Supplementary Material

Journal of Applied Volcanology

An Adaptive 6-Dimensional Floating-Search Multi-Station Seismic-Event Detector (A6-DFMSD) and its Application to Low-Frequency Earthquakes in the East Eifel Volcanic Field, Germany

By Konun Koushesh & Joachim R. R. Ritter

Karlsruhe Institute of Technology, Geophysical Institute, Karlsruhe, Germany

Correspondence: Konun Koushesh, konun.koushesh@kit.edu and koushesh@gmail.com

https://github.com/Koushesh/A6-DFMSD/tree/master

1 METHOD

The core idea of the detector approach is to find signals emitted from a certain seismic target zone. No a priori knowledge about the waveforms is required, however, narrow frequency bands can be used to tune the search algorithm This is important in volcanic areas where different kinds of seismic events occur: tectonic, magmatic, mono-frequent etc. events.

We define and determine the limits that separate the target part of the dataset from the rest. These limits are in six dimensions and they are determined by:

a) time,

b) space (two dimensions),

c) frequency (as “signal class”),

d) variation coefficient of recorded energy *, and

e) signal to noise power (as “sigClasPower”).

These parameters are described in the sub-sections below when they appear.

Users of A6-DFMSD first need to adjust the program’s control-file (Parameters) including the modeling and detection parameters (see Table S1 for a short description), then first execute the code part called "fromSingl2multi6obspyBased" and afterwards the part “multiPartMastStBackwardTimSel4". The role of the parameters for detection are explained step by step while explaining A6-DFMSD in its five major steps:

- Inputs
- Configuration of the Detection Model (CDM step)
- Detection Field preparation (DF preparation)
- Single Station Detection (SSD step)
- Multi Station Detection (MSD step).

An overview of these five steps is presented in Fig. S1 in a flow diagram. In the following, at first, we describe how a detection model is built in accordance with the control-file parameters, and how the model helps to find sets of station-based limits. Afterwards, we describe the application of the obtained limits through the detection by giving examples.

1.1 Inputs

In addition to the continuous three-components seismic velocity records, A6-DFMSD requires the following information to establish a detection model, accordingly:

- seismic stations codes and coordinates: stored as ”stInfo”,

- simple 1-D seismic P- and/or S-wave velocity models: stored as ”velS” and ”velP”, respectively, and

- a target zone expressed by a geographical center location stored as ”targCen”, a radius length stored as ”targRadi”, an upper depth limit stored as ”targDep1” and a lower depth limit stored as ”targDep2”. All lengths are in kilometers.

This information is given by the users in the control-file.

Before using the detector, the seismic records of each day of for each station should be converted into a Q-file format (https://www.seismic-handler.org/wiki/ShmDocMenuFileReadq) and saved in the same folder of the codes of the detector with the specific naming format: XXXXyyyymmdd. For example, seismic data of the day February 1, 2023 of the station code LAGB must be saved by the name LAGB20230201 and the seismic data of the station ABH for the same day by the name ABH 20230201. Note - if the station code has less than four characters, this gap must be filled with blanks.

1.2 Configuration of the Detection Model (CDM step)

1.2.1 Building up the synthetic seismic source positions on the upper and lower surfaces of the target zone

According to the control-file values, the seismic target zone has a cylindric shape as geometry, defined by the parameters: ”targCen”, ”targRadi”, ”targDep1” and ”targDep2” (center, radius, upper depth boundary, lower depth boundary).

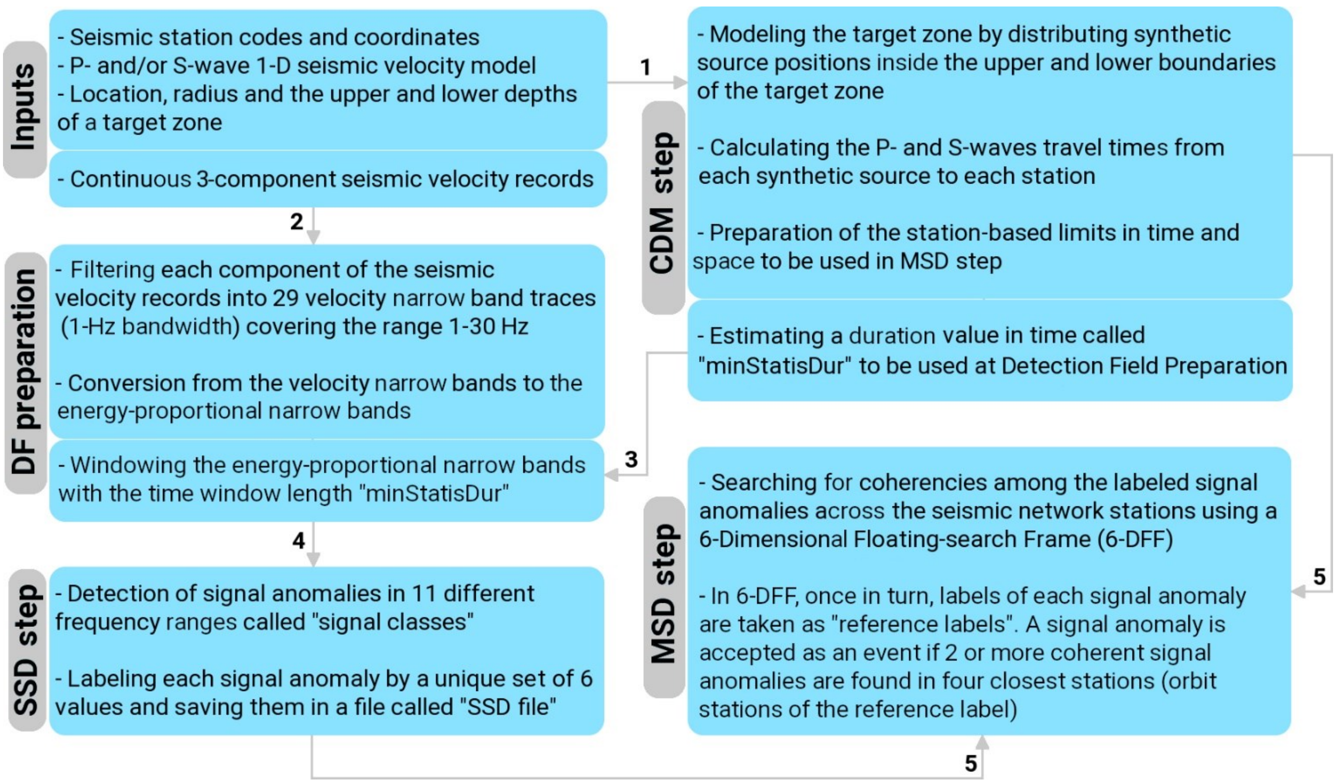


Fig. S1 Overview on the major steps of A6-DFMSD: Inputs, Configuration of the Detection Model (CDM step), Detection Field preparation (DF preparation), Single Station Detection (SSD step), and Multi-Station Detection (MSD step)

Then the program determines and sorts synthetic source positions on both, the upper and lower disk-shape surfaces of the target zone, by considering a uniform distance between the synthetic sources. The distance value is set in the control-file by the variable called ”minSor2SorDis”. In accordance with the test example of the program and its control-file values, Fig. S2a shows a top view of the sorted synthetic source positions for the case of the East Eifel Volcanic Field (EEVF), where the target zone has a radius of 25 km.

