https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Discussions

How can seismo-volcanic catalogues be improved or created using robust neural networks through weakly supervised approaches?

Manuel Titos[1,2], Carmen Benítez[1,2], Milad Kowsari[3,] Jesús M. Ibáñez[4,5]

[1]Department of Signal processing, Telematics and Communications, University of Granada, Granada, 18014, Spain

[2] Research Center on Information and Communication Technologies of the University of Granada (CITIC-UGR)

[3] University of Iceland, Faculty of Civil and Environmental Engineering, Reykjavík, 102, Iceland

[4] Instituto Andaluz de Geofísica, University of Granada, Granada, 18071, Spain

[5] Department of Theoretical Physics and the Cosmos, University of Granada, Granada, 18071, Spain

Correspondence to: Manuel Titos (mmtitos@ugr.es)

Abstract. Real-time monitoring of volcano-seismic signals is complex. Typically, automatic systems are built by learning from large seismic catalogues, where each instance has a label indicating its source mechanism. However, building complete catalogues is difficult owing to the high cost of data-labelling. Current machine learning techniques have achieved great success in constructing predictive monitoring tools; however, catalogue-based learning can introduce bias into the system. Here, we show that while monitoring systems recognize almost 90% of events annotated in seismic catalogues, other information describing volcanic behavior is not considered. We found that weakly supervised learning approaches have the remarkable capability of simultaneously identifying unannotated seismic traces in the catalogue and correcting mis-annotated seismic traces. When a system trained with a master dataset and catalogue is used as a pseudo-labeller within the framework of weakly supervised learning, information related to volcanic dynamics can be revealed and updated. Our results offer the potential for developing more sophisticated semi-supervised models to increase the reliability of monitoring tools. For example, the use of more sophisticated pseudo-labelling techniques involving data from several catalogues could be tested. Ultimately, there is potential to develop universal monitoring tools able to consider unforeseen temporal changes in monitored signals at any volcano.

1 Introduction

Understanding the dynamics of active volcanoes, and even more so, carrying out Early Warning protocols for volcanic eruptions require multiparametric observations focused on accomplish accurate and effective monitoring (Sparks, 2003). The objective of identifying precursors that warn of a possible volcanic eruption involves the analysis of long temporal series of data, characterizing and relating them with source models associated with the internal dynamics of the volcano (Witze, 2019; Palmer, 2020). Currently, the availability of multiparametric long-time data series, such as seismology, deformation, measurements of volcanic gases and fluids, space imaging, and other processes, is limited to a few volcanoes around the world. For this reason, volcanic seismology continues being the backbone of the analysis, both in real time and using data

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

from previous eruptive episodes (Chouet, 2003; McNutt, 2015). This is because the installation and acquisition of seismic data continues to be the most efficient procedure of volcanic monitoring, and because the existence of numerous open access repositories allows the scientific community reviewing consolidated databases to understand what occurred in the past for modelling future eruptions.

In volcanic seismology, the presence of various seismic signals—such as volcano-tectonic earthquakes (VT), long-period events (LP), ultra-long-period (VLP) events, hybrid (HY) events, explosions (EXP), and volcanic tremors (TR)—indicates the existence of multiple seismic sources, which can sometimes operate simultaneously and must be considered. Thus, models of brittle rock fracturing, conduit resonance, pressure transients in fluids, bubbles, cracking in viscoelastic mediums, elastic energy transfer by fluid flow, debris flows, and many others are used (Ibáñez et al., 2000; McNutt and Roman, 2015; Minakami, 1974)). Hence, this complex system of seismic sources results in varying interpretations of volcano dynamics based on the predominant signal type and its spatio-temporal evolution. Recently, Rey-Devesa et al. (2023, 2023) developed a series of algorithms based on the processing of seismic signals that allow efficient early warnings of volcanic eruptions. However, comprehending the underlying physics behind the eruptions, and thus understanding why eruptions occur, cannot be solely explained through such signal processing. It requires knowledge of the frequency and types of seismic events that occurred.

The introduction of automatic recognition procedures for earthquake-volcanic signals almost two decades ago (Ohrnberger 2001, Scarpetta et al., 2005; Alasonati et al., 2006; Benítez et al., 2006; Ibáñez et al., 2009, Curilem et al., 2009) has made the process of identifying and characterizing signals more efficient, faster, and comprehensive. The use of procedures based on the application of classic Machine Learning (ML) models as Support Vector Machine (SVM) or, Hidden Markov Models (HMM), and different types of Artificial Neural Networks as Baeysian Neural Networks (BNN), Deep Neural Networks (DNN) and Recurrent Neural Networks (RNN) among others- (e.g. Bhatti et al. 2016; Canario et al., 2020 ; Cortés et al., 2021; Bueno et al. (2021, 2022); Martínez et al. 2021; Titos et al. (2017, 2018, 2019), Bicego et al., 2022) allowed progress in a higher quality of reconnaissance procedures (unsupervised, semi-supervised and supervised) and their implementation in real time for the monitoring of active volcanoes. Although there are currently numerous and highly advanced works showcasing successes in this field, the results obtained have begun to reveal potential new problems. These outcomes raise open questions that should be efficiently addressed to adequately comprehend the derived source models. Some of these questions include: (a) Is the database used as a reference properly labelled? That is, are the signal names or labels adequately identified? (b) How do seismic attenuation processes or source radiation patterns influence changes in the appearance of a signal, thus confounding the associated source models? (c) How could background seismic noise affect the identification of seismic events?

For the second open question, there are potential solutions. In Titos et al. (2018), it was demonstrated that a signal interpreted as a VT at short distances appeared as an LP at another seismic station only a few kilometres away from the reference station due to the high seismic attenuation in volcanic environments. As potential solution to these ambiguities, Titos et al. (2023) propose using a system of multiple seismic stations for signal recognition and "establishing" rules or conditions for identifying signals simultaneously. Specifically, perform a second recognition procedure and assign the 'type of signal' based on the station with the highest high-frequency component.

The first and third open questions may potentially be more difficult to resolve. An incomplete or mislabelled seismic catalogue can result from various factors, such as noisy signals, human error, intense and overlapping seismic activity, among others. Most available seismic catalogues were manually developed based on time-frequency characteristics as well as wave-field properties obtained from the seismic signals. Human factors, including bias in manual labelling or influence from the prevailing scientific understanding at the time of labelling, may also contribute to incomplete or inaccurate catalogues. The occurrence of intense seismic activity can significantly impact the completeness of catalogues, with only the most energetic events being highlighted, or even combining several types of signals in the same label when the energies are not high enough and time pressures require quick responses. (e.g., labelling a superposition of signals as volcanic tremor). In some eruptive processes, such as the eruption of the island of El Hierro in 2011, the continuous occurrence of VT events led to a signal resembling volcanic tremor, but at a high frequency, due to the overlapping of hundreds of VTs per hour (Ibáñez et al., 2012; Díaz-Moreno et al., 2015). The use of denoising techniques can help reduce seismic noise levels in certain frequency bands, but in general, they will also affect the frequencies of the signals. Therefore, in many cases, they may not be efficient for improving labelling task. Another potential solution is to manually relabel the databases using new expertise and state-of-the-art techniques, but this is a highly time-consuming process in practice. Nevertheless, this issue is prevalent in many seismic databases worldwide. Furthermore, having a database with potential biases will interfere with the implementation of automatic recognition methodologies based on ML. Biased labels will propagate in the generation of models and feature vectors used for comparison and the implementation of these technologies.

In this work, we propose an alternative solution that avoids the need for manual relabelling and the use of denoising-based filtering techniques, while also addressing the correction of potential biases in databases and completeness in seismic catalogues. Based on our experience with transfer learning approaches (e.g., Bueno et al., (2021, 2021); López-Pérez et al., 2020; Titos et al., (2018, 2019, 2023)), we suggest the utilization of a high-quality, meticulously labelled volcano database (referred to as the 'Master database') as a foundational seed for retraining another seismic catalogue. We suggest employing ML approaches to assess the completeness and accuracy of the tested database.

