SnT2023
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE
HOFBURG PALACE - Vienna and Online
19 TO 23 JUNE

Classification of seismic events using a cross-correlation based approach

A. Atmani[1], E.H. Ait Laasri[2], D. Agliz[3], E. Akhouayri[4]
[1,2,3]GEMS Laboratory, National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco
[4]LETSMP Laboratory, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco

INTRODUCTION

Continuous seismic activity monitoring is one the methods used to detect nuclear explosions. Due to the large amount of recorded data, an automatic classification task is mandatory. In this study we have proposed a straightforward approach. The later is evaluated using four event classes.

METHODS/DATA

The proposed approach is based on the so-called cross-correlation function. Two variants of this method are presented. The former is used in the time domain and the later requires time-frequency domain.

The performance of this approach is examined using volcano data composed of four classes.

START

RESULTS

The method achieves a global accuracy of 59,7% in time domain and 92,48% in time-frequency domain using the four classes of the database.

When excluding the TC class, we obtain a global accuracy of 86,3% in time domain and 96.53% in the time-frequency domain.

CONCLUSION

The obtained results showed that the proposed approach can achieve high performance, particularly in time-frequency domain. In this study, the method was tested on a volcano data. Nevertheless, we are confident that the method can reach high performance in any other types of classes.

Please do not use this space, a QR code will be automatically overlayed

P3.5-758

Introduction

Among the most efficient methods used to survey continuously active volcanoes is monitoring its seismic activity.
The main goals of this monitoring are:

- Studying the behavior of the active volcano in order to understand different physical processes occurring inside it (explosions, rock fracturing, degasification, magmatic intrusion, eruptions, pressurization, and depressurization).

- Launching an early alarm when an eruption is about to take place

To achieve this goal, numerous seismic stations are deployed around the volcano, forming a seismic network. Every station transmits its recorded signal to a central observatory where they are retrieved, classified and stored for subsequent analysis and processing.

The classification task of volcano seismic events can be achieved using different methods. Some of them use more sophisticated techniques.

In this work, we propose an easy and straightforward method to classify volcano seismic events using the cross-correlation function in time and time-frequency domains.

Objectives

The main objective of this work is developing a classification approach for seismic signal classification using the cross-correlation function in time and time-frequency domains. Indeed, an automatic classification task becomes nowadays necessary due to the large amount of data recorded on a daily basis. This task can significantly help scientists and especially data seismic analysts to classify their databases to a predetermined number of classes depending on the diversity of the physical sources generating them.

The goal is to develop an easy and straightforward method, avoiding the complex processing steps generally needed in many other proposed methods in the literature.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Methods

Time Domain

To estimate the degree of similarity between two signals *u(t)* and *v(t)*, we use the Maximum Normalized Cross-Correlation (MNCC) function defined below :

$$MNCC_{uv} = \frac{max\ (|R_{uv}(k)|)}{\sqrt{R_{uu}(0)R_{vv}(0)}}$$

$R_{uv}(k)$ is the cross-correlation function of the sampled signals *u(m)* and *v(m)* deduced from the continuous signals *u(t)* and *v(t)*, respectively. It is defined as :

$$R_{uv}(k) = \frac{1}{N}\sum_{m=1}^{N} u(m)v(m-k)$$

N is the number of samples in the signal.

$R_{uu}(0)$ is the autocorrelation of the signal *u(m)*.

Time-Frequency domain

In this domain, we perform a 2D cross-correlation using the spectrogram *U(n,p)* of the signal *u(m)* :

$$U(n,p) = \left|\sum_{m=1}^{N} u(m)w(m-n)\, e^{-i2\pi pm}\right|^2$$

W(m-n) is a time-sliding Hamming window with a length of N/4, centered on discrete time *n* and normalized to unit energy.

$$R_{UV}(k,p) = \frac{1}{N}\sum_{n=1}^{N} U(n,p)V(n-k,p)$$

We then calculate the average in each row of the obtained cross-correlation matrix corresponding to frequency bins M :

$$R_{UV}(k) = \frac{1}{M}\sum_{p=1}^{M} R_{UV}(k,p)$$

We finally calculate the MNCC for the two spectrograms *U* and *V* :

