Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multifractal signatures reveal source-dependent differences between tectonic earthquakes and glacial tremors

Walid E. AboElnasr

walidelsayed@std.mans.edu.eg

Mansoura University

M. A. Zahran

Mansoura University

Mohamed M. Abdelsalam

Mansoura University

Elshaimaa Amin

Mansoura University

Article

Keywords: Seismic Discrimination, Multifractal analysis, Support Vector Machine, Seismic Waveform, Icequake, Earthquake

Posted Date: July 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9663926/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Multifractal signatures reveal source-dependent differences between tectonic earthquakes and glacial tremors

Walid E. AboElnasr*[1], M. A. Zahran[1], Mohamed M. Abdelsalam[2,3], and Elshaimaa Amin[4]

[1]Theoretical Physics Group, Physics Department, Faculty of Science, Mansoura University, Mansoura, Egypt
[2]Computers and Control Systems Engineering Dep., Faculty of Engineering, Mansoura University,Mansoura, 35511, Egypt
[3]Faculty of Engineering, Mansoura National University, Gamasa, 35712, Egypt
[4]Future Higher Institute of Engineering and Technology, Mansoura, Egypt

Abstract

Discriminating between local tectonic earthquakes and cryo-seismic events (icequakes) poses a significant challenge for automated seismic monitoring due to the similarity of their waveforms in the time and frequency domains. To address this, we present a robust classification framework utilizing Multifractal Detrended Fluctuation Analysis (MFDFA) combined with machine learning. Unlike traditional methods that rely on linear spectral features, this approach quantifies the intrinsic non-linear scaling properties and structural complexity of the seismic source. We analyzed seismic data from the Columbia Glacier region, Alaska, categorized into station groups based on Signal-to-Noise Ratio (SNR) and data availability. Analysis of a seismic network dataset across diverse SNR conditions reveals that tectonic earthquakes exhibit higher structural complexity and broader multifractal spectra in contrast to the narrower, quasi-monofractal signatures of cryo-seismic sources. This physical distinction enables robust discrimination across varying noise regimes. A Support Vector Machine (SVM) classifier achieved a global average accuracy of 93.95% with individual station performance peaking at 97.35%. These results demonstrate that the proposed multifractal framework provides a physically grounded and highly accurate tool for purifying seismic catalogs in complex glaciated environments.

Keywords— Seismic Discrimination, Multifractal analysis, Support Vector Machine, Seismic Waveform, Icequake, Earthquake

1 introduction

The reliability of seismic hazard assessment and risk mitigation strategies depends fundamentally on the integrity of earthquake catalogs. Seismic networks record a vast array of ground motion, ranging from tectonic earthquakes to non-tectonic events such as quarry blasts, traffic noise, and cryo-seismic phenomena (icequakes). While automated detection algorithms have significantly improved data

*Corresponding author: Walid E. AboElnasr; email: walidelsayed@std.mans.edu.eg

processing efficiency, the accurate classification of these events remains a persistent challenge (e.g., [1] [2] [3]). This issue is particularly acute in glaciated regions like Alaska, where cryo-seismic events occur frequently and often spatially overlap with active tectonic faults. Discriminating between local tectonic earthquakes and icequakes is difficult because they share similar time-domain characteristics, including overlapping magnitude ranges and impulsive onsets [4]. Since manual discrimination is no longer feasible given the exponential growth of seismic data volumes, the development of robust, automated discrimination techniques is essential for modern observational seismology.

1.1 Related work

Extensive research has been dedicated to automated event classification. Conventional approaches typically extract features from the time domain (e.g., amplitude ratio, kurtosis) or the frequency domain (e.g., spectral centroids)[5][6]. While effective for simple noise separation, these linear methods often struggle to distinguish between event types with subtle structural similarities. More recently, machine learning techniques have been employed successfully, including unsupervised approaches (clustering, dimensionality reduction [7]), supervised algorithms such as hidden Markov models [8], support vector machine (SVM)[9], and random forests [10] and [11] have demonstrated good performance on specific datasets. Furthermore, deep learning architectures, particularly convolutional neural networks
(CNNs)[12] have emerged as a powerful tool for seismic classification [13] and [14]. However, deep learning models often function as black boxes lacking physical interpretability and typically require massive labeled datasets for training, which are not always available for specific local phenomena such as icequakes. Moreover, standard spectral analysis techniques often assume signal stationarity, a property rarely found in the complex, non-linear dynamics of seismic rupture processes.

To address these limitations, this study proposes a classification framework based on fractal and multifractal analysis. Seismic signals are inherently non-linear and non-stationary, reflecting the complex dynamics of their source mechanisms. Previous work has successfully applied multifractal dimensions to distinguish earthquakes from quarry blasts [15]. Building on this, we hypothesize that the physical process of a tectonic fault rupture involving heterogeneous stress release and wave scattering generates a signal with a higher degree of complexity (multifractality) compared to the simpler brittle fracture mechanisms of icequakes. Multifractal Detrended Fluctuation Analysis (MFDFA) provides a rigorous mathematical toolkit to quantify these intrinsic fluctuations, offering a fingerprint of the source dynamics that captures information missed by linear methods.

1.2 The research contributions

In this paper, we develop a comprehensive framework to discriminate between tectonic earthquakes and cryo-seismic events using multifractal analysis combined with machine learning. The specific contributions of this work are as follows:

1. Large-scale diverse dataset: We compiled and processed a comprehensive dataset of 28,022 waveforms from the Columbia Glacier region, Alaska. The dataset was constructed to span a wide range of signal-to-noise ratio (SNR) conditions, including both high-quality stations and lower-SNR recordings, enabling a rigorous assessment of the robustness of the multifractal features across varying signal qualities.
2. Robust feature extraction framework: We implemented a well-defined and reproducible workflow for extracting multifractal features, with particular emphasis on the singularity spectrum width ($\Delta\alpha$), generalized dimensions (D_q), and signal persistence (α_0). In contrast to approaches based on qualitative descriptors, this framework provides a mathematically grounded methodology that captures the intrinsic nonlinear scaling behavior of seismic signals.
3. Physical interpretation: We provide a rigorous physical interpretation of the extracted multifractal features, directly linking their signatures to the underlying source processes. The

high structural complexity and persistence observed in earthquakes reflect heterogeneous fault rupture dynamics, whereas the simpler anti-persistent signatures of icequakes correspond to impulsive brittle fracture mechanisms.