1.2.2 P- and S- phase travel-time calculation

The seismic velocity model of the seismic target zone (”velS” and ”velP”) and the stations coordinates of the seismic network (”stInfo”) plus applying Snell's law, enable us to calculate the synthetic arrival times of the P phases (Tp) and the S phases (Ts) which are emitted from each synthetic source and recorded at each station. The results of this calculation are used in following steps.

1.2.3 Calculation of the typical minimum signal duration

Depending on the relative position of each pair of a synthetic source and a recording station, we calculate and assign a waveform duration to each pair. In this part, the program finds the typical half minimum signal duration, potentially recordable at the seismic network stations. For this calculation, the program considers only the synthetic sources located on the upper depth position of the target zone and traces the rays that emit from the synthetic sources to all the stations listed in stInfo. Afterwards, for all the pairs of sources and stations the $Ts - Tp$ values are calculated. The geometric-mean value minus the geometric standard deviation value of all the calculated $Ts - Tp$ values is saved in a variable called ”minStatisDur”. If we assume that the duration of an earthquake signal is at least twice longer than the time difference between its Ts and Tp, then the value, which is saved as minStatisDur, is a good approximation for representing the typical half minimum signal duration of the events that occurs on the upper depth position of the target zone and is potentially recorded by the seismic network stations. Below in section 1.3, the value of minStatisDur is used as a duration in time to window the record traces.

1.2.4 Calculation of the station-based relative phase-delay differences

The results of this part provide time limits specified for each station for several usages.

After this calculation, we know for example: how long before or/and after the detection of a phase x (can be a P-phase or a S-phase) at station A, another phase y (also can be a P-phase or a S-phase) should be expected at station B, if the source of both phases is the same source and located somewhere inside the seismic target zone. This is needed for the coincident detection of an event at several stations to minimize false detections due to locally emitted noise signals. Figs. S2b and c visually explain the role of this calculation for the detection and the phase-delay differences relative to station example DEP02 and also relative to station example ABH. The relative time differences are calculated and plotted, respectively.

For all the synthetic sources the following calculation is done: we take into account the Ts values at each station pair (for example stations A and B) and compare these with each other, assuming the Ts values are the S-phase travel times emitted from the same synthetic source position, recorded at stations A and B. If the Ts value at station B (BTs) is bigger than the Ts value at station A (ATs), we allocate the value of BTs to a MATLAB array called ”positivPhasDelayMax” and the value of $BTs - ATs$ to a MATLAB array called ”positivPhasDelayMin”. And, in addition, if the ATs value is bigger than the BTs value, we allocate the value of $ATs - BTp$ to a MATLAB array called ”negativPhasDelayMax” and the value of $ATs - BTs$ to a MATLAB array called ”negativPhasDelayMin”. Here, the ATp and BTp values are the P-phase travel time values emitted from the same synthetic source position and recorded at stations A and B, respectively. In the corresponding MATLAB codes arrays are built called ”positivPhasDelayMax”, ”positivPhasDelayMin”, ”positivPhasDelayMax”, and ”negativPhasDelayMin”.In the MATLAB arrays, the row, column, and slice indeces represent station A, station B and source number I, respectively.

Considering all the stations listed in the stInfo in the control-file, such kind of calculation is done for all possible pairs of stations and synthetic sources, and the results are stored in the same arrays with unique indeces as described. To find the maximum possible positive-phase-delay between each unique pair of stations, we determine the maximum value of the slices elements in the array ”positivPhasDelayMax” and build up a new array called ”maxPositivPhasDelayMax”. This array in its slice has only one value connected to each pair station. The rows and columns of the new MATLAB array have the same index configuration as the root array ”positivPhasDelayMax”. The same is done with the row, column and slice elements of the ”negativPhasDelayMax” and the results are saved in the array ”maxNegativPhasDelayMax”. On the other hand, to find the lower limits in the positive delays connected to each pair of station, we find the minimum value of the slice elements of the array ”positivPhasDelayMin” and build up a new array called "minPositivPhasDelayMin". This array in its slice has only one value connected to each pair of stations. The rows and columns of the new array have the same index configuration as the root array "positivPhasDelayMin". The same is done with the row, column and slice elements of the "negativPhasDelayMin" and the results are saved in the array "minNegativPhasDelayMin".