In the literature, as mentioned earlier, there are several successful methodologies that could satisfactorily address this transfer learning approach. Building on the successful applications demonstrated using new-generation Artificial Neural Networks (ANN) in sequence modelling tasks, this work will compare the performance of three architectures in achieving this goal. We will utilize a Recurrent Neural Networks with Long Short-Term Memory cells (RNN-LSTM) and a Temporal Convolutional Networks (TCN) previously developed by Titos et al. (2018, 2022), as well as an improved version of the RNN-LSTM based on multi-resolution dilated recurrent skip connections known as Dilated Recurrent Neural Networks (DRNN) (Chang et al. 2017; Titos et al. 2024). We aim to demonstrate that when properly fine-tuned, any of these neural networks will yield successful recognition results. However, the critical point lies in selecting the Master database for performing this tuning**.**

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Our Master catalogue was created using data from Deception Island volcano (Antarctica) and has already been successfully applied in different Deep Learning architectures such as BNNs, RNNs, DNNs, and TCNs, among others. The tested database belongs to Popocatépetl volcano (Mexico) and was created during a seismic experiment conducted in 2002. This database was selected not because there were doubts about its quality; quite the contrary, it is considered a carefully prepared database due to the quality of the human team that created it and the importance of the volcano. Therefore, controlling the possibility of improving it represents a challenge to our suggestion of how to further enhance, if possible, existing databases. Furthermore, the eruptive dynamics of both volcanoes are entirely different, so we also seek evidence of transferability capability – that is, if a high-quality master database is available, it is possible to export its knowledge to other volcanic systems. In this way, this work will not only be useful for potentially improving existing databases, but also for creating seismo-volcanic catalogues using ML in volcanoes where no catalogue has ever been previously made.

Therefore, ensuring appropriate seismic catalogues and support for developing monitoring tools should be a priority, as significant as applying new and more effective artificial intelligence (AI) techniques. It is important to note that the tested catalogue cannot be considered the official database used by the surveillance institution working on Popocatepetl volcano; therefore, we are not correcting any past or future work performed in this area. Our aim is to demonstrate how any seismic catalogue can be improved or even created with success and low time-consuming effort.

2 Seismic data and catalogues

In this work, we will use a labelled database obtained from Deception Island volcano in Antarctica, which we consider to be reliable and homogeneous. Deception Island (62º59'S, 60º41'W) is a horseshoe-shaped volcanic island that emerged during Quaternary period and placed in a marginal basin-spreading centre of the Bransfield Strait, where the South Shetland Islands and Antarctic Peninsula are being separated (Smellie, 1988; Martí et al., 2011; Carmona et al., 2012). Deception Island volcano has been extensively studied by our research group from various perspectives, providing a comprehensive understanding of its structure and dynamics (Ibáñez et al., 2000 ; Martínez-Arévalo et al., 2003 ; Zandomeneghi et al., 2009; Carmona et al., 2012; Ibáñez et al., 2017). Similarly, seismic data collected through numerous campaigns conducted on this volcanic island since 1994 have allowed us to create a highly reliable database. This database serves as the cornerstone for all our advancements in the field of recognizing and segmenting seismo-volcanic signals. It has served as the foundation for studies on the application of HMM models, various neural networks, parameter reduction algorithms, and many more.

By combining our understanding of the structural, source, and dynamic models of Deception Island volcano with our advancements in signal processing and ML, we can confidently assert that this database is highly reliable and ideal for our intended purpose of serving as a reference seed for building any other seismic catalogues. Therefore, we designate this database, named MASTER-DEC, as the 'Master database' to be utilized as a control and testing process for the Popocatepetl volcano database. While it is true that not all types of signals are present in this 'Master database'—especially those associated with ongoing eruptive processes—its primary objective aligns with our ML application, which aims to understand pre-eruptive processes. Hence, this database is ideal for preparing new ones or improving existing ones focused on serving as Early Warning tools for volcanic eruptions. For the current study, we extracted a subset of data considered the most reliable,

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences Discussions

consisting of 2,193 seismic events. The data is categorized into 5 classes (aligning with the volcano-seismic scientific labels and accompanying source models put forward by Ibañez et al. 2000, see Table 1), with the duration of the events varying: Background seismic noise (BGN)-1,222 events, Volcanic Tremor (TRE)-77 events, Long Period events (LPE)-765 elements, Volcano-Tectonic earthquakes (VTE)-75 elements, and Hybrid events (HYB)-54 events. It is important to note that although the number of selected events for some classes may seem limited for training ML procedures requiring a considerable number of elements, these events were chosen to be the most representative and of the highest quality from our Deception Island database.

The selected database for testing proposes belongs to Popocatepetl volcano, Mexico (19°1‘N, 98°37‘W). This volcano is placed within a different geodynamic framework and exhibits a different eruptive style compared to Deception Island; a subduction region in confront to a rift area. Popocatepetl is a large dacitic–andesitic stratovolcano covering > 500 km2 of the eastern Trans-Mexican volcanic belt (Alaniz-Álvarez et al., 2007; Siebe et al., 2017). It is surrounded by a densely populated area with a population of around 25 million people (Arango-Galván et al., 2020). The volcano is highly active, with the current active period beginning in December 1994 (Arango-Galván et al., 2020). The dataset used in this study was collected during a seismic experiment conducted between November and December 2002, using short-period seismic stations. Data labelling was manually performed by a group of geophysicists with extensive knowledge and experience of the volcano's dynamics. The Popocatepetl 2002 catalogue (named Popo2002) consists of 4,883 events, divided into similar classes as the MASTER-DEC catalogue (again aligning with the volcano-seismic scientific labels and accompanying source models proposed by Ibañez et al. 2000), but further subdivided based on spectral content. This includes three classes for TRE (T1, T2, T3)-273 events, six classes for LPE (LPA1, LPA2, LPB1, LPB2, LPC1, LPC2)-1155 events, and VTE distinguished between local and regional earthquakes (VTE, REG)-371 events. Additionally, the catalogue includes noisy events (labelled as GAR)-2739 events, and due to Popocatepetl's activity, there is a class for explosions (EXP)-4 events. Along with the event catalogue, we have continuous seismograms from this period that will be used for segmentation and identification processes.

Ibáñez, J.M et al. (2000)	McNutt, S. and Roman, D. (2015)	Minakami, T. (1974)	Frequency [Hz]	Example source models
Volcano Tectonic Earthquakes Tectonic Short Period Earthq.	High Frequency (HF)	A-Type	>5	Shear failure or slip along faults, usually as swarms within the volcanic edifice
Long Period Event Volcanic Long Coda Event Tornillo	Low Frequency (LF)	B-Type	1-5	Fluid driven cracks, pressurization processes (bubbles), and attenuated waves

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

Hybrid Event Medium Frequency	Mixed Frequency (MX)	-	1-12	Mixture of processes (e.g., cracks and fluids, frictional melting)
Explosion Volcanic Explosion	Explosion Quake (EXP)	Explosion Quake	>10	Accelerated emissions of gas and debris to the atmosphere
Volcanic Tremor Harmonic Tremor	Volcanic Tremor (TRE)	Volcanic Tremor	1-12	Pressure disturbance, gas emissions, debris processes, and pyroclastic flows

Table 1. Representative volcano-seismic scientific labels and associated source models proposed by Ibáñez, J.M. et al. (2000). Other labels and associated source models proposed by different authors have been included for comparison.

3 Methodology

The well-established procedure for creating a volcano-seismic database from scratch using supervised ML techniques involves selecting and segmenting a large, reliable set of well-labelled seismic events that cover the maximum possible range of events occurring in the studied volcanic area. As is widely known, this initial labelling process relies on expert/human decisions, which can be extremely time-consuming. These events serve as the initial seed for the training procedure. Once our initial selection of high-quality, differentiated types of event classes is completed, encompassing as many events as possible, various techniques can be employed for the training process. Several data-driven approaches have been utilized to uncover descriptive patterns that characterize the diverse types of events appearing in seismic signals. Ultimately, these advancements have resulted in the development of automatic early warning systems (Malfante et al., 2018; Lara et al., 2021). In the literature, seismic event recognition (detection and classification) has been classically explored from two distinct perspectives: (i) the classification of isolated events, where the classifier receives individual events (previously isolated by expert geophysicists) and assigns them to one of the available seismic categories (Hibert et al., 2017; Titos et al., 2018; Bueno et al., 2021; Titos et al., 2019; Canario et al., 2020; Bicego et al., 2022), and (ii) sequence modelling, where the model processes continuous sequences of seismic signals to detect, and classify the events present (Alasonati et al., 2006; Benítez et al., 2006; Köhler et al., 2010; Bhatti et al., 2016; Titos et al., 2018; Bueno et al., 2021; Titos et al., 2022). Considering real-time monitoring of seismo-volcanic signals as a sequence modelling problem, ANNs, including both recurrent and temporal convolutional approaches, have played an important role in recent years due to their temporal modelling capabilities. These approaches have yielded successful results in improving seismic catalogues and advancing in both the procedure of real time seismic monitoring and early warning protocols. However, as mentioned above, several databases can be biased by factors such as seismic attenuation, site effects, radiation patterns, and human decisions and experience. Furthermore, starting this procedure from scratch, regardless of the time it consumes, implies that creating a database in new scenarios could be discouraging. Moreover, in the case of new volcanic processes, old databases may not be suitable or adequate for seismic monitoring.