4. State-of-the-art classification: The proposed framework achieved a global average accuracy of 93.95%, validated through 10-fold stratified cross-validation across the seismic network. This result demonstrates that physics-based multifractal features capture the underlying physical processes of seismicity more effectively than traditional linear descriptors, enabling accurate discrimination while preserving clear physical interpretability essential for seismic hazard analysis.

2 Methodology

2.1 Multifractal detrended fluctuation analysis

MFDFA, proposed by Kantelhardt et al.[16], was used to extract the mean multifractal features. Indeed, the generalized MF-DFA procedure involves five distinct steps. For instance, let $x(i)$ for $i = 1, \ldots, N$ be a time series or a waveform, where N indicates its length.
Step 1: Construct the profile by integration after subtracting the mean$\langle$x$\rangle$ as defined.

$$Y(i) \equiv \sum_{k=1}^{i} \left[x_k - \langle x \rangle\right], i = 1, \ldots, N \tag{1}$$

$$\langle x \rangle = \frac{1}{N} \sum_{k=1}^{N} x_k \tag{2}$$

Step 2: Divide the profile $Y(i)$ into $N_s \equiv \frac{N}{s}$, non-overlapping segments of equal length s.
Step 3: Calculate the local trend for each of the $2N_s$ segments. To do this, a polynomial of degree m is assumed to fit the profile in each of the $2N_S$ segments. For this, the variance can be defined as follows.

$$F^2(s, \nu) \equiv \frac{1}{s} \sum_{i=1}^{s} \left\{Y[(\nu - 1)s + i] - y_\nu(i)\right\}^2, \tag{3}$$

for each segment $v, v = 1, \ldots, N_s$ and

$$F^2(s, \nu) \equiv \frac{1}{s} \sum_{i=1}^{s} \left\{Y\left[N - (\nu - N_s)\, s + i\right] - y_\nu(i)\right\}^2, \tag{4}$$

for each segment $v = N_s + 1, \ldots, 2N_s$. Hence, $y_v(i)$ is the fitting polynomial in segment v. Linear, quadratic, cubic, or higher-order polynomials can be applied in the fitting procedure.
Step 4: For $2N_s$ intervals, the q_{th} order fluctuation function can be computed by averaging over all segments as follows.

$$F_q(s) \equiv \left\{ \frac{1}{2N_s} \sum_{\nu=1}^{2N_s} \left[F^2(s, \nu)\right]^{\frac{q}{2}} \right\}^{\frac{1}{q}}, \tag{5}$$

When $q \to 0$, the fluctuation function can be reduced to the following form.

$$F_0(s) = \exp\left\{ \frac{1}{4N_s} \sum_{\nu=1}^{2N_s} \ln\left[F^2(s, \nu)\right] \right\} \tag{6}$$

The index variable q can take any real value except zero. When $q < 0$, the size of $F_q(s)$ depends on the size of the small fluctuation deviation $F^2(s, v)$, and when $q > 0$, the size of $F_q(s)$ depends on the size of the large fluctuation deviation $F^2(s, v)$. In this way, different values of q describe the effects of different degrees of fluctuation.
Step 5: Determine the scaling behavior of the fluctuation functions by analyzing $log - log$ plots $F_q(s)$ versus s for each value of q. If the series x_k is long-range power-law correlated, $F_q(s)$ increases with increasing S as a power law $F_q(s) \sim s^{H(q)}$.

In general, the exponent $H(q)$ depends on q. However, $H(q)$ is defined by the following relationship:

$$\tau(q) = qH(q) - 1 \tag{7}$$

Finally, the relationship between the generalized Hurst exponent $H(q)$ and the scale exponent $\tau(q)$ is defined via the well-known Legendre transformation[17],[18]

$$f(\alpha) = q\alpha - \tau(\alpha) \tag{8}$$

where α is defined as the Hölder exponent and $f(\alpha)$ indicates the dimension of the subset of the series that is characterized by it [19] If the time series under study is monofractal, the function $f(\alpha)$ is a certain or a constant value; if the time series is multifractal, the function $f(\alpha)$ consists of various fractal exponents.

This function effectively captures characteristics of the time series, revealing both its random fluctuations (stochastic behavior) and its complex, self-similar patterns (multifractal behavior). We analyze the behavior of $f(\alpha)$. The value of the function at the maximum is easy to determine by using the expression

$$\tau(q) = q\alpha(q) - f(\alpha(q))$$

or

$$(q-1)D_q = q\alpha(q) - f(\alpha(q))$$

At q = 0, we obtain that $f(\alpha_o) = D_o$. It is equal to the Hausdorff dimension. In the cases q=1 and q =2, applying the above equation gives the information and correlation dimensions.

In general, the multifractal dimension is extracted from each waveform, as shown in figure 1

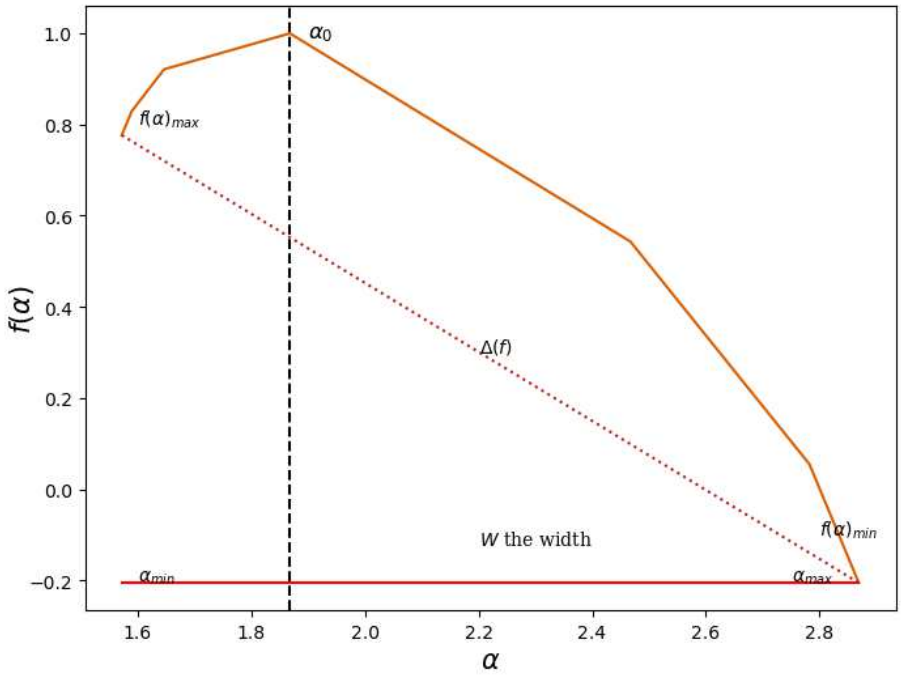


Fig. 1: multifractal spectrum

Table 1: Mathematical Definitions and Physical Interpretations of Extracted Features