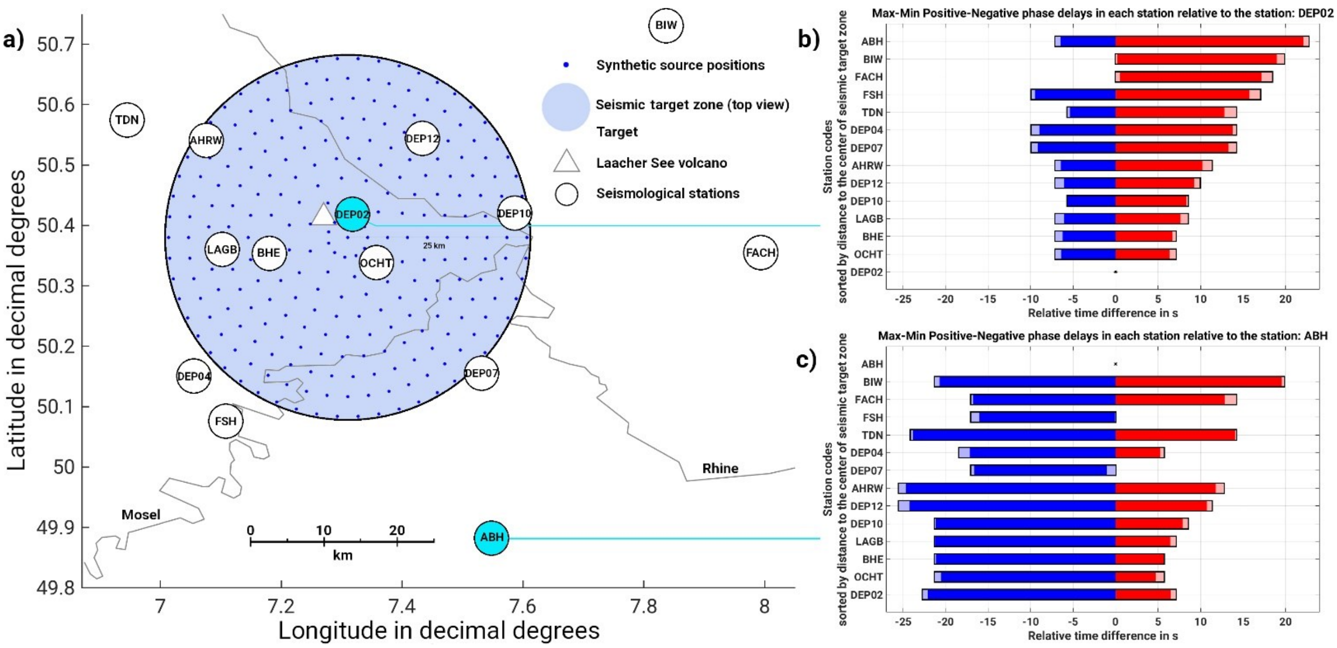


Fig. S2 Configuration of a detection model (at CDM step) in association with the test example in the EEVF. a) seismological station locations plus the top view of the seismic target zone considered in this study. In the CDM step each small blue dot is considered as a potential seismic source and the travel times for the first P- and S- phase arrivals are calculated for each source-station pair based on the input 1D seismic velocity model. b) subset of calculations in the CDM step to find the station-wise detection limits in the time domain. It is indicated how long before (in light blue) and/or after (in light red) the detection of a seismic phase at station DEP02 the routine searches for coherent seismic signals in the records of the other stations. c) the same b), but here the time differences are obtained relative to the recording station ABH. Note: in b) and c) the dark blue and dark red denote the exact time limits, and the light blue and light red indicate the time limits after involving time windowing (discretization) into the detection model

1.2.5 Calculation of the station-based typical negative phase delay for the local search

Results of this part are used in the MSD step section 1.5.4, where the program builds up a Global (net-wide) Detection list (GD-list) out of the Local Detection list (LD-list). The LD-list is indeed the result of the detection through the local coherency search and is built by writing information (date, time and signal class) of each detected event on separated lines in a text list. For building up the GD-list out of the LD-list (which is a kind of time clustering task), we need to know the maximum acceptable time difference between the declared events in the LD-list for clustering. The maximum acceptable time differences for clustering of the events in the LD-list should be obtained in a way that if, for example, for one event several detections are assigned in the LD-list, finally only one event is written in the GD-list. Also, this maximum acceptable time difference should be determined in a way that none of the events, which possibly occur inside the target zone, is omitted during clustering process.

To execute the above points, we do the following calculations: for each recording station X which contributes in the MSD step (section 1.5), first, we identify the four nearest stations to station X which are called the "orbit stations" (see Figure S3). Suppose these four nearest stations to the station X are Y1, Y2, Y3, and Y4 (we call them Yi). Then for each Yi, we calculate the mean plus the standard deviation value of the slice elements (which are relative phase-arrival time-differences between each station pair as explained above) in "negativPhasDelayMax" which correspond to the station X and stations Yi. The resulting value is allocated to a new array called "meanStdNegativPhasDelayMax", in which the row index points to station X and the column index to station Yi . When the calculation is finished for all stations Yi, we find the mean and its standard value of the elements in the array "meanStdNegativPhasDelayMax", in which the row points to the station X and the columns to the stations Yi. The result is stored in a one dimensional matrix called "netDelay", whose row index relates to station X. By applying the same operation to all stations, which contribute in MSD step, netDelay takes the values that specifically belong to each station.

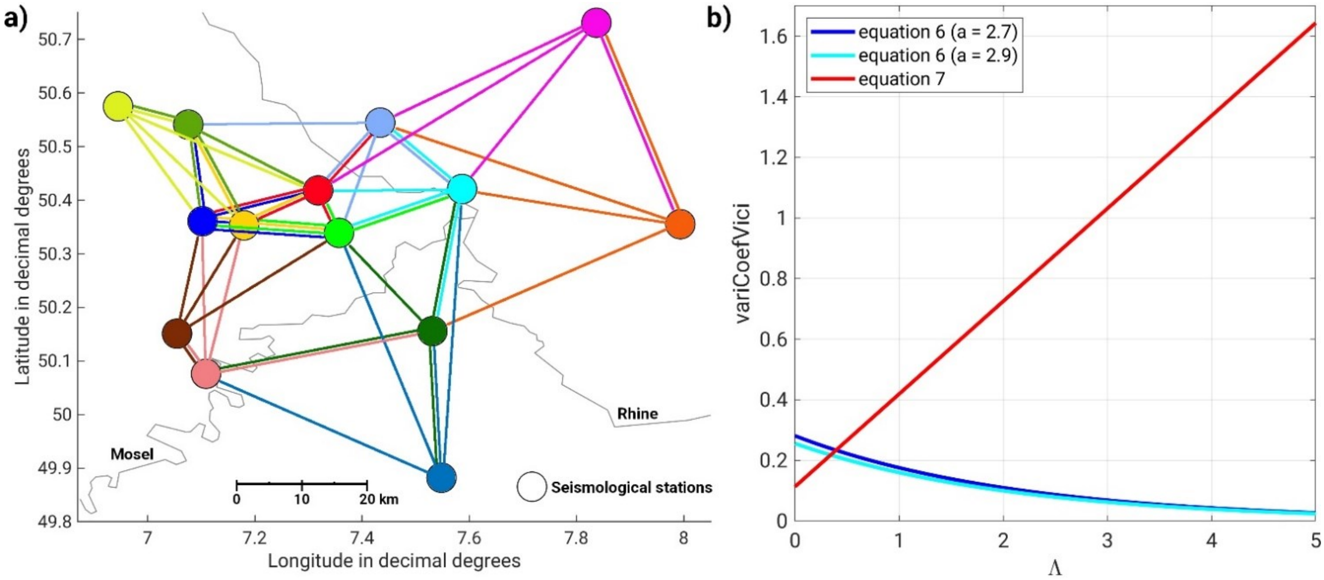


Fig. S3 a) Locations of recording stations and their relation with each other while searching and connecting coherent signals in the Multi-Station Detection step (MSD step). For each station, the search is limited to a radial distance in which only the four closest stations are involved. Here, each station is plotted with a specific color and the corresponding four nearest stations are linked by a line with the same color as the station, b) in MSD step, coherency search vicinities for the variation coefficient of DFs are obtained using empirical equations (6) and (7). Here this relationship is depicted in blue (dark and light) for equation (6), and in red for equation (7). The dark blue line shows this relationship when the goal is to search for tectonic and volcano-tectonic events (signal classes higher than 3) in the shallower zone (3-30 km depth), and the light blue line shows this relationship for possible tectonic and volcano-tectonic events in the deeper zone (30-50 km depth). The red line shows this relationship when searching for LF events (signal classes less than 4) in both seismic target zones

