VEVCC program for concatenation of volcanic events based on cross-correlation analysis

Dairoh Dairoh[2,3], *Anas* Fauzi Masykuri[2,4], *Rahmat* Setyo Yuliatmoko[2,4], *Afif Rakhman*[1], Sudarmaji Saroji[1], *Ahmad* Ashari[5] and *Wiwit* Suryanto[1*]

[1]Geophysical Laboratory, Department of Physics, FMIPA, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[2]Doctoral Program of Physics, Department of Physics, FMIPA, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.
[3]Department of Informatics Engineering, Politeknik Harapan Bersama, Tegal 52147, Indonesia.
[4]Meteorology Climatology and Geophysics Agency (BMKG)
[5]Department of Computer Science and Electronics, FMIPA,Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.

Abstract. Volcanic eruptions pose a significant risk to communities located near active volcanoes. Disaster mitigation and risk reduction efforts rely on detecting and monitoring volcanic activity as early as possible. This article introduces VEVCC, a MATLAB-based application designed to precisely identify and extract volcanic seismic events from continuous data streams. VEVCC's primary objective is to facilitate the creation of an Excel file containing the arrival times of detected events, which can then be used for various purposes, such as early warning disaster mitigation and automated event identification via machine learning techniques. VEVCC utilizes cross-correlation algorithms to identify volcanic seismic events. It separates these events from background noise and other sources of seismicity, allowing for the construction of a clean and informative dataset. The extracted data is a valuable resource for estimating the frequency of volcanic events and evaluating patterns of volcanic activity. VEVCC's time-stamped event data is indispensable for improving early warning systems, real-time surveillance, and automated event identification. We tested the program on the Merapi volcano datasets during a 1998 campaign for a broadband experiment with the capability to extract the events automatically. Further machine-learning models and algorithms enhance the automatic recognition of volcanic events.

1 Introduction

The significance of accurate event detection for volcano-tectonic events cannot be overstated. Volcano-tectonic events are a type of seismic activity that occur in volcanic regions, often associated with the movement and interaction of volcanic material inside and outside the edifice. These events can have major implications for the safety and well-being of nearby populations, as they can trigger volcanic eruptions and other harmful phenomena. To effectively mitigate the risks associated with volcano-tectonic events, it is crucial to develop robust and reliable methods for detecting and predicting these events [1-2]. One widely used method for monitoring and detecting volcano-tectonic events is through the analysis of seismic signals [3]. Seismic signals provide important information about the behavior and changes in activity of a volcano. More specifically, volcano seismic signals can be classified into six classes: long-period events, volcanic tremors, volcano-tectonic events, explosions, hybrid events, and tornillo [3-4]. Seismological methods have proven to be among the most useful techniques for monitoring volcanoes and detecting volcano-tectonic events [5-7]. These seismic signals can occur before and during an eruption, making their analysis and interpretation a crucial task for volcanic eruption forecasting. It is important to identify the event, even the number of events is significantly contribute to volcanic disaster mitigation and timely response. We apply cross-correlation techniques to analyze the time series of volcano seismic signals in order to detect and categorize volcano-tectonic events. Cross-correlation techniques have shown promise in accurately identifying and categorizing volcano-tectonic events. Cross-correlation techniques have shown promise in accurately identifying and categorizing volcano-tectonic events [8]. A common signal-processing method called cross-correlation is crucial to exploration and earthquake geophysics. In travel time tomography, seismic velocity estimation makes use of the cross-correlation between seismic records that were observed and those that were predicted [9-10].

Cross-correlation is a process for measuring the similarity of one time series (seismic trace) to another time series (seismic trace)[11]. In this case, broadband data from seismic recordings at station L57 were used to detect volcanic events from Mount Merapi. Where Lava 57 or L57 is one of the broadband seismometer stations

* Corresponding author: ws@ugm.ac.id

https://doi.org/10.1051/e3sconf/202346801006

installed on Mount Merapi. The location of L57 is in Boyolali Regency, specifically in the Selo District. The location of station L57 is as shown in **Fig. 1.** below.