Feature Symbol	**Physical Interpretation**
α_0	Represents the most probable or dominant regularity (texture) of the seismic signal. Higher values indicate a smoother underlying process.
$\alpha_{\min}$	Corresponds to the most singular (roughest) components of the signal, such as abrupt onsets or spikes.
$\alpha_{\max}$	Corresponds to the most regular (smoothest) components of the signal, often associated with the coda or background noise.
$\Delta\alpha$	Quantifies the geometric dominance of fluctuations. $\Delta f < 0$ implies that the signal structure is dominated by strong singularities (roughness), whereas $\Delta f > 0$ implies dominance of smooth features.
A	Describes the skewness of the multifractal spectrum. $A > 0.5$ indicates a left-skewed spectrum (strong complexity), while $A \approx 0.5$ indicates a symmetric spectrum.
D_0	Characterizes the fractal (capacity) dimension of the signal.
D_1	Quantifies the information content and entropy of the signal distribution (information dimension).
D_2	Measures the correlation properties of the signal; often used to estimate the attractor dimension in phase space (correlation dimension).

Upon computing the multifractal spectrum $f(\alpha)$ and the generalized fractal dimensions D_q, we extracted a set of 10 robust features to characterize the seismic waveform. These features quantify the complexity, regularity, and geometric structure of the signal. Table 1 summarizes the mathematical definitions and physical interpretations of each extracted feature.

3 The Dataset

the seismic data were retrieved from the alaska Earthquake center(AEC) catalog spanning the period from january 1 2005 to january 1 2022, following the data collection strategy outlined by [20], we focused on a circular region with a 50 KM radius centred on the columbia glacier (latitude 61.219722°N, longitude of 146.895278°W) as shown in Fig 2

we applied the following criteria:

1. Icequake :we retrieved all available events labelled as glacial within the spatial and temporal window. This yielded a total of 2650 icequake events.
2. Earthquakes: we queried for local tectonic earthquakes with magnitudes between 0 and 3.0. the raw catalog contained over 10.000 such events

a significant challenge in seismic classification is the inherent imbalance between tectonic seismicity and rarer cryo-seismic events.to address this, we retained all 2650 icequake but restricted the tectonic dataset to the most recent earthquake (occuring between may 2018 and april 2022). This near-parity ensures that the machine learning model is trained on a balanced representation of both physical phenomena.

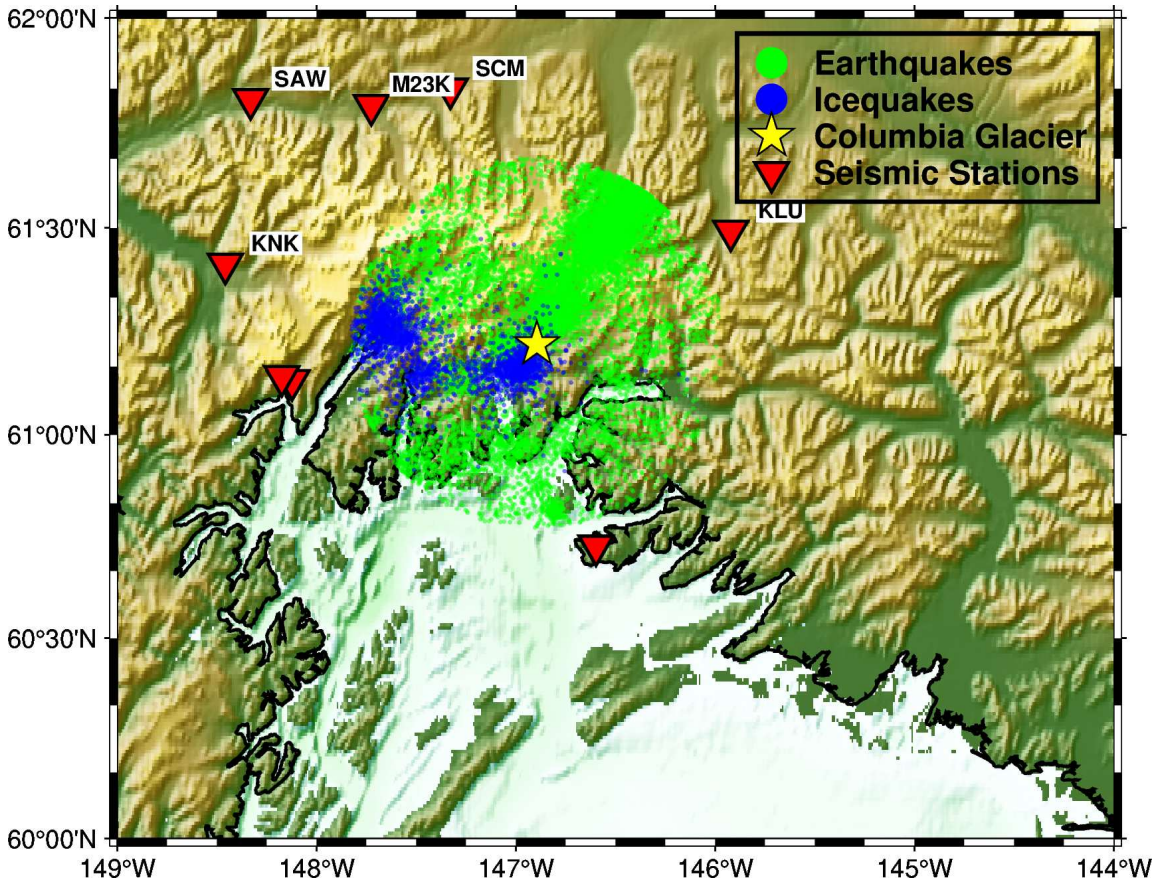


Fig. 2: Spatial distribution of the seismic dataset used in this study, located in the region of Alaska. The background topography illustrates the rugged terrain surrounding the Columbia Glacier. Red inverted triangles mark the locations of the 15 broadband seismic stations. The scatter plot shows the epicenters of the selected events: green circles represent the 3,000 tectonic earthquakes ($M_L \leq 3.0$) and blue circles represent the 2,650 icequakes concentrated near the glacier terminus.