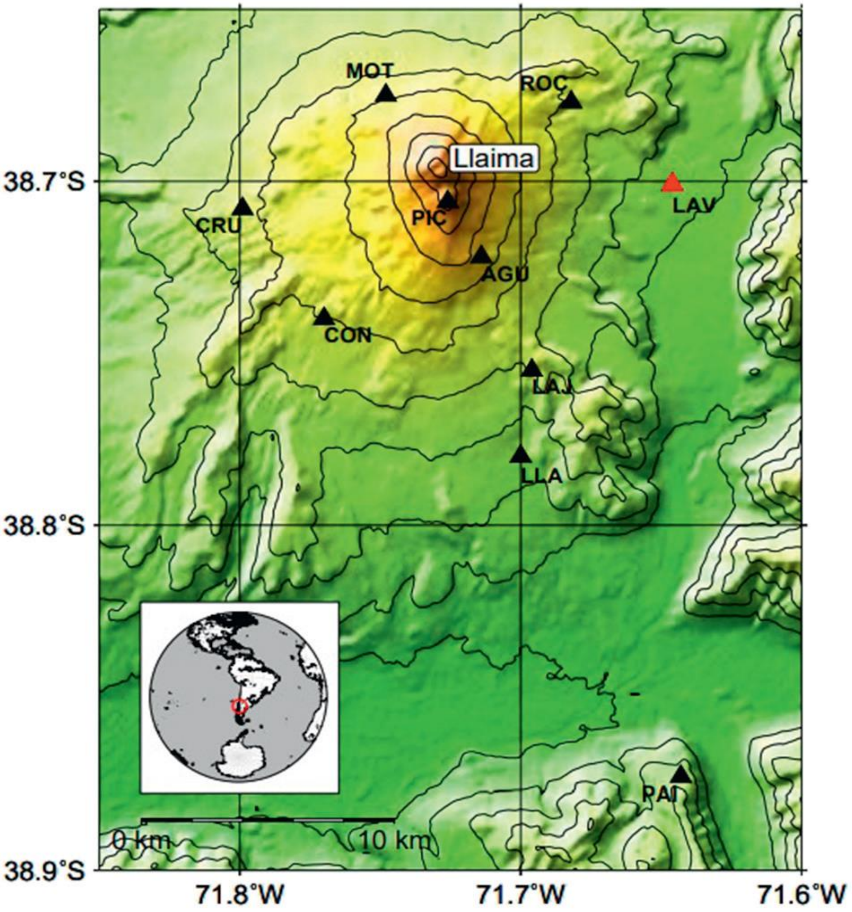
$$MNCC_{UV} = \frac{max\ (|R_{UV}(k)|)}{\sqrt{R_{UU}(0)R_{VV}(0)}}$$

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Data

To evaluate the performance of the proposed approach, we have used a dataset of the Llaima volcano located in Chile. These data are composed of vertical component of the LAV station (marked by a red triangle in the image on the right) of the OVDAS (*Observatorio Vulcanológico de los Andes Sur)* seismic monitoring network.

- Recording period : between 2010 and 2016.
- Sampling frequency : 100 Hz.
- Frequency range : 1 and 10 Hz (Filtered with a numerical *I0*th order Butterworth bandpass).
- Normalization : maximum value.
- Data classes : Long Period (LP), Tremor (TR), Volcano-Tectonic (VT), and Tectonic (TC).



