Improving Seismic Signal Classification through Novel Similarity-Based Techniques

A. Atmani, E. H. Ait Laasri, M. Derife, N. Abboune, S. Afnzar, W. Dinne

National School of Applied Sciences – Ibn Zohr University (ENSA – UIZ))

INTRODUCTION AND MAIN RESULTS

This study introduces a novel, simple, and computationally efficient approach for the classification of seismic events, based on the Euclidean distance function applied in both the frequency and time–frequency domains. When tested on volcanic data comprising four distinct classes, the proposed method achieved an overall classification accuracy of approximately 80% in the frequency domain and 93.7% in the time–frequency domain.

SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

Improving Seismic Signal Classification through Novel Similarity-Based Techniques

A. Atmani, E. H. Ait Laasri, M. Derife, N. Abboune, S. Afnzar, W. Dinne

P3.5-596

Introduction

Nowadays, in order to save time and resources, it is essential to automatically classify the large volume of events continuously recorded by global seismic networks. The primary objective of this classification is to identify the source of each event (e.g., earthquake, volcanic activity, quarry blasts, nuclear explosions, or human activities). Advanced seismic event classification can even provide insights into the physical processes occurring at the source. This is particularly relevant in the case of volcanic seismic activity, where the classification of signals recorded by local seismic networks can reveal processes such as explosions, rock fracturing, degassing, magmatic intrusions, eruptions, pressurization, and depressurization.

A wide range of techniques has been developed for the classification of volcanic seismic events. Many of these approaches rely on artificial intelligence neural networks, typically preceded by a feature extraction step. Previous work carried out by our research team focused on classifying the same database using the inter-correlation function in both time and time–frequency domains [1]. However, this approach was computationally expensive in terms of both processing time and memory requirements.

In this study, we propose a simple, fast, and memory-efficient method for the classification of volcanic seismic events, based on the Euclidean distance function in both frequency and time–frequency domains.

Methods/Data

In the frequency domain, the similarity between two signals u(i) and v(i) is estimated using the Euclidean distance (***ED***), defined as :

$$ED_{uv} = \sqrt{\sum_{i=1}^{n} \left(u(i) - v(i)\right)^2}$$

For one-dimensional signals, as in our database, the ***ED*** can be simplified to the sum of absolute differences:

$$ED_{uv} = \sum_{i=1}^{n} |u(i) - v(i)|$$

In the time–frequency domain, the ED is first computed for each frequency component j, and then integrated over all components:

$$ED_{uv} = \sum_{j=1}^{m} \sum_{i=1}^{n} |u_j(i) - v_j(i)|$$

The dataset used in this study consists of seismic signals from Llaima volcano (Chile), recorded by the vertical component (Z) of the LAV station of the OVDAS monitoring network between 2010 and 2016. The signals were sampled at 100 Hz, band-pass filtered with a 10th-order Butterworth filter in the range 1–10 Hz, normalized to their maximum amplitude, and classified into four categories [2] : Long-Period (LP), Tremor (TR), Volcano-Tectonic (VT), and Tectonic (TC).

Class	LP	TR	VT	TC
Events number	1310	490	304	1488

The ED_{uv} value corresponding to events u and v in the database is placed in the element located at row u and column v of a square, symmetrical matrix, which is then represented as a heat map in Figure 1. The latter illustrates the expected ideal case, where events belonging to the same class are perfectly similar (ED=0), while events from different classes are completely dissimilar (ED=max).

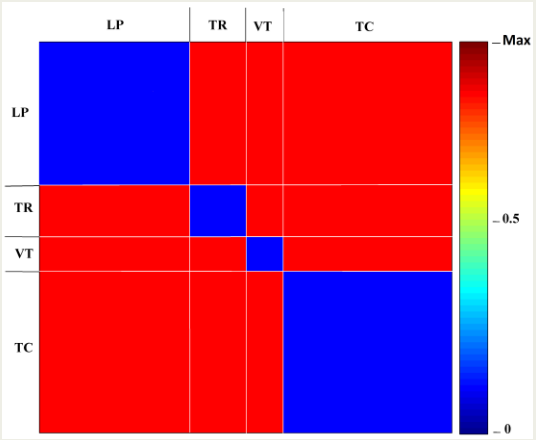

Figure 1 : Expected ideal heat map when the Llaima volcano data-base events are perfectly classified

Improving Seismic Signal Classification through Novel Similarity-Based Techniques

A. Atmani, E. H. Ait Laasri, M. Derife, N. Abboune, S. Afnzar, W. Dinne

Results

Figure 2 presents the heat maps obtained with this method in the frequency domain (left) and the time–frequency domain (right), illustrating the similarity between the Llaima volcano database events. The results confirm that the classification based on the ED method is more effective in the time–frequency domain.