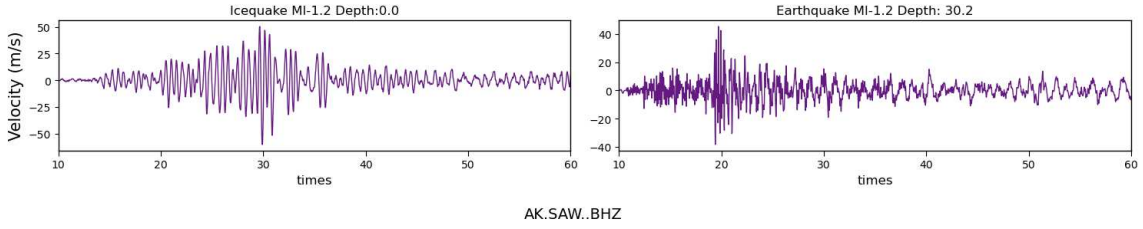


Fig. 3: Comparison of representative vertical-component (BHZ) waveforms recorded at station AK.SAW. (Left) An icequake occurring at the Columbia Glacier surface (Depth 0.0 km). (Right) A tectonic earthquake ($M_L = 1.2$) located at a depth of 30.2 km. Waveforms have been preprocessed (0.5-25 Hz bandpass filter) and are shown in a zoomed 50-second window to highlight the structural differences. Note the impulsive onset and rapid attenuation of the icequake compared to the complex coda of the tectonic earthquake.

For each event, a 5-minute (300 s) window of the vertical-component (BHZ) waveform was extracted. The raw data were resampled to 50 Hz to standardize the sampling rate across all stations. Instrument responses were removed using the ObsPy library (BeyreuthBeyreuther et al., 2010), and the signals were detrended and band-pass filtered between 0.5 and 23 Hz to isolate the frequency band of interest and suppress noise.

glacial calving events and icequakes exhibit waveform signatures that are distinct from typical tectonic siesmicity. as illustrated in fig 3,icequake are generally characteristic by emergent onsets,lack of distinct body-wave phases(P and S wave) and a quasi-monochromatic frequency content,typically concentrated in the 1-3Hz range.in contrast tectonic earthquakes or comparable magnitude usually display impulsive onset,clear phase arrivals and a broader frequency spectrum.

historically, classifying these events has relied on visual inspection of these qualitative traits by human analyst. however,this process is subjective and computationally insufficient for large dataset.the objective of this study is to more beyond qualitative visual assessment by identifying robust,quantitative multifractal features, that can automatically distinguish between these two event types with high precision.

4 Support Vector Machines (SVMs)

Support Vector Machines (SVMs) are a type of supervised machine learning algorithm used for classification and regression tasks [21]. The core principle of SVMs is to find the optimal hyperplane that best separates data points into distinct classes in a multidimensional space. This hyperplane is defined by a subset of data points known as support vectors. The diagram below visualizes a decision boundary, or hyperplane, separating two different classes 4.

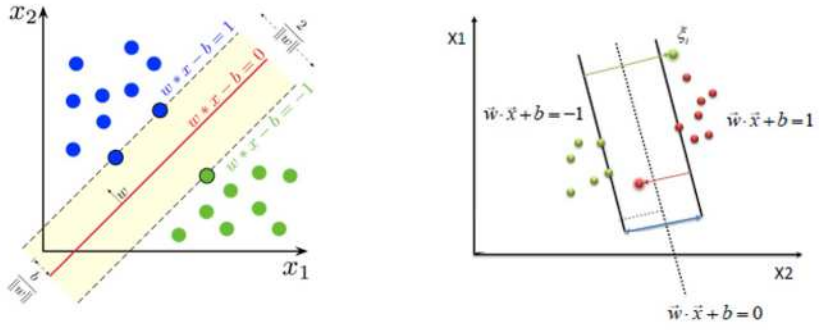


Fig. 4: Maximum-margin hyperplane and margins for an SVM; this margin is known as support vectors

The concept of a separating hyperplane that separates the classes is extended through soft-margin classification, which adds a C penalty to the objective function for observations in the training set that are misclassified. The SVM algorithm chooses a decision boundary that optimizes the trade-off between a wider margin and a lower total error penalty. The η variable allows some observations to fall on the wrong side of the margin, but penalizes them by parameter C, as shown in the functions below.

$$\min_{w,b} \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}\xi_i$$

$$y_i\left(\sum_{k=1}^{K} w_k\phi(x_{i,k}) + b\right) \geq 1-\xi_i, \quad \xi_i \geq 0 \quad \forall i$$

The implementation used the scikit-learn Python package for machine learning SVC and optimization, available at:
https://github.com/scikit-learn/scikit-learn

In this study, discrimination between icequakes and earthquakes can be performed using support vector machine algorithms. The classification technique is based on extracting the ten most variable multifractal features of the waveform using multifractal analysis. Based on the extracted multifractal features, as shown in table 1, we employed a support vector machine (SVM) classifier. The SVM was chosen for its effectiveness in high-dimensional feature spaces and its robustness when using non-linear kernels.

Hyperparameter optimization: To maximize model performance and ensure generalization, we utilized a systematic grid search approach (GridSearchCV) with 5-fold cross-validation. We optimized the following hyperparameters:

- Kernel: radial basis function (RBF), to capture non-linear relationships in the multifractal feature space.
- Regularization Parameter (C): We searched the space [10, 100,1000]. A higher C was found to be effective, indicating that a decision boundary with a tighter margin was necessary to separate the classes.
- Kernel coefficient (Γ): We searched ['scale', 0.01, 0,001]

Final model selection: The optimal model configuration was identified as C = 1000m Γ = 0.01 and kernel ='rbf'. This optimized model achieved a test accuracy of 93.95%.

5 Result and discussion

To evaluate the discriminatory power of the proposed framework, we analyzed a refined dataset comprising 28,002 labeled waveforms recorded across nine seismic stations in the Columbia Glacier region. We employed a rigorous 10-fold stratified cross-validation scheme to ensure the statistical robustness of our results. For each waveform, the multifractal detrended fluctuation analysis (MFDFA) algorithm was applied to extract a suite of singularity spectrum features, quantifying the non-linear scaling behavior of the source mechanism. Fig. 5 illustrates the multifractal singularity spectra for earthquakes (red curves) and icequakes (blue curves) for some waveforms.

Broader Multifractal Spectra in Earthquakes:
Earthquake spectra (red curves) span a wider range of singularity exponents (α) than those of icequakes (blue curves). The spectrum width $\Delta\alpha$ quantifies the signal multifractality; wider $\Delta\alpha$ values imply a more intricate heterogeneous process with a diverse array of scaling behaviors. Consequently, earthquake waveforms demonstrate greater multifractal complexity than icequake waveforms.