We will assess the efficacy of three ANNs to demonstrate the transferability of knowledge from one volcanic scenario to another, aiming to: a) verify the reliability of the existing database and; b) create a new database without the initial time-consuming human supervision. Specifically, we will employ a RNN-LSTM, a Dilated-RNN, and a TCN.

RNN–LSTM is a RNN with Long Short-Term Memory (LSTM) cells (Hochreiter, S., Schmidhuber, J., 1997). LSTM cells are a special kind of RNN units that are capable of learning long-term dependencies and mitigating the vanishing/exploding gradient problems, which are common in traditional RNNs (Schmidhuber, J., 2015). LSTM cells improve their performance through a complex architecture that includes a gates mechanism (input, forget, and output gates) which regulate the flow of information, allowing the network to maintain and update a cell state over extended periods of time (Schmidhuber, J., 2015). This makes LSTMs particularly effective for tasks that involve long-term sequential data, such as speech recognition, language translation, and time series prediction within emerging research geosciences fields such as climate change or remote sensing (Racic et al., 2020; Yan et al., 2020). This framework is especially useful for monitoring temporal evolution of volcanic systems in quasi real time.

Dilated Recurrent Neural Network (Dilated-RNN) is a type of RNN designed to handle long-term dependencies more efficiently by using "dilations" in the sequence processing steps (Chang et al. 2017). Analogous to the dilated CNN proposed in (Aäron van den Oord et al. 2016, Fisher et al., 2017), this architecture modifies the standard RNN structure by skipping certain time steps allowing the network to introduce multi-resolution dilated recurrent skip connections between layers, therefore enhancing efficiency and reduced training parameters (Chang et al. 2017). The dilation technique allows for the selection and analysis of non-correlative temporal segments in each network layer during the learning process. This enables an advanced trainable selection of individual features based on their relevance to the recognition task. The flexibility to select non-correlative temporal segments offers the advantage of focusing on the search and detection of seismic events with a broader perspective of the previous temporal steps. Additionally, it demonstrates comparable state-of-the-art performance in sequence modelling tasks involving very long-term dependencies, while helping to mitigate issues of vanishing and exploding gradients.

Regarding TCNs (Lea et al., 2017), unlike traditional RNNs which process sequences sequentially, they leverage dilated 1D convolutional layers and residual connections to process entire temporal sequences simultaneously. The dilated convolutions (similar to the multi-resolution dilated, recurrent skip connections between layers in the Dilated-RNN) introduce gaps between the input elements of the convolutional layers, enabling the network to cover a larger temporal context without a significant increase in computational complexity. Residual connections (Fisher et al., 2027) help mitigate the vanishing gradient problem and facilitate the training of deeper networks by allowing gradients to flow smoothly. This approach efficiently captures long-range dependencies and temporal patterns, offering several advantages over traditional RNNs, such as better handling of the exploding/vanishing gradient problem, improved memory retention, and enhanced parallelization during training (Lea et al., 2017).

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Building on these architectural strengths and by integrating the ML's advanced temporal modelling capabilities, this work proposes a weakly supervised transfer learning (TL) algorithm to improve the reliability of existing databases and create new databases with minimal initial human supervision.

3.1 Weakly supervised learning as a pseudo-labelling approach

Weakly supervised learning is a branch of ML covering the construction of predictive models by learning with weak supervision (Zhou, 2018). Such techniques focus on learning with incomplete, inexact, and/or inaccurate information derived from noisy, limited, or imprecise supervision processes. The objective is to automatically provide supervision for labelling large amounts of data using labelling functions derived from domain knowledge. This approach replaces the costly and impractical hand-labelled process with inexpensive weak labels with the understanding that although imperfect, they can be used to create a strong predictive model.

In this study, the source domain (denoted as D_s) is the MASTER-DEC dataset (based on refined physical models and a strong revision process). The target domain (denoted as D_t) is the Popo2002 dataset (whose available seismic catalogue will not be considered). The goal of this study is to address a domain adaptation task (Kouw and Loog, 2019; Farahani et al., 2021) to reduce the cost of developing a reliable seismic catalogue and database with minimal initial human supervision. That is, automatically provide supervision for labelling large amounts of data from D_t using labelling functions derived from domain knowledge D_s.

In a domain adaptation framework, typically D_s and D_t have the same feature space but different distributions. However, in this study, for the pseudo-labelling task we assumed that:

- The marginal distributions of D_s and D_t are the same: $P_s(X_s) = P_t(X_t)$, where X_s and X_t are the input feature vectors associated with different seismic windows or frames in both domains. As such, the pseudo-labelled samples do not need to contain any domain information and the occurrence of different seismic events is equally likely in both domains.
- The conditional distributions of D_s and D_t are the same: $Q_s(Y_s|X_s) = Q_t(Y_t|X_t)$. As such, the pseudo-labelled samples are valid in both domains.

Such assumptions have important implications since in the target domain, while the marginal distributions of D_s and D_t are the same [$P_s(X_s) = P_t(X_t)$], the conditional distributions could be different [$Q_s(Y_s|X_s)$ /= $Q_t(Y_t|X_t)$]. This shows how similar feature vectors taken as the input could output different probabilistic event detection matrices. That is, the description or characterisation of seismic categories could change between domains, or D_t could contain seismic categories unforeseen in D_s.

Therefore, by leveraging the probabilistic detection matrices output by the system trained in D_s, we can apply a weakly supervised learning technique as a pseudo-labeller in D_t to construct a new dataset from which to train a new system in a supervised way. Those parts of the unlabelled dataset with high per-class probability, and then high confidence, are added to the new training set. Although imperfect, this method guarantees that at least events showing characteristics similar to those

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

annotated in the master catalogue will be included in the new training dataset. As a result, after re-training phase, the target catalogue is enlarged and updated.

Our proposed weakly supervised transfer learning (TL) algorithm is outlined as follow and depicted in Fig. 1. It is important to highlight that our methodology is just one possible approach; many other weakly supervised methods could also be applied to achieve similar results.

1. **Subset Analysis:** A subset of the Popo2002 dataset (20%-40% of the total) is analyzed using a ML system (RNN-LSTM, Dilated-RNN, TCN) trained on MASTER-DEC.
2. **Event Detection and Confidence Analysis:** Ignoring the information contained within the available Popo2002 seismic catalogue, we analyze the confidence of each detected event using a probabilistic event detection matrix with per-class membership output from the softmax layer. This step allows us to quantify the severity of drift between datasets (‘concept drift detection’) (Lu et al. 2018). High or extremely high per-class recognition probabilities for each event type indicate the systems are well-fitted to the master database. Low per-class probabilities indicate a change in the description of the analyzed information. Accurate and robust dissimilarity measurement and statistical hypothesis evaluation are not strictly necessary, given the well-known dissimilarity between the volcanic environments.
3. **Drift Adaptation Mechanism:** An adaptive threshold mechanism is adopted. Detected events with an average per-class probability above a certain threshold are selected and included as training instances.
4. **Re-training process:** Finally, the pre-trained ML systems used in the step 1 are re-trained using the selected instances and labels obtained in the step 3.
5. **Iterative Refinement:** Repeat steps 2 to 4 iteratively until the desired result is achieved.