Class	LP	TR	VT	TC
Number of events	1310	490	304	1488

Total of 3592 events in the database

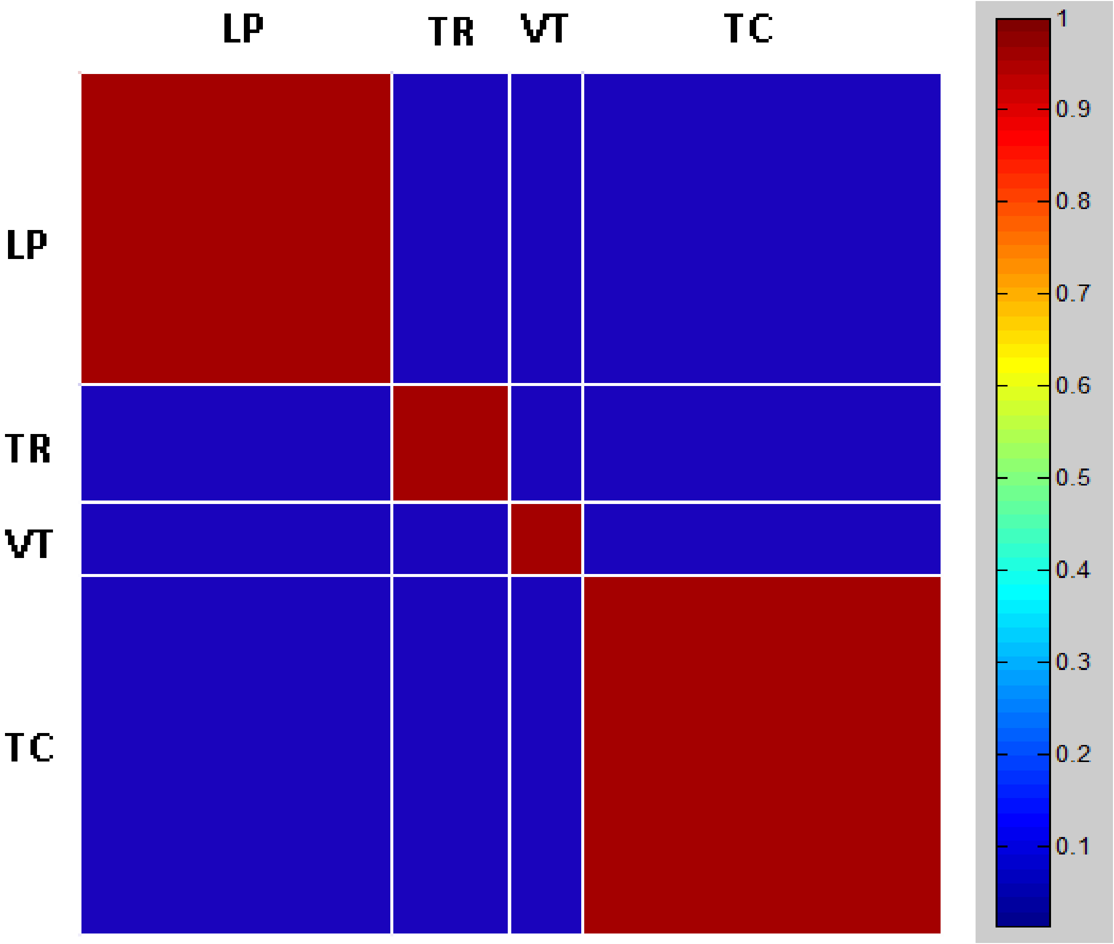
INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Please do not use this space, a QR code will be automatically overlayed

Results

To clearly present the results, the obtained $MNCC_{ij}$ values are presented in a squared and symmetrical matrix (Figure on the right), where the index i and j correspond to the MNCC value between events *i* and *j*.

As shown by the colorbar, events that are perfectly correlated (MNCC = 1) are indicated by brown color. These events should be in the same class. whereas uncorrelated events (MNCC = 0) are presented in a blue color. These events should be in different classes.