Shift in Peak Singularity (α_0):
The modal exponent α_o, the peak of each spectrum, differs between earthquakes and icequakes, indicating a shift in their dominant scaling regimes. This displacement reflects a fundamental change in the typical regularity of the seismic process for each event type.

Spectral Asymmetry and Signal Characteristics:
The shape of each spectrum's tails provides insight into the nature of its fluctuations:

- Left-Heavy Earthquakes: The red curves display a pronounced left tail, signaling an abundance of sharp high-frequency singularities (lower α values). This asymmetry corresponds to abrupt transient features in the waveform.
- Right-Heavy Icequakes: In contrast, the blue curves often have a longer right tail, indicating prevalent smooth, low-frequency components (higher α values). This reflects more gradual persistent oscillations in the icequake signals.

Symmetry of the Spectrum around the Peak:

- Left-Skewed Earthquakes: As mentioned before, the earthquake spectra appear left-skewed. This indicates that the multifractal structure is more sensitive to large fluctuations. This is consistent with the physics of earthquakes, which involve sudden, high-energy releases.
- More Symmetric Icequakes: The icequake spectra seem to be more symmetric around their peak. This suggests a more balanced distribution of singularities and a less event-driven or explosive character compared to earthquakes.

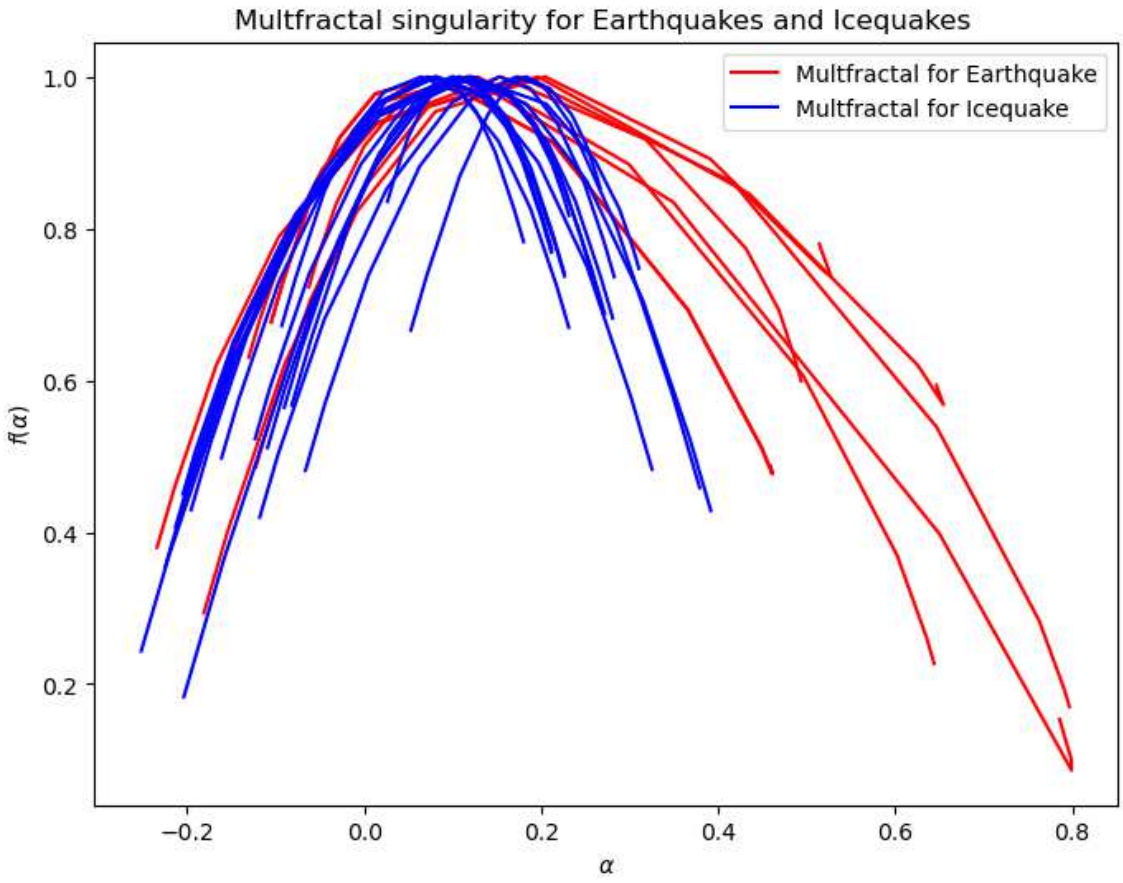


Fig. 5: Multifractal singularity spectra for earthquakes and icequakes

To investigate the physical validity of the proposed features, we analyzed the probability density functions of the key multifractal parameters across the nine high-performance stations. Figure 6 illustrates the distributional differences between tectonic earthquakes (orange) and cryo-seismic icequakes (blue).

The visual analysis reveals three consistent physical patterns across the network:

1. Complexity and Information Content ($D_1 and D_2$): As shown in 6(a)(b), the distributions for D_1 (Information Dimension) and D_2 (Correlation Dimension) show that tectonic earthquakes consistently exhibit a shift toward higher values compared to icequakes. This indicates that earthquake ruptures possess a higher structural complexity and information density. In contrast, icequakes tend to cluster at lower dimensions, reflecting the simpler, more impulsive nature of tensile fracturing.

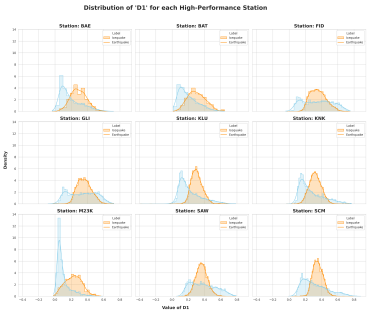
2. Signal Persistence and Regularity α_o: As shown in 6(c), the most distinct separation is observed in this parameter (spectrum peak). Earthquakes generally center around $\alpha_o \approx 0.5 - 0.6$, indicating long-range correlation (persistence) in the waveform. Conversely, icequake distributions are shifted to lower values ($\alpha_o < 0.4$), characteristic of anti-persistent or random signals. This separation is particularly sharp at high-fidelity stations like KNK and GLI.