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

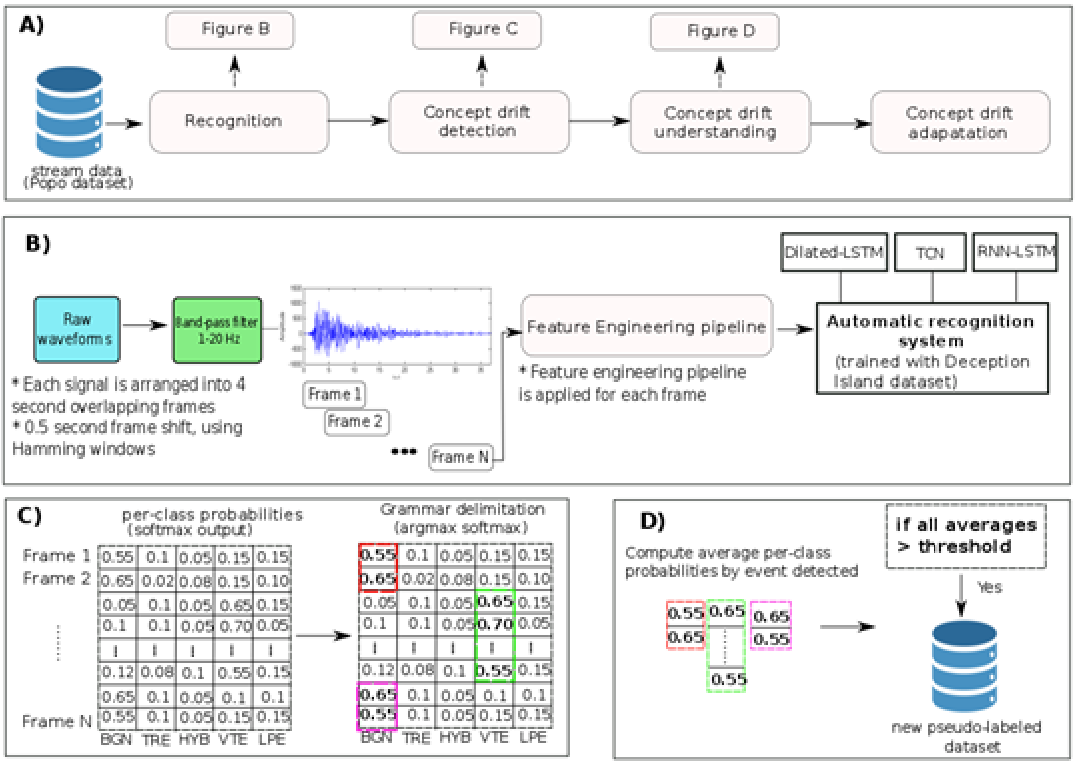

Figure 1 a) Weakly supervised event selection algorithm applied to Popo2002 dataset. A subset of the dataset (in our case 40% of the total) is used as a training set by the previously trained ML models. The rest of the data is used as a test set. Only high per-class probability recognized events are selected as new training instances. b) Workflow structure and the specific preprocessing steps employed in this article. The specific preprocessing elucidates the feature engineering process, which relies on frequency analysis within the logarithmic filter bank domain (ref). This processed information serves as the input for the different volcano-seismic recognition systems. c) For each detected event, the confidence of the detection is analysed using a probabilistic event detection matrix with per-class probabilities output by the softmax layer. d) Drift adaptation mechanism based on an adaptive threshold was then adopted. Those events whose average number of per-class probability was greater than a given threshold were selected and included as training instances.

4 Results

This section presents the results supporting the objectives outlined in the previous section: a) verifying the reliability of the existing database and, b) creating a new database without the initial time-consuming human supervision. The reliability of the dataset will be verified through the self-consistency results of the Popo2002 database. Although self-consistency could be achieved through two different approaches, training the system from scratch or applying transfer learning, in this work we focus on transfer learning approaches since they offer significant advantages (they reduce training time and resource usage by leveraging pre-trained models and they enhance performance with limited data by utilizing knowledge gained from diverse datasets, leading to better generalization and mitigating overfitting). Therefore, the self-consistency results correspond to the automatic recognition performance obtained by conducting a classical supervised Transfer Learning where a model previously trained in one volcanic environment (MASTER-DEC) is used as a starting point to build a new system in

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

a different volcanic environment (Popo2002) using as training the available original catalogue. Finally, to verify the feasibility of creating a new database without human supervision, we present the recognition results obtained using our weakly supervised learning proposal.

4.1 Checking the self-consistency of Popo2002 dataset: Traditional transfer of knowledge from MASTER-DEC to Popo2002

The most efficient procedure to check the self-consistency of a given database is to tune an automatic system encompassing all types of available events. As mentioned before; to accelerate development and achieve high-quality results compared to training from scratch, we will employ a classical Transfer Learning approach. If the events classified as elements of the same class are not homogeneous, the model's performance will be low (usually below 70% accuracy).

The self-consistency of the Popo2002 database is checked using the three proposed ML architectures (RNN-LSTM, Dilated-RNN and TCN). We apply a Leave One Out cross-validation method with 4 random partitions. Each time, we select T% of the entire database as training set, and the remaining (100-T) % as test set to check the performance of the systems. Given that using a pre-trained model allows for more efficient utilization of computational resources and that the fine-tuning phase generally requires fewer resources than training a model from scratch, we conducted two experiments with T = 20 and T = 40. This means that the results were obtained using 80% and 60% of the data in the test partition, respectively.

Table 2 presents the results of the preliminary self-consistency test for the three proposed architectures. As indicated, the confidence matrix indices are high, suggesting good self-consistency. Table 3 shows the averaged normalised confusion matrices of the Leave One Out cross validation process for the Popo2002 dataset.

The results of these tests indicate that inside of each seismic class there is a coherence of the elements selected to belong to them. However, as noted by Titos et al. (2018), propagation and source effects can influence the characterization of seismic events. For example, according to the Table 3 and the original Popo2002 catalogue, VTE events are not properly identified, as the confusion level is over 60% in some cases, meaning only 40% of the VT events are correctly classified. In all cases, the highest levels of confusion are observed with the LPE class. It could be explained because LPE events could share similar characteristics than high attenuated VTs, producing potential biases in the allocation of events in the right category. Consequently, different seismic categories may have elements located in overlapping areas of the representation space, where they share similar projected parameters. These events, despite being assigned to a specific cluster, could easily transition to another. Therefore, a high performance on the self-similarity test does not guarantee an unbiased classification. While this self-consistency suggest uniformity in the training procedure, it's important to acknowledge that potential biases in the labelling process may not be evident in this measure of goodness.

	5 seismic categories	7 seismic categories

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

	Training percentage		Training percentage	
	20%	40%	20%	40%
RNN-LSTM	77.38	88.99	84.01	84.39
Dilated-RNN	82.88	84.70	84.05	85.21
TCN	82.46	88.30	85.77	83.27

Table 2 Self-consistency results using 5 and 7 seismic categories, with 20% and 40% of the data for training and 80% and 60% for testing, respectively. The results correspond to the average accuracy over the four partitions.

	RNN-LSTM					Dilated-RNN					TCN				
	BGN	TRE	HYB	VTE	LPE	BGN	TRE	HYB	VTE	LPE	BGN	TRE	HYB	VTE	LPE
BGN	0.97	0.02	0	0	0.01	0.96	0.02	0	0	0.02	0.98	0.01	0	0	0.01
TRE	0.06	0.78	0	0.05	0.11	0.13	0.69	0	0	0.18	0.11	0.68	0	0.09	0.12
HYB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VTE	0.08	0.13	0	0.51	0.28	0.12	0.17	0	0.31	0.4	0.14	0.09	0	0.59	0.18
LPE	0.05	0.07	0	0.03	0.85	0.04	0.18	0	0	0.78	0.05	0.05	0	0.04	0.86

Table 3 Averaged normalized confusion matrices associated with the Leave One Out cross validation process for the Popo2002 dataset. These results belong to the test using 5 seismic categories.

4.3 Weakly supervised Transfer Learning

The direct application of classic TL could be considered a kind of blind test because, the automatic system acquires the pre-existing knowledge in the catalogue, since they used the original labels during training, thereby propagating the existing bias. According to section 3.1, we propose a solution to improve the TL approach through a weakly supervised method. By leveraging an existing master database, we can impose prior knowledge on the database being tested. The seismic categories of the testing database will be re-reviewed under the supervision of the system trained with the 'unbiased' dataset. In this way, the potential bias of the tested database will not be propagated to either domain, as the Master catalogue imposes its knowledge.

Table 4 displays the self-consistence of the Popo2002 dataset when the MASTER-DEC knowledge is imposed. We observe a noticeable decrease in accuracy, attributable to the aforementioned bias in event classification. The first column of Table 4 shows the recognition results obtained when directly applying the models trained on the MASTER-DEC dataset to recognize

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

events in Popo2002 compared to the original catalogue. The second column presents the results obtained after applying our weakly supervised TL approximation, comparing the new Popo2002 catalogue with the original one. These results highlight how the new POPO data base has been modified in comparison with the original.

