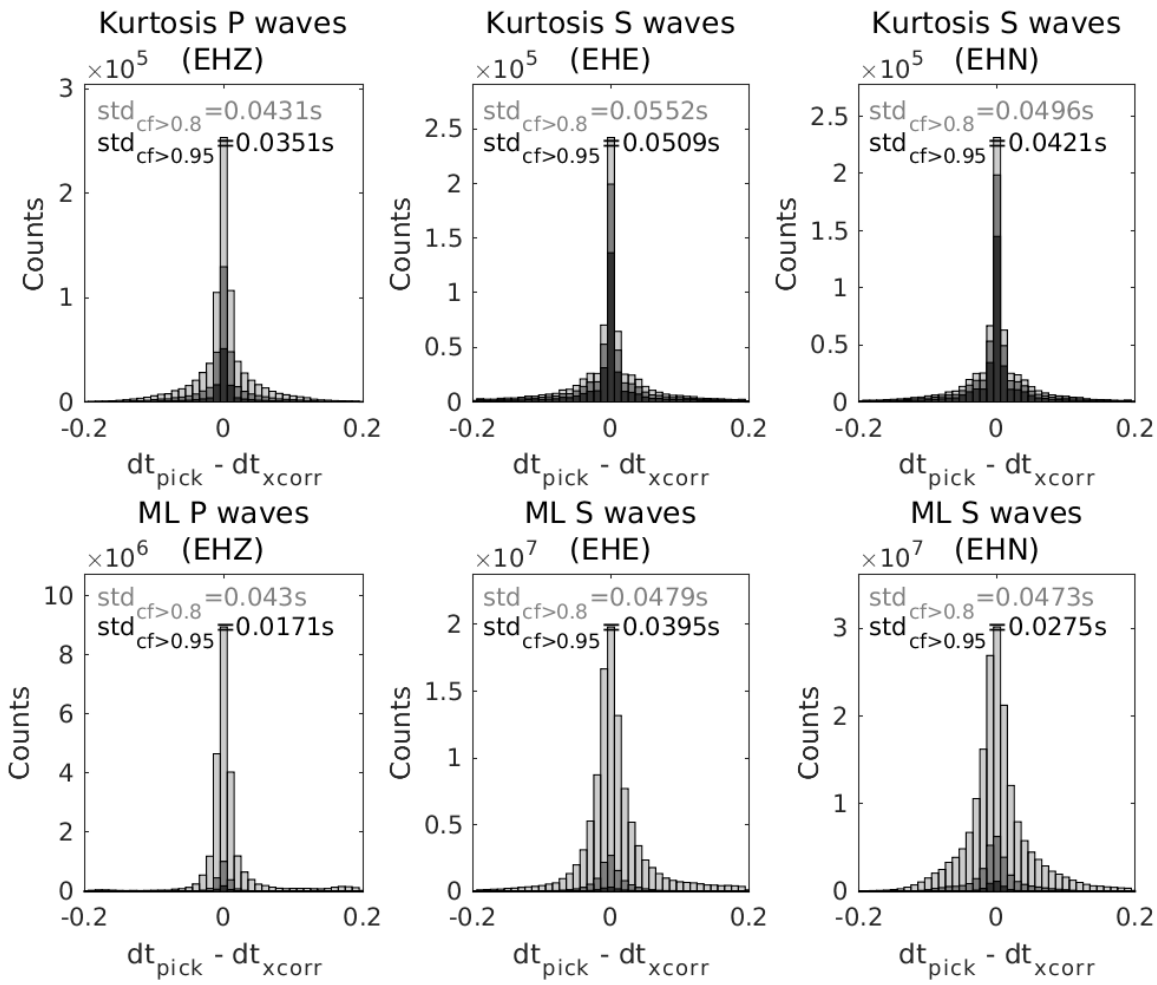
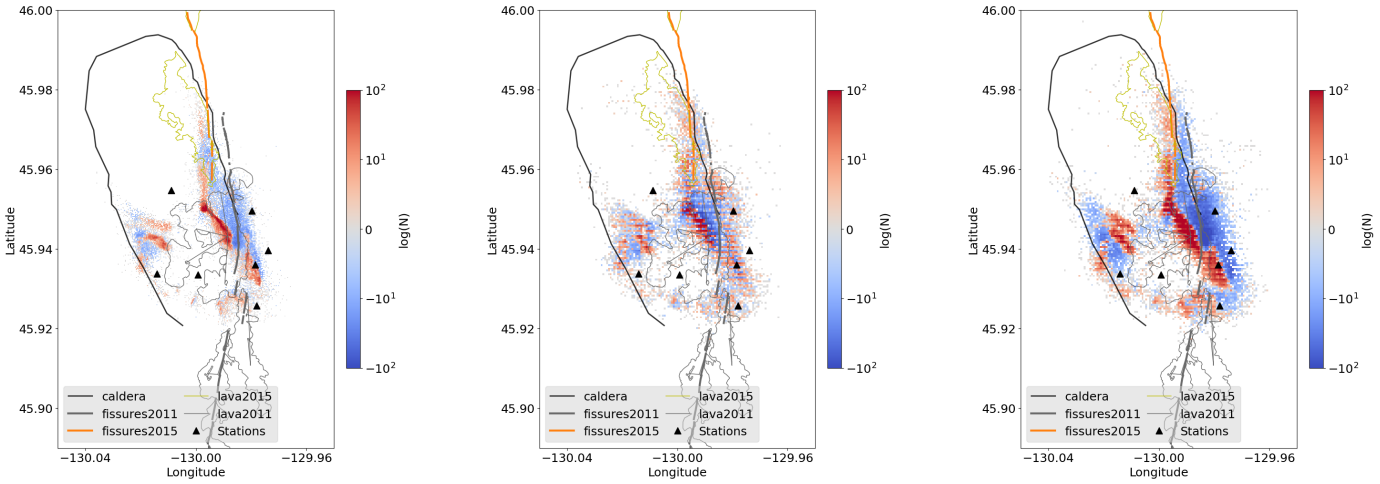