3. Singularity Strength (α_{min}): As shown in 6(d), the α_{min} distributions show that icequakes frequently exhibit lower values than earthquakes. Physically, this suggests that cryo-seismic

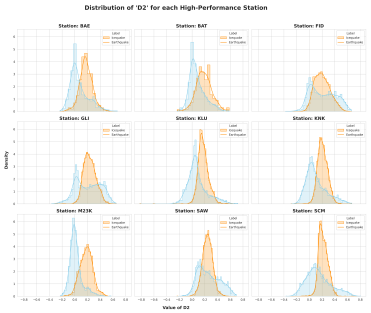
events are dominated by stronger, more abrupt singularity fluctuations (shocks) compared to the relatively more regular rupture propagation of tectonic events.

4. Multifractal Width ($\delta\alpha$):

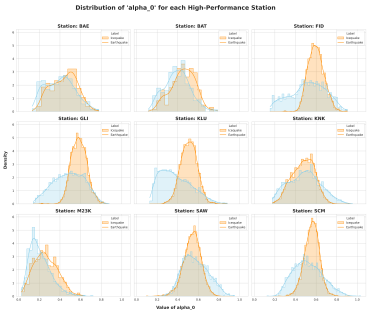
   As shown in 6(e), while there is some overlap due to signal attenuation, tectonic earthquakes generally display a broader distribution in Spectrum Width ($\Delta\alpha$). This confirms that tectonic sources exhibit a richer, more heterogeneous multifractal structure, whereas icequakes often appear quasi-monofractal (narrower width), indicating a more uniform source process.



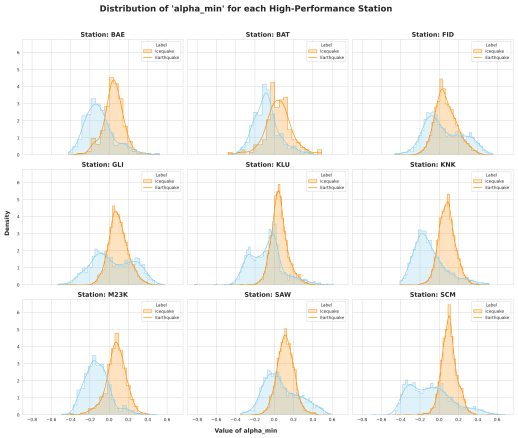
(a) PDFs of D_1



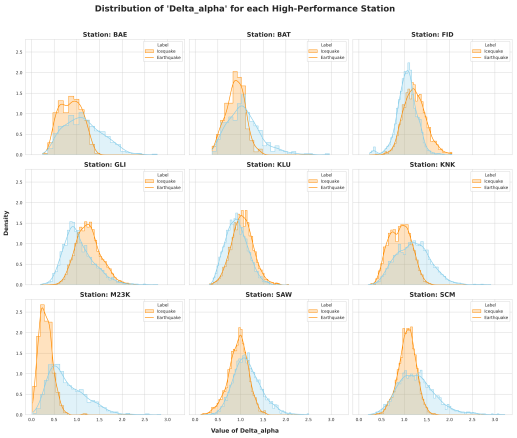
(b) PDFs of D_2



(c) PDFs of α_o



(d) PDFs of α_{min}



(e) PDFs of $\Delta\alpha$

Fig. 6: Probability density functions (PDFs) of key multifractal features across the nine seismic stations. The plots illustrate the distribution of (a) Information Dimension (D_1), (b) Correlation Dimension (D_2), (c) Spectrum Peak (α_o), (d) Singularity Strength (α_{min}), and (e) Spectrum Width ($\Delta\alpha$). Tectonic earthquakes (orange) consistently exhibit higher fractal dimensions and stronger signal persistence compared to cryo-seismic icequakes (blue), which display lower complexity and anti-persistent characteristics. This statistical separation between the two classes reflects the distinct physical source mechanisms, providing the discriminatory basis for the machine learning framework.

While the degree of overlap varies slightly by station (e.g., station BAE shows more overlap than KNK due to local site effects), the fundamental modal separation between the two classes remains stable across the network. This statistical divergence provides the physical basis for the high classification accuracy (93.95%) achieved by the machine learning model.

The proposed method is based on extracting the multifractal features from the seismic waveforms, then feeding all these features to the machine learning model. In this case, the support vector machine is chosen. To evaluate the discriminatory power of the proposed framework, we evaluated the classification accuracy across the nine stations using a 10-fold stratified cross-validation scheme. The results summarized in Figure 7 and Table 2 demonstrate the superior discriminative capability of multifractal features.

The framework achieved a global average accuracy of 93.95% (indicated by the red dashed line in Figure 7). Performance was remarkably consistent across the network, with every station exceeding a 90% accuracy threshold. Station KNK yielded the highest performance, achieving an exceptional accuracy of 97.35%, likely due to the high signal-to-noise ratio (SNR) and clear spectral separation at this site. Stations FID and GLI also showed superior results with accuracies of 95.91% and 95.45%, respectively.

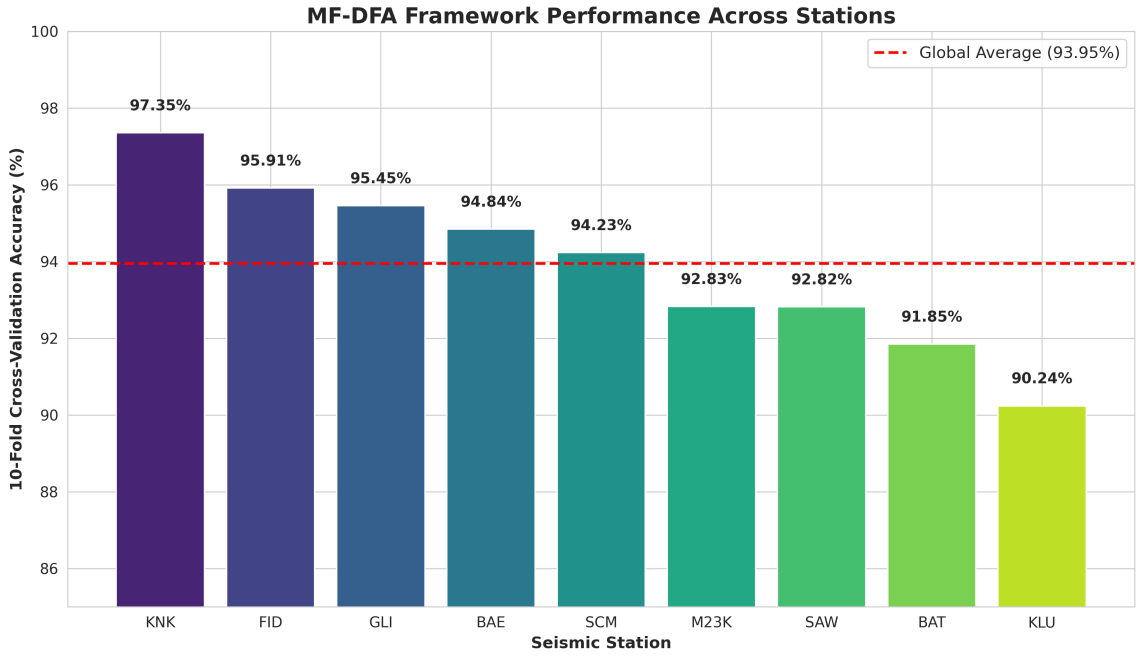


Fig. 7: Station-wise classification accuracy of the MFDFA/SVM framework using 10-fold cross-validation