1.3 Detection Field Preparation (DF preparation)

As shown in the examples in Fig. S4a, b, c, and d, for each station, each seismogram component is at first decomposed into 29 1-Hz wide (narrow-) bands. Before filtering, seismograms are demeaned and detrend and after filtering one minute of the beginning and one minute of the end of the records are cut to remove the filtering artifacts. The upper and lower corner frequencies in Hz are listed in control-file as "monoBanList" and the order of tapering which we use is 3 in the frequency domain using the MATLAB signal processing toolbox. They cover the frequency range between 1 Hz and 30 Hz. Then correspondingly for each record sample i of each narrow-band j, a conversion proportional to the kinetic energy (called energy-proportional or E^*) is obtained (Fig. S3e) according to the following equation:

$$E^*_{j,i} = Z^2_{j,i} + N^2_{j,i} + E^2_{j,i} \tag{1}$$

where *Z*, *N* and *E* are the seismic velocity records for each component in e.g., m/s. Afterwards, the energy-proportional narrow bands (E^*_j) are windowed (Fig. S4f) by the window length of minStatisDur.

Windowing of the E^*_j traces is done when the mean value of the samples in each window is written in a new array called DFj (Detection Field of the band j). By this means, we efficiently mask/smooth traces to prevent redundant time-consuming operations in the next steps. Note: the number and width of the frequency bands may be chosen differently for other applications, e.g., for very low frequency events with f << 1 Hz.

1.4 Single Station Detection (SSD step)

During the SSD step, signal anomalies are detected and labeled with a unique set of six values which represent the properties of the signal anomalies.

1.4.1 Signal Anomaly Detection and Labeling

The tasks of the detection program in this part are:

- picking the signal anomalies in the detection field (DF) of each individual station,
- labeling their features, and
- saving the featues in a list.

Signal anomalies are independently detected in different frequency ranges called signal classes. In our example 11 signal classes are chosen (Fig. S4f). Each signal class is defined by two values denoting the upper and lower corner frequency of the signal class in Hz. For each signal class a Short Time Average (STA) amplitude to Long Time Average (LTA) amplitude ratio is determined with all DFs, which are within the upper and lower corner frequencies of the signal class. This leads to the detection criterion. A signal anomaly is detected, if all DFs within a signal class reach to the sample that has amplitude values higher than the mean value of the three previous sample amplitudes plus a multiplication of 0.7 (an empirical constant) to their mean absolute deviation. As an example, in Fig. S4g, a signal anomaly detection in signal class $n = 6$ is presented. The signal class $n = 6$ includes the DFs that are indexed from 8 to 16. In the example, the red dot (at sample i) is found as a signal anomaly, because at all DFs the amplitude values at sample i have a higher value than the mean plus a multiplication with 0.7 to the mean absolute deviation of the reference samples (shown as three green dots right before the sample i in the DFs). For each signal anomaly, which is detected in a signal class, a unique set of 6 values are saved as labels of the signal anomaly in a file called "SSD file". These six values are:

- the station code (implicitly, the latitude and longitude of the station location),
- the sample number of the detection point (i) representing the detection time,
- the signal class (n) which is defined by the upper border (j_2) and the lower border (j_1), and
- a value called "variation coefficient of DFs (Λ)" which is obtained by equation (2) and represents the energy* distribution status at the point of detection in the corresponding signal class. The value of Λ is indeed the standard deviation of the DF amplitudes located in the signal class n at the sample of detection i divided by the mean value of the DF amplitudes (μ) located at the same signal class n and at the same sample of detection i. The term $\mu_{n,i}$ used in equation (2) is obtained by equation (3).

$$\Lambda_{n,i} = \frac{\sqrt{\frac{1}{j_2 - j_1 + 1}\sum_{j=j_1}^{j_2}(DF_{j,i} - \mu_{n,i})^2}}{\mu_{n,i}} \tag{2}$$

$$\mu_{n,i} = \frac{1}{j_2 - j_1 + 1}\sum_{j=j_1}^{j_2} DF_{j,i} \tag{3}$$

- a value called "sigClasPower (Γ)" which is obtained by equation (4) and represents the amplitude level of energy* in the signal class at the detection point sample relative to the 3 previous samples just before the detection point sample (Fig. S4g). This value is equal to the result of the following division operation: the numerator is the mean value of the DF amplitudes of the signal class n at the detection point i ($\mu_{n,i}$) minus the mean value of the DF amplitudes located three samples right before the samples at the detection point (in Fig. S4h, this is equal to the length of the red line). The denominator is the mean absolute deviation of the mean values of the DF amplitudes for the three samples right before the sample at the detection point (in Fig. S4h, this is equal to the length of the blue line). The term $M_{n,i}$, which is used in the equation (4), is obtained by equation (5) and it is equal to the mean value of the mean value of the DF amplitudes three samples right before the sample of detection i in the signal class n.