Figure 7: Histograms of the phase picks error distribution as compared to cross-correlation delay time measurements. Top panel are the Kurtosis P and S wave picks and bottom panel are the PhaseNet P and S wave picks. Light gray, gray, and dark gray bars are the distributions of different cross-correlation coefficient groups: $cf > 0$, 8, $cf > 0.9$, and $cf > 0.95$. The standard deviations of different cross-correlation coefficient groups are labeled on each subplot.

concentrated in the southern part of the caldera near the hydrothermal field International District (Kelley et al., 2014). The automatic near-real-time analysis of the continuous waveform data has identified 31,710 earthquakes since the end of the base catalog, 5,773 whale calls (65% of them from fin whales), 1 misidentified impulsive events, and 10,975 unlabeled events. We set a relatively tight threshold in 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 threshold Chadwick et al. (2023). However, its inflation rate has gradually decreased over the past few years and levelled out in the last year, deviating from the steady inflation pattern observed in previous cycles (Chadwick Jr 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



(a) ML-DD catalog with the Waldhauser et al. (2020) catalog

(b) ML-NLL initial locations with the Kurtosis-NLL initical locations in the Waldhauser et al. (2020) catalog

(c) ML-NLL initial locations with the Kurtosis-HINV initical locations in the Wilcock et al. (2016) catalog

Figure 8: Comparison of earthquake locations in this study and catalogs in Wilcock et al. (2016) and Waldhauser et al. (2020). The three plots are earthquake density differences between the catalogs in log-scale counts. Red and blue color shows where the ML catalog has more events and less events compared with the other catalog. The bin size is 25 m × 25 m in (a) and 50 m × 50 m in (b) and (c).

events, including short term precursory events, in real-time. Our current real-time system is set to operate on a daily basis during the low-activity period. The real time ML-DD catalog can be accessed at `https://axialdd.ldeo.columbia.edu/index.html.proto`. 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 minutes, when 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 relocation step, further saving computational time for the purpose of 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.

Furthermore, we are expecting the deployment of a new temporary OBS array which 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 robust and accessible operation.