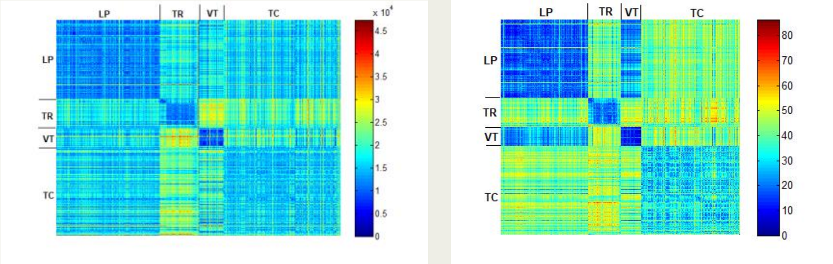

Figure 2 : ED heatmaps in frequency domain (in the left) and in time-frequency domain (in the right) of the Llaima volcano data-base events

To classify each event into one of the four classes, a representative model event is selected for each class. The chosen model corresponds to the event with the minimum mean ED relative to all other events within the same class. The ED of any new event is then computed with respect to the four models, and the predicted class is the one yielding the minimum ED. Figure 3 displays the corresponding confusion matrices. These tables highlight higher diagonal values, representing correctly classified events, in the time–frequency domain (right) compared to the frequency domain (left). The performance evaluation of the proposed classifier is presented in Figure 4 for both the frequency domain (top table) and the time–

		Target Class			
		LP	TR	VT	TC
Predicted class	LP	1085	33	14	247
	TR	7	441	0	51
	VT	48	0	276	119
	TC	170	16	14	1071

		Target Class			
		LP	TR	VT	TC
Predicted class	LP	1256	9	23	23
	TR	1	454	0	39
	VT	39	0	278	48
	TC	14	27	3	1378

Figure 3 : Confusion matrixes obtained by classifying the Llaima volcano data-base events with the ED method in frequency domain (in the left) and in time-frequency domain (in the right)

	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	Error (%)
LP	82,82	87,12	78,68	85,55	14,45
TR	90,00	98,13	88,38	97,02	2,98
VT	90,79	94,92	62,30	94,57	5,43
TC	71,98	90,49	84,26	82,82	17,18

	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	Error (%)
LP	95,88	97,59	95,80	96,97	3,03
TR	92,65	98,71	91,90	97,88	2,12
VT	91,45	97,35	76,16	96,85	3,15
TC	92,61	97,91	96,91	95,71	4,29

Figure 4 : erformance evaluation of the classifier on the base of ED method in frequency domain (top table) and in time-frequency domain (bottom table) applied to the Llaima volcano data-base events

frequency domain (bottom table). All performance indicators exhibit higher values, while error rates are lower in the time–frequency domain. These results demonstrate that the ED-based classifier is more effective in the time–frequency domain, achieving an overall accuracy of 93.71%, compared to 79.98% in the frequency domain.

Conclusions

The Euclidean distance (ED) function represents a simple yet effective approach for the classification of seismic events. The method successfully classifies the four event types in the seismic database, achieving an overall accuracy of approximately 80% in the frequency domain. In the time–frequency domain, the classification performance is further improved, reaching a global accuracy of 93.7%.

References

1. A. Atmani, E. H. Aït Lassri, D. Agliz and E. Akhouayri '*Classification of seismic events using a cross-correlation based approach*', **SnT2013, 19 to 23 June 2023**, Hofburg Palace, Vienna, Austria.
2. J. P. Canario, R. Mello, M. Curilem, F. Huenupan , and R. Rios, "*Llaima Volcano Dataset: In-Depth Comparison of Deep Artificial Neural Network Architectures on Seismic Events Classification*", *Data in Brief*, **30**, **105627**, **2020**.