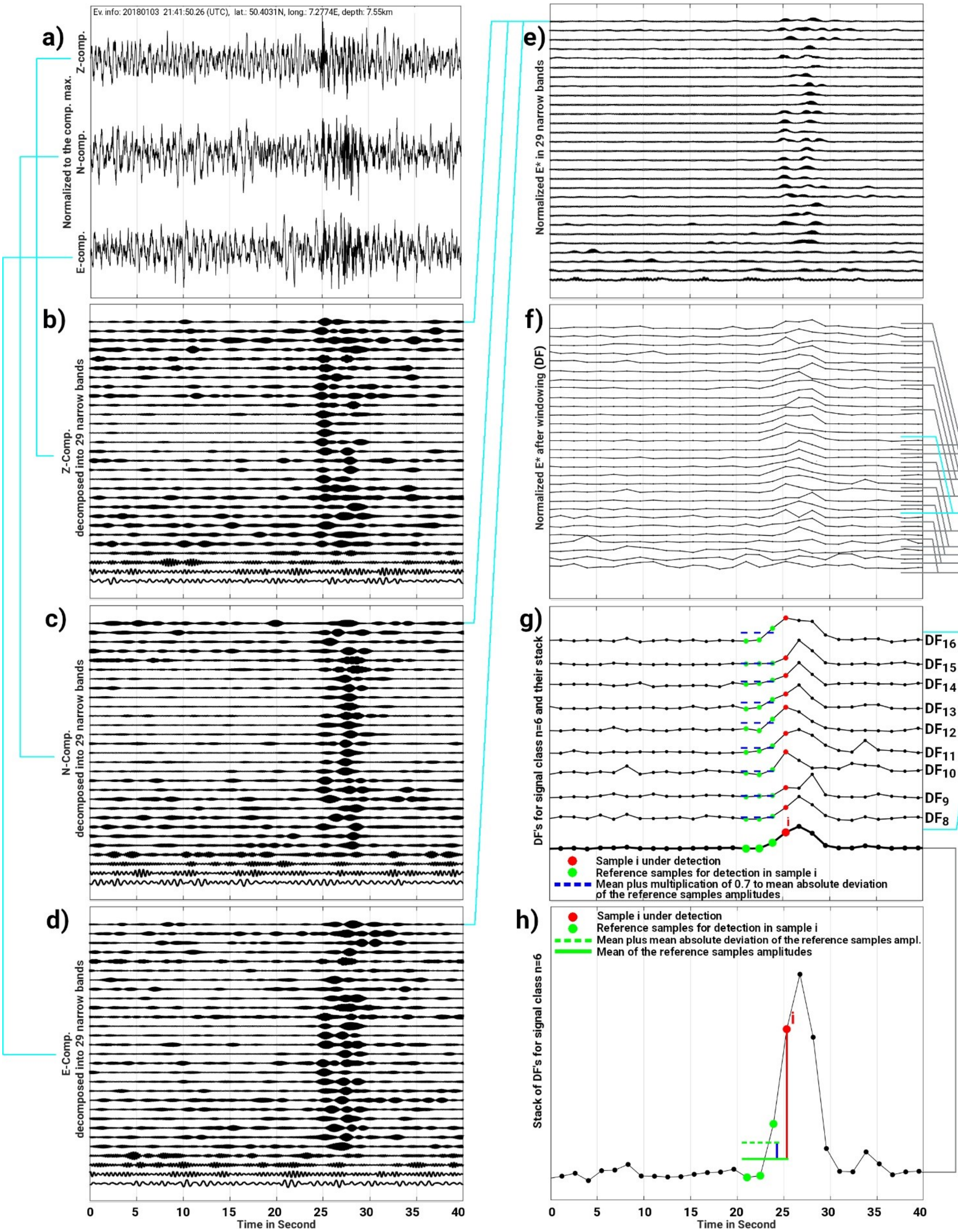
$$\Gamma_{n,i} = \frac{\mu_{n,i} - M_{n,i}}{\frac{1}{3}\sum_{i=i-3}^{i-1} \mu_{n,i} - M_{n,i}} \tag{4}$$

$$M_{n,i} = \frac{1}{3}\sum_{i=i-3}^{i-1} \mu_{n,i} \tag{5}$$

When the single station detection for one station is finished, the results are stored in a file with a unique name including the date and station code called SSD-file. Then it continues to find signals at the remaining stations listed in stInfo.

Next page:

Fig. S4 Examples for Detection Field preparation and Single Station Detection (SSD step). a) three-component band-pass filtered (1-30 Hz) ground motion velocity of a microearthquake recorded at station AHRW, b), c), and d) decomposition of each component Z, N, and E by filtering with 29 1-Hz width frequency bands, e) conversion of the ground motion velocity in each frequency band into energy-proportional narrow bands (E^*) following equation (1), f) the traces from e) are windowed using the window length of "minStatisDur" to prepare the Detection Fields (DFs) g) the detection status of a signal anomaly at the sample i in signal class *n* = 6 is shown; the red dot (sample i) is found as a signal anomaly because at all DFs located in signal class 6 the amplitude values at sample i have a higher value than the mean plus a multiplication with 0.7 to the mean absolute deviation of the reference samples (shown as three green dots right before the sample i in the DFs). The bold line is the stack of the DFs located in this signal class (DFs from DF_8 to DF_{16}), h) close-up view of the stacked trace in g). The stored value of $\Gamma_{6,i}$ for the signal anomaly at sample i in signal class *n* = 6 is equal to the ratio of the length of the red line to the blue line. The technique used for detecting signal anomalies (g and h) is similar to the STA/LTA method with a very sharp criterium (here we suppose that the window length in the LTA part is only 3 times longer than in the STA part)



Ev. info: 20180103 21:41:50.26 (UTC), lat.: 50.4031N, long.: 7.2774E, depth: 7.55km
a)
Z-comp.
N-comp.
E-comp.
Normalized to the comp. max.
b)
Z-Comp. decomposed into 29 narrow bands
c)
N-Comp. decomposed into 29 narrow bands
d)
E-Comp. decomposed into 29 narrow bands
e)
Normalized E* in 29 narrow bands
f)
Normalized E* after windowing (DF)
n=11
n=10
n=9
n=8
n=7
n=6
n=5
n=4
n=3
n=2
n=1
g)
DF's for signal class n=6 and their stack
DF16
DF15
DF14
DF13
DF12
DF11
DF10
DF9
DF8
Sample i under detection
Reference samples for detection in sample i
Mean plus multiplication of 0.7 to mean absolute deviation of the reference samples amplitudes
h)
Sample i under detection
Reference samples for detection in sample i
Mean plus mean absolute deviation of the reference samples ampl.
Mean of the reference samples amplitudes
Stack of DF's for signal class n=6
0
5
10
15
20
25
30
35
40
Time in Second

1.5 Multi Station Detection (MSD step)

During the MSD step the saved SSD-files of all the stations are loaded and the coherent signals are searched by applying a 6-Dimensional Floating-search Frame (6-DFF). To increase the reliability of the final detection list, before doing a global (net-wide) search, the program follows two strategies simultaneously:

1- In the 6-DFF, the search for coherent signals is limited within a certain radial distance (called local search distance) in which only the four closest stations relative to each station are involved in the search (see Fig. S3a). This radius differs from station to station depending on the network configuration and it is adjusted dynamically according to the location of each station in the seismic network. This strategy avoids the identification of coincident non-relevant signals (noise etc.) by the detection procedure. For example, the signals detected at the stations located at the two opposite edges of the network can be connected to each other, only if coherent signals at the stations, which are geographically located in between these two stations, were detected, too. This procedure make use of the wave propagation properties of direct seismic waves that waves arrive earlier at the stations closer to the source than at the more distant stations.