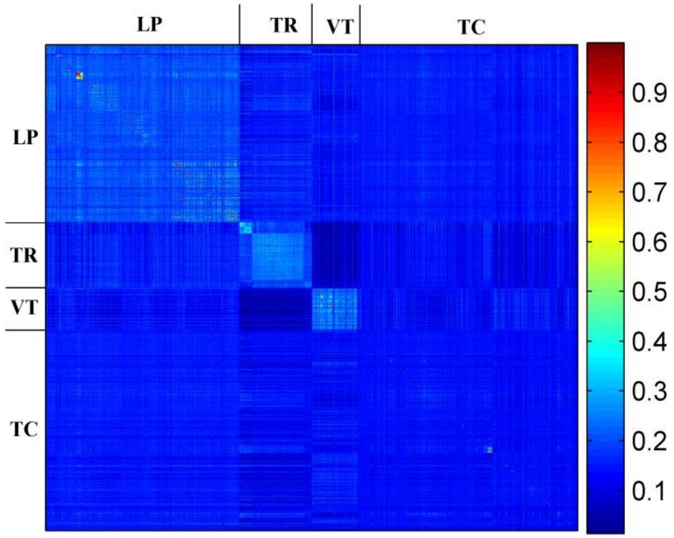


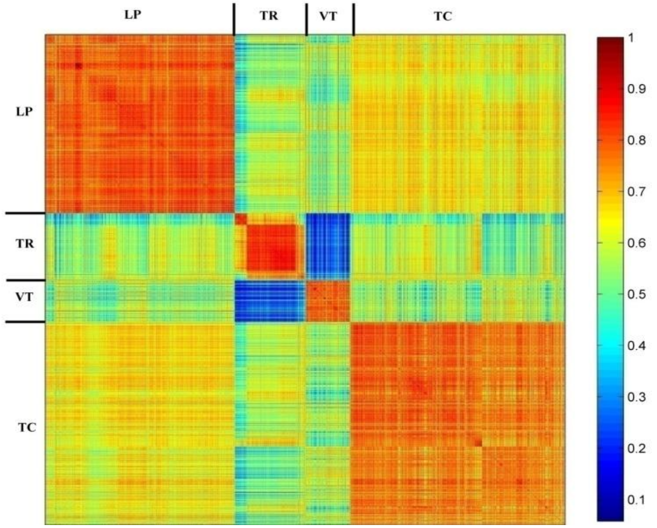
Results

Time Domain

Time-Frequency domain

MNCC
heat map





MNCC mean

MNCC mean

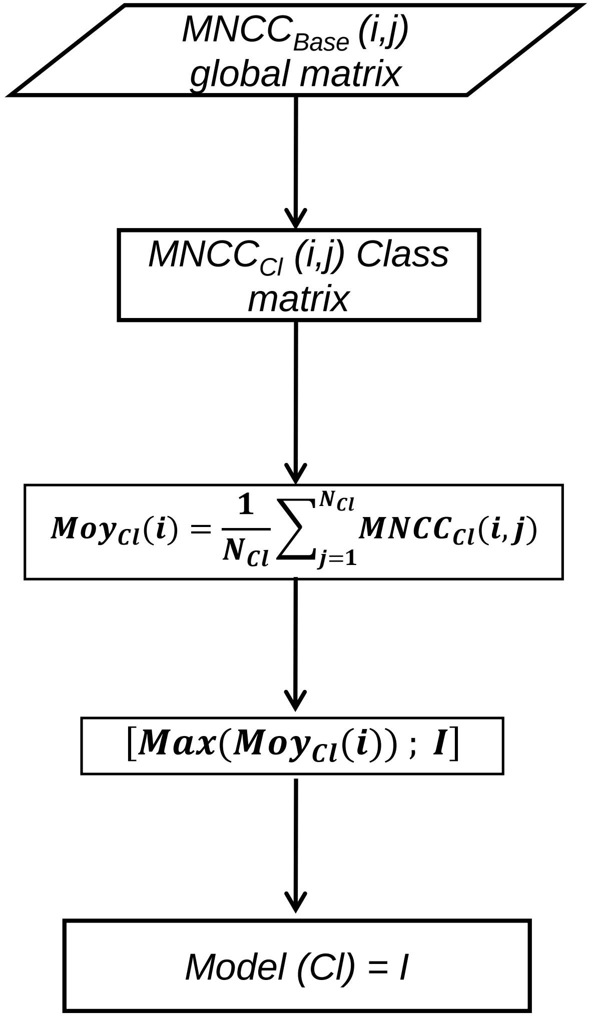
MNCC Global
mean values

	LP	TR	VT	TC
LP	0,21	0,15	0,14	0,15
TR	0,15	0,22	0,07	0,13
VT	0,14	0,07	0,23	0,13
TC	0,15	0,13	0,13	0,13

	LP	TR	VT	TC
LP	0,74	0,51	0,53	0,61
TR	0,51	0,72	0,25	0,51
VT	0,53	0,25	0,74	0,52
TC	0,61	0,51	0,52	0,71