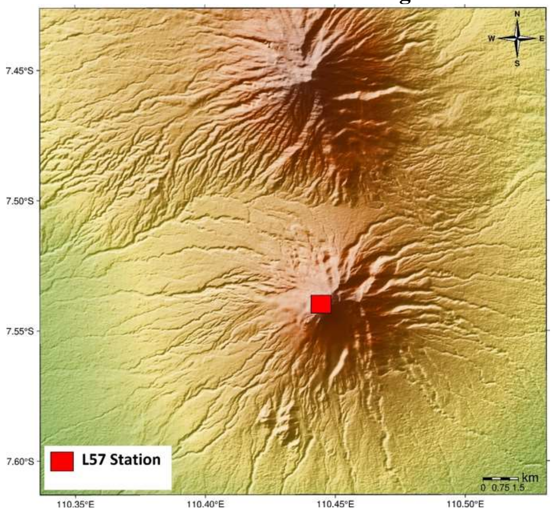


Fig. 1. The location of the broadband seismic station L57.

This paper presents VEVCC, an application built on MATLAB that was developed specifically for the identification and extraction of volcanic seismic events from continuous data streams. The major objective of the VEVCC is to provide a dataset that is spotless and well-organized, and which includes the arrival times of all observed events. These types of data can be of tremendous assistance in the development of early warning systems, real-time surveillance, and the use of machine learning algorithms for the purpose of automating event recognition [5, 12-13].

2 Methods

VEVCC identifies and extracts volcanic seismic events by utilizing cross-correlation techniques, as described in the following procedures:

Initialize indeks to 0
Initialize sta_idx to sta * 100
Initialize lta_idx to lta * 100
Initialize event01 to master_event
Initialize lev to the length of event01

if lev is greater than lta_idx, then
Set akhir_idx to lev
else
Set akhir_idx to lta_idx
end If

initialize indeks to 0

for i in the range from 1 to the length of data minus akhir_idx minus 1:
increment indeks by 1
compute cross-correlation between data and master_event, normalize the result
Set cc(indeks) to the maximum value of p
set the mean absolute value of sta data
set the mean absolute value of lta data

calculate the envelope
find the intercept points cc and treshold. s\tore the intercept points in (X0, Y0).

Set namafile to 'event_terpilih.xlsx'

Write the values in X0 to an Excel file named namafile (e.g., using writematrix with appropriate arguments).

(1) Acquisition of Data: The VEVCC obtains its input, which consists of continuous seismic data, from a real-time monitoring or an archived data system.
(2) Analysis of Cross-Correlation: The program makes use of techniques for analyzing cross-correlation in order to discover characteristic seismic wave patterns that are connected with volcanic occurrences. These patterns cannot be attributed to the background noise [8].
(3) Event Extraction: Seismic events that have been detected are timestamped and recorded in an Excel file, which provides a structured dataset that may be used for further investigation [3, 14].

Overall, the VEVCC processing is depicted as shown in the following flowchart:

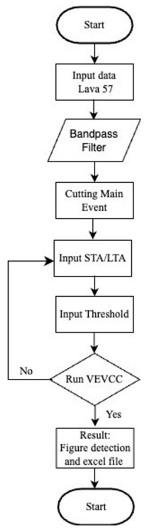


Fig. 2. Flowchart of VEVCC program

During a broadband experiment campaign in 1998, we tested VEVCC using seismic data from the Merapi volcano [5]. The results showed that VEVCC worked as expected. A dataset that is acceptable for study was produced as a result of the program's successful identification and time stamping of volcanic occurrences (**Fig.3**). This dataset has the potential to be extremely helpful in determining the frequency and patterns of volcanic activity. As a result, it can make a contribution to early warning systems and real-time surveillance.

The cross-correlation of the two signals *xp* and *yp* can be represented in the following form. The similarity between the events *xp(t)* and *yp(t)* is quantified using the cross-correlation function *rxy*(τ)[15]:

$$r_{xy}\left(\tau\right) = \int x_p\left(t\right) y_p\left(t+\tau\right) d\tau \qquad (1)$$

where τ is the delay between the two signals. The change in τ changes the relative position of the signal *x* relative to the signal *y*. It should be noted that the correlation function $\boldsymbol{rxy}$ measures only the similarity of the waveform of the signal, and not the amplitude of events. Thus, the amplitudes can vary for events with similar waveforms. This means that the waves moved along an almost identical trajectory, but were not necessarily created by a source with a constant force. The quantitative correlation parameter is the correlation coefficient r. The value of r is in the range of numbers from -1 to 1. The closer r is to 1, the stronger the direct relationship between the variables, the closer r is to -1 – the inverse. When r = 0, there is no significant relationship between the two variables [15-16].

