

Recursive Entropy Method of Segmentation for Seismic Signals

A. Bueno¹, A. Díaz-Moreno², S. De Angelis², C. Benítez¹, J.M.Ibañez³

¹Department of Signal Theory, Telematic and Communications, University of Granada, Spain.

²Department of Earth, Ocean and Ecological Sciences, University of Liverpool, UK

³Instituto Andaluz de Geofísica, University of Granada, Spain

Key Points:

- Entropy-based segmentation
- Automatic Detection
- Data visualization

Corresponding author: Angel Bueno, angelbueno@ugr.es

Abstract

A wealth of data collected over the past three decades have demonstrated that volcanic unrest is often associated with elevated levels of seismicity. Volcano seismic networks commonly record intense swarms of earthquakes in the weeks to months before eruptions; peak rates of more than one event per minute are not uncommon. The ability to readily detect and classify these signals is crucial to effective monitoring operations, and hazard assessment. The sheer volume of information collected, however, poses a challenge to volcano observatories due to the unrealistically large number of staff required for manual inspection of these data. Here, we present REMOS (Recursive Entropy Method of Segmentation), a computationally efficient Python workflow to detect, extract, and classify volcanic earthquakes starting from raw, continuous, waveform data. Within REMOS, seismograms are first analyzed using the well-established Short-Term Average/Long-Term Average method to identify trigger times of candidate earthquakes. A new algorithm based on measurements of seismic energy and minimum entropy is then employed to investigate large amounts of earthquake triggers, and to discriminate and parse events into individual waveforms for further analyses. REMOS also includes a facility for classification of the extracted waveforms, based on simple frequency-domain metrics. Finally, the results can be visualized using t-Distributed Stochastic Neighbor Embedding (t-SNE), a technique for dimensionality reduction that is particularly well-suited to inspection of high-dimensional datasets. In this work we demonstrate the use of REMOS with seismic data recorded in 2007 during a period of unrest and eruption at Bezyminyany volcano. Our results show that REMOS can efficiently detect, segment and classify earthquakes at scale, and at very low computational cost.

Introduction

Volcano and earthquake observatories routinely archive large amounts of seismic data that contain informative evidence on the likelihood of forthcoming eruptions. Arising from this plethora of observations there is a growing appreciation of how earthquake swarms reflect a variety of physical processes underneath volcanoes (e.g., *McNutt et al.* [2015]). Earthquake catalogs associated with intense periods of unrest and elevated seismicity at volcanoes, however, remain largely incomplete. The sheer volume of data recorded by volcano observatories poses a challenge due to the unrealistic amount of human resources that would be required for manual inspection in real-time. Similarly, efforts to retrospectively analyse decade-long time-series, and re-assess catalogs, are rarely undertaken. On the other hand, complete earthquake catalogs in near real-time would allow computation of parameters to form the input for, or allow validation of, advanced statistical and physical models of volcanic activity (*Bell and Kilburn* [2012]). Understanding the time history of seismicity at volcanoes can, thus, provide valuable constraints for the assessment of the likelihood and timing of eruptions, and the prognosis of their related hazards. During volcanic crisis, the timeliness of response to unrest is critical, and large volumes of continuous data must be rapidly processed and interpreted (*Tepp, G.* [2018]).

Here, we introduce REMOS (Recursive Entropy Method of Segmentation), a complete workflow for detection, segmentation, classification, and visualization of seismic data. Starting from continuous data streams and the activation times of STA/LTA (Short-Term Average/Long-Term Average (*Withers et al.* [1998])), we identify a set of energy and entropy-based criteria to parse data into individual waveforms (segmentation); segmented data are then classified according to their frequency content, using well-established frequency domain metrics (*Buurman and West* [2010]). In order to assess the quality of the resulting catalog, and to provide direct knowledge of the high-dimensional structure of the dataset, we extend the capabilities of REMOS by including an exploratory data visualization tool based on t-Distributed Stochastic Neighbor Embedding (t-SNE) and frequency attributes (*Maaten and Hinton* [2008]). REMOS is developed in Python, and leverages its rich scripting syntax and scientific libraries, including Numpy, Scikit and Obspy (*Beyreuther et al.* [2010]). Here, the applicability of REMOS is tested using data collected at Bezymianny volcano (Kamchatka, Russia) (*Thelen et al.* [2010]). In particular, we analyze data recorded during a pre-eruption sequence between September-December 2007, which includes variable rates of seismic activity, and a comprehensive range of waveform types. This example demonstrates the use of REMOS, and provides a preliminary benchmark for its scalability when applied to comparatively large datasets.