This procedure could be applied iteratively until a reliable catalogue is achieved. In our case, after three iterations, the improvements only reached 2% compared to the first iteration. However, it could change when using a different test dataset. In the next section, we will demonstrate how this new catalogue dramatically improves the original one, confirming that the new Popo database could be better adapted.

	Five seismic categories blind test	**'Weakly supervised TL' using five seismic categories TL**
RNN-LSTM	55.95	64.89
Dilated-RNN	50.13	55.72
TCN	58.27	66.16

Table 4 Classification accuracy (acc. %) on the test set obtained when directly applying the models trained on the MASTER-DEC dataset to recognize events in Popo2002 compared to the original catalogue and after applying our weakly supervised TL approximation keeping only five seismic categories and using 40% of the total dataset as the training set.

	RNN-LSTM					Dilated-RNN					TCN				
	BGN	TRE	HYB	VTE	LPE	BGN	TRE	HYB	VTE	LPE	BGN	TRE	HYB	VTE	LPE
BGN	0.88	0.09	0	0	0.03	0.67	0.32	0	0	0.01	0.8	0.19	0	0	0.01
TRE	0.29	0.36	0.03	0.02	0.03	0.29	0.5	0	0	0.21	0.19	0.7	0	0	0.11
VTE	0.27	0.41	0.08	0.03	0.21	0.46	0.28	0	0.03	0.23	0.36	0.46	0.03	0.06	0.09
LPE	0.36	0.19	0.06	0.06	0.33	0.47	0.18	0	0.01	0.34	0.41	0.33	0.01	0.01	0.24

Table 5 Normalised confusion matrices related to the weakly supervised approach implemented for the three architectures, using the Popo2002 catalogue as reference. The results are over the whole test set using 40% of the whole set for training and five seismic categories. The y-axis corresponds to the real label or ground-truth and the x-axis corresponds to predicted labels.

5 Discussion

This work emphasizes in the use of Weakly supervised approaches to improve seismic-volcanic catalogues. We have verified that, when effectively use, these procedures could significantly enhance the detection and identification capabilities

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

of various earthquake-volcanic signals. Consequently, early warning protocols for volcanic eruptions could also being improved.

What stands out in the Table 2 and Table 3 is that not all TL techniques are universally valid for improving seismic-volcanic catalogues. Our initial process involved directly applying self-consistency study techniques to the initial database. To achieve this, we utilized traditional TL approaches with various ANN architectures and analysed the confusion matrix obtained to determine the quality of the database. As shown in such tables, the recognition performances for Popo2002 dataset are high, with accuracy levels up to 85%. This result is often used in many works as the sole criterion for validating the quality of the database. However, a closer inspection of Fig. 1 (Uniform Manifold Approximation and Projection for dimension reduction (UMAP) (McInnes et al., 2018) visualization of the MASTER-DEC) as well as the study by Bueno et al. (2021) for the Bezymianny volcano indicate that the identification of classes of seismic events is not free of uncertainties. Different seismic events share many parameters in the representation space, and as a result, the definition of the event type depends on the training process, which could be influenced by human factors. Therefore, while systems' performances range from approximately 85% to 90%, it does not always reflect that the seismic catalogue is complete or unbiased, but rather that the training process has been self-consistent. More than learning to characterise the volcano dynamics describing the latent physical model, the systems are learning the information contained in the catalogue. For instance, in a tableware classification process, if a trained class includes both spoons and forks, the result of the accuracy matrix will be optimal, although the bias is evident. That is, catalogue-induced learning could bias the generalisation capability of a system, with information or dynamics relevant to improving background knowledge about the volcano remaining hidden.

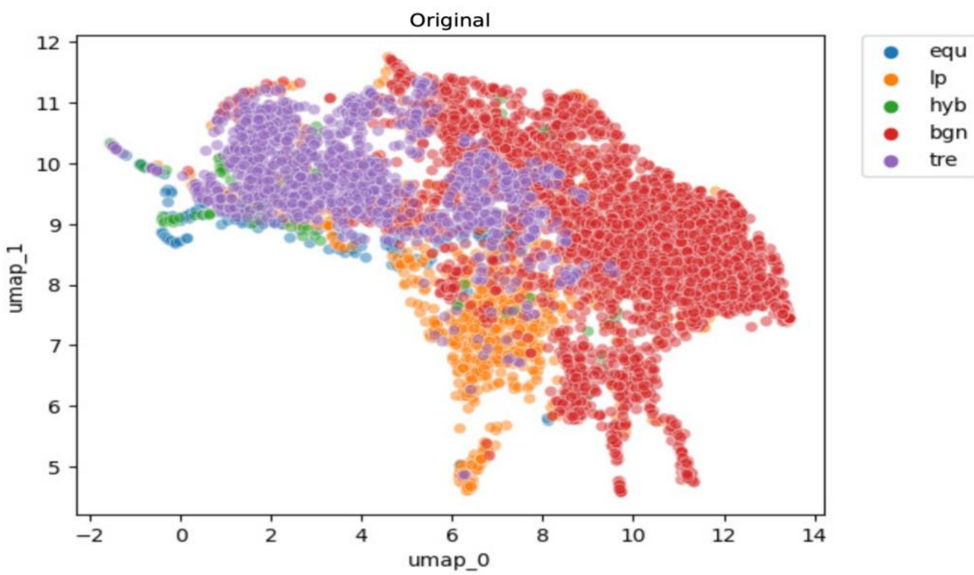

Figure 2 Uniform Manifold Approximation and Projection (UMAP) obtained for the input vector forming the original data of MASTER-DEC dataset. Different seismic categories may have elements located in overlapping areas of the representation space, where they share similar projected parameters.

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences Discussions

Our proposal to address this trend and potentially improve the system's generalization capabilities as well as the completeness of the seismic dataset is based on weakly supervised procedures. By leveraging an existing unbiased database, we can impose prior knowledge on the database being tested. This involves using an automatic system, which we consider well-trained, to review the seismic categories of the new dataset and retrain itself using such pseudo-labels.

According to Table 4, when using this approach, the performance decreased substantially compared with the classical TL approaches (Table 2). However, closer inspection of the results shows other aspects of the performance being very encouraging. First, the models detected events that were originally not recognised by expert geophysicists during data-labelling. Table 6 summarizes the comparison between the events initially annotated in the catalogue and the events detected by the three automatic systems. The vast majority of undetected events, which were not annotated in the preliminary catalogue, were discovered within segments labelled as GAR (Fig. 3a)). Second, the data-labelling process carried out by geophysical experts located the ends and beginnings of some events in positions very different from those provided by the models, which decreased per-frame recognition. This can clearly be seen during earthquake recognition (Fig. 3b)), where an average of just 5% of frames were correctly recognised (see Table 5 for detail): While segments with high spectral content were correctly detected and classified as VTE, those with low spectral content were classified as BGN or TRE. However, after a posterior supervision by a geophysical expert, we consider that the outputs obtained by the models were better, as the earthquakes were generally well recognised.

Attenuation and sources effects also affected the final recognition. The new dataset contains high intra-class variability in some categories (those composed of different subcategories as LPE or TRE). Again, the nature of the seismic data played an essential role. Within the feature space, the representation of events belonging to a given subcategory in the new domain (Popo2002) was closely related to the representation of events belonging to a different category in the source domain (MASTER-DEC). For example, similar to what occurs with some events in Fig. 1, the representations of some LPE subcategories in Popo2002 were very close to the representation of TRE in MASTER-DEC (Fig. 4a)). As such, the weakly supervised selection and labelling algorithm assigned the TRE label during the training phase. This decreased the systems performance since many frames (33%, 19%, and 18% for TCN, RNN–LSTM, and Dilated–LSTM, respectively) were detected as TRE, resulting in an 'error' that was not really an error. The same issue arose for some attenuated earthquakes, which were labelled as LPE in the original seismic catalogue but 'mis-classified' as VTE or TRE in the new one. After a posterior supervision by a geophysical expert, we consider the output obtained by the models to be correct (Fig. 4b) since the seismic traces correspond to earthquakes with attenuated high frequencies. Finally, low-energy TRE events were clearly mis-classified as BGN because the peak-to-peak amplitude degradation of the signals was related to attenuation effects.