Conclusions

In this study, we presented an ML-based processing workflow for Axial Seamount that effectively discriminats events from various types of source processes and constructed a high-resolution earthquake catalog. We addressed challenges unique to the Axial OBS dataset, such as station-dependant noise signatures and separation of non-earthquake signals with diverse spectral patterns. We implemented this workflow for real-time seismic monitoring and event classification, showing its capa-

bility to track pre-eruption precursors and co-eruption volcanic signals as they evolve. This may help improve the short-term eruption forecasting for Axial Seamount's next eruption with the potential to generalize to seismic activity at other submarine volcanoes.

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 pre-trained semi-supervised 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 (Figure 4), including precursory MFEs (Figure 5e) that potentially indicate the preparation of an eruption (Wang et al., 2024) and seafloor impulsive events (Figure 5f) that can be used to track magma outflows during an eruption (Caplan-Auerbach et al., 2017).

Data and Resources

The seismic data used in this study are downloaded from the Incorporated Research Institutions for Seismology (IRIS) Data Management Center (DMC).

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.[1]

Acknowledgments

We would like to thank the reviewers for providing helpful comments and suggestions. This work is funded by NSF awards OAC-2103741 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**(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. *Bulletin of the Seismological Society of America* **104**(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. *Journal of Geophysical Research: Solid Earth* **124**(12), 12997–13020.

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**(6433), eaau0323.

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

Beroza, G. C., M. Segou, and S. Mostafa Mousavi (2021). Machine learning and earthquake forecasting—next steps. *Nature communications* **12**(1), 4761.

Caliński, T. and J. Harabasz (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods* **3**(1), 1–27.

Caplan-Auerbach, J., R. Dziak, J. Haxel, D. Bohnenstiehl, and C. Garcia (2017). Explosive processes during the 2015 eruption of a xial s eamount, as recorded by seafloor hydrophones. *Geochemistry, Geophysics, Geosystems* **18**(4), 1761–1774.

Chadwick, W. W., 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. *AGU23*.

Chadwick Jr, W., 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. *Geophysical Research Letters* **43**(23), 12–063.

Chadwick Jr, W. W., 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. *Geochemistry, Geophysics, Geosystems* **23**(1), e2021GC010153.

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. *Geochemistry, Geophysics, Geosystems* **23**(11), e2022GC010475.

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

Davies, D. L. and D. W. Bouldin (1979). A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence* (2), 224–227.

Holtzman, B., N. Groebner, T. Sawi, T. Xing, M. Pec, H. Ghaffari, U. Mok, R. Skarbek, J. Paisley, T. Mittal, et al. (2021). Unsupervised spectral feature extraction applied to acoustic emissions during brittle creep of basalt under dry and wet conditions. In *AGU Fall Meeting Abstracts*, Volume 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. *Science advances* **4**(5), eaao2929.

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. *Marine Geology* **352**, 426–450.

Klein, F. W. (2002). User's guide to hypoinverse-2000, a fortran program to solve for earthquake locations and magnitudes. Technical report, US Geological Survey.

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. *Geochemistry, Geophysics, Geosystems* **21**(9), e2020GC009136.

Lomax, A., A. Michelini, A. Curtis, and R. Meyers (2009). Earthquake location, direct, global-search methods. *Encyclopedia of complexity and systems science* **5**, 2449–2473.

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. *Advances in seismic event location*, 101–134.

Mousavi, S. M. and G. C. Beroza (2023). Machine learning in earthquake seismology. *Annual Review of Earth and Planetary Sciences* **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. In *Journal of Physics: Conference Series*, Volume 1361, pp. 012015. IOP Publishing.

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 and Planetary Science Letters* **259**(3-4), 500–516.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* **20**, 53–65.

Rubin, K. H., S. A. Soule, W. W. Chadwick Jr, 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. *Oceanography* **25**(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. *Journal of Geophysical Research: Earth Surface*, e2022JF006909.

Schaff, D. P. and F. Waldhauser (2005). Waveform cross-correlation-based differential travel-time measurements at the northern california seismic network. *Bulletin of the Seismological Society of America* **95**(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, 17–19 sj geophys. res. 107 (b6).

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. *Oceanography* **31**(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**(7632), 261–265.