The REMOS algorithm

Segmentation of continuous seismic data streams relies on signal processing methods that exploit the internal structure of the data. From an information theory perspective, when earthquakes are generated by sources within a volcano the system is characterized by low-levels of chaos; in contrast, when volcano-seismic sources are not active seismometers record the continuous Earth's background noise, characterized by high entropy levels. Based on these principles, REMOS calculates the pseudo-energy of the seismic signal within specified time windows along the continuous waveforms in order to discriminate the boundaries between events (low entropy) and background noise (high entropy). The pseudo-energy of a seismogram, on the component being analysed, is calculated as

$$E = \langle s(n) \cdot s(n) \rangle = \sum_{n=1}^N |s(n)|^2 \quad (1)$$

where s is a vector of ground velocities, and N the number of samples in s . Note that E is proportional to seismic energy but it does not have physical dimensions of energy. In the next sections we present a description of the four main steps in REMOS: pre-processing, segmentation, classification, and visualization.

Pre-processing

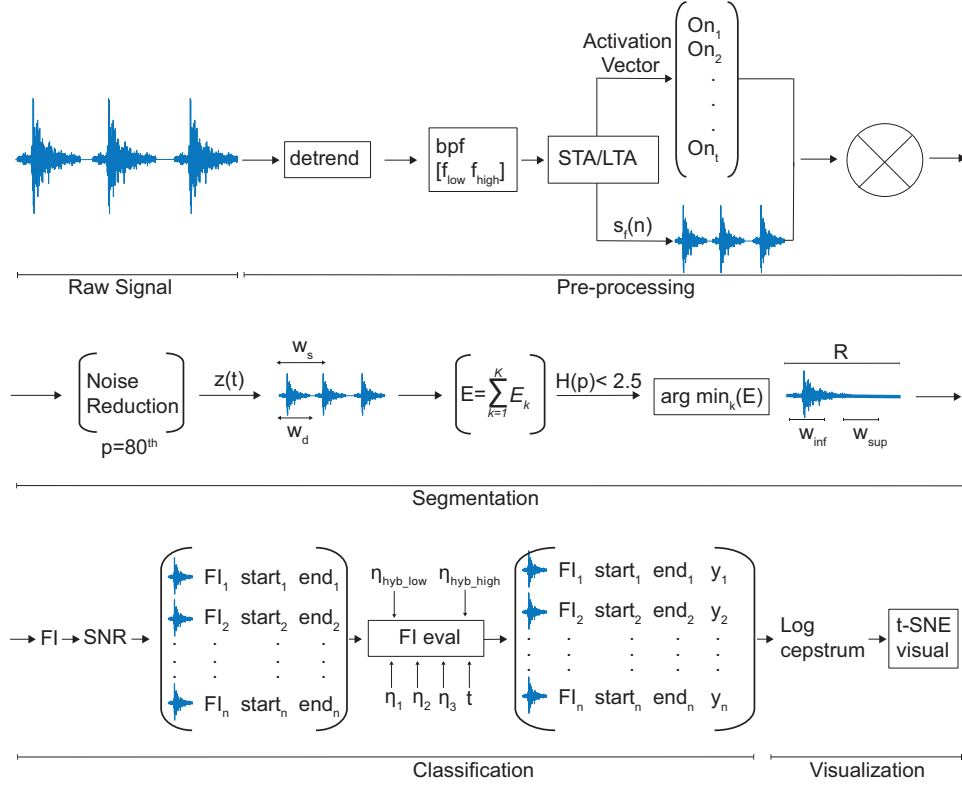
Figure 1 illustrates the workflow of REMOS. Preliminary processing steps are applied to the continuous data before segmentation. Linear trends associated with effects such as very long-period instrument drift are removed. A bandpass filter in a user-selected frequency range [f_{lo} f_{high}] ($Hz.$) is applied to enhance the presence of earthquake signals; during our tests we observed that for most volcanoes the frequency band [1-15] $Hz.$ is appropriate. The filtered data are then scanned using a recursive STA/LTA algorithm to obtain trigger times of potential events (Withers *et al.* [1998]). The parameters for the STA/LTA algorithm are selected by the user according to the specific dataset (Trnkoczy [1999]). For volcano-seismic signals recorded at relatively close distance from the source ($< 10km$), a short-term window of $0.5 - 2s$ and a long-term window of $8 - 15s$ are, frequently, appropriate; the STA/LTA threshold to declare a trigger is dependent on signal-to-noise ratio. This procedure yields a one-dimensional activation vector, $on = [on_1, on_2, on_3, \dots, on_n]$, of earthquake trigger times.

Segmentation

In the second step of REMOS the vector of activation times, on , is used in combination with the filtered signal to investigate regions within the continuous data stream that contain earthquakes, and to parse these events into individual waveforms. Data segmentation in REMOS depends on two parameters: the maximum search window, W_s , and the minimum duration window, W_d . The maximum search window W_s represents the time window that REMOS explores to detect an event; W_d captures the minimum duration that REMOS uses to calculate signal energy. A noise reduction procedure is applied to the data stream (Fig. 1) to mitigate the influence of external noise sources (e.g: electronic spikes). Data are corrected according to:

$$z(n) = \begin{cases} 0 & s_f(n) \leq T \\ \frac{s_f(n)}{T} & s_f(n) > T \end{cases} \quad (2)$$

where $s_f(n)$ is the filtered signal, and T a threshold value that corresponds to the n -th percentile of the data stream. This threshold is selected to be large enough to represent the background noise levels for the data stream. We, then, consider all activation times in on , and extract segments of data (exploration regions) with duration W_s starting at these times. Each of these regions is then windowed into k frames with duration W_d . The energy of the signal is computed according to (1) within all exploration regions, for each individual frame. The entropy, $H(p)$, is then calculated from the pseudo-probability normalized vector, p , as:



88 **Figure 1.** REMOS workflow. The algorithm is divided into four different stages: pre-processing, segmen-
 89 tation, classification and visualization. From continuous raw data streams, initial processing aims to increase
 90 signal quality and produce an activation vector (on) from application of STA/LTA. Each element in on , is consid-
 91 ered to explore a segmentation region for potential earthquakes. Candidate events are segmented according to a
 92 minimum energy criterion within these regions. When all events are segmented, frequency-based classification
 93 and t-SNE visualization are performed.

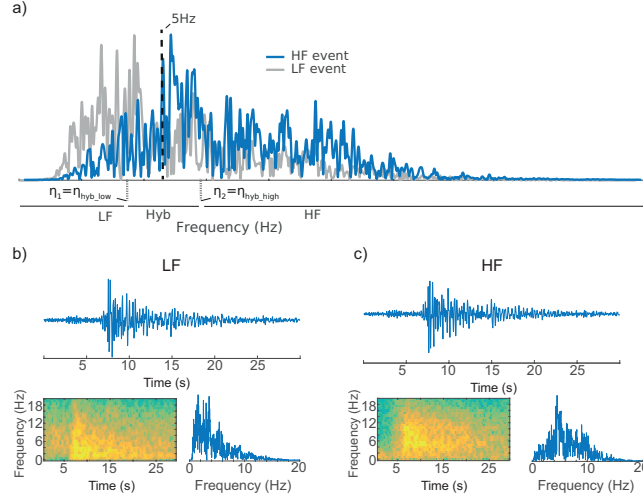
$$110 \quad H(p) = \sum_{i=1}^k -p_i \log(p_i) \quad (3)$$

111 where p_i is the normalized energy of each frame. If the entropy is low, below a pre-defined
 112 threshold, ϵ , then the onset of an event is declared. In this instance, the candidate event is extracted
 113 from the exploration region and parsed into an individual waveform according to the minimum of
 114 the energy distribution:

$$115 \quad \arg \min_k E_k \quad (4)$$

116 The end of an event is declared when the energy reaches a minimum. If the exploration region
 117 contains multiple earthquakes, using the minimum energy as a factor to discriminate the end of an
 118 event ensures correct segmentation. To overcome long segmentations when taking the lowest energy,
 119 we compute the ratio R of the energies of the initial and final parts of the segmented signal. This
 120 ratio would be close to zero if no earthquake signal was present in the final section of the segmented
 121 data. In that case, the original candidate is segmented again. Otherwise, for ratios closer to one, the
 122 original candidate is selected, as it could be for instance an episode of long-duration tremor. Once
 123 the candidate is segmented, results are stored into a matrix. REMOS continues segmenting signals

124 along the main trace until the activation vector *on* is exhausted. As a final step, candidates with low
 125 SNR can be erased from the final dataset. In REMOS, the SNR is computed as the ratio of the peak
 126 amplitude within a time window containing the surface wave signals to the root-mean-square of the
 127 noise trailing the signal arrival window.



128 **Figure 2.** Comparison of high-frequency and low-frequency spectra and threshold selection for frequency
 129 index calculation. In order to include hybrids, we select $\eta_1 = \eta_{hybrid\ low}$ and $\eta_{hybrid\ high} = \eta_2$ as frequency
 130 threshold, but other configurations are possible. Figure 2.b and 2.c show the seismograms, spectrograms and
 131 power spectrum for two segmented signals at BELO station, classified as low-frequency (2.b) and high-frequency
 132 (2.c) events.

133 Classification

134 The individual traces extracted from the continuous data are then classified based on a simple
 135 metric, the frequency index (FI) (*Buurman and West* [2010]). Automated data classification has the
 136 advantage of producing objective classifications at scale (*Titos et al.* [2018]), (*Ibáñez et al.* [2009]).
 137 The time required to analyze data is greatly reduced, and any mismatch in event classification due to
 138 the bias introduced by a human analyst are potentially eliminated. REMOS computes the FI as:

$$139 \quad FI = \log_{10} \left(\frac{A_{high}}{A_{low}} \right) \quad (5)$$

140 where A_{high} and A_{low} are spectral energies above and below a certain threshold value, respec-
 141 tively. We define a set of adjustable parameters $\eta = \{\eta_1, \eta_{hybrid\ low}, \eta_{hybrid\ high}, \eta_2, \eta_3\}$ as the FI
 142 thresholds that control how events are categorized according to their FI. Events are classified as
 143 Low-Frequency (LF) if the FI is below than a given threshold η_1 . Similarly, if the FI its greater
 144 than the given threshold η_2 , earthquakes are classified as High-frequency (HF). The events with FI
 145 between η_1 and η_2 deliver energy at both, low and high frequencies, and could be categorized as
 146 hybrids. The parameters η_1 and η_2 can be set with the same values, thus acting as a single threshold
 147 for HF and LF events.

148 Tremors and Rockfalls are discriminated according to their frequency index and duration t .
 149 Additionally, η_3 is a threshold to control when an event is declared to be a rockfall or a tremor. The
 150 minimum duration t over which to consider an event either low frequency or tremor depends of the
 151 user and the volcanic-environment (*McNutt et al.* [2015]). Figure 2.a illustrates the frequency content
 152 of a HF and LF events at Bezymianny volcano. In the case of LF events, most of the frequency

content is below 5 Hz, making the FI from equation 5 to be below zero or negative. Similarly, HF events result in FI to be greater than zero, or positive. Thus, LF's events have negative FI's, whereas HF will have positives FI's. The range of values between η_1 and η_2 could represent hybrid events.

REMOS supports the configuration of $\eta_{\text{hybrid low}}$ and $\eta_{\text{hybrid high}}$ parameters to quantify hybrid events and overcome the limitation of a single threshold for HF and LF events. Hybrid earthquakes could be labeled by selecting any intermediate values between η_1 and η_2 . However, one can set the values of $\eta_{\text{hybrid low}}$ and $\eta_{\text{hybrid high}}$ to zero to not categorize any hybrid events, or set $\eta_{\text{hybrid low}}$ and $\eta_{\text{hybrid high}}$ to any value between η_1 and η_2 in order to categorize hybrids with FI between the threshold for low and high frequency events. The value of t could be set by the analysts based on geophysical knowledge of the specific study volcano. Typically, 30s should suffice for most volcanoes. Volcanic tremors are often observed as a precursor to and during eruptions; tremor that can last from minutes to days, or even longer, and delivers energy at frequencies, typically, between 1 and 10 Hz. In the case of harmonic tremor, multiple, equally spaced, spectral peaks can be observed. Classification of these signals is beyond the scope of REMOS, which focuses on the segmentation of seismic signals at a higher level; specific detectors exist to discriminate sub-categories within the general classification scheme (D. C. Roman [2017]).

Data exploration and visualization

Exploratory analysis holds potential to provide insights about topological hierarchies in high dimensional data. Here, for the purpose of visualization of REMOS outputs, we adopt t-SNE, a dimensionality reduction method that has proved very useful to discover subtle differences in highly complex data, and to obtain meaningful information about how data points are distributed in high dimensional spaces. The algorithm t-SNE aims to find a representation (*embedding – vector*) that can minimize the distance between multidimensional data (with hundred of features) and a lower representation. This distance can be inferred using the KL-divergence as it measures the difference between two probability distributions. Therefore, the *embedding – vectors* could be found as the minimization of the *KullbackLeibler* divergence (KL-divergence hereafter) between the joint probability distribution from data points in high dimensional space, P , and the joint probability distribution Q from the embeddings in the low dimensional space:

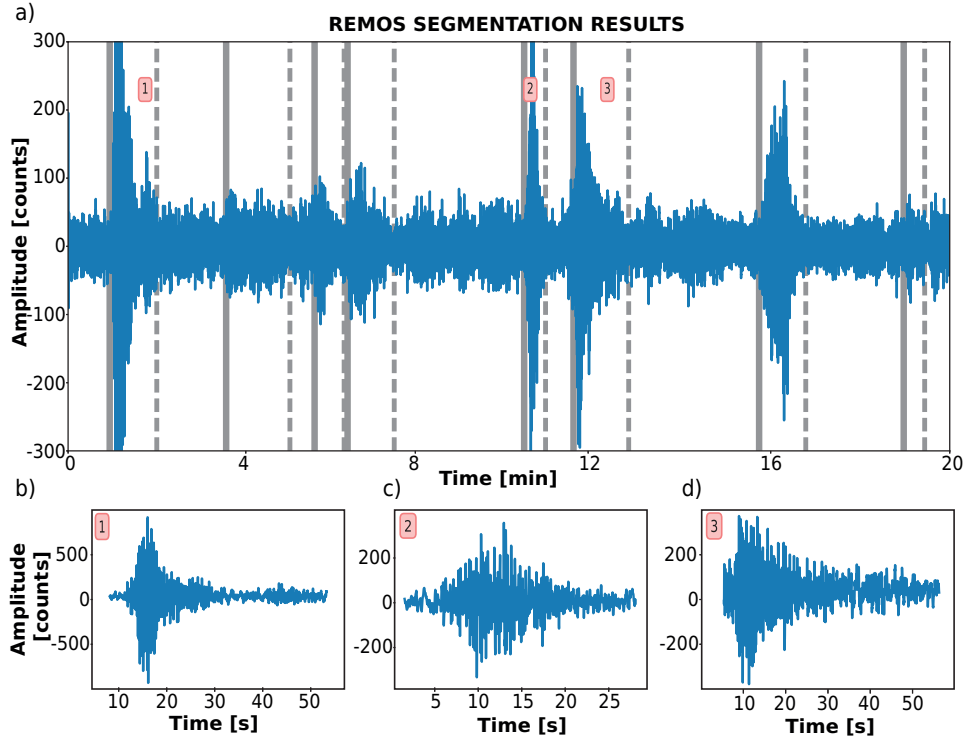
$$KL[P||Q] = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (6)$$

with P and Q the probability distributions between the high- and low-dimensional space, p_{ij} and q_{ij} the pairwise similarities across data points in the high- and low-dimensional space. This minimization process helps to preserve topology and information from complex data by projecting representative embeddings into a low dimensional space that faithfully capture hidden similarities. The reader is referred to *Maaten and Hinton* [2008] for further details on the t-SNE algorithm and visualization methodology. We use log-cepstrum frequency features (cepstral coefficient on a logarithmic scale of the given data *Beyreuther et al.* [2010]), as the input data for the t-SNE algorithm. The natural logarithms of the filter-bank energies are calculated, producing a 16-parameter feature vector. The Discrete Cosine Transform (DCT) is used to de-correlate the features and reduce the number of components from the feature to 13 coefficients. Therefore, log-cepstrum coefficients would enhance lower frequencies whilst providing highly refined information for the visualization.

Application of REMOS

In this section we demonstrate the use of REMOS with data collected at Bezymianny volcano (Russia) during 2007; this dataset contains a large number earthquakes, including high frequency (HF), low frequency (LF), rockfalls (R), and tremors (T). Unrest at Bezymianny during 2007 and the dataset used here have been previously investigated by *Thelen et al.* [2010], thus providing a reliable benchmark for the performance of REMOS and geophysical insight for the configuration of W_d , W_s and ϵ parameters. A Python script and notebook examples, along with seismic data from this

200 example, are included as supplementary material. We perform our experiments on a 64-bit computer
 201 with a i7-8700k CPU (3.70GHz), 32GB RAM and Ubuntu 16.04. The unprocessed trace is filtered
 202 between $f_{low} = 1\text{Hz}$ and $f_{high} = 12\text{Hz}$. The recursive STA/LTA algorithm is used with STA of 3
 203 seconds, LTA of 15 seconds, 2.0 as the on trigger, and 1.0 as off trigger thresholds, respectively. A
 204 vector of activation times of length 818 (the number of regions to explore) is generated. Individual
 205 events are then extracted using the segmentation method implemented in REMOS:



206 **Figure 3.** (a) Segmentation results with REMOS triggers on the continuous trace (17th October 2007) from
 207 BELO station prior to classification and visualization stage, filtered between 1 and 12 Hz. Different examples
 208 of segmented waveforms by REMOS, with (b) a signal with a small energetic coda, (c) a short event and (c)
 209 exponential decay segmented waveform.

```
210 >>> W_d = 10
211 >>> W_s = 80
212 >>> windows = [W_d, W_s]
213 >>> remos_segmentation(original, processed, on, [W_d, W_s], epsilon=2.5,
214 plot=False, cut="original", delay_in=4.5 )
```

215 The REMOS short window (W_d) is set as 10 seconds, whereas the region of exploration W_s is 80
 216 seconds. The parameter $delay_{in}$ can be configured to segment the signal before the detection time.
 217 In this work, and following *Thelen et al.* [2010], we define $delay_{in}$ as 4.5 seconds before the trigger
 218 time, given our STA/LTA is configured to detect very impulsive events. Being completely optional,
 219 the parameter $delay_{in}$ can be set to zero. If the parameter is included, the difference between the *on*
 220 time and $delay_{in}$ is included in the segmented waveform. The normalized 80th percentile from the
 221 seismic trace is used to reduce noise, and processed trace along with activation vector *on* are used
 222 for discriminative segmentation.

223 Regions of exploration with duration W_s are segmented and windowed with 10 seconds duration
 224 (W_d). The entropy $H(p)$ is computed using the short-term energy as in equation 3. If $H(p)$ is below

225 $\epsilon = 2.5$, the minimum of the energy is taken (see equation 1) within the W_s region and potential
 226 candidates are cut. The value of the ϵ threshold is selected to minimize the excessive influence of
 227 background noise, and can be adjusted by the analyst. Further, each candidate is windowed by half
 228 value of W_d , and energy ratios computed. If the parameter *cut* is set to *original*, the candidate is taken
 229 from the raw data, and a 0.5 Hz high-pass filter is applied to reduce background noise. Otherwise,
 230 the event is selected from the filtered trace. When all the events are segmented, signals with low
 231 SNR are erased from the final data matrix. Start and end times (in samples) and frequency index of
 232 each event are also stored.

233 Figure 3 shows the results for the segmentation of a 20-minute seismogram recorded on 17th
 234 October, 2007 at station BELO. All events are taken from the filtered data by setting the parameter
 235 *cut* to *filtered*. The entropy criteria, used jointly with the minimum energy search, results in very
 236 effective data segmentation. Figure 3.b illustrates how REMOS preserve the coda decay of seismic
 237 signals. Figure 3.c, for instance, shows how REMOS can detect short-duration events.

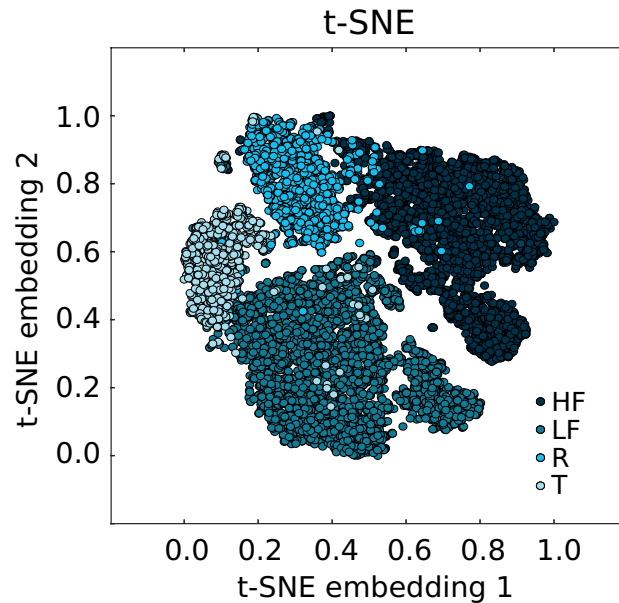
238 The next step is FI classification. Events at Bezymianny volcano are characterized by a frequency
 239 range between 1 and 10 Hz. As the parameter *cut* was set to *original*, we use the high-pass filtered
 240 signal (0.5 Hz) to perform classification. Therefore, FI is calculated by considering a low-frequency
 241 band between 1 to 5 Hz, and high frequencies between 6 and 10 Hz. The FI cutoffs η_1 (for LF),
 242 η_2 for (HF) and η_3 (for R, T) are selected as $\eta_1 = -0.2$ and $\eta_2 = 0.2$, and $\eta_3 = 0.2$. These cutoffs
 243 values were empirically defined after manual data inspection. As no hybrid events were present in
 244 the study by *Thelen et al.* [2010], $\eta_{\text{hybrid low}}$ and $\eta_{\text{hybrid high}}$ were set to zero. Figure 2 shows the
 245 classification results for two segmented signals. Notice that the ratio of spectral energy above and
 246 below 5Hz suffices to discriminate events at BELO station.

247 We note that the FI thresholds must be fine-tuned and calibrated according to the specific dataset.
 248 We refer the reader to (*Buurman and West* [2010]) for further discussion on the choice of the correct
 249 FI parametrization. In simple words, when the frequency spectrum is dominated by energy below
 250 5Hz, the event is labeled as LF (figure 2.b). When the frequency content is mostly above 5Hz, the
 251 events are labeled as HF (figure 2.c). The duration of the segmented waveform serves to discriminate
 252 rockfalls and tremor from other events. REMOS was able to detect 3259 HF, 3796 LF, 1442 R and
 253 864 T. The classification matrix is stored in Numpy format (*.npy*). The average processing time for
 254 these steps, (segmentation and classification) was 7.04 seconds.

256 In the final step of the REMOS workflow we combine t-SNE and log-cepstrum features to
 257 assess the quality of the labeled dataset. The non-linear nature of t-SNE reduces the dimensionality
 258 of the log-cepstrum features whilst providing a good understanding of global and local data structures
 259 (*Maaten and Hinton* [2008]). Log-cepstrum features enhance frequencies below 20 Hz, especially
 260 at the lower end of the spectrum (*Ibáñez et al.* [2009]). A total of 13 log-cepstrum coefficients are
 261 computed using the *Obspy* signal processing package. A normalized t-SNE visualization map is
 262 shown in figure 4. At BELO station, t-SNE reveals the presence of four distinct clusters. Notice that
 263 few tremors (clear blue) fall within the low-frequency cluster (turquoise). Similarly, some rockfalls
 264 (blue) appear within the high frequency region (darkest blue). This map allows rapid visualization
 265 of what type of events are present in a database, and serves as an exploratory data analysis that could
 266 be used to refine posterior statistical analysis. Thus, given a specific classification criteria, REMOS
 267 produces an objective data-driven classification consistent with time and different observed features.

268 Conclusions

269 We have introduced REMOS, a Python module for the detection, segmentation, classification
 270 and visualization of continuous volcano-seismic data streams. REMOS makes extensive use of
 271 optimized scientific libraries, including Obspy, SciPy and NumPy. REMOS has a highly modular
 272 structure that can be easily adapted to different case studies. The results of our tests suggest that
 273 REMOS is capable of consistently segmenting and classifying large amounts of volcano-seismic
 274 data, with minimum supervision. REMOS provides a computationally cost-effective solution to
 275 enhance volcano monitoring systems, allowing rapid assessment of volcano-seismic unrest. Here,



255

Figure 4. t-SNE visualization on BELO station

276 we have demonstrated the use of REMOS with data gathered during a period of unrest at Bezmyianny
 277 volcano in 2007. The robustness and scalability of REMOS were tested against previous analyses
 278 of the same dataset. Our results show that REMOS can effectively detect, segment and classify
 279 volcano-seismic events, even when the signal-to-noise ratio is poor or event occurrence rates are
 280 high. The methodology described here is easily transferable to other domains, such as infrasound or
 281 hydroacoustic data. We suggest that real-time implementation of algorithms like REMOS is crucial
 282 to improving our ability to monitor unrest at active volcanoes.