This complex scenario was widely discussed by Titos et al. (2018); therefore, to correctly deal with these errors, further information from several seismic stations is needed. Additional illustrative examples can be found in the Weakly Supervised Learning section of the supplementary material (refer to Fig. F1, F2, and F3). The results suggest that per-frame recognition can be strongly biased by the intrinsic limitations addressed when developing the seismic catalogue and from which the comparative metrics were obtained. Therefore, if labelling criteria between datasets differ, per-frame recognition results will vary widely. Until now, the development of new monitoring systems has focused primarily on improving existing

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

recognition rates. However, our findings confirm that weakly supervised learning approaches have the remarkable capability of simultaneously identifying unannotated seismic traces in the catalogue and help to correct the labels of mis-annotated seismic traces. When a weakly supervised approach based on a master seismic catalogue guides the re-training phase, although the general performance of the system decreases, previously hidden information that can improve knowledge of the volcanic dynamic background can be obtained.

	Popo2002 catalogue	**RNN-LSTM**	**Dilated-LSTM**	**TCN**
BGN	340	>20,000	>20,000	17,206
TRE	273	3,291	2,538	3,204
VTE	371	1,741	1,032	94
LPE	1,155	2,230	2,250	2,159

Table 6 Comparison between the events initially annotated in the catalogue and the events detected by the three distinct automatic systems following the implementation of a weakly supervised learning TL approach.

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Discussions

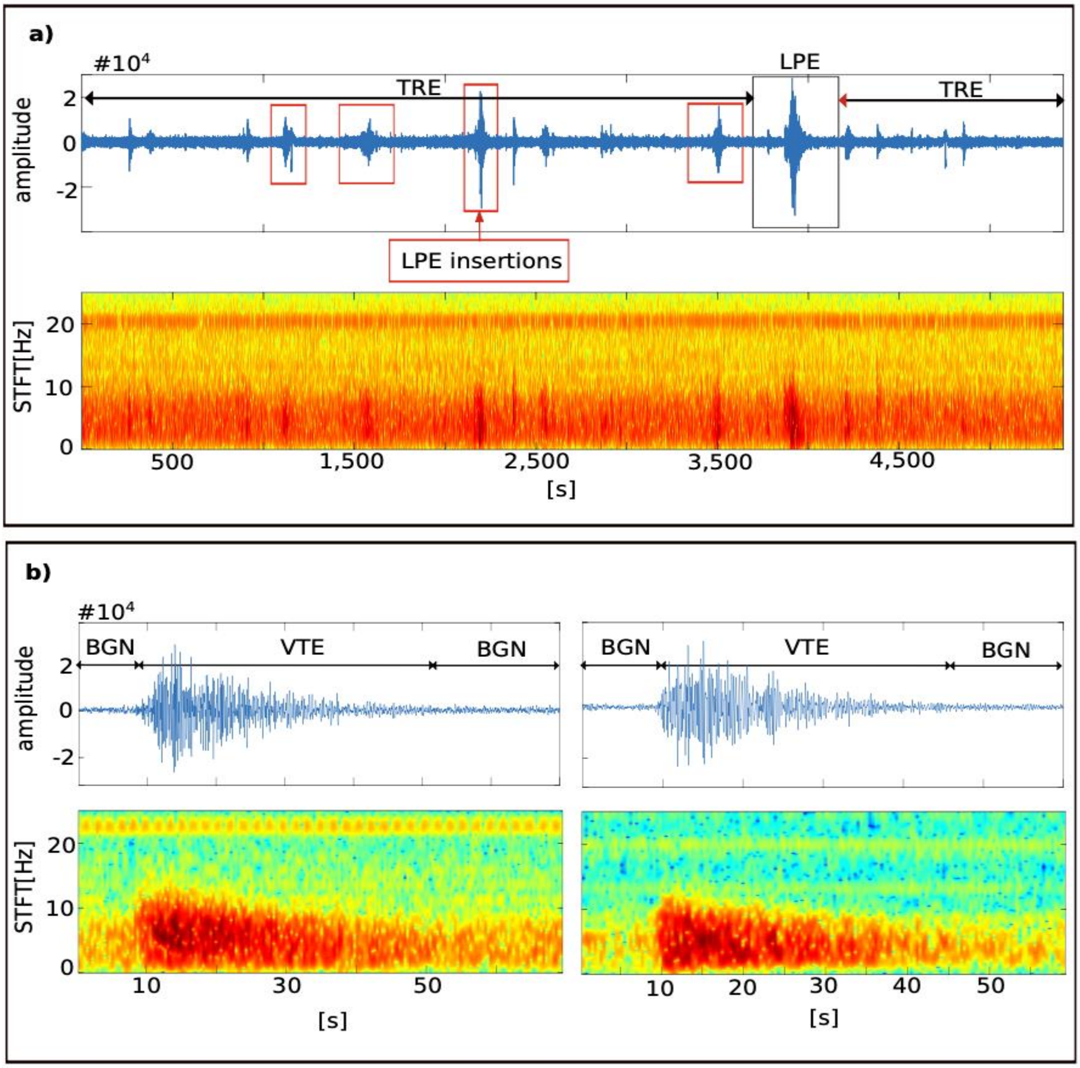

Figure 3 Insertion-based errors when applying a weakly supervised transfer learning (TL) approach. (a) Detection of events previously not labelled: long period events (LPE; red boxes) that were not originally taken into account during the labelling process have been added to the seismic trace labelled as TRE (tremor)–LPE–TRE. This scenario occurs many times throughout the dataset and these insertions decrease per-frame recognition with respect to the original labelling; however, they do not always correspond to errors. (b) Event delimitation: examples of the labelling process obtained for the models. Instead of recognising entire seismic traces, such as volcano-tectonic earthquakes (VTE), background noise (BGN) segments are detected before and after the earthquakes. Again, these insertions reduce per-frame recognition; however, after a posterior revision, they should not be considered as an error. The current colormap in the spectrogram represents the energy levels. The blue color corresponds to the minimum energy, while the red color corresponds to the maximum energy.

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Discussions

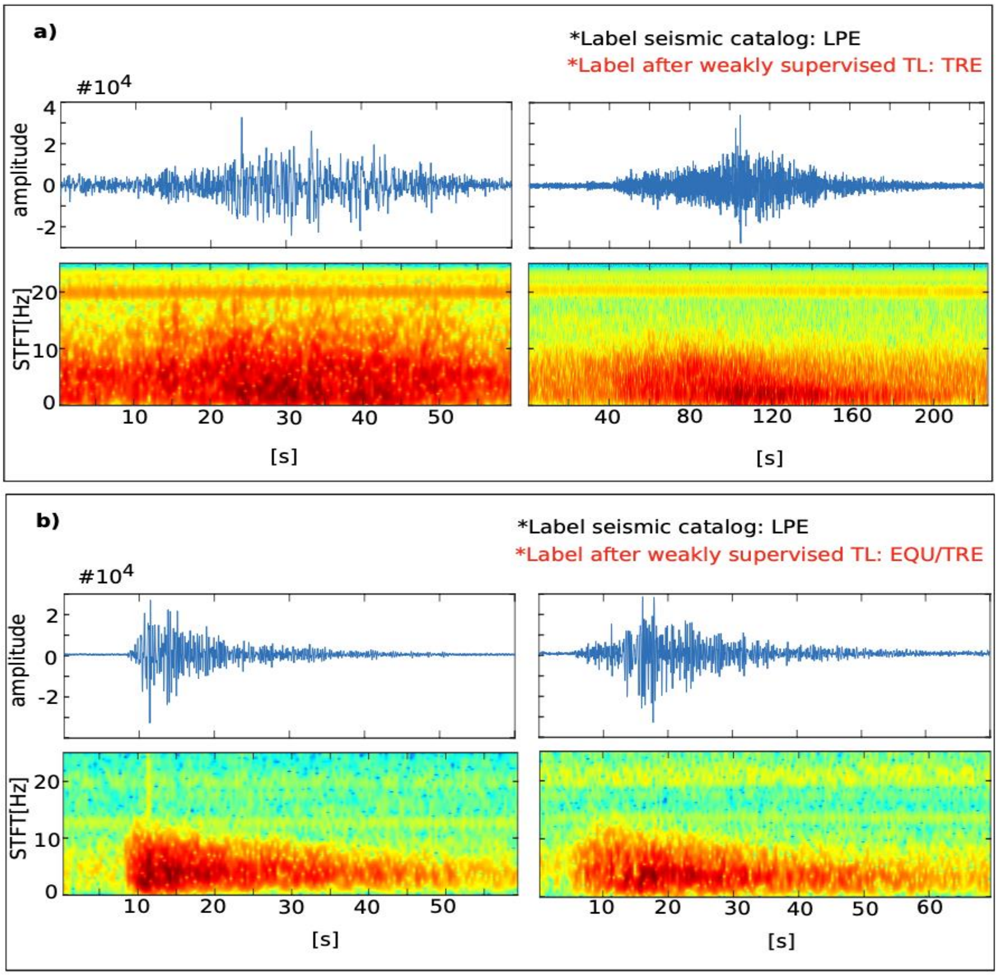

Figure 4 Detailed analysis of intra-class variability and attenuation-based errors when applying a weakly supervised transfer learning (TL) approach. (a) Intra-class variability-based errors: some long period event (LPE) subcategories in Popo2002 are very close to the representation of tremor (TRE) in MASTER-DEC. Thus, the weakly supervised selection and labelling approach will assign the TRE label during the training phase; this decreases per-frame recognition. (b) Two attenuated earthquakes labelled as LPE in the seismic catalogue, but 'miss-classified' as volcano-tectonic earthquake (VTE) or TRE. After a posterior supervision, we consider the model outputs to be correct since the seismic traces correspond to earthquakes with attenuated high frequencies. The current colormap in the spectrogram represents the energy levels. The blue color corresponds to the minimum energy, while the red color corresponds to the maximum energy.