Even at station KLU, which recorded the lowest accuracy within this high-fidelity subset, the model maintained a robust performance of 90.24%. This tight performance range (90.24% - 97.35%) indicates that the multifractal signatures of tectonic and cryo-seismic events are physically distinct and remain detectable across different recording environments, provided that extreme environmental noise artifacts are appropriately filtered.

Table 2: Multifractal Performance Accuracy for Different Stations

Station	MF Accuracy	MF Sensitivity	MF Specificity
BAE	0.948413	0.976744	0.933735
BAT	0.918478	0.843137	0.947368
FID	0.959128	0.959184	0.959052
GLI	0.954545	0.941504	0.968563
KLU	0.902352	0.905714	0.899004
KNK	0.973541	0.982256	0.963504
M23K	0.928279	0.911392	0.936364
SAW	0.928188	0.931790	0.923767
SCM	0.942295	0.962963	0.917044
Average	**0.9395**		

Unpacking the Exceptional Results: The Impact of Multifractal Features

The classification performance presented in Table 2 is driven by the unique capability of multifractal metrics to resolve the underlying physical difference between shear-slip (earthquake) and tensile (icequake) mechanisms. Specifically, the machine learning model exploits four critical structural signatures:

- Intrinsic Complexity ($\Delta\alpha$): The model leverages the distinct separation in spectral width. It identifies that tectonic earthquakes exhibit a broad, heterogeneous multifractal structure, distinguishing them from the simpler, quasi-monofractal nature of cryo-seismic events.
- Information Density (D_1): The information dimension (D_1) serves as a proxy for the entropy of the seismic source. The model learns that tectonic ruptures generate signals with higher information content (higher D_1) due to the complex interaction of fault asperities, whereas icequakes exhibit lower D_1 due to the complex interaction of fault asperities, whereas icequakes exhibit lower.
- Signal Persistence (α_o): By analyzing the dominant regularity α_o, the model detects that earthquake waveforms possess stronger long-range correlations. In contrast, icequakes display a more random, anti-persistent texture characteristic of localized fracturing.
- Structural Asymmetry (A): The model utilized spectral asymmetry as a discriminator of the source process; it recognizes that tectonic signals are structurally dominated by strong singularity fluctuations (negative asymmetry), a feature that provides a powerful fingerprint for separating them from symmetric background signals.

In essence, the proposed method succeeds because it transforms the raw waveform into a multidimensional signature of the source physics. By feeding the SVM a feature set that quantifies the non-linear dynamics of the rupture, we achieve the robust discrimination results demonstrated in this study.

Implications for Seismic Hazard and Risk Management.

The performance of the MFDFA framework, evidenced by the high-confidence metrics in Table 2, translates directly into significant improvements for seismic hazard and risk management in cryospheric regions. The contribution of this study is the ability to automate the production of high-purity seismic catalogs, which serve as the fundamental input for probabilistic seismic hazard assessment.

- Preventing Catalog Contamination: In glaciated regions, catalog contamination is a primary source of error in hazard models. Our framework achieved a global average specificity of

93.87%, with station GLI reaching as high as 96.86%. This high level of performance ensures that non-tectonic events are effectively excluded, preventing the artificial inflation of seismicity rates.

- The completeness of seismic records is essential for accurately identifying and mapping active tectonic faults. The framework demonstrated an average sensitivity of 93.49%, with station KNK achieving 98.23%. This ensures that the vast majority of true seismic events are correctly identified and preserved. By capturing these events, our method guarantees that hazard models are built upon a comprehensive dataset of regional stress release.
- Scalability for Real-Time Monitoring: By validating the method on a large-scale dataset of over 28,000 waveforms with consistent results (Global Accuracy: 93.95%), we have demonstrated that this framework is stable enough for automated deployment. This capability significantly reduces the burden of manual analysis, enabling real-time catalog cleaning in data-rich environments like Alaska.

6 Challenges, limitations and Future work

While the proposed model achieved a state-of-the-art global accuracy of 93.95% on the Columbia Glacier dataset, it is crucial to acknowledge that seismic and cryo-seismic signatures can vary regionally. A critical next step for future research is to validate the transferability of this model to different geological settings, such as strike-slip faults versus subduction zones, and different glaciological environments. Cross-regional validation will determine if the multifractal features identified here are universal indicators of the source physics or if region-specific retraining is required.

The current study utilized a quality control protocol based on stations that represent a variety of Signal-to-Noise Ratio (SNR) conditions across the network. Our results indicate that signal quality is a significant factor in the visibility of multifractal signatures. Stations with high SNR, such as KNK and FID, yielded exceptional accuracies exceeding 95%, while stations with lower SNR still maintained robust performance above 90%. However, as observed in our preliminary analysis of the broader network, extremely low SNR environments still pose a challenge, as background noise can begin to distort the singularity spectrum. Future studies should systematically evaluate the model's performance degradation as a function of SNR to define precise operational limits for real-time monitoring.

Additionally, the extraction of multifractal features is computationally more intensive than standard linear filtering. While the current workflow is highly effective for offline catalog curation and post-processing (processing 28,000 waveforms efficiently), its integration into millisecond-latency early warning systems presents a challenge. Future work could focus on optimizing the MFDFA algorithm or developing a lightweight version of the feature set using only the most critical parameters, such as D_1 and $\Delta\alpha$, to balance high accuracy with the speed required for real-time decision making.

7 Conclusion

In this study, we presented a physics-based machine learning framework for discriminating between tectonic earthquakes and icequakes. By moving beyond traditional linear descriptors and leveraging Multifractal Detrended Fluctuation Analysis (MFDFA), we successfully quantified the intrinsic structural differences between the complex rupture dynamics of earthquakes and the simpler fracture mechanisms of icequakes.