Waldhauser, F. (2009). Near-real-time double-difference event location using long-term seismic archives, with application to northern california. *Bulletin of the Seismological Society of America* **99**(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. *Bulletin of the Seismological Society of America* **90**(6), 1353–1368.

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. *Journal of Geophysical Research: Solid Earth* **125**(5), e2019JB018796.

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 2003-2013. *Plos one* **12**(10), e0186127.

Wilcock, W. S., R. P. Dziak, M. Tolstoy, W. W. Chadwick Jr, S. L. Nooner, D. R. Bohnenstiehl, J. Caplan-Auerbach, F. Waldhauser, A. F. Arnulf, 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. *Oceanography* **31**(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**(6318), 1395–1399.

Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods. In *33rd annual meeting of the association for computational linguistics*, pp. 189–196.

Zhu, W. and G. C. Beroza (2019). Phasenet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International* **216**(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. *Geophysical Journal International* **232**(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. *Journal of Geophysical Research: Solid Earth* **127**(5), e2021JB023249.

Mailing address

Kaiwen Wang: 201A Seismology, 61 Route 9W, Palisades, NY 10964

Felix Waldhauser: 210 Seismology, 61 Route 9W, PO Box 1000, Palisades, NY 10964

David Schaff: 108C Seismology, P.O. Box 1000, 61 Route 9W, Palisades, NY 10964

Maya Tolstoy: 1492 NE Boat St, Suite 200 Seattle, WA 98105

William Wilcock: 1503 NE Boat Street, Box 357940, Seattle, WA 98195-7940

Yen Joe Tan: Room 317, 3/F, Mong Man Wai Building, The Chinese University of Hong Kong Shatin, N.T., Hong Kong

List of Figure Captions

Figure 1: ML-based workflow for catalog construction and event discrimination.

Figure 2: (a) Examples of event spectrograms at a broad band station AXCC1 and two short period stations AXAS1 and AXEC1. Initial arrival and reflections in the water column are marked by arrows. Contingency matrix of cluster labels at two OBS stations AXCC1 and AXID1 for (b)single-station models and (c)retrained models.

Figure 3: Examples of spectrogram and corresponding fingerprints for different clusters. (a) spectrogram examples. (b) fingerprint examples.

Figure 4: Histograms of different types of sources in seismic monitoring and their locations. (a) Histogram shows earthquake total rate. The bin size is 5 days. Bars of earthquake rate is colored by MFE count percentage in each bin. Main figure shows activity before and during the 2015 eruption. Dashed gray line marks the eruption onset. Inset figure shows a zoom in view of post eruption period from May 2015 to the end of 2021. (b) Blue and brown bars show whale call and impulsive seafloor events rate. (c) Map of earthquake density in log scale. Eruptive fissure and lava flow locations of 2015 and 2011 eruption is plotted with colored contours, showing approximate locations of the impulsive seafloor events.

Figure 5: Examples of spectrogram and corresponding waveform for different species of whale calls ((a)Fin whale, (b)Sei whale, and (c)Blue whale), volcanic earthquakes, precursory MFEs, and co-eruption impulsive seafloor events. The spectrograms are plotted with 100 samples (0.5 s) window length and 90 samples overlap.

Figure 6: Histograms of the differences between matched phase picks. Pick time difference is plotted as Kurtosis pick time - PhaseNet pick time.

Figure 7: Histograms of the phase picks error distribution as compared to cross-correlation delay time measurements. Top panel are the Kurtosis P and S wave picks and bottom panel are the PhaseNet P and S wave picks. Light gray, gray, and dark gray bars are the distributions of different cross-correlation coefficient groups: $cf > 0, 8$, $cf > 0.9$, and $cf > 0.95$. The standard deviations of different cross-correlation coefficient groups are labeled on each subplot.

Figure 8: Comparison of earthquake locations in this study and catalogs in Wilcock et al. (2016) and Waldhauser et al. (2020). The three plots are earthquake density differences between the catalogs in log-scale counts. Red and blue color shows where

the ML catalog has more events and less events compared with the other catalog. The bin size is 25 m × 25 m in (a) and 50 m × 50 m in (b) and (c).

Manuscript Received 00 Month 0000