3 Discussion

The VEVCC application is a useful resource for organizations working to lessen the impact of natural disasters [3]. It helps with early warning and monitoring in real time by providing accurate data on volcanic occurrences, which can be monitored in real time [17]. In addition, this dataset is extremely helpful in the process of training machine learning models for automated event recognition, which might potentially improve the response time as well as the accuracy during times of volcanic emergency [9]. In addition to its application in emergency management, the VEVCC can also be utilized in the field of scientific investigation to better understand the behavior of volcanoes [5] .

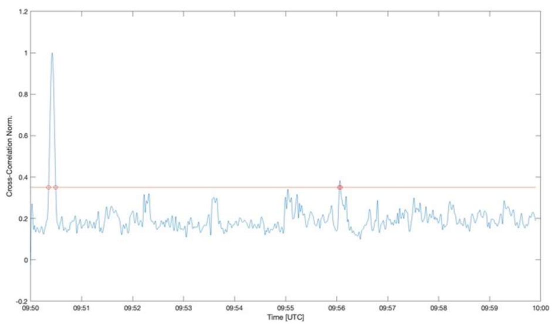
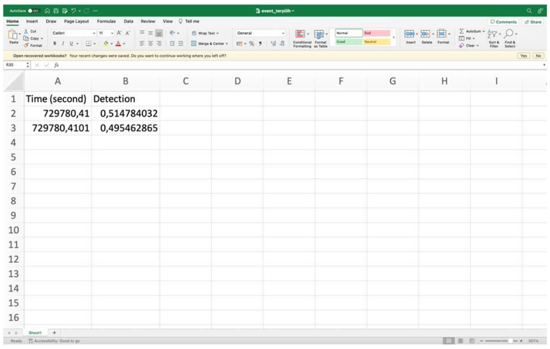


Fig. 3. An example of the cross-correlation results with the threshold line that crossing at the specific time window for the events.

And the results of the cross-correlation master event (VTA) value are used to detect signals in a 10-minute event length for event determination with a specified threshold of 3.5 (**Fig.3.)** and then his detection process was obtained 2 events with cross-correlation values presented in Excel **(Fig.4.)**



	A	B
1	Time (second)	Detection
2	729780,41	0,514784032
3	729780,4101	0,495462865

Fig. 4. Cross-correlation detection value in a 10-minute event length

The cross-correlation results obtained using VEVCC yielded 2 detections for the master event (VTA) with detection values of 0.514784032 and 0.495462865.

The scientific community has recently placed a greater emphasis (in recent years) on the significance of early warning systems and real-time surveillance in the context of the management of volcanic risk [10, 18, 19]. The accurate identification, as well as continuous monitoring of volcanic activity, are both essential components of these systems. VEVCC makes a contribution to this field by offering a tool for reliably identifying and extracting seismic events linked with volcanic activity. This makes it possible to take safety precautions and evacuation protocols in a timely manner, which is an important aspect of the subject. A common signal-processing method called cross-correlation is crucial to exploration and earthquake geophysics. In travel time tomography, seismic velocity estimation makes use of the cross correlation between seismic records that were observed and those that were predicted [9, 20].

In addition to this, the usefulness of VEVCC extends into the field of applications that use machine learning [11, 21]. Researchers are able to create and train machine learning models to automatically recognize volcanic eruptions if they have access to the dataset that was generated by the VEVCC. Because machine learning models can handle massive datasets much more quickly than people can [12] this has the potential to drastically shorten the amount of time it takes for authorities to respond to volcanic emergencies [22]. The effective recognition of volcanic eruptions with the use of machine learning can contribute to improved hazard preparedness, response, and recovery efforts [13, 23].

In addition, the dataset that was produced by VEVCC has the potential to be an extremely helpful resource for scientific investigation [14]. Researchers are able to get insights into the frequency and patterns of volcanic activity by studying the data that was collected from the volcano [24]. This information has the potential to contribute to a better understanding of volcanic processes, which might then lead to improved models for predicting volcanic eruptions [25].