6 Conclusions

This study provides the first comprehensive analysis of seismic catalogue-induced bias when developing automatic recognition systems. We evaluated the ability of several monitoring systems trained using a master seismic catalogue from Deception Island volcano to adapt to a new seismic catalogue from Popocatépetl volcano through our novel, proposed weakly supervised framework.

Our results confirm the robustness of data-driven approaches as a basis for the construction of short-term early-warning systems. However, quantitative and qualitative analysis confirmed that the reliability of a system is strongly biased by the undetailed coverage of the seismic catalogue. While systems performance reached almost 90% per-frame recognition accuracy, intrinsic limitations when developing seismic catalogues led to extremely useful information describing the volcanic behaviour being ignored. Instead of simply learning to characterise volcanic dynamics by describing the latent physical model, catalogue-induced learning can bias the system by discarding useful data describing volcanic dynamics. However, when a weakly supervised learning approach based on a master seismic catalogue is applied, an unknown amount of information related to volcano dynamics is revealed.

This study raises important questions about the relevance of catalogue-induced learning when developing new monitoring systems. Our results demonstrate that systems based on iterative weakly supervised or even unsupervised learning techniques could offer a more successful approach than supervised techniques under crude seismic catalogues. Therefore, we conclude that ensuring appropriate seismic catalogues and support for developing monitoring tools should be a priority to the same extent as applying new and more effective AI techniques. The use of more sophisticated pseudo-labelling techniques involving data from several catalogues could help to develop universal monitoring tools able to work accurately across different volcanic systems, even when faced with unforeseen temporal changes in monitored signals.

Acknowledgements

This work has been partially funded by the Spanish Project PID2022-143083NB-I00, “LEARNING”, funded by MCIN/AEI /10.13039/501100011033 and by FEDER (EU), DIGIVOLCAN (PLEC2022-009271), SDAS (B-TIC-542-UGR20), and PROOF-FOREVER (EUR2022-134044). The authors would also like to acknowledge Instituto Andaluz de Geofísica, Raúl Arámbula-Mendoza and CENAPRED for providing the data from Deception Island and Popo volcanoes supporting this work. Any opinions, findings, conclusions, or recommendations expressed in this research are those of the authors and do not necessarily reflect the view of the funding agencies.

Author contributions statement

All authors contributed to the conception of the study. M.T., C.B., conceived and conducted the experiments. All authors analysed and interpreted the results. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests

Additional information

Correspondence and requests for materials should be addressed to M.T.

References

Sparks, R. S. J. Forecasting volcanic eruptions. Earth and Planetary Science Letters, 210(1-2), 1-15 (2003). Doi: 10.1016/S0012-821X(03)00124-9

Witze, A. How ai and satellites could help predict volcanic eruptions. Nature 567, 156–158 (2019). doi: 10.1038/d41586-019-00752-3

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Palmer, J.: The new science of volcanoes harnesses ai, satellites and gas sensors to forecast eruptions. Nature 581, 256–260 (2020), doi: 10.1038/d41586-020-01445-y

Chouet, B.: Volcano seismology. Pure applied geophysics 160, 739–788 (2003), doi: 10.1007/PL00012556

McNutt, S. R., Roman, D. C.: Volcanic seismicity. In The encyclopedia of volcanoes, 1011–1034(2015), doi: 10.1016/B978-0-12-385938-9.00059-6

Minakami, T.: Prediction of volcanic eruptions. Dev. Solid Earth Geophys.6, 313–333 (1974), doi: 10.1016/B978-0-444-41141-9.50020-6.

Rey-Devesa, P., Prudencio, J., Benítez, C. et al.: Tracking volcanic explosions using Shannon entropy at Volcán de Colima. Sci Rep 13, 9807 (2023). doi:10.1038/s41598-023-36964-x

Rey-Devesa, P., Benítez, C., Prudencio, J. et al.: Volcanic early warning using Shannon entropy: Multiple cases of study. Journal of Geophysical Research: Solid Earth, 128, e2023JB026684 (2023). doi:10.1029/2023JB026684

Ohrnberger, M. Continuous automatic classification of seismic signals of volcanic origin at Mt. Merapi, Java, Indonesia. Ph.D. thesis, Potsdam, Univ., Diss., 2001 (2001)

Scarpetta, S., Giudicepietro, F., Ezin, E. C., et al.: Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks. Bulletin of the Seismological Society of America, 95(1), 185-196 (2005).doi: 10.1785/0120030075.

Alasonati, P., Wassermann, J., Ohrnberger, M.: Signal classification by wavelet-based hidden Markov models: application to seismic signals of volcanic origin. Statistics in Volcanology, H. M. Mader, S. G. Coles, C. B. Connor, L. J. Connor (2006). doi: 10.1144/IAVCEI001.13

Benítez, M. C., Ramírez, J., Segura, J. C., et al.: Continuous HMM-based seismic-event classification at Deception Island, Antarctica. IEEE Transactions on Geoscience and remote sensing, 45(1), 138-146 (2006). doi:10.1109/TGRS.2006.882264

Ibáñez, J. M., Benítez, C., Gutiérrez, L. A., et al.: The classification of seismo-volcanic signals using Hidden Markov Models as applied to the Stromboli and Etna volcanoes. Journal of Volcanology and Geothermal Research, 187(3-4), 218-226 (2009). doi:10.1016/j.jvolgeores.2009.09.002

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Curilem, G., Vergara, J., Fuentealba, G., et al.: Classification of seismic signals at Villarrica volcano (Chile) using neural networks and genetic algorithms. Journal of volcanology and geothermal research, 180(1), 1-8 (2009). doi: 10.1016/j.jvolgeores.2008.12.002

Bhatti, S. M., Khan, M. S., Wuth, J., et al.: Automatic detection of volcano-seismic events by modeling state and event duration in hidden Markov models. Journal of Volcanology and Geothermal Research, 324, 134-143 (2016). doi: 10.1016/j.jvolgeores.2016.05.015

Canario, J. P., Mello, R., Curilem, M., et al.: In-depth comparison of deep artificial neural network architectures on seismic events classification. J. Volcanol. Geotherm. Res. 401, 106881 (2020). doi: 10.1016/j.jvolgeores.2020.106881

Cortés, G., Carniel, R., Lesage, P., et al.: Practical volcano-independent recognition of seismic events: Vulcan. ears project. Front. Earth Sci. 8, 616676 (2021). doi: 10.3389/feart.2020.616676

Bicego, M., Rossetto, A., Olivieri, M., et al.: Advanced knn approaches for explainable seismic-volcanic signal classification. Math. Geosci. 1–22 (2022). doi: 10.1007/s11004-022-10026-w

Bueno, A., Benítez, C., Zuccarello, L., et al.: Bayesian monitoring of seismo-volcanic dynamics. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-14 (2021). doi: 10.1109/TGRS.2021.3076012

Bueno, A., Balestriero, R., De Angelis, S., et al.: Recurrent scattering network detects metastable behavior in polyphonic seismo-volcanic signals for volcano eruption forecasting. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-23 (2021). doi:10.1109/TGRS.2021.3134198

Martínez, V. L., Titos, M., Benítez, C., et al.: Advanced signal recognition methods applied to seismo-volcanic events from Planchon Peteroa Volcanic Complex: Deep Neural Network classifier. Journal of South American Earth Sciences, 107, 103115 (2021). doi: 10.1016/j.jsames.2020.103115