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Results



MNCC$_{Base}$ (i,j) global matrix

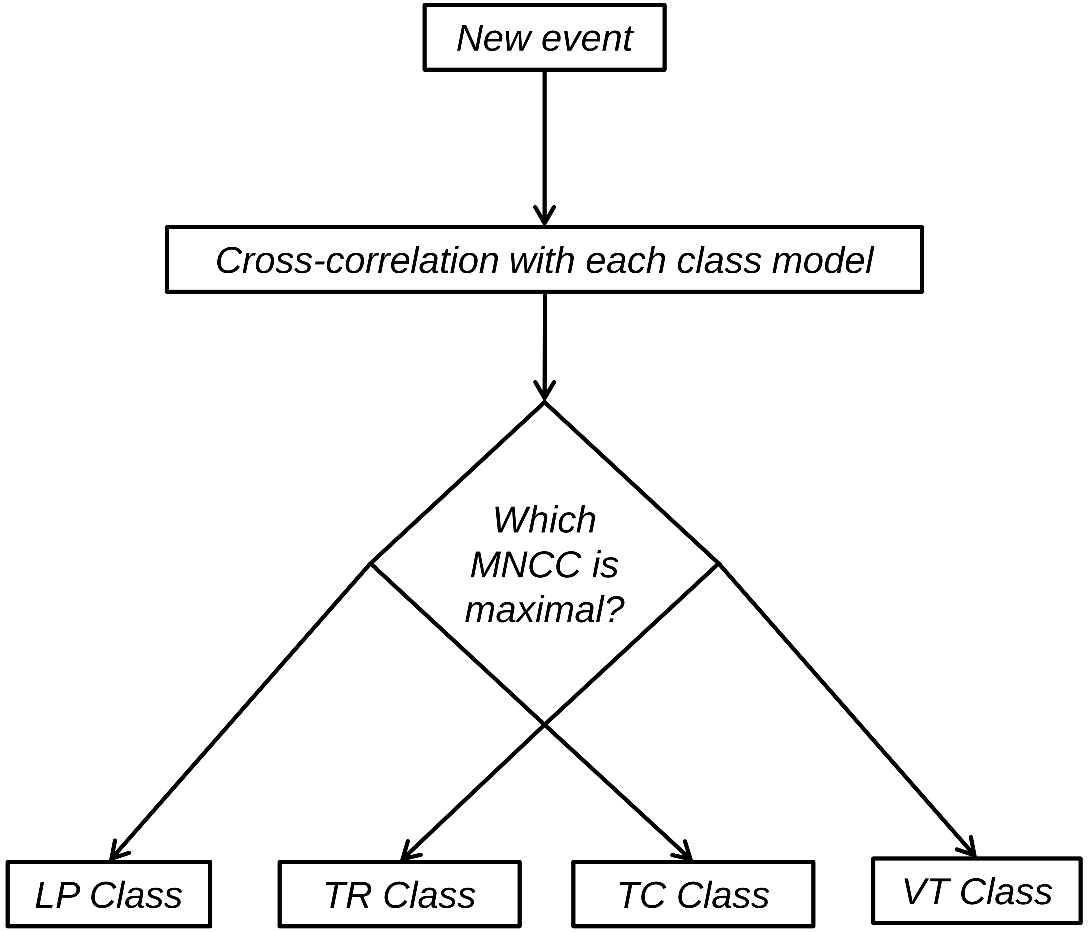
MNCC$_{Cl}$ (i,j) Class matrix

$$Moy_{Cl}(i) = \frac{1}{N_{Cl}} \sum_{j=1}^{N_{Cl}} MNCC_{Cl}(i,j)$$

$$[Max(Moy_{Cl}(i))\ ;\ I]$$

Model (Cl) = I

In each class, we extract the most correlated event with all the others to represent the class



New event

Cross-correlation with each class model

Which MNCC is maximal?