The Volcanic Earthquake and Volcanic Component Catalog, or VEVCC, is a robust program that can recognize and extract volcanic seismic events from continuous data streams [2]. The data that is extracted is guaranteed to be accurate and reliable if the appropriate cross-correlation methods are used [11, 20]. This dataset is essential for the development of prevention strategies, early warning systems, and applications of machine learning [3, 12]. The efficiency of the algorithm is demonstrated by the fact that the test on the dataset from the Merapi volcano was successful [5]. Its skills in automated event recognition will be enhanced by further research and integration of machine learning models, adding to improvements in volcanic risk management [11, 26].

4 Conclusion

There is a lot of potential in the application of VEVCC to the subject of volcanic risk management. It is able to recognize and extract volcanic events from continuous data streams, making it a vital tool in efforts to lessen the impact of volcanic eruptions on human communities. Additionally, it is useful in applications using machine learning and contributes to the advancement of scientific knowledge, making it a valuable tool overall. These tools are used to construct robust and responsive systems for disaster preparedness and response. VEVCC utilizes cross-correlation algorithms to identify volcanic seismic events. The time-stamped event data from VEVCC is crucial for enhancing early warning systems, real-time monitoring, and automated event identification, thus enabling further machine learning models and algorithms to improve the automatic recognition of volcanic events.

The Author thanks to Geophysics Laboratory at Universitas Gadjah Mada for the computing facilities. The paper is part of the Research Methodology Class at Doctoral Graduate of Physics at UGM.

References

1. A. Rakhman, Wahyudi, A. B Santoso, H. Humaida, and W. Suryanto, “Ambient Seismic Noise Analysis Associated with the 2010 Eruption of Merapi Volcano Using Temporal Variations of Randomness and Background Noise,” *Int. J. Geophys,* **2020**,1–10, (2020), doi: 10.1155/2020/8853376.
2. S. Nowé, T. Lecocq, C. Caudron, K.Jónsdóttir, and F. Pattyn, “Permanent, seasonal and episodic seismic sources around Vatnajökull, Iceland, from the analysis of correlograms Sylvain,” *Volcanica*, **4**, 2,135–147, (2021), doi: 10.30909/vol.04.02.135147.
3. S. Yu and J. Ma, “Deep Learning for Geophysics: Current and Future Trends,” *Rev. Geophys.*, **59**, 3,1–36, (2021), doi: 10.1029/2021RG000742.
4. J. Caldera and S. Wirasinghe, “Analysis and Classification of Volcanic Eruptions,” in *International Conference of the International Institute for Infrastructure Resilience and Reconstruction (I3R2)*, **May**, 128–133, (2014) doi: 10.5703/1288284315372.
5. M. Singh *et al.*, *Machine learning for Earth System Science (ESS): A survey, status and future directions for South Asia*. **2021**, 1–17.
6. J. Li, M. He, G. Cui, X. Wang, W. Wang, and J. Wang, “A Novel Method of Seismic Signal Detection Using Waveform,” *Appl. Sci.*,**10,** 2919, 1–17, (2020), doi: 10.3390/app10082919.
7. H. Darmawan, P. Yuliantoro, A. Budi, H. Humaida, and W. Suryanto, “Dynamic velocity and seismic characteristics of gravitational rockfalls at the Merapi lava dome,” *J. Volcanol. Geotherm. Res.*, **404**, 107010, 1–10, (2020), doi: 10.1016/j.jvolgeores.2020.107010.
8. J. Barrière *et al.*, “Single-Station Seismo-Acoustic Monitoring of Nyiragongo’s Lava Lake Activity (D.R. Congo),” *Front. Earth Sci.*, **6,** June, 1–17, (2018), doi: 10.3389/feart.2018.00082.
9. Y. Zhao, F. Niu, Z. Zhang, X. Li, J. Chen, and J. Yang, “Signal Detection and Enhancement for Seismic Crosscorrelation Using the Wavelet-Domain Kalman Filter,” *Surv. Geophys.*, **42**. 1, 43–67, (2021), doi: 10.1007/s10712-020-09620-6.
10. J. Han *et al.*, “Seismic Event and Phase Detection Using Deep Learning for the 2016 Gyeongju Earthquake Sequence,” *Geosci. J.*, **27**, 3, 285–295, (2023), doi: 10.1007/s12303-023-0004-y.
11. D. Bobrov, I. Kitov, and L. Zerbo, *Perspective of Cross Correlation in Seismic Monitoring at the International Data Center*, **171**, 3–5. 439–468,(2014), doi: 10.1007/s00024-012-0626-x.
12. T. Glade and F. Nadim, “Early warning systems for natural hazards and risks Thomas Glade,” *Nat. Hazards*, **70,** 3, 1669–1671, (2014), doi: 10.1007/s11069-013-1000-8.
13. S. M. Mousavi and G. C. Beroza, “Machine Learning in Earthquake Seismology,” *Annu. Rev. Earth Planet. Sci.*, **51**, November, 105–129, (2023), doi: 10.1146/annurev-earth-071822-100323.
14. B. Chouet, “Volcano Seismology,” *Pure Appl. Geophys.*, **160**, 3–4, 739–788, (2003), doi: 10.1007/PL00012556.
15. E. Onajite, *Un derstanding Seismic Wave Propagation*. 2014.
16. I. A. Zavedevkin, A. A. Shakirova, and P. P. Firstov, “DrumCorr Program for Selecting Volcanic Earthquake Multiplets Based on Cross-Correlation Analysis,” in *Journal of Physics: Conference Series*, **2094,** 3, 1–8, (2021), doi: 10.1088/1742-6596/2094/3/032048.
17. M. Malfante, M. Dalla Mura, J. P. Metaxian, J. I. Mars, O. Macedo, and A. Inza, “Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives,” *IEEE Signal Process. Mag.*, **35**, 2, 20–30, (2018), doi: 10.1109/MSP.2017.2779166.
18. S. G. Mosher and P. Audet, “Automatic Detection and Location ofSeismic Events From Time-Delay Projection Mapping and Neural Network Classification,” *J. Geophys. Res. Solid Earth*, **125,** 10, 1–18, (2020), doi: 10.1029/2020JB019426.