283

Data and Resources

284 The data used in this study can be obtained from IRIS DMC. The facilities of IRIS Data
 285 Services, and specifically the IRIS Data Management Center, were used to access the waveforms,
 286 related metadata, and/or derived products used in this study. IRIS Data Services are funded through
 287 the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal
 288 of the National Science Foundation under Cooperative Agreement EAR-1261681.

289 The REMOS implementation, along with example data, extensive documentation and conda
 290 environment can be found at <https://github.com/srsudo/remos>. Scientific libraries, including NumPy,
 291 SciPy and ObsPy are accessible via their web-pages: <https://www.numpy.org>, <https://www.scipy.org>
 292 and <https://www.obspy.org>. Scikits.talkbox can be accessed at <https://github.com/cournape/talkbox>,
 293 and installed via Python package manager, *pip*. Last access on all websites on August 2018.

Acknowledgments

294 We would like to thank all the members at the Seismic Laboratory at University of Liverpool for
 295 continuous support and advice when developing REMOS algorithm. We would like to thank Weston
 296 Thelen and an anonymous reviewer for their detailed and constructive comments, which were very
 297 helpful in improving the original article. This research was funded by the "Spanish Ministerio
 298 de Economiay Competitividad" grant TEC2015-68752 (MINECO/FEDER) and by the UK Natural
 299 Environment Research Council (NERC) Grant NE/P00105X/1.
 300

References

- 301
- 302 Bell, A. F., and C. R. J. Kilburn (2012), Precursors to dyke-fed eruptions at basaltic volcanoes:
 303 insights from patterns of volcano-tectonic seismicity at Kilauea volcano, Hawaii, *Bulletin of*
 304 *Volcanology*, 74(2), 325–339, doi:10.1007/s00445-011-0519-3.
- 305 Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010), Obspy: A
 306 python toolbox for seismology, *Seismological Research Letters*, 81 (3), 530–533.
- 307 Buurman, H., and M. West (2010), Seismic precursors to volcanic explosions during the 2006
 308 eruption of Augustine volcano, 1769.
- 309 Roman, D. C. (2017), Automated detection and characterisation of harmonic tremor in continuous
 310 seismic data: Harmonic tremor detection, 44.
- 311 Ibáñez, J.M., C. Benítez, L. Gutiérrez, G. Cortés, A. García-Yeguas, and G. Alguacil (2009), The
 312 classification of seismo-volcanic signals using hidden Markov models as applied to the stromboli
 313 and etna volcanoes, *Journal of Volcanology and Geothermal Research*, 187, 218–226.
- 314 Maaten, L. v. d., and G. Hinton (2008), Visualizing data using t-sne, *Journal of machine learning*
 315 *research*, 9 (Nov), 2579–2605.
- 316 McNutt, S. R., G. Thompson, J. Johnson, S. D. Angelis, and D. Fee (2015), Chapter 63 - seis-
 317 mic and infrasonic monitoring, in *The Encyclopedia of Volcanoes (Second Edition)*, edited
 318 by H. Sigurdsson, second edition ed., pp. 1071 – 1099, Academic Press, Amsterdam, doi:
 319 <https://doi.org/10.1016/B978-0-12-385938-9.00063-8>.
- 320 Thelen, W., M. West, and S. Senyukov (2010), Seismic characterization of the fall 2007 eruptive
 321 sequence at Bezymianny volcano, Russia, *Journal of Volcanology and Geothermal Research*, 194
 322 (4), 201–213.
- 323 Tepp, G. (2018) A Repeating Event Sequence Alarm for Monitoring Volcanoes. *Seismological*
 324 *Research Letters*, 89 (5), 1863–1876. doi:<https://doi.org/10.1785/0220170263>.
- 325 Titos, M., A. Bueno, L. García, and C. Benítez, (2018), A deep neural network approach to automatic
 326 recognition systems for volcano-seismic events, *IEEE Journal of Selected Topics in Applied Earth*
 327 *Observations and Remote Sensing*, 11 (5), 1533–1544, doi:10.1109/JSTARS.2018.2803198.
- 328 Trnkoczy, A. (1999), Topic understanding and parameter setting of sta/lta trigger algorithm, *New*
 329 *Manual of Seismological Observatory Practice*, 2.
- 330 Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998), A
 331 comparison of select trigger algorithms for automated global seismic phase and event detection,
 332 88, 95–106.