LP Class

TR Class

TC Class

VT Class

The obtained models are then used to classify any new event

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Please do not use this space, a QR code will be automatically overlayed

Results

Time Domain

The confusion matrix

Target classes

Predicted classes	LP	TR	VT	TC
LP	969	46	16	408
TC	153	441	0	329
VT	20	1	203	220
TR	168	2	85	531

Performance evaluation of the classifier

	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	Error (%)
LP	74	79,40	67,34	77,42	22,58
TR	90	84,46	47,78	85,22	14,78
VT	66,8	92,67	45,72	90,48	9,52
TC	35,7	87,88	67,56	66,26	33,74

Global accuracy of 59,7%

Time-Frequency domain

Target classes

Predicted Classes	LP	TR	VT	TC
LP	1264	5	25	20
TR	5	475	0	105
VT	25	0	274	54
TC	16	10	5	1309

	Sensitivity (%)	Specificity (%)	Precision (%)	Eccuracy (%)	Error (%)
LP	96,49	97,81	96,19	97,33	2,67
TR	96,94	96,45	81,20	96,52	3,48
VT	90,13	97,60	77,62	96,97	3,03
TC	87,97	98,53	97,69	94,15	5,85

Global accuracy of 92,48%

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Conclusion

The classification of seismic events is the first processing step before any further analysis. The use of cross correlation function is one of the simplest but also quite efficient approach especially in time-frequency domain. The principal results of this study are summarized below :

- **In time domain :**
  - The tectonic events (TC) are not well identified by this method. A global accuracy of about 60% was reached using the four classes.
  - Excluding the TC events improves the global accuracy to 86,3%.

- **In time-frequency domain :**
  - The classification task is significantly improved compared to time domain. The method allows the classification of all types of events in the seismic database including the TC class.
  - We have obtained a global accuracy reaching 92,48% using all classes of the database, and 96.53% if the TC class is excluded.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

References

1. A. Atmani, E. H. Ait Laasri, D. Agliz, E. Akhouayri, "Llaima's Volcano Seismis Event *Classification Using The Cross Correlation Function*", International Journal of Computer Sciences and Engineering, Vol. **10**, Issue. **3**, pp. **1–7**, **2022**.
2. Theodorus Permana, Takeshi Nishimura, Hisashi Nakahara and Nikolai Shapiro. " *Classification of volcanic tremors and earthquakes based on seismic correlation: application at Sakurajima volcano, Japan*". Geophysical Journal International. 229, 1077–1097, **2022**.
3. E.H. Ait Laasri, A. Atmani, D. Agliz, E. Akhouayri, "Application of Cross-Correlation to Seismic Signal Database of Agadir". International Journal of Computer Sciences and Engineering. Vol. 9, Issue.6. **2021.**
4. 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**.
5. M. Curilem, R.F. de Mello, F. Huenupan, C. San Martin, L. Franco, E. Hernández, R.A. Rios. "*Discriminating seismic events of the Llaima volcano (Chile) based on spectrogram cross-correlations*", Journal of Volcanology and Geothermal Research **367**, pp.**63–78**, **2018**.
6. Soto, R., Huenupan, F., Meza, P., Curilem, M., & Franco, L. Spectro-temporal features applied to the automatic classification of volcanic seismic events. Journal of Volcanology and Geothermal Research, 358, 194–206. doi:10.1016/j.jvolgeores.2018.04.025, **2018**,
7. David P. Schaff, Won-Young Kim, Paul G. Richards, Eunyoung Jo, and Yonggyu Ryoo. Using Waveform Cross Correlation for Detection, Location, and Identification of Aftershocks of the 2017 Nuclear Explosion at the North Korea Test Site. Seismological Research Letters, Volume 89, Number 6 pp. 2113-2119. **2018**.
8. B.R. Julian, "Volcanic tremor: nonlinear excitation by fluid flow". Int. J. Comput. Sci. Secur. 99, 11859–11877. **1994**.
9. B. Chouet, "*A seismic model for the source of long-period events and harmonic tremor*". Volcanic Seismology. Berlin, Heidelberg. pp. 133–156. **1992**.
10. S.R. McNutt, "*Volcanic tremor*". Encyclopedia of Earth System Science. 4. pp. 417–425. **1992**.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION