Zali, Z., Ohrnberger, M., Scherbaum, F., Cotton, F., Eibl, E. P. S. (2021): Volcanic Tremor Extraction and Earthquake Detection Using Music Information Retrieval Algorithms. - Seismological Research Letters, 92, 6, 3668-3681.

https://doi.org/10.1785/0220210016

Volcanic Tremor Extraction and Earthquake Detection using Music Information Retrieval Algorithms

Zahra Zali[1,2], Matthias Ohrnberger[1], Frank Scherbaum[1], Fabrice Cotton[2,1], and Eva P. S. Eibl[1]

[1]University of Potsdam, Institute of Geosciences, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

[2]GFZ German Research Centre for Geosciences, Potsdam, Germany

Zahra Zali (zali@uni-potsdam.de). Campus Golm, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Declaration of Competing Interests: The authors acknowledge there are no conflicts of interest recorded.

Abstract

Volcanic tremor signals are usually observed before or during volcanic eruptions and must be monitored to evaluate the volcanic activity. A challenge in studying seismic signals of volcanic origin is the coexistence of transient signal swarms and long-lasting volcanic tremor signals. Separating transient events from volcanic tremors can therefore contribute to improving upon our understanding of the underlying physical processes. Exploiting the idea of harmonic-percussive separation in musical signal processing, we develop a method to extract the harmonic volcanic tremor signals and to detect transient events from seismic recordings. Based on the similarity properties of spectrogram frames in the time-frequency domain, we decompose the signal into two separate spectrograms representing repeating (harmonic) and non-repeating (transient) patterns, which correspond to volcanic tremor signals and earthquake signals, respectively.

We reconstruct the harmonic tremor signal in the time domain from the complex spectrogram of the repeating pattern by only considering the phase components for the frequency range where the tremor amplitude spectrum is significantly contributing to the energy of the signal. The reconstructed signal is, therefore, clean tremor signal without transient events.

Furthermore, we derive a characteristic function suitable for the detection of transient events (e.g., earthquakes) by integrating amplitudes of the non-repeating spectrogram over frequency at each time frame. Considering transient events like earthquakes, 78% of the events are detected for Signal to Noise Ratio (SNR) = 0.1 in our semi-synthetic tests. In addition, we compared the number of detected earthquakes using our method for one month of continuous data recorded during the Holuhraun 2014-2015 eruption in Iceland with the bulletin presented in Ágústsdóttir et al. (2019). Our single station event detection algorithm identified 84% of the bulletin events. Moreover, we detected a total of 12619 events, which is more than twice the number of the bulletin events.

Plain Language Summary

Volcanic tremors are important signals in volcano seismology because they usually precede or accompany volcanic eruptions, and might be considered as a forecasting signal for them. While there are unsolved questions about the origin of these signals, they are usually recorded along with many earthquake signals during periods of unrest. This makes the study of volcanic tremors more complicated. A reliable signal processing scheme is therefore required to extract volcanic tremor signals from seismological records. Inspired by the algorithms for separating harmonic and percussive components in musical signal processing, we have developed a method to separate volcanic tremor signals from earthquake signals within seismic waveforms. As by-product, we have obtained a new approach for transient signal detection (e.g., earthquakes) that allows for the detection of smaller seismic events.

1 Introduction

Volcanic tremors are long-lasting low-frequency seismic signals that frequently precede or accompany volcanic eruptions (McNutt, 1992). They can reveal information about eruptive activities (Alparone et al., 2003; Eibl et al., 2017a, 2017b) and are one of the most commonly studied signals in volcano seismology (Falsaperla et al., 2005) for use in eruption forecasting as well as investigating the physics of the underlying volcanic processes (Chouet, 1996; Yukutake et al., 2017).

Despite different hypothesis about the generation mechanisms of volcanic tremors, the details are not yet well understood (Eibl et al., 2017b; Davi et al., 2012) and a variety of physical processes may explain the seismological evidence observed so far (Hellweg, 1999). Volcanic tremor signals are usually seen in the seismic records alongside many tectonic earthquakes or other transient signals occurring during a period of volcanic unrest (Dmitrieva et al., 2013; Eibl et al., 2017a; Hotovec et al., 2013), affecting the observability of the tremor signal. Both volcanic tremors and earthquakes may help to better understand the underlying physical processes of volcanic eruptions, however, the superposition of signals makes it challenging to study the details of each signal separately. A reliable signal processing operation is thus required to separate earthquakes as well as other transient signals from the volcanic tremor signals in the recorded seismic waveforms during periods of volcanic unrest. There have been attempts in terms of the detection and discrimination of volcanic tremor and tectonic earthquake signals in previous studies. For example, an automatic P-and S-wave detection was used in Rouland et al. (2009) in order to identify volcanic tremors as events containing only P-type wave, and tectonic earthquakes, containing both P- and S-waves. However, this study introduces for the first time the topic of extracting tremor signals from seismic waveforms and reconstructing the volcanic tremor signal with related phase information.

Inspired by similarities of seismic and acoustic signals, we take advantage of the expertise developed in the field of Music Information Retrieval (MIR) and audio signal processing. A seismic waveform is the record of Earth vibrations, which, in terms of signal properties and generation mechanism, can be seen to be similar to sound signals generated by musical instruments (including the human voice) (Johnson & Watson, 2019; Schlindwein et al. 1995). Exploiting the extensive research results in MIR (e.g., Müller, 2015), we have developed a seismological data processing scheme for the purpose of separating volcanic tremor signals from transient signals generated during a volcanic crisis.

The separation of harmonic and percussive components of sound is of great interest in musical signal processing (e.g., Rafii & Pardo, 2011). Pop music, for example, often consists of a repetitive percussive background and a vocal foreground, which is locally non-repetitive (FitzGerald 2012). In this type of music, the different characteristics of harmonic and percussive sounds in the spectrogram domain (see Müller, 2015) allow a separation of foreground vocals from the more percussive background sound (FitzGerald & Gainza, 2010). Similarly, a seismic waveform during an eruption may consist of (harmonic) volcanic tremor signals over which transient seismic signals are superimposed. The long-duration volcanic tremor signal that lasts minutes to days with a restricted frequency range (1-9 Hz according to McNutt, 1992) contrasts with transient seismic signals such as earthquakes with a wider range of frequencies (0.1-30 Hz in this study). In particular, harmonic volcanic tremor signals with distinct spectral lines are readily distinguishable from transient, short-duration (seconds) seismic events in the time-frequency domain. In musical signal processing, the goal of harmonic-percussive source separation (HPSS) is to decompose an input signal into the sum of two signals, one consisting of all harmonic components and the other of all percussive components (Müller, 2015). The same algorithms could be implemented in the seismology domain to decompose a seismic signal into its harmonic components (harmonic volcanic tremors) and percussive components (transient events such as earthquakes). In musical signal processing, several methods for Harmonic-Percussive Separation (HPS) have been suggested (Müller, 2015).

Here, in the first step of our method, we adopted the repetition/similarity (REPET-SIM) method (Rafii & Pardo 2012; Rafii et al., 2014) to separate volcanic tremors from transient earthquakes. The advantage of this method is its ability to process music pieces with quickly-varying repeating structures without the need to identify periods of the repeating structure beforehand. The approach evaluates the underlying repeating structure by looking for the similarities in the spectrogram time frames. This repeating part of the signal is then subtracted from the original spectrogram. The remaining time frames contain the percussive events. We use this approach and apply it to seismic data collected from a volcano. In this setting, repeating structures, which result in a harmonic spectrum, correspond to volcanic tremors and percussive (non-repeating and impulsive) elements correspond to transient events like earthquakes. Another method similar to REPET-SIM for HPS was proposed by FitzGerald (2010), which we use in the second step of our method in order to remove remaining percussive components in the repeating spectrogram and vice versa.

The remainder of this paper is organized as follows. In section 2, we describe existing methods in MIR for our problem (section 2.1) and explain how we developed our method based on these algorithms. Modifications to and the application of the REPET-SIM method (Rafii & Pardo 2012; Rafii et al., 2014) and the HPS using median filtering (FitzGerald ,2010) for extracting seismic tremor signals are outlined in section 2.2, while section 2.3 describes the detection and timing of the remaining transient events (e.g., earthquakes). Section 2.4 outlines the selection of the method's parameters. Section 3 presents the generation of semi-synthetic data (3.1), an evaluation of the proposed method based on a semi-synthetic test on tremor extraction (3.2) and earthquake detection (3.3), as well as real data tests (3.4). The feasibility of the method with respect to processing speed is discussed in section 3.5. In section 4 we discuss the results and provide our conclusions about the applicability of the method.

2 Method

2.1 Harmonic-percussive separation algorithms

Harmonic-Percussive Separation (HPS) as an application of musical source separation (Cano et al., 2018) has attracted significant attention in MIR research in recent years (Rafii et al., 2018). HPS algorithms are based on the different characteristics of harmonic and percussive components in a music signal.

Harmonicity expresses the situation in which the complete signal can be seen as the superposition of spectral components (partials) whose frequencies are all integer multiples of a fundamental frequency. Harmonics form stable horizontal ridges in a Short Time Fourier Transform (STFT) spectrogram, which means constant frequencies exist along the time axis. A percussive (impulsive) sound is short and similar to the sound of hitting a drum. Percussive signals form vertical ridges in a STFT spectrogram, corresponding to the existence of different frequencies in an instant, i.e., a broadband characteristic of short duration.

In order to separate harmonic and percussive elements, one simple approach is to apply a median filter to the STFT spectrogram of the signal (FitzGerald, 2010). Median filters are usually used to remove noisy parts of a signal by replacing each sample by the median value determined from the neighboring samples within a specific window. Within the HPS, a median filter applied along the horizontal axis of the spectrogram (time) suppresses ‘short-lived’ broadband percussive components interrupting the long-lasting horizontal narrowband ridges. This results in a ‘de-noised’ harmonic spectrogram. Similarly, applying a median filter along the vertical axis of a spectrogram (frequency) emphasizes short-lived broadband features while suppressing long-lasting narrowband horizontal frequency lines (harmonic components) and results in a ‘denoised’ percussive spectrogram. These two median filters are used separately in order to generate the related spectrograms with dominant harmonic or percussive content, respectively.

Another promising approach for our purpose is REPET-SIM, which treats repetition as a basic property in generating and perceiving structure in music (Rafii & Pardo 2012; Rafii et al., 2014). The main step in this method is to identify similar patterns using a calculated similarity matrix. Given a music signal, first its complex STFT is calculated, which is named **X** here. Considering **V** as the amplitude spectrogram **V** = |**X**|, the similarity matrix **S** is calculated to measure the cosine similarity (the similarity between two vectors of an inner product space) between time frames of the spectrogram **V**. As shown in equation (1), the cosine similarity is calculated through the multiplication of the transposed **V** by **V** with normalization of the **V** time frames.

$$S(j_a,j_b) = \frac{\sum_{i=1}^{n} V(i,j_a)V(i,j_b)}{\sqrt{\sum_{i=1}^{n} V(i,j_a)^2}\sqrt{\sum_{i=1}^{n} V(i,j_b)^2}}. \qquad (1)$$

where $\forall j_a, j_b \in [1,m]$, where m is the number of time frames and n equaling $N/2+1$ is the number of frequency channels for each time frame of length N (samples). $\mathbf{S}(j_a, j_b)$ is then the cosine similarity between the time frames j_a and j_b of the spectrogram **V**.

For all the frames j in **V**, similar frames are identified using a threshold in the similarity matrix and stored in an array **J**. A repeating spectrogram model (**W**) is then derived using the similar frames. For all the frames j, the corresponding frame in **W** is derived by taking the median of **J** for each frequency. Repeating time-frequency bins are captured by the median and build the repeating spectrogram model **W**. A refined repeating spectrogram model W' is created by taking the minimum between **W** and **V**. The rationale is that the non-negative spectrogram **V** is the sum of two non-negative spectrograms of repeating and non-repeating patterns, hence, **W** is less than or at most equal to **V**.

In the following, a time-frequency mask **M** is derived by normalizing W' by **V**. Time-frequency bins with repeating patterns will have values close to 1 in **M** and time-frequency bins without repeating patterns will have values close to 0. The mask **M** is applied to STFT **X** and the repeating spectrogram will be created. Finally, the harmonic signal in music is obtained by inverting the repeating spectrogram into the time domain. The percussive signal is obtained by subtracting the harmonic signal from the input signal (Rafii & Pardo 2012).

2.2 Volcanic tremor extraction approach

Among the different tremor observations in volcanic seismology, the so-called harmonic tremor is a special signal showing a band-limited harmonic spectrum. It has been observed at many volcanoes and has been reported often during times of increased volcanic activity, and is thought to be connected to fluid flow or (de-) pressurization of the volcanic system (e.g., Montegrossi et al., 2019). This is the motivation for using HPS algorithms in order to separate harmonic volcanic tremor signals from earthquake signals representing the percussive event type. Being able to extract this special kind of tremor signal from seismic waveforms provides the opportunity to improve the observations and analyses of harmonic tremors. In particular, extracting low-amplitude harmonic tremor signals that are hidden in the background seismic noise or overprinted by earthquake sequences accompanying volcanic activity may allow new insights into the generation processes and their relationships to volcanic eruptive activity.
In this study we analyze the seismic waveforms of the Holuhraun 2014-2015 eruption in Iceland (FLUR station from network 7Z (White, R. 2010)) to separate the harmonic and percussive components. Figure 1 shows the eruption site and the station location in Iceland with an example of one day of seismic waveforms (Figure1a & b), the PSD (power spectral density), and the spectrogram (Figure1c & d). The PSD and spectrogram of the extracted harmonic components are shown in Figure 1e & f.

Our method is derived from a combination of the REPET-SIM method (Rafii & Pardo 2012; Rafii et al., 2014) and the HPS algorithm given by FitzGerald (2010) after tuning parameters to adapt it to seismic data. For building our method, we used Librosa, a Python package for audio and music signal processing (McFee et al., 2020). Furthermore, we implement a phase reconstruction procedure for the volcanic tremor signal. A detection algorithm for earthquakes as transient signals has been derived as a by-product of the applied processing.

The REPET-SIM, as described in section 2.1, is used to create a similarity matrix and to derive a time-frequency model of repeating patterns. We derive the non-repeating spectrogram model by subtracting W' from V. Once the model spectrograms are calculated, they are used to derive two time-frequency masks for repeating and non-repeating patterns.

We modified the REPET-SIM algorithm by using a soft mask via Wiener filtering (Vaseghi, 1998) instead of a binary mask. The calculation of the soft mask **M1** and **M2** are shown below as equations (2) and (3):

$$M1_{n.m} = \frac{{W'_{n,m}}^{P}}{{W'_{n,m}}^{P} + \left(V_{n,m} - W'_{n,m}\right)^{P}} \tag{2}$$

$$M2_{n.m} = \frac{\left(V_{n,m} - W'_{n,m}\right)^{P}}{\left(V_{n,m} - W'_{n,m}\right)^{P} + {W'_{n,m}}^{P}} \tag{3}$$

where **M1** is a repeating mask and **M2** is a non-repeating mask. A power factor *P* is applied to the model spectrograms to further enhance the signal-to-noise ratio. We use a power factor of 2 in our calculations.

Once we have constructed the masks, we multiply them with the input amplitude spectrograms to separate the components. Equation (4) shows the element-wise multiplication of the repeating mask (**M1**) and the input amplitude spectrogram (**V**).

$$R = M1 \otimes V, \qquad (4)$$

where **R** denotes the repeating amplitude spectrogram. The same element-wise multiplication operation is applied for the non-repeating mask and the input amplitude spectrogram as it is shown is equation (5):

$$NR = M2 \otimes V, \qquad (5)$$

where **NR** denotes the non-repeating amplitude spectrogram.

From this we obtain two spectrograms, one for repeating patterns and one for non-repeating patterns. The harmonic and percussive components of the signals are separated into their respective masked spectrograms, although small traces of percussive components are still visible in the repeating spectrogram, and remnants of the harmonic components can be recognized in the non-repeating spectrogram. Therefore, a second HPS approach is subsequently applied to the resulting spectrograms from the first processing step by using the median filtering method of FitzGerald (2010). In particular, we use median filtering along the time axis, enhancing the harmonic components within the spectrogram. Applying another median filtering along the frequency axis results in a denoised spectrogram of the percussive components. Following the above notation, each spectrogram of **R** and **NR** will be decomposed into two spectrograms of their harmonic and percussive components. Equation (6) and (7) show this separation:

$$R = H1 + P1, \quad (6)$$

$$NR = H2 + P2, \quad (7)$$

where **H1** and **P1** are harmonic and percussive components of the repeating spectrograms, and **H2** and **P2** are harmonic and percussive components of the non-repeating spectrograms. We create a soft mask using **H1** and multiply it the **R** spectrogram, which results in the final harmonic spectrogram, which we name **HARM**. Another soft mask is created using **P2** and is multiplied by the **NR** spectrogram to derive the final transient spectrogram that we have named **TRAN** (see Figure 2).

Figure 2 shows the flow-chart of the method with an example of a seismic waveform from 3 September 2014 during the Holuhraun 2014-2015 eruption in Iceland (FLUR station from network 7Z (White, R. 2010)). On this day we were 4 days into a 6-month long fissure eruption accompanied by tremors and long-period and volcano-tectonic earthquakes (Eibl et al. 2017a). For further details on the background of the Holuhraun eruption event, the reader is referred to Sigmundsson et al. (2015) and Gudmundsson et al. (2016). For details on the events on 3 September 2014, the reader is referred to Eibl et al. (2017a) and Woods et al. (2018).

Besides describing the processing steps (Figure 2a), we show an input waveform and its spectrogram, which is decomposed in two steps (Figure 2b). In the first step using the modified REPET-SIM algorithm, we decompose the X spectrogram into a ‘repeating’ spectrogram (R spectrogram) and a ‘non-repeating’ spectrogram (NR spectrogram). Each of these two spectrograms are then decomposed into their harmonic and percussive components in the subsequent step, following the algorithm of FitzGerald (2010). The harmonic component of the repeating spectrogram shows the final result for the harmonic spectrogram (HARM spectrogram) and the percussive component of the non-repeating spectrogram shows the final result for the transient or percussive spectrogram (TRAN spectrogram). The HARM spectrogram corresponds to the tremor spectrogram according to our assumptions of the generating process. From the tremor spectrogram in the frequency domain, the tremor signal can be reconstructed in the time domain. The problem of reconstructing a signal from its modified STFT has varieties of applications in audio signal processing, where modifications are applied to the amplitude STFT and the phase information is lost (Sturmel & Daudet, 2011). The standard phase reconstruction Griffin-Lim algorithm (Griffin & Lim., 1984) which is based on random phase initialization followed by the minimization of the squared error between the STFT of the estimated signal and the modified STFT, shows poor performance for our seismological test signals. The random initialization of phase is an inadequate starting model for the inversion procedure and results in an unreliable signal estimate. The problem of this inadequate signal reconstruction is illustrated by an example (Figure 3d) and is described at the end of this section.

We must use phase information of the original STFT **X** in order to reconstruct the signal in the time domain. Considering the notation in section 2.1, we calculate the similarity matrix based on **V** as the amplitude spectrogram. Therefore, we need to separate the complex-valued spectrogram **X** into its amplitude (**V**) and phase components using equation (8).

$$X = V * exp(1j * \varphi), \tag{8}$$

where φ denotes the phase of **X** and j is the imaginary unit. The procedure of using the initial phase matrix is more problematic than it might seem at first glance. Simply using the phase information of **X** can lead to a noisy reconstructed signal due to the noise contributions in the phase matrix of the seismic waveform. Therefore, we use the values of the phase matrix only in the dominant frequency band of the **HARM** spectrogram. We do so by integrating the **HARM** spectrum amplitude squared for all time frames and determine the starting frequency as the 5% quantile of the total energy in the spectrum and the stop frequency as the 95% quantile, respectively. The dominant frequency band is between the start and stop frequencies. Then, we add this modified phase information (weighted phase information) named φ_t to the **HARM** spectrogram using equation (9).

$$T = HARM * exp(1j * \varphi_t), \tag{9}$$

where **T** is the complex tremor spectrogram and **HARM** is the harmonic amplitude spectrogram. Finally, we reconstruct the tremor signal time series from the complex spectrogram **T**, using the inverse short-time Fourier transform. The inversion process is done using the Griffin-Lim algorithm (Griffin & Lim., 1984) for converting a complex-valued spectrogram to a time-series by minimizing the mean squared error between the complex STFT of the estimated signal and the modified STFT **T**. Note that using a part of the phase information sees the Griffin-Lim algorithm converging to a reasonable time domain signal, whereas it won't if starting with randomly selected phases.

Figure 3 shows the seismic signal (Figure 3a blue) and a comparison of the reconstructed volcanic tremor signal for one minute of seismic waveform from 3 September 2014 using our approach (Figure 3a green and 3b) and two other methods (Figure 3c and d) described below. As shown in Figure 3 (b) the reconstructed tremor signal using our method is not noisy and shows almost no trace of transient signals. Figure 3 (c) shows the reconstructed signal using the inverse short-time Fourier transform, after applying horizontal median filtering (FitzGerald 2010) on the STFT spectrogram with the aim of separating and extracting the harmonic tremor signal. In this case, the tremor signal is reconstructed by adding the phase of the original seismic waveform to the modified STFT. Transient signal energy still exists in the reconstructed harmonic signal, which demonstrates that horizontal median filtering is not sufficient for extracting a clean tremor signal without signs of transient events.

In Figure 3 (d), we show the estimated tremor signal using the original Griffin-Lim algorithm for phase reconstruction. The effect of earthquake signals is almost eliminated, as in Figure 3 (b), which is reasonable as both Figure 3 (b) and (d) are extracted from the **HARM** spectrogram. However, a significant difference compared to the seismic signal is visible in Figure 3 (d) in terms of the shape of the signal. Also, the phase is not reconstructed correctly. Therefore, this signal (Figure 3d) is not applicable for seismological purposes. This shows the importance of using appropriate phase information for reconstructing a seismic signal in the time domain. We note that a pre-filtering of the original seismic data is necessary to remove microseismic signals before applying our algorithm. Indeed, microseisms are harmonic signals, which may have a dominant energy in the tremor spectrogram. Therefore, the amplitude and the phase of the reconstructed tremor signal could be significantly affected by such microseism signals if they are not filtered out beforehand. We applied a high-pass filter with a cut-off frequency of 0.5 Hz on our real dataset.

2.3 Transient signal detection and timing estimation

In a second step, we use the transient spectrogram to locate the occurrence of transient signals in time. We do so by integrating the spectral amplitudes over the full frequency band at each time frame, thus deriving a characteristic function suitable for detecting transient events. At the time of transient events, this function has large values compared to zero or very small values in other parts of the function.

Most observed transient signals in the seismic recordings can be interpreted as seismic wave arrivals of earthquakes. A standard task in observational seismology is then to estimate arrival times of wave groups from timing the onset of transient signals. Proposing the detection of transient signals with the characteristic function described above, we further aim to extract an accurate onset time of the transient signals. For detection, we use a local maximum (peak) search on the transient characteristic function. Two thresholds are applied to the characteristic function; the upper threshold and the lower threshold. The upper threshold is used for transient signal detection and the lower threshold is used for accurate onset timing. The upper threshold prevents picking up minor local maxima representing coda waves or other fluctuations in the earthquake records. This threshold is determined by visually analyzing the peak value distribution on some smaller test set in the data. The local maxima, which are larger than the threshold, are then considered to represent detected earthquakes. The maximum peak of the characteristic function corresponds mostly to S-wave arrivals, while the P-wave onset can be associated with the earliest break in the characteristic function. We have therefore developed a straightforward procedure to find the first arrival onset of the transient events by considering amplitude and amplitude derivatives of the characteristic function for the pre-peak interval time window from the largest local maximum found in the characteristic function (Figure 4).

We used a 5 second pre-peak interval time window because most of the earthquakes in this study are local and t_S-t_P difference times are less than 5 seconds. This time window is shown in Figure 4. It is recommended to use a larger pre-peak interval time window for regional earthquakes. We shorten this pre-peak interval time window preceding each peak using the following criterion. First, we adjust the lower threshold visually to the level of residual signal energy from the harmonic signal component remaining after the separation process. The lower threshold is the smallest non-zero number in the characteristic function, which does not correspond to the transient signals. This allows the removal of minimal amount of residual energy due to the separation process. Using the lower threshold improves the accuracy of onset time picking. We set all values of the characteristic function below the lower threshold to zero. Second, we check if there are some neighboring zero samples in the time window and change the starting point of the window to one sample after the last zero sample in order to prevent mixing with a very close preceding event. Indeed, neighboring zero samples means that there is no transient signal and shortening the window avoids confusion with a close preceding event. We skip the samples following a local maximum within the window if there are any. Then, we calculate the slope between each two neighboring samples and we skip the samples following a slope reduction if there are any. Finally, the starting point of the transient signal (P-wave arrival) is the point showing the maximum slope increase (see Figure 4).

2.4 Parameters selection

Although the separation process creates a harmonic and percussive spectrogram, the process must be repeated twice with different FFT window lengths if both tremor signals and the timing of the transient events are to be determined. Due to the uncertainty principle in Fourier analysis, it is impossible to increase both the temporal resolution and the frequency resolution. A better frequency resolution requires a longer time window for the spectral analysis (longer FFT length), which implies a reduced temporal resolution. Similarly, using a shorter FFT window increases the temporal resolution, while the frequency resolution will be reduced. For extracting the tremor signal, we need a high resolution in the frequency domain and therefore a large number of FFT points is chosen. We use a FFT window length of 81.92 seconds with an overlap of 75%, corresponding to an FFT size of 8192 at a sampling frequency of 100 Hz. To detect transient events, a high resolution in the time domain is needed and a small number of FFT points and short hop size (number of samples between each successive FFT window) are chosen. We use a FFT size and FFT window length of 1.28 seconds, with an overlap of 75%. Considering the data's 100 Hz sampling frequency, neighboring FFT windows are spaced in time by an interval of 0.32 seconds (3.125 samples per second). Fourier transforms with a narrower FFT size are not recommended for our algorithm due to the resulting limited frequency resolution.

There are two sets of median filter procedures used in our method. The first one, which is described in section 2.1, is part of the REPET-SIM algorithm and is depicted in the flow-chart of Figure 2(a). After identifying the similar frames and storing them in the array $\mathbf{J}$, the median of $\mathbf{J}$ is taken for each frequency in order to construct $\mathbf{W}$.

The second median filter procedure is described in the section 2.2 where a second harmonic-percussive separation approach is applied by using the median filtering method of FitzGerald (2010). Both a horizontal and a vertical median filter are applied separately to the spectrograms of **R** and **NR** (see the flowchart in Figure 2(a)). We use a standard kernel size of 31 for both the horizontal and vertical median filters, as it has been shown by Driedger et al. (2014) that the choice of this parameter is not critical if not choosing extreme values. Both **R** and **NR** are decomposed into two spectrograms, i.e., containing harmonic and percussive signal components . The harmonic component of the **R** spectrogram is the final harmonic spectrogram (**HARM,** see figure 2b). The percussive component of the spectrogram is the final spectrogram of the transient components (**TRAN,** see Figure 2b).

3 Data sets and testing

3.1 Generation of semi-synthetic data

We created a synthetic harmonic signal, convolving equally-spaced spikes with a real-valued Morlet wavelet (Figure S1 a). In this way, we can model the basic features of a harmonic spectra (Schlindwein et al. 1995). Instead of using exact constant repetition intervals and a fixed amplitude, which produces a perfect harmonic tremor signal, we varied the interval times as well as the amplitude of the spikes according to a normally distributed random variable around some mean value with about 10% variance. This results in slightly broadened peaks of the harmonic spectrum and reproduces the variation that we observe in seismic records of volcanic tremors (Eibl et al. 2017a) (Figure S1 b). After creating the harmonic signal, colored noise resembling Peterson's low noise model (LNM, Peterson 1993) is added to the signal. The colored noise is synthesized by computing coefficients of a zero phase FIR (Finite-Impulse-Response) filter via inverse FFT from the spectral representation of the LNM. Then, we apply the FIR filter to a random time series of arbitrary length and multiply it with an amplitude factor to adjust the SNR of the tremor versus colored noise (Figure S2). Finally, we add real earthquake recordings randomly in time to the resulting time series of synthetic tremor and noise (Figure S3). Each earthquake signal, which is used for semi-synthetic data creation, is cut from the beginning of the P wave until the signal amplitudes returns to the pre-event noise level after the S- or Surface wave coda part. We used different types of the earthquakes' signals, i.e., both long period and volcano-tectonic events within the time period from 15 September to 20 September 2014 show significantly different signal durations. In total, we created 24 hours of semi-synthetic data by combining 500 real earthquake recordings with synthetic harmonic waveform and a seismic noise series. More details about the semi-synthetic data generation can be found in the supplementary Figures (S1 to S3). Figure 5 (a, b, & c) show the components of the semi-synthetic signal and Figure 5 (d) shows the created semi-synthetic signal.

We applied our method to this semi-synthetic dataset. The synthetic harmonic signals were extracted and the earthquakes were detected via the characteristic function. Figure 5 (e) shows the semi-synthetic signal after subtracting the extracted tremor signal from it and we name it the de-tremored signal. As shown in Figure 5 (e), this signal has a larger earthquake Signal to Noise Ratio (SNR) and an improvement in the first-motion piking is seen. This is useful when we need to remove a harmonic noise from the seismic waveform. Figure 5 (f & g) show the extracted harmonic signal and the earthquake characteristic function as outputs of the method.

3.2 Testing the tremor extraction algorithm using semi-synthetic data

To evaluate the ability of the method for tremor signal extraction, we use the created semi-synthetic data with different SNR of the harmonic signal. In order to set different SNRs, we normalize each component of the semi-synthetic data by dividing it by its standard deviation and then we weight them based on the desired SNR. Our harmonic signal extraction process is performed on the semi-synthetic data and the harmonic signal is then reconstructed. The cross-correlation of the synthetic harmonic signal and the reconstructed harmonic signal using our method is measured (Figure 6). Cross-correlations measure the similarity of two time series, so we calculate them to evaluate how similar the reconstructed harmonic signal is to the synthetic harmonic signal. If the two-time series are identical, the cross-correlation coefficient will be 1 and if they are completely different, the cross-correlation coefficient will be 0. We can reconstruct the tremor signal for a SNR of at least 0.4 with a cross-correlation of more than 0.8. The synthetic harmonic signal and the reconstructed signal match well in both phase and shape (see Figure 5b and f). The differences between these two signals is usually related to small fluctuations in the input harmonic signal, which shows a random pattern instead of a repetitive pattern. The similarity matrix is not able to identify random patterns, therefore, they are not reconstructed in the output signal. Figure 6 shows the SNR and related cross-correlation of input and output harmonic signal.

3.3 Testing the earthquake detection algorithm using semi-synthetic data

To evaluate the capacity of our method for earthquake detection, we use the created semi-synthetic data with different earthquake SNR. We report the local SNR here, which refers to the ratio between the variance of the earthquake signal and the variance of the local related segment of the semi-synthetic data. The local related segment is the time window, which contains the earthquake signal as well as synthetic tremor signals and seismic noise in the background. The segment has a variable length that corresponds to the earthquake signal duration. The advantage of the semi-synthetic signals is that we can measure and control the individual components. The results show that for SNR = 0.1, we can detect more than 78% of the events, however, below SNR = 0.3, there is a significant number of false picks (up to 30% of all events), while the average percentage of false picks is 6% for SNR between 0.3 to 1. For SNR higher than 0.1, 42% of the missed events are LP events. The SNR and related detection rates are reported in Figure 7 (a). Some examples of semi-synthetic data with different earthquake SNR and different SNR of harmonic signal component are presented in the supplementary figures (Figures S4 to S7).

Most missed events are similar to that shown in Figure 7 (b), which are classified as long period (LP) events (Woods et al., 2018). Figure 7 (c) shows a typical volcano-tectonic event for comparison. That LP events are often not detected can be explained by the properties of the detection characteristic function. This function is derived from summing all frequencies in the transient spectrogram for each time frame. Thus, the characteristic function is sensitive to broadband signals. However, LP events are narrow band, which results in a poor performance, although the signals are contained in the transient spectrogram. Also, if LPs persist longer, it becomes more difficult to detect them because of the basic structure of the method. Indeed, to create the repeating spectrogram, for all time frames, we derive the corresponding frame (in the repeating spectrogram) by taking the median of the similar frames (which are identified using the similarity matrix) for each frequency bin. For a transient (short duration in time) event, there are a few numbers of similar frames in the spectrogram, so it is identified as a non-repeating pattern. Therefore, it will show a short-lasting sharp peak in the transient characteristic function. In contrast, for a long-lasting event, there are some adjacent similar frames, which will be replaced in the repeating spectrogram by the median of them. Therefore, it shows some long lasting, less sharp, adjacent peaks in the transient characteristic function, which is less likely to be detected by the local maximum finder compared to sharper peaks.

3.4 Real data tests

In a final step, we applied the method to a dataset of the Holuhraun 2014-2015 eruption and extracted volcanic tremor signals from the seismological records. As discussed in section 2.2 and showed in Figure 3 (a & b), the reconstructed tremor signal matches well with the original seismological records and has no trace of transient, earthquakes-related signals. This dataset consisted of one month (September 2014) of recordings by the FLUR station and we use a single vertical component to detect earthquakes. We compared our detected earthquakes with the bulletin presented in Ágústsdóttir et al. (2019). For the station location with respect to the eruption fissures, please see Figure 1 and Woods et al. (2018). 84% of the total of 5071 events listed by Ágústsdóttir et al. (2019) were detected by our proposed approach.

We detected a total of 12619 events, which is more than twice the number of listed events in the bulletin. The bulletin is made based on an automatic detection method using Coalescence Micro-seismic Mapping (Drew et al., 2013) with the velocity model used in Ágústsdóttir et al. (2016) (their Figure S2 (c)). The bulletin earthquakes were relocated (Ágústsdóttir et al., 2019) using cross-correlated, sub-sample relative travel times following the method of Woods et al. (2019). A dense local seismic network comprising 72 three-component broadband instruments was used to create the one-year bulletin. Our detection process currently uses only one component of seismic recording from a single station. In the future, the result could be improved using three-component signals and additional stations because some of the smaller events may have larger amplitudes on the other components or stations. An event with a larger amplitude shows a larger peak in the characteristic function, hence the probability of its detection using our algorithm will increase.

Our method can detect two adjacent earthquakes with a minimum interval of around 10 seconds. This interval is defined by the number of samples, which must be waited after picking a peak in the local maximum finder. The interval value depends on the number of FFT (Fast Fourier Transform) points, the hop size, and the type of earthquake. In our dataset, earthquakes are mostly local, where shorter waiting time values will result in the detection of more than one peak for one event.

Using the algorithm described in section 2.3, we are able to find P-wave arrival times using the detected peaks via the local maximum finder. The uncertainty in the example shown in Figure 4 is 0.1 second through visual inspection. The pattern of the characteristic function for different types of events is, however, not always similar to the simple shape we have assumed, which mostly corresponds to the energy shape of a local event and could have more fluctuations, thus the uncertainty in detecting the P arrivals could be higher. We compared the P-arrival time residuals of our method and those given by the bulletin of Ágústsdóttir et al. (2019) for one month. For 52% of the events, the time difference is less than one second, while 48% of the events show a time difference of between one to six seconds. A significant part of large time differences is related to LP events, where the duration of the event is long compared to volcano-tectonic events in the characteristic function, where the first arrival is outside of the pre-peak interval time window. In this case, the algorithm is able to send the first selected peak back in time to the starting point of the window and shorten the time difference, however the emergent onset of the LP event is still earlier in the time axis. This algorithm (finding P-wave arrival times using the detected peaks) could be improved upon by assigning different parameters for different event types.

The algorithm which is proposed here is a simple way to attribute the peaks to the starting point of changes in the characteristic function. This could be applied in different fields when a function has rather stable values, but also experiences sudden changes, and finding the first point of the starting changes is important. One could develop the algorithm by adding more criteria based on the information about the phenomena that are attributed to the changes to decrease the uncertainty in finding the starting point of change.

3.5 Feasibility of the method with respect to processing speed

The average computation time for the tremor extraction of a one-day long record with a FFT window length of 81.9 seconds, overlap of 75%, and a sampling frequency of 100 Hz, is about 70 seconds, when implemented in Python using Librosa on a PC with an Intel core i7 (six-core) processor of 2.2 GHz and 16 GB of RAM. For transient signal detection with an accuracy of 0.32 seconds, the computation time is about 34 minutes with a FFT window length of 1.28 seconds and an overlap of 75%. The significant difference in the computation time between the tremor extraction and transient signal detection is due to the different FFT window lengths of the two processes. Reducing the FFT length and using the same overlap of 75% increases the number of FFT windows for the overall data time range and the associated computation time.

4 Conclusions and outlook

In this work we have developed a method to extract and reconstruct volcanic tremor signals, as well as to detect transient signals from seismic waveforms. We used a combination of two harmonic-percussive separation algorithms from the field of music information retrieval to separate harmonic and percussive elements of the seismic waveform in the time frequency domain. This combination leads to a better separation of the components and results in clean tremor and transient spectrograms. The tremor signals are reconstructed in the time domain using weighted phase information of the initial seismic complex spectrogram at each time frame through the energy contribution of the tremor spectrogram. We showed that it is important to use phase information to reconstruct a signal in the time domain for seismological purposes to provide an accurate phase reconstruction. We also discussed how to use a weighted phase matrix based on the dominant frequency band of the tremor spectrogram that can almost eliminate the noise contributions in the phase matrix of the seismic waveform. The reliability of the reconstructed signal was shown using semi-synthetic tests. The cross-correlation between the synthetic harmonic signal and the reconstructed harmonic signal using our method was higher than 0.8 for SNRs of the synthetic harmonic signal above 0.4. In addition, more than 78% of earthquake signals in the semi-synthetic data with SNR = 0.1 can be detected using oue method.

The capability of the method for earthquake detection was also evaluated in comparison to a real earthquake catalog. The detection of more than twice the number of the Ágústsdóttir et al. (2019) bulletin events demonstrates the ability of the proposed method for detecting smaller seismic events, even when only a single station and component is available.

The developed method is able to extract harmonic tremor signals and is applicable to other volcanoes that exhibit such phenomena. A possible application of the proposed method is to extract volcanic tremor signals using a network or an array during a period of heightened volcanic activity. In particular, the clean tremor signal can be used for tremor source location using array analysis given that the tremor signal reconstruction provides the true phase of the signal. This may provide an improved analysis of the spatial and temporal evolution of volcanic tremors during active volcanic periods.

Another application of this method is in the field of earthquake analysis research. Here, we suggest using the seismic waveform after subtracting the tremor signals (if tremors are present). We named this signal as the “de-tremored” signal in section 3.1 (see Figure 5e). The advantage of using the de-tremored signal is the resulting increase in the earthquake SNR and improvements in the first-motion picking.

In our opinion, the transient signal detection algorithm introduced in this study is a useful tool for detecting seismic events and is especially applicable for detecting small events during an earthquake swarm. While we used one component of one station for earthquake detection in this study, the results could be improved using three components and additional station because some events with low amplitude on the current component and station may show larger amplitude on the other components or stations.

In conclusion, the presented method could provide a basis for tremor source investigations as well as research into eruptive activity since it provides simultaneous information about tremors and earthquakes and allows the extraction of a clean signal of the tremor for detailed investigations.

Data and Resources

All data used in this paper is openly available at IRIS (network code 7Z, White 2010). A Jupyter notebook with all the Python codes and parameters related to the proposed method is available as an electronic supplement. The supplementary material related to this article also contains illustrations of the semi-synthetic data generation. The application of the method using some examples of semi-synthetic data with different earthquake SNRs and different SNRs of the harmonic signal component are also presented in the supplementary material.

Acknowledgments

Zahra Zali is grateful for the support by the German Academic Exchange Service (DAAD) through the Graduate School Scholarship Programme under reference number 91721165. This work was also supported by the German Research Foundation (DFG MU 2686/13-1, SCHE 280/20-1) and the Daimler Benz Foundation (32-02/18). The dataset is openly available at IRIS (network code 7Z, White, 2010). We thank Tom Winder for sharing the P-wave arrival times of the bulletin in Ágústsdóttir et al. (2019). We thank Diana Roman and an anonymous reviewer for their constructive comments. We thank Kevin Fleming for proofreading this article and for his spelling corrections and editorial improvements. The Python code related to the proposed method is freely available from https://gitup.uni-potsdam.de/zali/harmonic-tremor-extraction-and-transient-signal-detection

References

Ágústsdóttir, T., Winder, T., Woods, J., White, R. S., Greenfield, T., & Brandsdóttir, B. (2019). Intense Seismicity During the 2014–2015 Bárðarbunga-Holuhraun Rifting Event, Iceland, Reveals the Nature of Dike-Induced Earthquakes and Caldera Collapse Mechanisms, *J . Geophys. Res.: Solid Earth*, **124**(8), 8331-8357. doi:10.1029/2018JB016010.

Ágústsdóttir, T., Woods, J., Greenfield, T., Green, R. G., White, R. S., Winder, T., ... & Soosalu, H. (2016). Strike-slip faulting during the 2014 Bárðarbunga-Holuhraun dike intrusion, central Iceland, *Geophys. Res. Lett*. **43**(4), 1495-1503. doi:10.1002/2015GL067423.

Alparone, S., Andronico, D., Lodato, L., & Sgroi, T. (2003). Relationship between tremor and volcanic activity during the Southeast Crater eruption on Mount Etna in early 2000, *J . Geophys. Res.: Solid Earth*, **108**(B5). doi:10.1029/2002JB001866.

Cano, E., FitzGerald, D., Liutkus, A., Plumbley, M. D. & Stöter, F. R. (2018). Musical source separation: An introduction*, IEEE Signal Process Mag*. **36**(1), 31-40. doi: 10.1109/MSP.2018.2874719.

Chouet, B. A. (1996). Long-period volcano seismicity: its source and use in eruption forecasting, *Nature*, **380**(6572), 309-316.

Davi, R., O'Brien, G. S., De Barros, L., Lokmer, I., Bean, C. J., Lesage, P., ... & Soto, G. J. (2012). Seismic source mechanisms of tremor recorded on Arenal volcano, Costa Rica, retrieved by waveform inversion, *J. Volcanol. Geotherm.* Res. **213**, 1-13. doi:10.1016/j.jvolgeores.2011.10.008.

Dmitrieva, K., Hotovec-Ellis, A. J., Prejean, S. & Dunham, E. M. (2013). Frictional-faulting model for harmonic tremor before Redoubt Volcano eruptions, *Nat. Geosci*. **6**(8), 652-656. doi: 10.1038/NGEO1879.

Drew, J., White, R. S., Tilmann, F. & Tarasewicz, J. (2013). Coalescence microseismic mapping, *Geophys. J. Int*. **195**(3), 1773-1785. doi: 10.1093/gji/ggt331.

Driedger, J., Müller, M. & Disch, S. (2014, October). Extending Harmonic-Percussive Separation of Audio Signals. In ISMIR (pp. 611-616).

Eibl, E. P., Bean, C. J., Jónsdóttir, I., Höskuldsson, A., Thordarson, T., Coppola, D. & Walter, T. R. (2017b). Multiple coincident eruptive seismic tremor sources during the 2014–2015 eruption at Holuhraun, Iceland, *J . Geophys. Res.: Solid Earth*, **122**(4), 2972-2987. doi:10.1002/2016JB013892.

Eibl, E. P., Bean, C. J., Vogfjörd, K. S., Ying, Y., Lokmer, I., Möllhoff, M., ... & Pálsson, F. (2017a). Tremor-rich shallow dyke formation followed by silent magma flow at Bárðarbunga in Iceland, *Nat. Geosci*. **10**(4), 299-304. doi: 10.1038/NGEO2906.

Falsaperla, S., Alparone, S., D'Amico, S., Grazia, G., Ferrari, F., Langer, H., ... & Spampinato, S. (2005). Volcanic tremor at Mt. Etna, Italy, preceding and accompanying the eruption of July–August, 2001, *Pure Appl. Geophys*. **162**(11), 2111-2132. doi: 10.1007/s00024-005-2710-y.

Fitzgerald, D. (2010, September). Harmonic/percussive separation using median filtering, In *Proceedings of the International Conference on Digital Audio Effects (DAFx)* (Vol. 13).

FitzGerald, D. (2012). Vocal separation using nearest neighbours and median filtering, *23rd IET Irish Signals and Systems Conference*, Maynooth. 28-29th. June 2012. doi: 10.1049/ic.2012.0225.

FitzGerald, D., & Gainza, M. (2010). Single channel vocal separation using median filtering and factorisation techniques*, ISAST Transactions on Electronic and Signal Processing,* No. 1, Vol. 4,2010 (ISSN 1797-2329), pages: 62 - 73, 2010.

Griffin, D., & Lim, J. (1984). Signal estimation from modified short-time Fourier transform, *IEEE Trans. Acoust. Speech Signal Process*. **32**(2), 236-243. doi:10.1109/icassp.1983.1172092.

Gudmundsson, M. T., Jónsdóttir, K., Hooper, A., Holohan, E. P., Halldórsson, S. A., Ófeigsson, B. G., ... & Einarsson, P. (2016). Gradual caldera collapse at Bárdarbunga volcano, Iceland, regulated by lateral magma outflow. *Science*, **353**(6296). doi: 10.1126/science.aaf8988.

Hellweg, M. (2000). Physical models for the source of Lascar's harmonic tremor, *J. Volcanol. Geotherm. Res.* **101**(1-2), 183-198. doi: 10.1016/S0377-0273(00)00163-3.

Hotovec, A. J., Prejean, S. G., Vidale, J. E., & Gomberg, J. (2013). Strongly gliding harmonic tremor during the 2009 eruption of Redoubt Volcano, *J. Volcanol. Geotherm.* Res. **259**, 89-99. doi:10.1016/j.jvolgeores.2012.01.001.

Johnson, J. B., & Watson, L. M. (2019). Monitoring volcanic craters with infrasound "music", *Eos, 100*. doi:10.1029/2019EO123979.

McFee, B., Lostanlen, V., Metsai, A., McVicar, M., Balke, S., Thomé, C., … Kim, T. (2020, July 22). librosa/librosa: 0.8.0 (Version 0.8.0). Zenodo. doi: 10.5281/zenodo.3955228.

McNutt, S. R. (1992). Volcanic tremor, *Encyclopedia of earth system science*, **4**, 417-425.

Montegrossi, G., Farina, A., Fusi, L., & De Biase, A. (2019). Mathematical model for volcanic harmonic tremors. *Scientific reports*, **9**(1), 1-14.

Müller, M. (2015). Fundamentals of music processing: Audio, analysis, algorithms, applications. Cham, Switzerland: *Springer International Publishing*. doi: 10.1007/978-3-319-21945-5.

Peterson, J. R. (1993). Observations and modeling of seismic background noise (No. 93-322). *US Geological Survey*. doi:10.3133/ofr93322.

Rafii, Z., Liutkus, A. & Pardo, B. (2014). REPET for background/foreground separation in audio. In Blind Source Separation (pp. 395-411). *Springer, Berlin, Heidelberg*.

Rafii, Z., Liutkus, A., Stöter, F. R., Mimilakis, S. I., FitzGerald, D., & Pardo, B. (2018). An overview of lead and accompaniment separation in music*, IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **26**(8), 1307-1335. doi: 10.1109/TASLP.2018.2825440.

Rafii, Z., & Pardo, B. (2011, May). A simple music/voice separation method based on the extraction of the repeating musical structure, In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 221-224). IEEE. Doi: 10.1109/ICASSP.2011.5946380.

Rafii, Z., & Pardo, B. (2012, October). Music/Voice Separation Using the Similarity Matrix, In *ISMIR* (pp. 583-588).

Rouland, D., Legrand, D., Zhizhin, M., & Vergniolle, S. (2009). Automatic detection and discrimination of volcanic tremors and tectonic earthquakes: An application to Ambrym volcano, Vanuatu, *J. Volcanol. Geotherm.* Res. **181**(3-4), 196-206. doi: 10.1016/j.jvolgeores.2009.01.019.

Schlindwein, V., Wassermann, J., & Scherbaum, F. (1995). Spectral analysis of harmonic tremor signals at Mt. Semeru volcano, Indonesia, *Geophys. Res. Lett*. **22**(13), 1685-1688. doi:10.1029/95GL01433.

Sigmundsson, F., Hooper, A., Hreinsdóttir, S., Vogfjörd, K. S., Ófeigsson, B. G., Heimisson, E. R., ... & Drouin, V. (2015). Segmented lateral dyke growth in a rifting event at Bárðarbunga volcanic system, Iceland, *Nature*, **517**(7533), 191-195. doi: 10.1038/nature14111.

Sturmel, N., & Daudet, L. (2011, September). Signal reconstruction from STFT magnitude: A state of the art, In *International conference on digital audio effects (DAFx)* (pp. 375-386).

Vaseghi, S. V. (1996). Advanced signal processing and digital noise reduction. Vieweg+ Teubner Verlag.

White, R. (2010). Northern Volcanic Zone [Data set], *International Federation of Digital Seismograph Networks*. doi: 10.7914/SN/Z7_2010.

Woods, J., Donaldson, C., White, R. S., Caudron, C., Brandsdóttir, B., Hudson, T. S. & Ágústsdóttir, T. (2018). Long-period seismicity reveals magma pathways above a laterally propagating dyke during the 2014–15 Bárðarbunga rifting event, Iceland, *Earth Planet. Sci. Lett*. **490**, 216-229. doi: 10.1016/j.epsl.2018.03.020.

Woods, J., Winder, T., White, R. S. & Brandsdóttir, B. (2019). Evolution of a lateral dike intrusion revealed by relatively-relocated dike-induced earthquakes: The 2014–15 Bárðarbunga–Holuhraun rifting event, Iceland, *Earth Planet. Sci. Lett*. **506**, 53-63. doi: 10.1016/j.epsl.2018.10.032.

Yukutake, Y., Honda, R., Harada, M., Doke, R., Saito, T., Ueno, T., ... & Morita, Y. (2017). Analyzing the continuous volcanic tremors detected during the 2015 phreatic eruption of the Hakone volcano, *Earth Planets Space*, **69**(1), 164. doi: 10.1186/s40623-017-0751-y.

Mailing address for each author

Zahra Zali: Campus Golm, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Matthias Ohrnberger: Campus Golm, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Frank Scherbaum: Campus Golm, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Fabrice Cotton: Helmholtzstraße 6/7, 14467 Potsdam, Germany

Eva P. S. Eibl: Campus Golm, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

List of Figure Captions

Figure 1. Aspects of the Holuhraun 2014-2015 eruption data and the application of the proposed method. (a) The eruption site and the station location. (b) An example of real data from 3 September 2014 (HHZ component of FLUR station from network 7Z (White, R. 2010)). (c) The PSD and the (d) spectrogram of this day for the raw seismic data. (e) The PSD and the (f) spectrogram for the extracted tremor signal using the proposed method.

Figure 2. Method flowchart. (a) Processing steps of the method and (b) illustration of the processing steps with a real data example.

Figure 3. Comparison of the extracted tremor signal using the proposed method and two other methods visualized for a short time window of data from 3 September 2014 (HHZ component of

FLUR station from network 7Z (White, R. 2010)). (a) The raw seismic signal (blue) and the reconstructed tremor signal using our method (green). (b) Same as the green trace in (a). (c) The reconstructed tremor signal using horizontal median filtering. The traces of transient events still exist in this signal. (d) The estimated tremor signal using the Griffin-Lim algorithm for phase reconstruction. The vertical red line is drawn to illustrate the phase alignment of the signals.

Figure 4. Flowchart for backtracking the peaks to the arrival time. The example shows an earthquake time history and its characteristic function. The vertical green line in the top left figure shows the first selected peak, which is sent back in time to the P arrival time step by step. In the top left the pre-peak interval time window is demonstrated as [start, end). The bracket means including the start point in the time window and the parentheses means excluding the end point from the time window. The uncertainty of the P arrival time in this example is 0.1 second through visual inspection.

Figure 5. Testing the method with semi-synthetic data. (a) Earthquake signals, (b) synthetic tremor signal, and (c) seismic noise signal are the elements for creating semi-synthetic data. Each of these three signals is normalized by dividing by their standard deviation. (d) Weighted sum of the data in subfigures a-c, which is used as an input for our method. The SNR of the earthquakes is 0.2 and the harmonic SNR is equal to 2. (e) The de-tremored signal derived by subtracting the extracted tremor signal from the semi-synthetic signal. (f) Extracted tremor signal and (g) transient characteristic function as outputs of our method.

Figure 6. Cross-correlation of the semi-synthetic harmonic signal and the reconstructed harmonic signal versus the SNR of harmonic signal.

Figure 7. Detection rate of earthquakes in the semi-synthetic data as well as two earthquakes as samples of detected and not-detected events by our method. (a) Detection rates for semi-synthetic data as a function of the SNR. (b) Seismic waveform and spectrogram of a not-detected long period (LP) event on 16 September 2014. (c) Seismic waveform and spectrogram of a detected volcano-tectonic event on 16 September 2014.

Figures

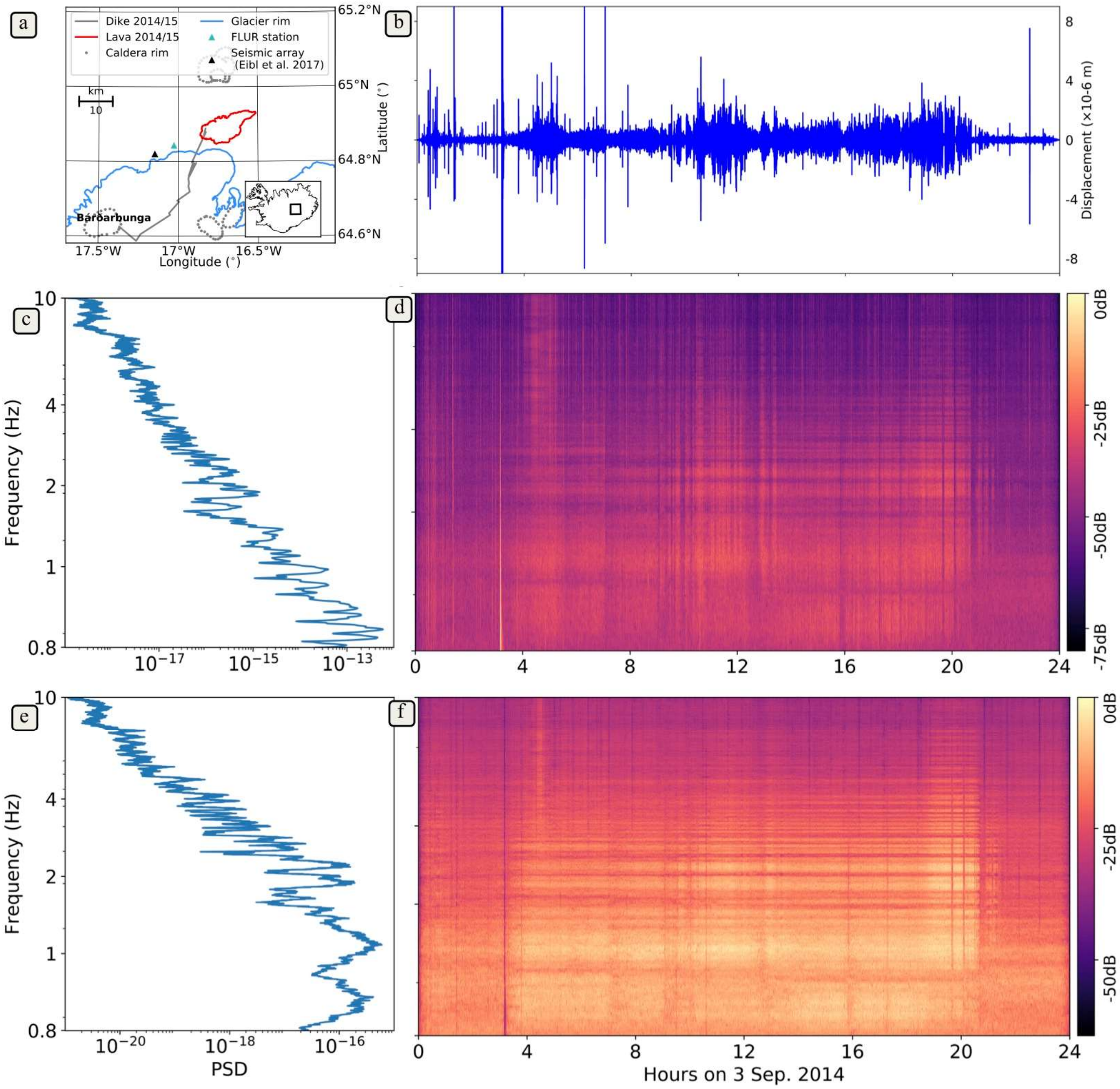

Figure 1. Aspects of the Holuhraun 2014-2015 eruption data and the application of the proposed method. (a) The eruption site and the station location. (b) An example of real data from 3 September 2014 (HHZ component of FLUR station from network 7Z (White, R. 2010)). (c) The PSD and the (d) spectrogram of this day for the raw seismic data. (e) The PSD and the (f) spectrogram for the extracted tremor signal using the proposed method.

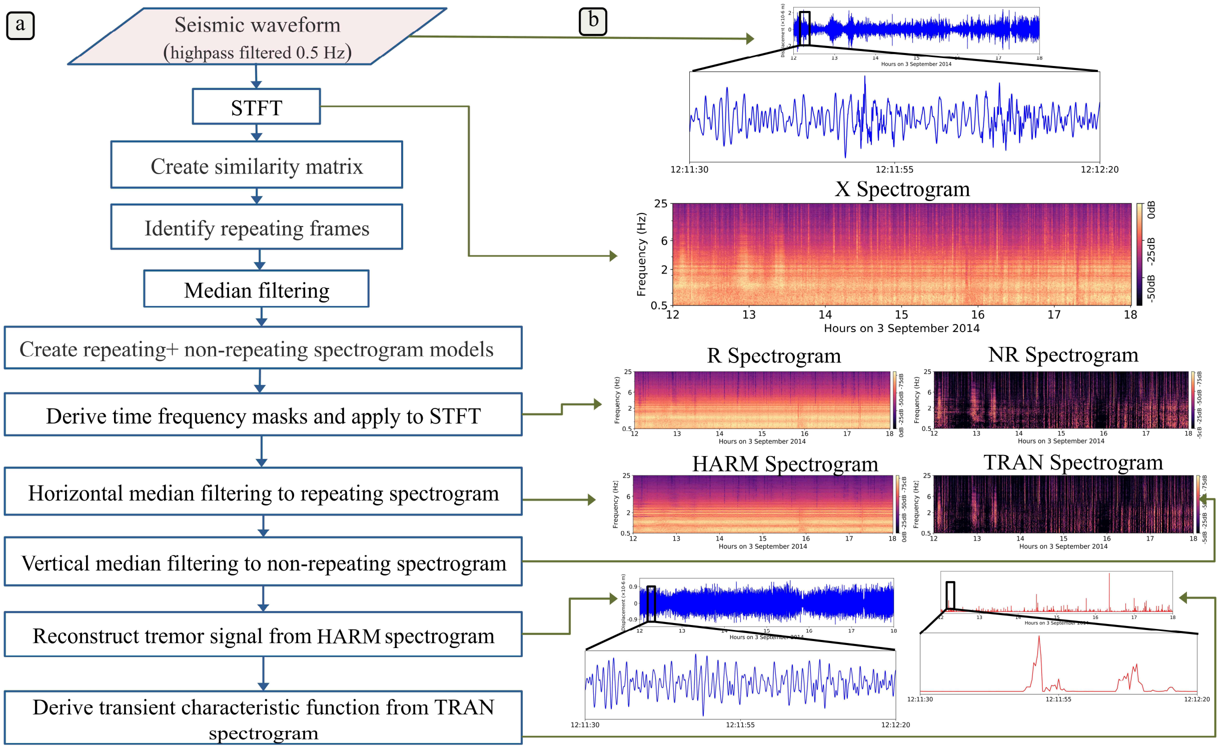

Figure 2. Method flowchart. (a) Processing steps of the method and (b) illustration of the processing steps with a real data example.

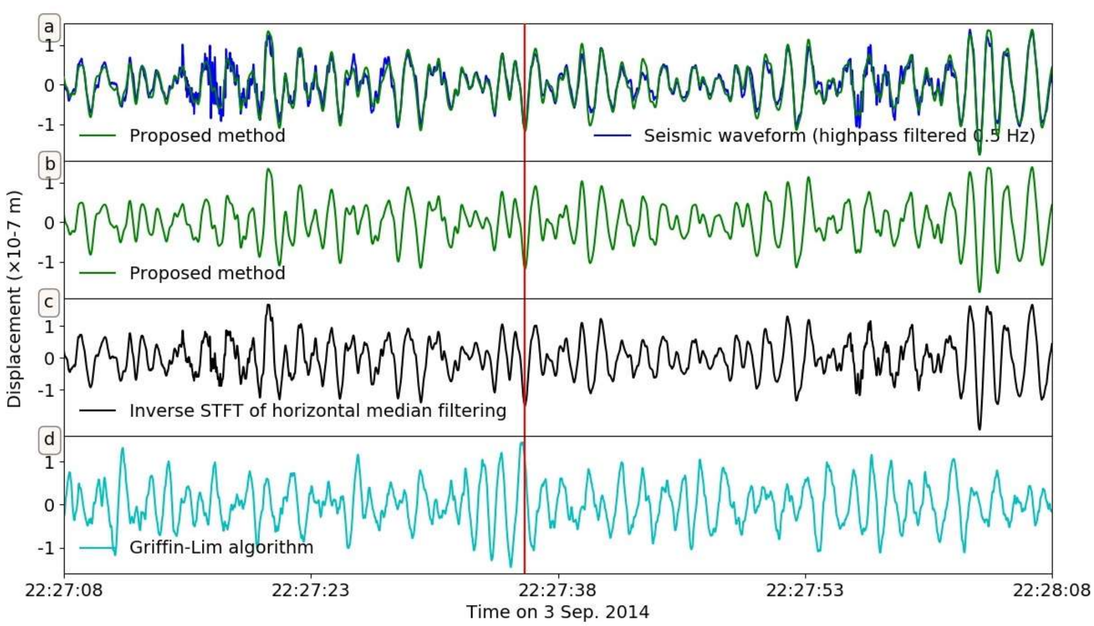

Figure 3. Comparison of the extracted tremor signal using the proposed method and two other methods visualized for a short time window of data from 3 September 2014 (HHZ component of FLUR station from network 7Z (White, R. 2010)). (a) The raw seismic signal (blue) and the reconstructed tremor signal using our method (green). (b) Same as the green trace in (a). (c) The reconstructed tremor signal using horizontal median filtering. The traces of transient events still exist in this signal. (d) The estimated tremor signal using the Griffin-Lim algorithm for phase reconstruction. The vertical red line is drawn to illustrate the phase alignment of the signals.

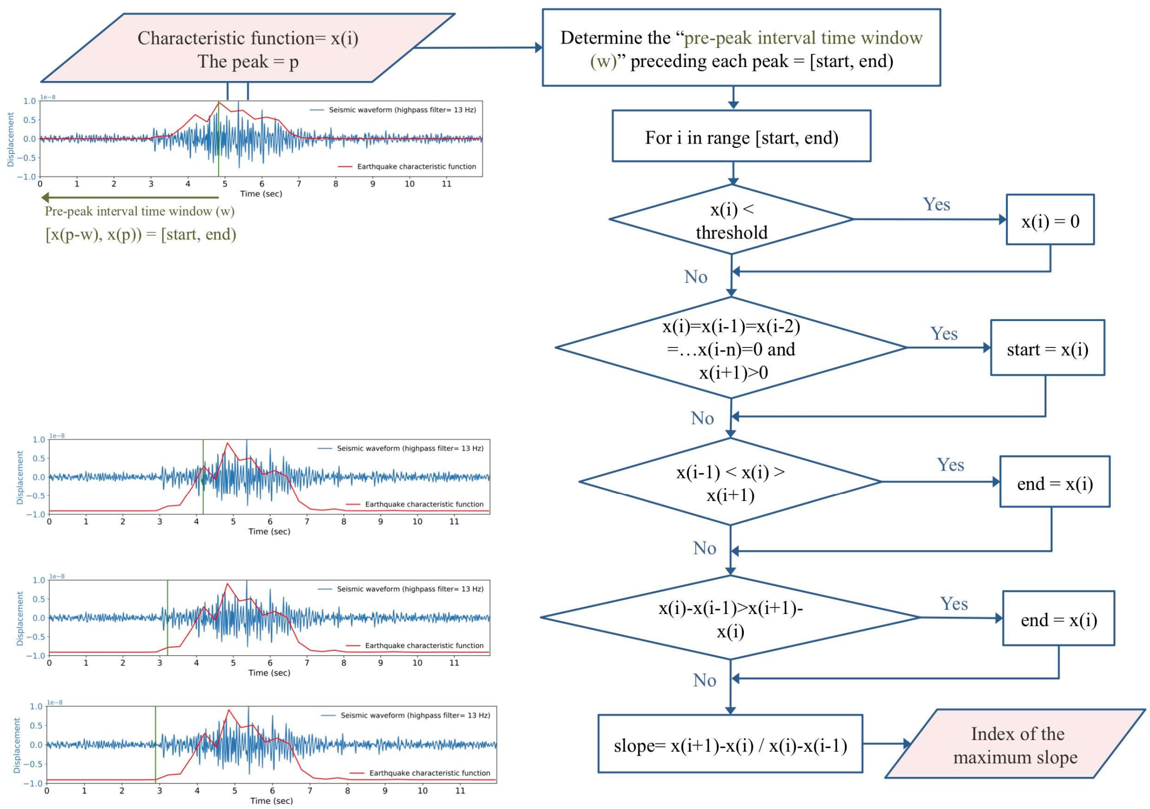

Figure 4. Flowchart for backtracking the peaks to the arrival time. The example shows an earthquake time history and its characteristic function. The vertical green line in the top left figure shows the first selected peak, which is sent back in time to the P arrival time step by step. In the top left the pre-peak interval time window is demonstrated as [start, end). The bracket means including the start point in the time window and the parentheses means excluding the end

point from the time window. The uncertainty of the P arrival time in this example is 0.1 second through visual inspection.

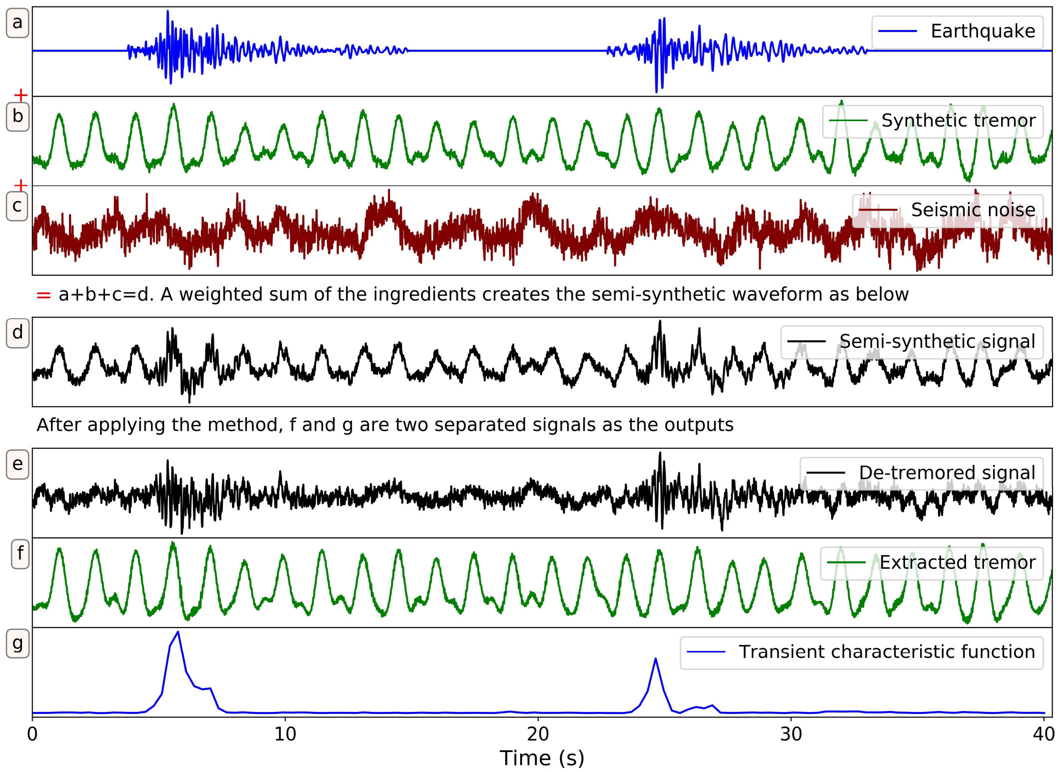

Figure 5. Testing the method with semi-synthetic data. (a) Earthquake signals, (b) synthetic tremor signal, and (c) seismic noise signal are the elements for creating semi-synthetic data. Each of these three signals is normalized by dividing by their standard deviation. (d) Weighted sum of the data in subfigures a-c, which is used as an input for our method. The SNR of the earthquakes is 0.2 and the harmonic SNR is equal to 2. (e) The de-tremored signal derived by subtracting the extracted tremor signal from the semi-synthetic signal. (f) Extracted tremor signal and (g) transient characteristic function as outputs of our method.

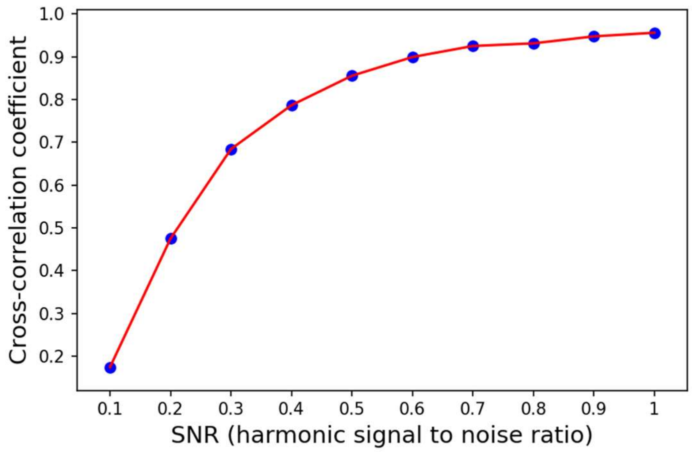

Figure 6. Cross-correlation of the semi-synthetic harmonic signal and the reconstructed harmonic signal versus the SNR of harmonic signal.

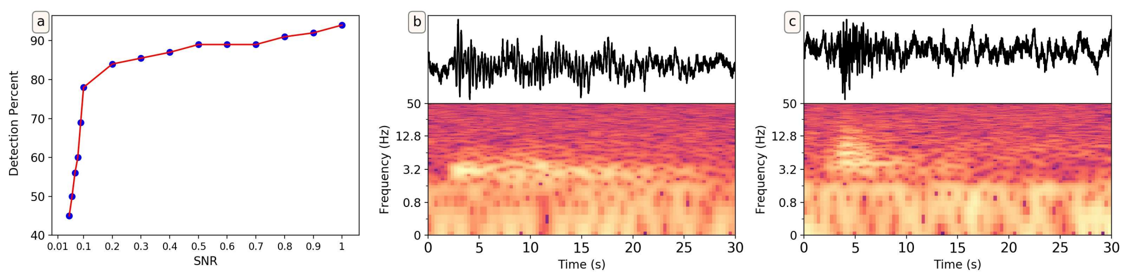

Figure 7. Detection rate of earthquakes in the semi-synthetic data as well as two earthquakes as samples of detected and not-detected events by our method. (a) Detection rates for semi-synthetic data as a function of the SNR. (b) Seismic waveform and spectrogram of a not-detected long period (LP) event on 16 September 2014. (c) Seismic waveform and spectrogram of a detected volcano-tectonic event on 16 September 2014.

Volcanic Tremor Extraction and Small Earthquakes Detection using Music Information Retrieval Algorithms

Zahra Zali, Matthias Ohrnberger, Frank Scherbaum, Fabrice Cotton, and Eva P. S. Eibl

Contents of the supplemental material:

- Supplementary figures 1 to 7
- A Jupyter notebook of all the codes and parameters

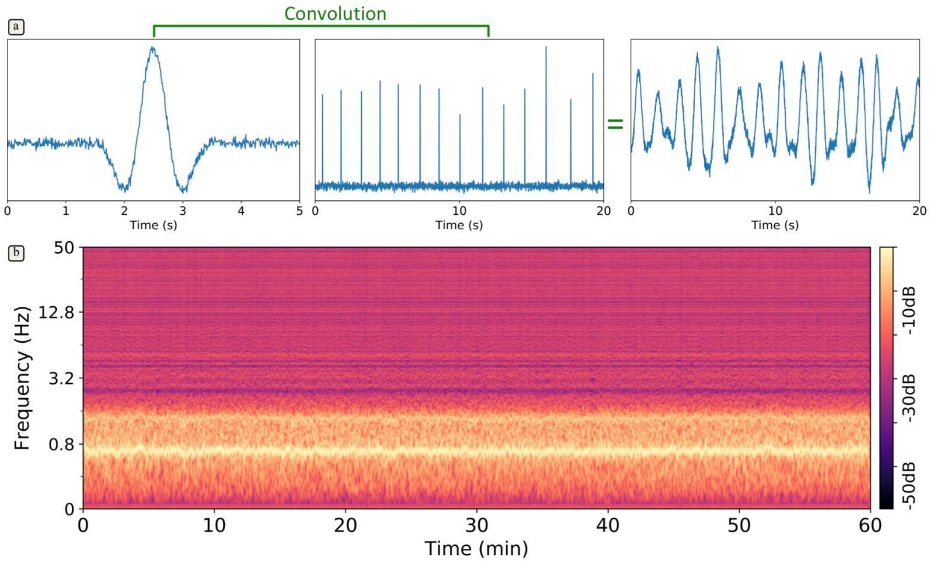

Figure S1. The process of synthetic harmonic signal generation. (a) shows the convolution of equally spaced spikes with a real-valued Morlet wavelet which results in the synthetic harmonic signal. The spikes are separated with inter-event times that are fluctuating around a mean value of 1.45 seconds with about 10 % variance. The spectrogram of the synthetic harmonic signal is shown in (b).

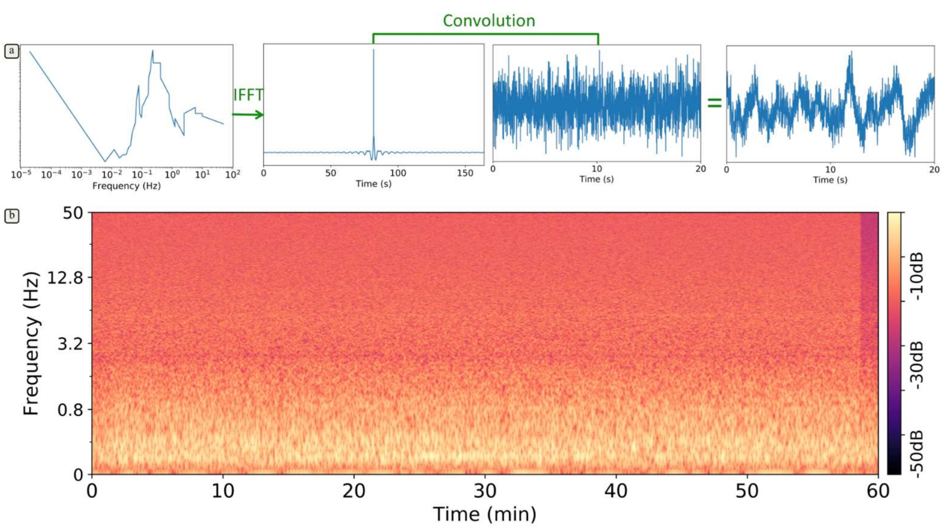

Figure S2. The process of the colored noise (Peterson 1993) generation. (a) Peterson's original Low Noise Model (LNM) amplitude spectrum is used to create FIR (Finite-Impulse-Response) filter coefficients by inverse FFT operation. The zero phase filter is applied to a random time series resulting in a synthesized colored noise resembling LNM's amplitude spectrum. The spectrogram of the colored noise signal is shown in (b).

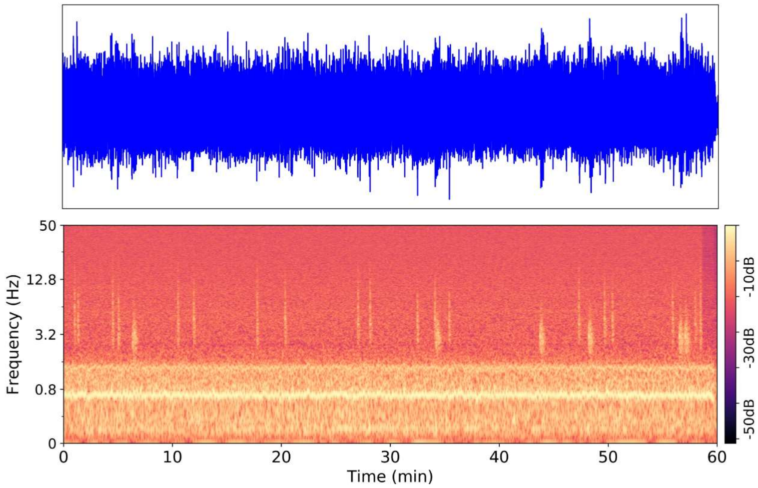

Figure S3. One hour of semi-synthetic signal and corresponding spectrogram.

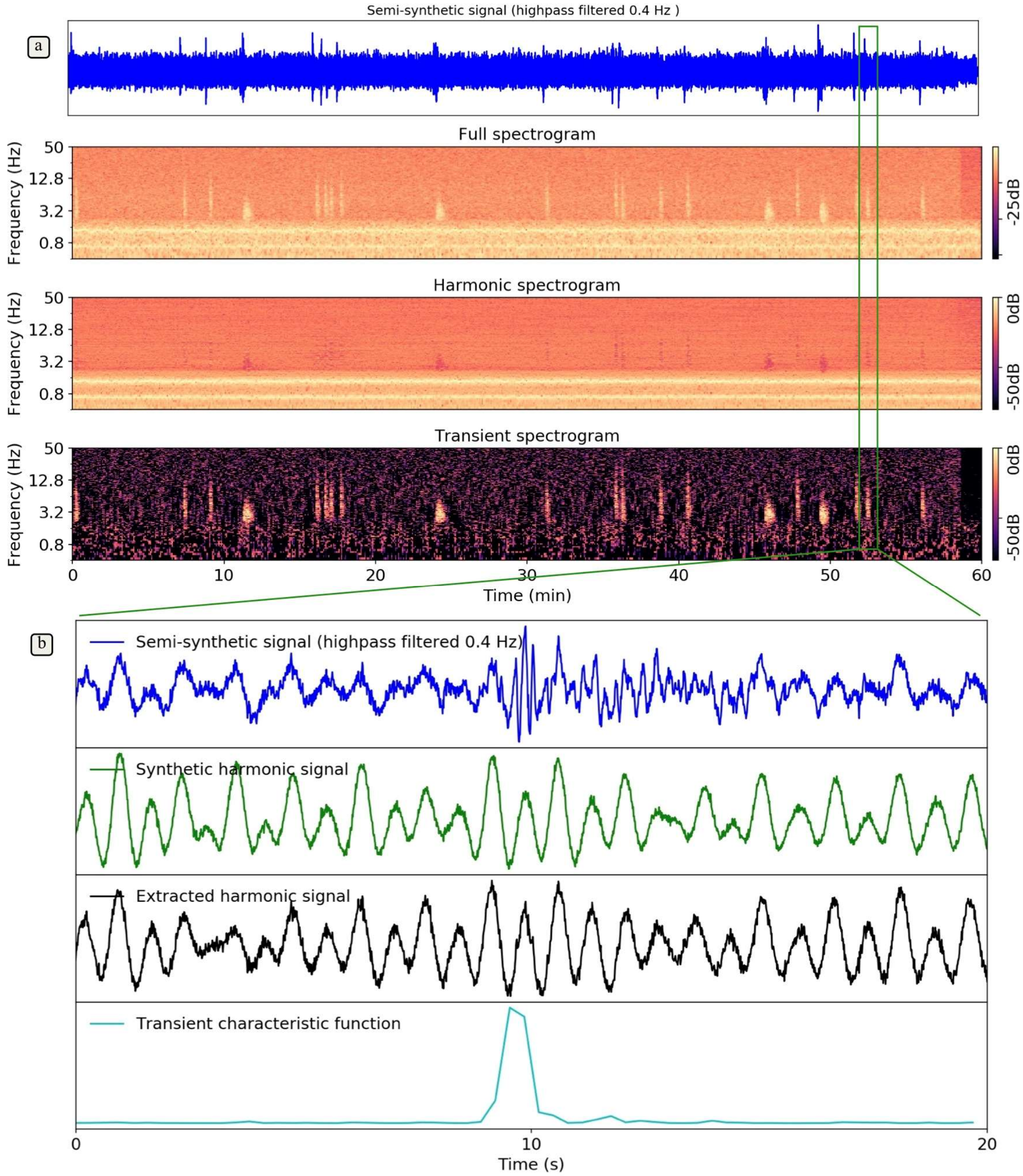

Figure S4. An example of a semi-synthetic signal with an earthquake signal (SNR=0.7) and harmonic signal (SNR=1). (a) shows one hour of the semi-synthetic signal, the spectrogram of semi-synthetic signal, the tremor spectrogram and the transient spectrogram which are derived after applying the method. (b) shows 20 seconds of the semi-synthetic signal, the synthetic harmonic signal which is used and the extracted harmonic signal through the method. We can see how the extracted harmonic signal and the synthetic harmonic signal are similar. The transient characteristic function is shown at the bottom. We see a clear peak in the characteristic function at the time of earthquake.

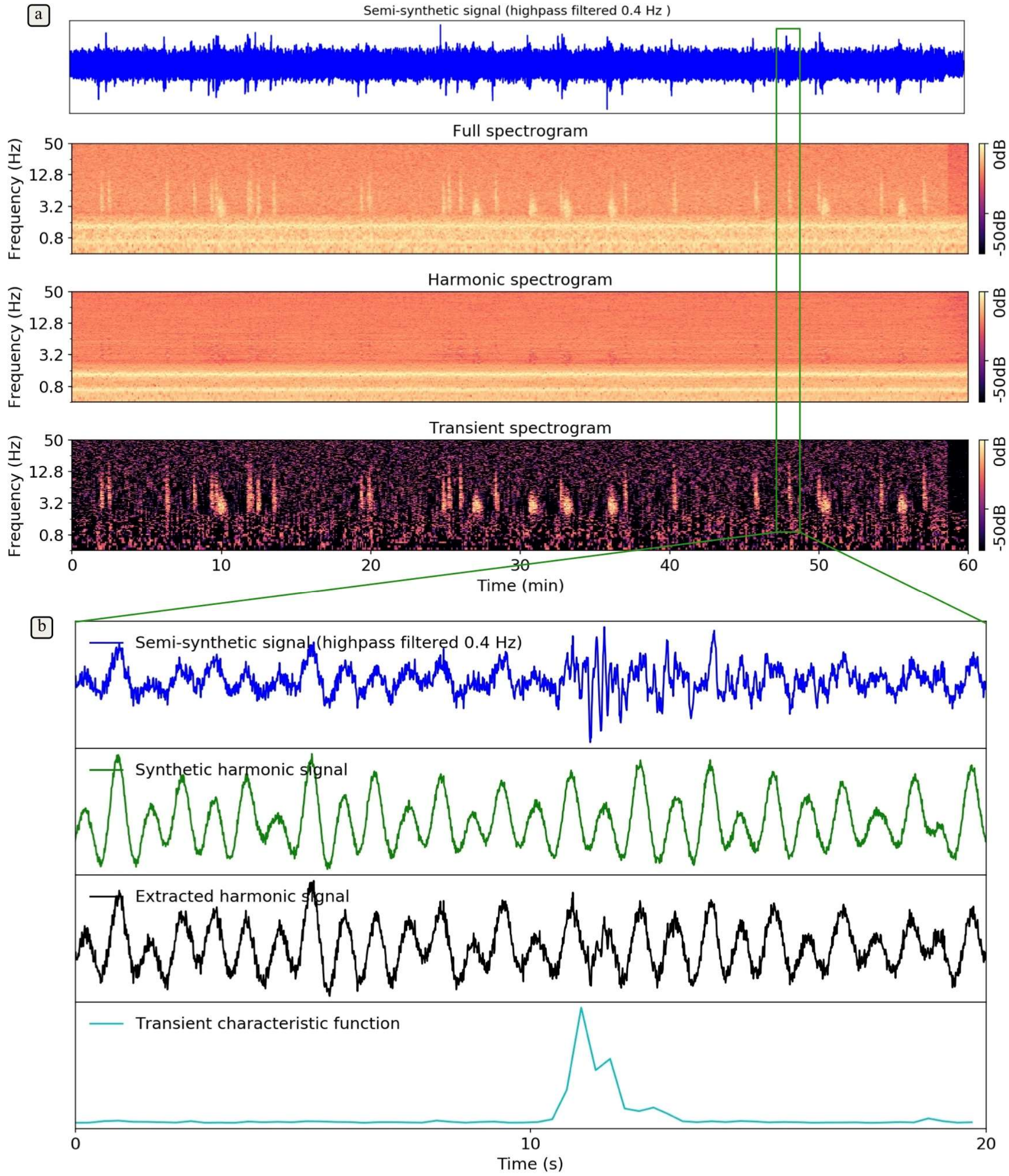

Figure S5. An example of a semi-synthetic signal with earthquake signal (SNR=1) and harmonic signal (SNR=0.7). (a) same as subfigure S4a. (b) shows 20 seconds of the semi-synthetic signal, the synthetic harmonic signal which is used and the extracted harmonic signal through the method. We can see that for this harmonic signal (SNR=0.7) the extracted harmonic signal is almost similar to the synthetic harmonic signal. The transient characteristic function is shown at the bottom. We see a clear peak in the characteristic function at the time of earthquake.

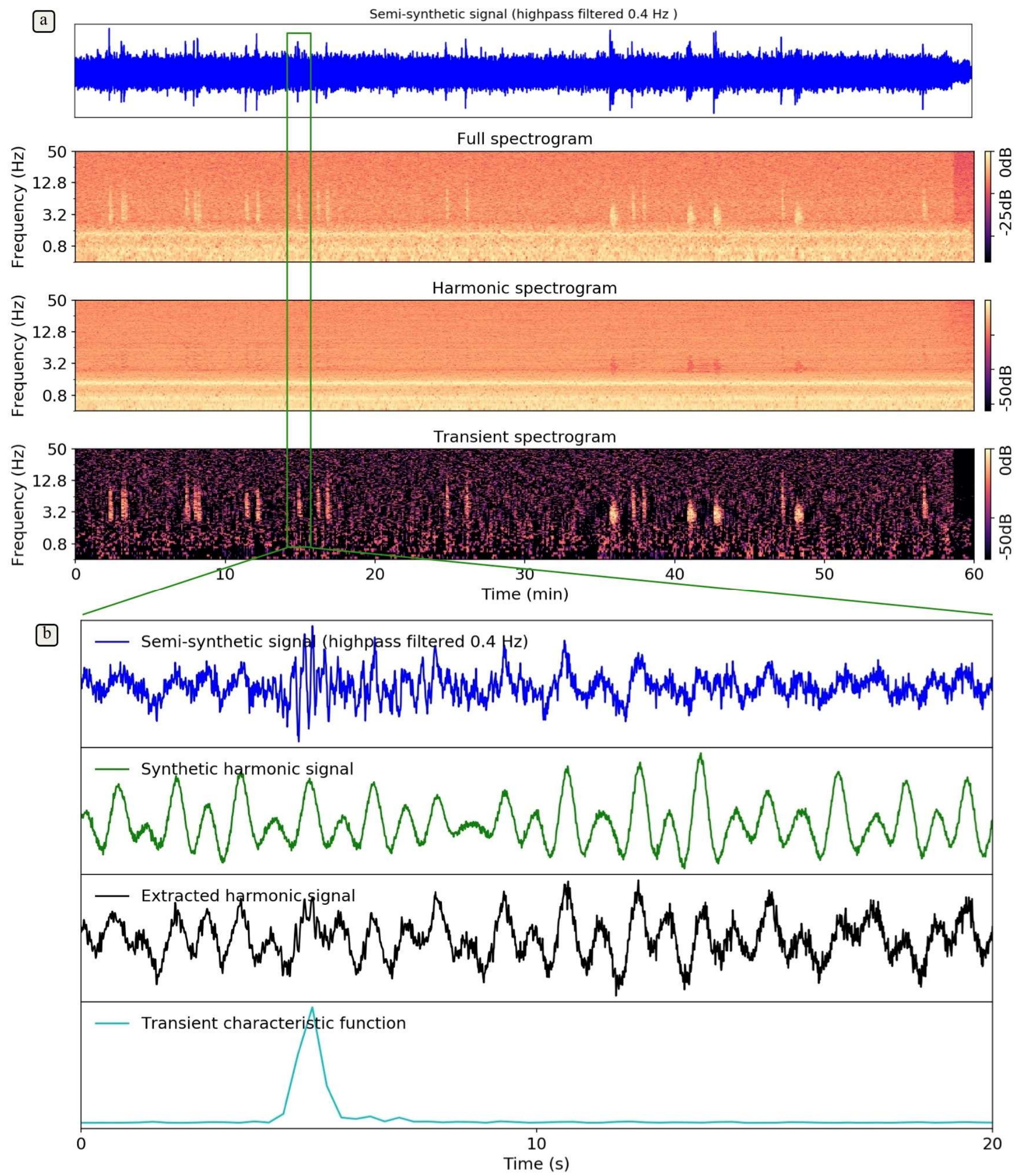

Figure S6. An example of a semi-synthetic signal with earthquake signal (SNR=0.5) and harmonic signal (SNR=0.5). (a) same as subfigure S4a. (b) shows 20 seconds of the semi-synthetic signal, the synthetic harmonic signal which is used and the extracted harmonic signal through the method. We can see the extracted harmonic signal is following the general shape of the synthetic harmonic signal for this harmonic signal (SNR=0.5), but some differences are visible. The transient characteristic function is shown at the bottom. We see a clear peak in the characteristic function at the time of earthquake.

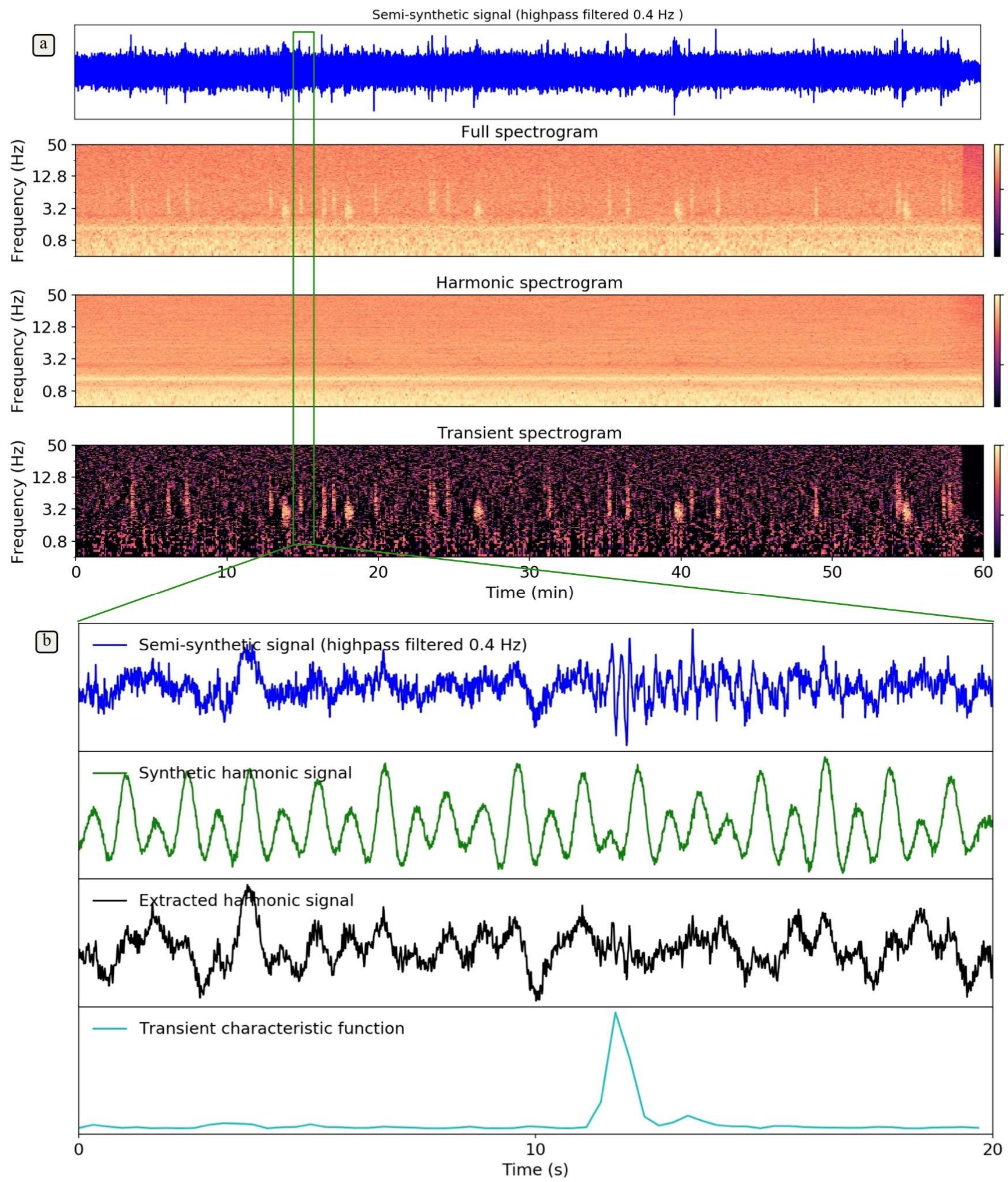

Figure S7. An example of a semi-synthetic signal with earthquake signal (SNR=0.3) and harmonic signal (SNR=0.3). (a) same as subfigure S4a. (b) shows 20 seconds of the semi-synthetic signal, the synthetic harmonic signal which is used and the extracted harmonic signal through the method. We can see that the harmonic signal is not well reconstructed here because of low SNR of harmonic signal. The transient characteristic function is shown at the bottom. Although the earthquake SNR is low, we see a clear peak in the characteristic function at the time of earthquake.