Titos, M., Bueno, A., Garcia, L., Benitez, C.: A deep neural networks approach to automatic recognition systems for volcano-seismic events. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 11, 1533–1544 (2018). doi: 10.1109/JSTARS.2018.2803198

Titos, M., Bueno, A., García, L., Benítez, C., Ibáñez, J. M.: Detection and classification of continuous volcano-seismic signals with recurrent neural networks. IEEE Transactions on Geosci. Remote. Sens. 57, 1936–1940 (2019). doi: 10.1109/TGRS.2018.2870202

Titos, M., García, L., Kowsari, M., Benítez, C.: Toward knowledge extraction in classification of volcano-seismic events: Visualizing hidden states in recurrent neural networks. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 15, 2311–2325, DOI: 10.1109/JSTARS.2022.3155967 (2022). doi: 10.1109/JSTARS.2022.3155967

Titos, M., Gutiérrez, L., Benítez, C., et al.: Multi-station volcano tectonic earthquake monitoring based on transfer learning. Frontiers in Earth Science, 11, 1204832 (2023). doi: 10.3389/feart.2023.1204832

Ibáñez, J. M., De Angelis, S., Díaz-Moreno, A., et al.: Insights into the 2011–2012 submarine eruption off the coast of El Hierro (Canary Islands, Spain) from statistical analyses of earthquake activity. Geophysical Journal International, 191(2), 659-670 (2012). doi: 10.1111/j.1365-246X.2012.05629.x

Díaz-Moreno, A., Ibáñez, J. M., De Angelis, S., et al.: Seismic hydraulic fracture migration originated by successive deep magma pulses: The 2011–2013 seismic series associated to the volcanic activity of El Hierro Island. Journal of Geophysical Research: Solid Earth, 120(11), 7749-7770 (2015). doi: 10.1002/2015JB012249

Chang, S., Zhang, Y., Han, W., et al.: Dilated recurrent neural networks. Advances in neural information processing systems, 30. (2017).

Titos, M., Carthy J., García, L., et al.: Dilated-RNNs: a deep approach for continuous volcano-seismic events recognition. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.

Smellie, J. L. Recent observations on the volcanic history of deception island, south shetland islands. Br. Antarctic Surv. Bull. 83–85 (1988).

Carmona, E., Almendros, J., Serrano, I., et al.: Results of seismic monitoring surveys of deception island volcano, antarctica, from 1999–2011. Antarctic Sci. 24, 485–499 (2012). doi: 10.1017/S0954102012000314

Martí J., Geyer, A., Aguirre-Díaz, G. et al.: Deception Island (south shetland islands, antarctica): an example of a tectonically induced collapse caldera. (2011). URI: http://hdl.handle.net/10261/162691

Ibañez, J.M., Pezzo, E. D., Almendros, J,. et al. Seismovolcanic signals at Deception Island volcano, Antarctica: Wave field analysis and source modeling. J. Geophys. Res. Solid Earth 105, 13905–13931 (2000). doi: 10.1029/2000JB900013

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Natural Hazards and Earth System Sciences
Open Access
Discussions

Martínez-Arévalo, C., Bianco, F., Ibáñez, J. M., et al..: Shallow seismic attenuation and shear-wave splitting in the short period range of Deception Island volcano (Antarctica). Journal of Volcanology and Geothermal Research, 128(1-3), 89-113 (2003). doi: 10.1016/S0377-0273(03)00248-8

Zandomeneghi, D., Barclay, A., Almendros, J., et al.: Crustal structure of Deception Island volcano from P wave seismic tomography: Tectonic and volcanic implications. Journal of Geophysical Research: Solid Earth, 114(B6) (2009). doi: 10.1029/2008JB006119

Ibáñez, J.M., Díaz-Moreno, A., Prudencio, J., et al.: Database of multi-parametric geophysical data from the TOMO-DEC experiment on Deception Island, Antarctica. Scientific data, 4(1), 1-18 (2017). doi: 10.1038/sdata.2017.128

Arango-Galván, C., Martin-Del Pozzo, A. L., Flores-Márquez, E., et al. Unraveling the complex structure of popocatepetl volcano (central mexico): New evidence for collapse features and active faulting inferred from geophysical data. J. Volcanol. Geotherm. Res. 407, 107091 (2020). doi:10.1016/j.jvolgeores.2020.107091

Alaniz-Álvarez, S. A., Nieto-Samaniego, Á. F., Mexicana, S. G. Geology of México: Celebrating the centenary of the Geological Society of México, vol. 422 (Geological Society of America, 2007)

Siebe, C., Salinas, S., Arana-Salinas, L., Macías, J. L., et al. The~ 23,500 y 14C BP white pumice plinian eruption and associated debris avalanche and tochimilco lava flow of popocatépetl volcano, méxico. J. Volcanol. Geotherm. Res. 333, 66–95 (2017). doi: 10.1016/j.jvolgeores.2017.01.011

Malfante, M., Dalla Mura, M. et al.: Machine learning for volcano-seismic signals: Challenges and perspectives. IEEE Signal Process. Mag. 35, 20–30 (2018). doi: 10.1109/MSP.2017.2779166

Lara, F., Lara-Cueva, R., Larco, J. C., et al.: A deep learning approach for automatic recognition of seismo-volcanic events at the cotopaxi volcano. J. Volcanol. Geotherm. Res. 409, 107142 (2021). doi: 10.1016/j.jvolgeores.2020.107142

Hibert, C., Provost, F., Malet, J. P. et al. Automatic identification of rockfalls and volcano-tectonic earthquakes at the piton de la fournaise volcano using a random forest algorithm. J. Volcanol. Geotherm. Res. 340, 130–142 (2017). doi: 10.1016/j.jvolgeores.2017.04.015

Köhler, A., Ohrnberger, M. & Scherbaum, F. Unsupervised pattern recognition in continuous seismic wavefield records using self-organizing maps. Geophys. J. Int. 182, 1619–1630 (2010). doi: 10.1111/j.1365-246X.2010.04709.x

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

Hochreiter, S. & Schmidhuber, J. Long short-term memory. Neural computation 9, 1735–1780 (1997). doi: 10.1162/neco.1997.9.8.1735

Schmidhuber, J. Deep learning in neural networks: An overview. Neural networks 61, 85–117 (2015). doi: 10.1016/j.neunet.2014.09.003

Lea, C., Flybbn, M. D., Vidal, R., et al.: Temporal convolutional networks for action segmentation and detection. In Proc. 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 1003–1012 (2017).

Yan, J., Mu, L., Wang, L., et al.: Temporal convolutional networks for the advance prediction of enso. Nat. Sci. Reports 10 (2020). doi: 10.1038/s41598-020-65070-5

Racic, M., Oštir, K., Peressutti, D., et al.: Application of temporal convolutional neural network for the classification of crops on sentinel-2 time series. ISPRS - Int. Arch. Photogramm. Remote. Sens. Spatial Inf. Sci. XLIII-B2-2020, 1337–1342, (2020). doi: 10.5194/isprs-archives-XLIII-B2-2020-1337-2020

Aäron van den Oord, S. D., Heiga Z., Karen S., et al.: A generative model for raw audio. CoRR abs/1609.03499, (2016). doi: 10.48550/arXiv.1609.03499

Fisher Y., Vladlen K., and Thomas F.: Dilated residual networks. arXiv preprint arXiv:1705.09914, (2017). doi: 10.48550/arXiv.1705.0991

Zhou, Z. H. A brief introduction to weakly supervised learning. Natl. science review 5, 44–53 (2018). doi: 10.1093/nsr/nwx106

Kouw, W. M. & Loog, M.: A review of domain adaptation without target labels. IEEE transactions on pattern analysis machine intelligence 43, 766–785 (2019). doi: 10.1109/TPAMI.2019.2945942

Farahani, A., Voghoei, S., Rasheed, K., Arabnia, H. R.: A brief review of domain adaptation. Adv. data science information engineering 877–894 (2021). doi: 10.1007/978-3-030-71704-9_65

Lu, J., Liu, A., Dong, F.: Learning under concept drift: A review. IEEE transactions on knowledge and data engineering, 31(12), 2346-2363 (2018). Doi: 10.1109/TKDE.2018.2876857

https://doi.org/10.5194/nhess-2024-102
Preprint. Discussion started: 17 June 2024

McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018). doi: 10.48550/arXiv.1802.03426