2- Local searches are done in two steps. During the first step, only those stations are involved that are characterized by a relatively low level of noise. The output of this step provides a list of events that is called ”confident list”. During the second step, the stations with a moderate level of noise are added in 6-DFF. In this step, the 6-DFF uses the time search limits only around the time points listed as events in the confident list. We explain more details about the station categorization regarding the level of noise, and the mechanism of the search in the following sections.

1.5.1 Station categorization

Before applying 6-DFF on the saved labeled signal anomalies, based on the number of detected signal anomalies in SSD-files, the program categorizes stations into three categories: removSt (removed stations), mastSt (master stations) and seconSt (secondary stations).

removSt are the stations whose number of detected signal anomalies in their SSD-files exceeds from a certain predefined level (this level is written at the control-file as ”maxEvPerDay” and we set it to 600 which means the maximum acceptable number of events per day for the records of each station is 600). The sites of these stations are very noisy or there are possible instrument failures (e.g.s the recording instruments generate pulses that make the records useless). As a part of data quality control, the program does not involve SSD-file of these stations into the search process.

mastSt are the stations located at the sites experiencing a relatively low level of transient noise. SSD-files of these stations have a relatively low number of signal anomalies. The confident list of events is generated using only these stations during the first step of the local search. Optionally, users of the program can force the detector to keep specific stations in the list of mastSt by adding their codes in the parameter called ”morImpoSt” in the control-file. This option is useful, if the location of a station (e.g. being close to the known active seismic sources) is more important than any possibly bad effect due to an increased noise level of the station.

seconSt are the stations with a moderate transient noise level. The information saved in the SSD-file of these stations are used only during the second step of the local search. Optionally, users of the program can force the detector to keep specific stations in the list of seconSt by adding their codes in the parameter called ”lesImpoSt”in the control-file. This option is useful when users want to reduce the influence of some stations to contribute in the detection process, due to a priori knowledge regarding, e.g., the noise conditions at those stations.

The separation of mastSt and seconSt is determined by using a simple statistical approach regarding the number of detected signals in all SSD-files. After loading all the saved SSD-files, the program sorts the stations according to the number of signal anomalies in the SSD-files of each station. Then it selects the stations whose number of signal anomalies is between the top 70% and 90% number of signal anomalies. Afterwards, it considers the rate of change in the number of detected signal anomalies among the selected stations. It finds the point where a relatively clear change of numbers of signal anomalies happens. The stations with a higher number of detected signal anomalies than the point of clear change are listed as the seconSt and the stations with a lower number are listed as mastSt.

1.5.2 The first step of the Local Search (LS-1)

The SSD-file of each station is a matrix with six columns, containing the detection information for each signal anomaly such as the time sample, the latitude and longitude of the station, the signal class, the variation coefficient of DFs in the signal class and the signal class power (as described in SSD step). In this part, the program focuses only on SSD-files of the mastSt to provide the confident list of events. For each mastSt, each row (line) of the corresponding SSD matrix can be listed as an event, if at least one of the four nearest mastSt in the vicinity also has a line in its SSD-file that satisfies the specific conditions of the 6-DFF. For example, suppose the mastSt are stations A, B, C, D, and E. Consider the program takes the first line of SSD-file of the station A as reference label. It searches in the lines of the SSD-file of the station B to find which of the following conditions are satisfied:

1- station B should be a member of the four nearest mastSt to station A

2a- time sample in station B <= time sample in station A + maxPositivPhasDelayMax(A,B)

2b- time sample in station A - maxNegativPhasDelayMax(A,B) <= time sample in station B

3- signal class in station A = signal class in station B

4a- variation coefficient of DFs in station B <= variation coefficient of DFs in station A + variCoefVici

4b- variation coefficient of DFs in station A – variCoefVici <= variation coefficient of DFs in station B

5- signal class power in both stations A and B >= sigClas2NoisTher

Here ”sigClas2NoisTher” takes its value from the control-file as the predefined minimum limit. The sigClas2NoisTher enables the program to eliminate the weak signal anomalies for which their relative amplitudes are less than a certain limit. Those signal anomalies are prevented from being involved into the detection.

If all the above five conditions logically become true, the program adds one to a counter variable, then it continues the search with station C (later, D and E) by applying the same procedure. If the counter has a value of 2 or more after searching with stations C, D, and E, the program writes the first line of SSD-file of station A into a new matrix called ”confEvList” as a confident list of events. This process is done for all the lines at SSD-file of stations A, B, C, D, and E. After building up the new array of ”confEvList”, the program continues with the second step of the local search.

The value of variCoefVici in the above conditions is determined by empirical equations (6) and (7).

$$\text{variCoefVici} = 1.6^{-(\Lambda+a)} \tag{6}$$

a is a constant which can take a minimum value of 2.7. Following our trial error tests concerning the ratio of false detections to total number of detections, we found values 2.7 and 2.9 are suitable choices for *a* when the upper border of the target zones is set to 3 km depth and 30 km depth, respectively. Fig. S3b shows the relationship between variCoefVici and Λ while detecting events in different signal classes in different defined seismic target zones in our test example. In the control file of detector, the constant variable *a* is renamed to "*variCoefViciEffectFactor"* for having clearer code. Note *a* may change for other application and possibly should be testedas outlined above.

Note: In the main manuscript on the detector variCoefVici is called b for simplicity.