19. A. Nieto-Torres, L. F. Guimarães, C. Bonadonna, and C. Frischknecht, "A New Inclusive Volcanic Risk Ranking, Part 1: Methodology," *Front. Earth Sci.*, **9**, August, 1–22, (2021), doi: 10.3389/feart.2021.697451.
20. L. Z. Dmitry Bobrov, Ivan Kitov, "Perspectives of Cross Correlation in Seismic Monitoring at the International Data Centre." 1–48, (2014), doi :10.1007/s00024-012-0626-x
21. J. Hu, Y. Ding, S. Lin, H. Zhang, and C. Jin, "A Machine-Learning-Based Software for the Simulation of Regional Characteristic Ground Motion," *Appl. Sci.*, **13**, 14, 1–18, (2023), doi: 10.3390/app13148232.
22. M. Ripepe *et al.*, "Infrasonic Early Warning System for Explosive Eruptions M.," *J. Geophys. Res. Solid Earth*, **123**, 11, 9570–9585, (2018), doi: 10.1029/2018JB015561.
23. R. Scarpa and R. I. Tilling, "Sesimic Monitoring and Eruption Forecasting of Volcanoes: A Review of the State-Of-the-Art and Case Histories," *Monit. Mitig. Volcano Hazards*, **15,** January, 99–146, (1996), doi: 10.1007/978-3-642-80087-0.
24. P. Mathieu and C. Aubrecht, *Earth Observation Open Science and Innovation*. (2018), Springer Nature, Washington, USA.
25. R. S. Jakka and S. Garg, "Understanding and Parameter Setting of STA/LTA Trigger Algorithm," *Earthq. Eng. Eng. Vib.*, **14**, 1, 27–35, (2015), doi: 10.1007/s11803-015-0004-7.
26. M. Malfante, M. Dalla M, J. I. Mars, J. P. Métaxian, O. Macedo, and A. Inza, "Automatic Classification of Volcano Seismic Signatures Marielle," *J. Geophys. Res. Solid Earth*, **123**, 12, 1–14, (2018), doi: 10.1029/2018JB015470.