We validated this approach on a large-scale dataset of 28,022 waveforms from varaity seismic network. The proposed method achieved a global average accuracy of 93.95%, with individual station performance reaching as high as 97.35%. This result, confirmed through rigorous 10-fold cross-validation, offers a direct, practical solution for automated seismic catalog cleaning. By

effectively filtering cryo-seismic noise, this study ensures that Probabilistic Seismic Hazard Assessments are physically grounded in crustal deformation processes, reducing uncertainty in regional hazard models and strengthening risk mitigation strategies in glaciated environments.

Declarations

Availability of data and materials

The seismic waveform data used in this study were obtained from the IRIS Data Management Center. The earthquake parametric data, produced by the Alaska Earthquake Center, were retrieved from the ANSS Comprehensive Catalog (https://earthquake.usgs.gov/data/comcat/). The processed datasets generated during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Authors' contributions

The authors contributed to this work as follows:

WEA: Led the overall study, including conceptualization of the research, development of the methodology, software implementation, formal analysis of the data, and writing of the original manuscript draft.
MAZ: Supervised the project, contributed to the investigation by validating the results, and provided critical input on the physical interpretation of the multifractal features. He also participated in the review and editing of the manuscript.
MMA: Contributed to the investigation by validating the results and provided key conceptual guidance. He also contributed to the final review and editing of the manuscript.
EA: Contributed to the investigation by validating the results and provided key conceptual guidance. She also contributed to the final review and editing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to express their gratitude to the providers of the seismic data used in this work. We also thank the editorial team and reviewers for their time and valuable feedback.

References

[1] B. Julian. Volcanic termor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research*, 991, 06 1994.

[2] Jacques Deparis, D. Jongmans, Fabrice Cotton, Laurent Baillet, François Thouvenot, and Didier Hantz. Analysis of rockfall and rockfall avalanche seismograms in the french alps. *Bulletin of The Seismological Society of America - BULL SEISMOL SOC AMER*, 98:1781–1796, 08 2008.

[3] Clément Hibert, David Michéa, Floriane Provost, J.-P Malet, and Marten Geertsema. Exploration of continuous seismic recordings with a machine learning approach to document 20 years of landslide activity in alaska. *Geophysical Journal International*, 219, 07 2019.

[4] Timothy Draelos, Matthew Peterson, Hunter Knox, Benjamin Lawry, Kristin Phillips-Alonge, Abra Ziegler, Eric Chael, Christopher Young, and Aleksandra Faust. Dynamic tuning of seismic signal detector trigger levels for local networks. *Bulletin of the Seismological Society of America*, 108, 03 2018.

[5] Hanae Morioka, Hiroyuki Kumagai, and Takuto Maeda. Theoretical basis of the amplitude source location method for volcano-seismic signals. *Journal of Geophysical Research: Solid Earth*, 122:6538–6551, 08 2017.

[6] Changbin Wang, Guangyao Si, Chengguo Zhang, Anye Cao, and Ismet Canbulat. A statistical method to assess the data integrity and reliability of seismic monitoring systems in underground mines. *Rock Mechanics and Rock Engineering*, 54, 11 2021.

[7] William Jenkins, Michael Bianco, and Peter Bromirski. Unsupervised deep clustering of seismic data: Monitoring the ross ice shelf, antarctica. *Journal of Geophysical Research: Solid Earth*, 126, 09 2021.

[8] Moritz Beyreuther, Roberto Carniel, and Joachim Wassermann. Continuous hidden markov models: Application to automatic earthquake detection and classification at las canãdas caldera, tenerife. *Journal of Volcanology and Geothermal Research*, 176:513–518, 10 2008.

[9] Jari Kortström, Marja Uski, and Timo Tiira. Automatic classification of seismic events within a regional seismograph network. *Comput. Geosci.*, 87(C):22–30, February 2016.

[10] Alexis Falcin, Jean-Philippe Métaxian, Jérôme Mars, Éléonore Stutzmann, Jean-Christophe Komorowski, Roberto Moretti, Marielle Malfante, François Beauducel, Jean-Marie Saurel, Céline Dessert, Arnaud Burtin, Guillaume Ucciani, Jean-Bernard de Chabalier, and Arnaud Lemarchand. A machine-learning approach for automatic classification of volcanic seismicity at la soufrière volcano, guadeloupe. *Journal of Volcanology and Geothermal Research*, 411:107151, 2021.

[11] Clément Hibert, David Michéa, Floriane Provost, J.-P Malet, and Marten Geertsema. Exploration of continuous seismic recordings with a machine learning approach to document 20 years of landslide activity in alaska. *Geophysical Journal International*, 219, 07 2019.

[12] S. Mostafa Mousavi, Weiqiang Zhu, William Ellsworth, and Gregory Beroza. Unsupervised clustering of seismic signals using deep convolutional autoencoders. *IEEE Geoscience and Remote Sensing Letters*, 16(11):1693–1697, 2019.

[13] João Paulo Canário, Rodrigo Mello, Millaray Curilem, Fernando Huenupán, and Ricardo Rios. In-depth comparison of deep artificial neural network architectures on seismic events classification. *Journal of Volcanology and Geothermal Research*, 401:106881, 05 2020.

[14] Nadège Langet and Fred-Marcus-John Silverberg. Automated classification of seismic signals recorded on the Åknes rock slope, western norway, using a convolutional neural network. *Earth Surface Dynamics*, 11:89–115, 02 2023.

[15] Alexey Lyubushin, Zdeněk Kaláb, Markéta Lednická, and Haggag Mohamed. Discrimination of earthquakes and explosions using multi-fractal singularity spectrums properties. *Journal of Seismology*, 17:975–983, 07 2013.

[16] Jan W. Kantelhardt, Stephan A. Zschiegner, Eva Koscielny-Bunde, Shlomo Havlin, Armin Bunde, and H.Eugene Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316(1):87–114, 2002.

[17] Benoit B. Mandelbrot. *Self-similarity and panorama of self-affinity*, pages 146–197. Springer New York, New York, NY, 1997.

[18] Feder J . Fractals. *New York, NY, USA: Plenum Press*, volume 44. *New York:Springer.*, 1988.

[19] Kantelhardt, Jan and Zschiegner, Stephan and Koscielny-Bunde, Eva and Bunde, Armin and Havlin, Shlomo and Stanley, H. Multifractal detrended fuctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316, 02 2002.

[20] Akash Kharita, Marine Denolle, and Michael West. Discrimination between icequakes and earthquakes in southern alaska: an exploration of waveform features using random forest algorithm. *Geophysical Journal International*, 237, 03 2024.

[21] Marti A. Hearst. Support vector machines. *IEEE Intelligent Systems*, 13(4):18–28, July 1998.