$$\text{variCoefVici} = 0.306 \times \Lambda + 0.113 \tag{7}$$

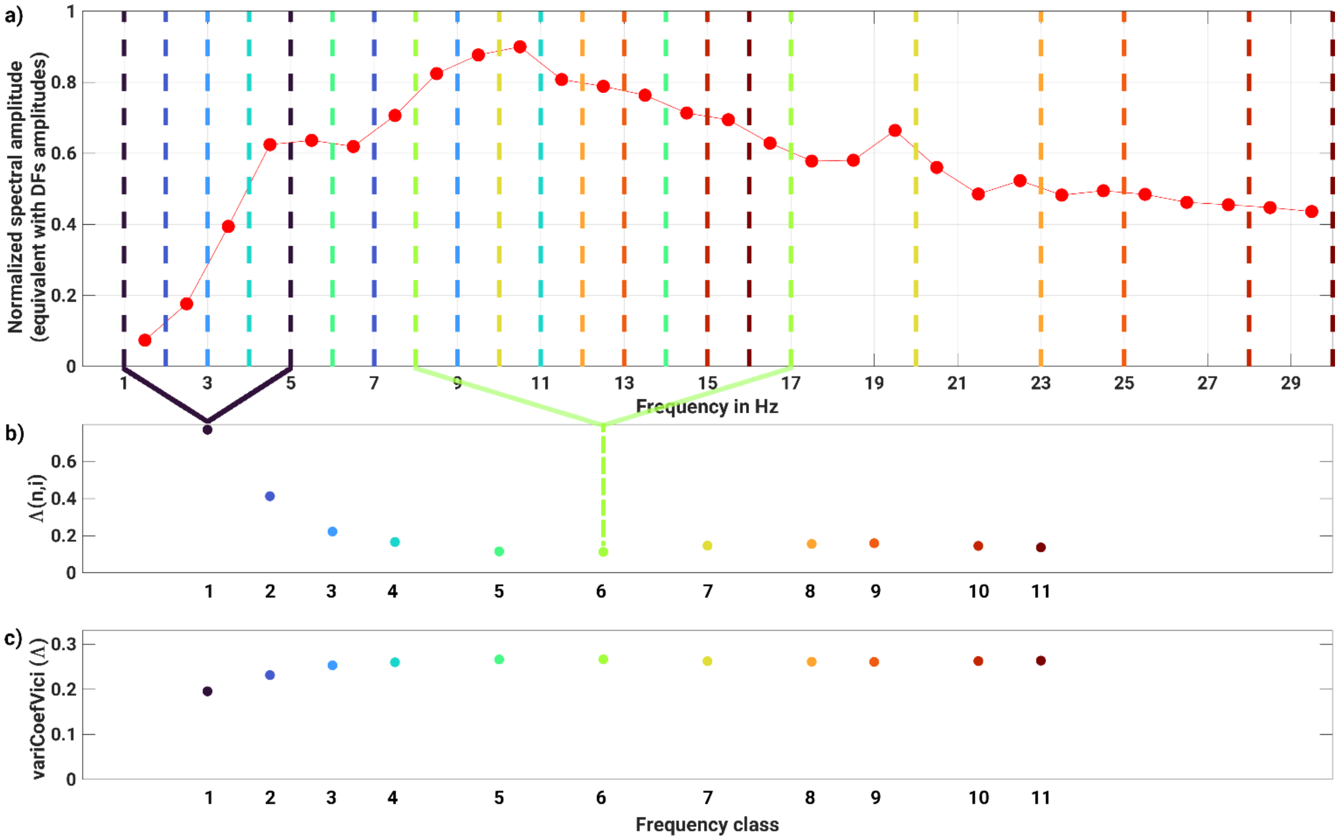


Fig. S5 Effect of equation (6) on the determination of *variCoefVici*. a) normalized spectral amplitude of a typical tectonic microearthquake in EEVF and borders of defined signal classes. Borders of signal classes are marked by coloured dashed lines. Each colour represents one signal class. b) value of Λ obtained for each signal class is plotted. c) *variCoefVici* obtained for each signal class in connection with corresponding Λ is plotted. Signal classes having max. and min. values of Λ are selected and marked to demonstrate extreme effects of using equation (6)

In equation (6), the coherency search gets a wider border when Λ at the reference label has smaller values. In contrast, in equation (7) the coherency search gets a wider border when Λ at the reference label has bigger values. The borders determined by equation (6) are suitable for the detection of the signals which show less fluctuations in the amplitude of their frequency spectra around their dominant frequencies, like tectonic type events. In contrast, the borders determined by equation (7) are suitable for the detection of signals which show large fluctuations in the amplitude of their frequency spectra around their dominant frequencies. This occurs for example for LF type seismic events and their signals.

Figs. S5 and S6 respectively present the effect of equations (6) and (7) on the determination of *variCoefVici*. In Figs. S5a and S6a normalized spectral amplitudes are presented from typical tectonic and LF microearthquakes in the EEVF, respectively. In A6-DFMSD, the spectral amplitude of the events represents DFs of the events, especially when the spectrum is windowed with the window length of 1 Hz.

In Figs. S5 and S6, the borders of 11 frequency classes are highlighted by coloured dashed lines in which each colour indicates borders of a signal class. In Figs. S5b and S6b the value of Λ for each signal class is plotted. Correspondingly, the value of *variCoefVici* for each signal class in connection with Λ is plotted in Figs. S5c and S6c. In Figs. S5b and S6b maximum and minimum values of Λ are marked. As it is shown in Fig. S5a, where the spectral amplitude experiences the most fluctuation (at signal class 1) the smallest value is selected for *variCoefVici* (Fig. S5c) to reduce the sensitivity of the coherency search in this signal class. In contrast, where the spectral amplitude experiences the least fluctuations (at signal class 6) the largest value is selected for *variCoefVici* (Fig. S5c) to increase the sensitivity of the coherency search in this signal class. In Fig. S6a, where the spectral amplitude experiences the most fluctuation (at signal class 3) the largest value is selected for *variCoefVici* (Fig. S6c) to increase the sensitivity of the coherency search in this signal class. In contrast, where the spectral amplitude experiences the least fluctuations (at signal class 7) the smallest value is selected for *variCoefVici* (Fig. S6c) to reduce the sensitivity of the coherency search in this signal class.

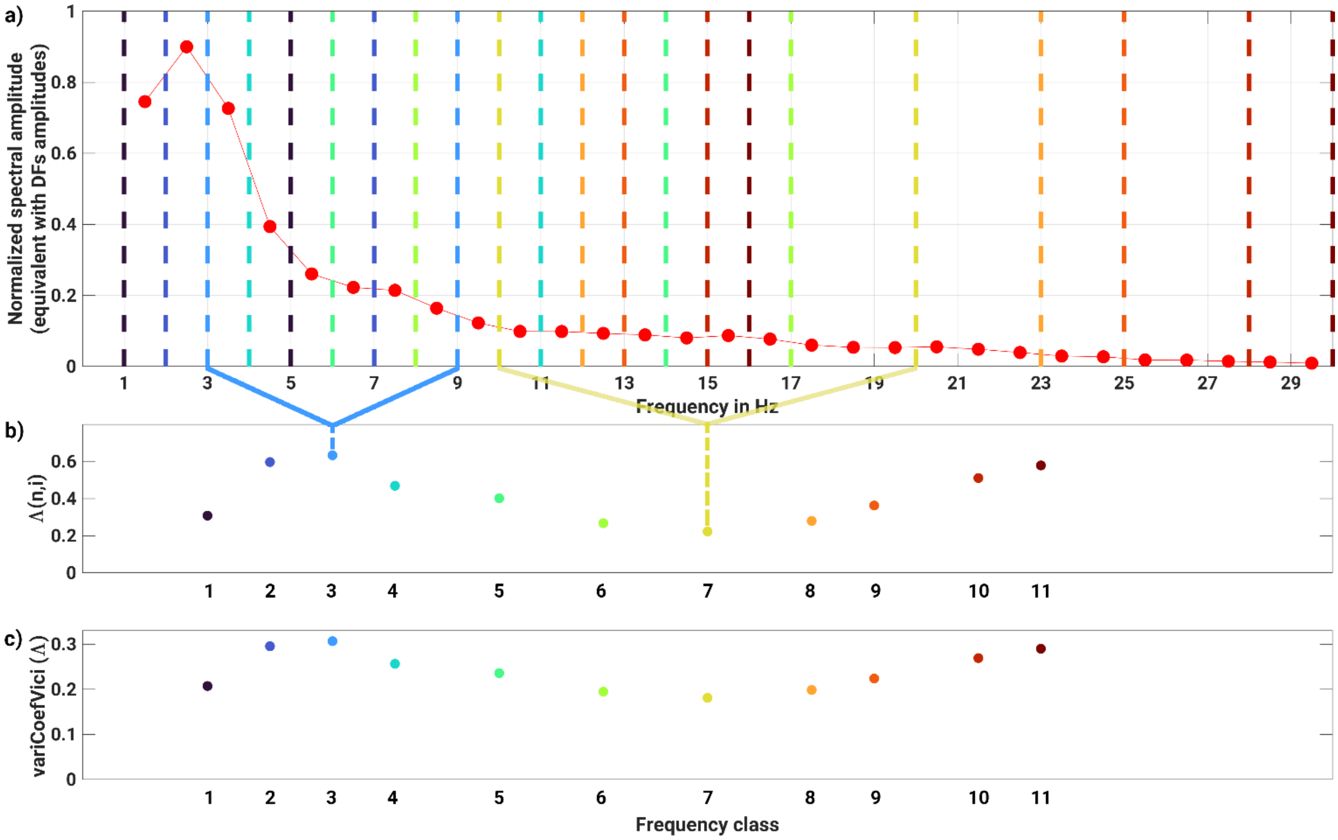


Fig. S6 Effect of equation (7) on determination of *variCoefVici*. a) normalized spectral amplitude of typical LF microearthquake in EEVF and borders of defined signal classes. Borders of signal classes are marked by colourful dash lines. Each colour represents one signal class. b) value of Λ obtained for each signal class is plotted. c) *variCoefVici* obtained for each signal class in connection with corresponding Λ is plotted. Signal classes having max. and min. values of Λ are selected and marked to demonstrate extreme effects of using equation 7

1.5.3 The second step of the Local Search (LS-2)

After LS-1, the program takes each line of the confident list as a reference line and search in the SSD-files of the secondSts for any coherent signal anomaly. It uses the same conditions through a local search on the SSD-files like the ones which were used in the LS-1, but this time it involves both categories of stations including mastSt and secondSt (not only the mastSt) and it declares a detected event, if the counter variable exceeds a value of 2.

This selection process can be regulated by the parameters: ”minLocSearCohStNum” and “maxLocSearStNum“ in the control-file. By this means, all the secondary stations, which have recorded signals of events (e.g., not disturbed due to high noise), get a chance to be involved in the detection. At the end of this step, the local-search-detection-list (LSD-list) is prepared. There, for each detected event, station codes of all the involved stations in event declaration, either mastSt or secondSt, are written next to the event date and time point, and event signal class.

1.5.4 Global (net-wide) Search (GS)

Each line of the LSD-list includes date, time, signal class, and the station codes of the stations that have contributed to the detection of the event. The task of this part of the program is time clustering of the events from the LSD-list to build up a global detection list. To describe how the program works, we explain this by an example: consider the first two lines of a LSD-list. We need to know, whether the event at the second line is connected to the first line or not. If the sources of the events given in both lines are the same and located in the seismic target zone, the time difference between the declared events would never exceed a certain limit of time. This time limit can be estimated by referring to the netDelay elements which have already been provided in section 1.2.5. If, for example, stations A, B, and C are the stations in the second line of the LSD-list, we correspondingly find the elements in netDelay which are pointed by the stations A, B, and C. Regarding the seismic target zone, the maximum value of those elements is estimated to be the maximum acceptable time difference between the second and the first line of the LSD-list. If the time difference between the first and the second line is less than the estimated value, the program accepts both lines as one event and writes the following information into a line in a new list called ”GD-list” (Global Detection list): the date and the time of the first line in the LSD-list, concatenated station list resulting from the first and the second line, the time difference between the first and the second line as the network-wide signal duration, and the signal class of the event (if the event is the result of merging several lines, the signal class, which is written in the GD-list, is the value of the signal class that is most often found as signal class within the collected lines).

1.6 Output files:

Detected events for each day of data is listed in a set of txt files. Each txt file contains a certain part of the detected events. For example, the detection results in connection with the target zone-1 are found in a txt file which its file name contains the word “Zone-1”. Also, the events which were detected by a minimum number of X stations (X is a number) are found in a txt file which its file name contain the word “StX”. In this way, for example, if one wants to see the events detected by minimum number of 5 stations in the data of the date February 1, 2023 and in connection with the target zone-1, must see the txt file name “date20230201St5Zon-1.txt” in the set of the outputs.

2 CONTROL-FILE ADAPTATION IN ASSOCIATION WITH THE TEST EXAMPLE

We adapted the control-file of the program (Table S2) to search a wide range of signals. This is due to the previous studies, concerning the variety of the observed signals in the East Eifel Volcanic Field (Hensch et al., Geophys. J. Int., 216, pp 2025–2036, 2019). The search covers the frequency ranges belonging to the observed LF events and tectonic events by considering 11 band width search limits (11 signal classes). We set the seismic target zone to have a center between the epicenter of the reported deep LF events and the center of the Laacher See crater, with the radius of 25 km. The control-file parameters are adjusted to search events possibly occurring in two different depth zones limited between: 3 km and 30 km, and 30 km and 50 km. The used velocity model is KIT5 by Ritter et al., J. Seism. (in revision).

Table S1. Control file parameters

Short Description	**Parameter Name**
Station codes and coordinates in decimal degree	stInfo
P-wave velocity in km/s per depth	velP
S-wave velocity in km/s per depth	velS
Minimum source distance between synthetic source positions	minSor2SorDis
Location of the target seismic zone center in decimal degree	targCen
Target seismic zone radius in km	targRadi
Target seismic zone upper depth in km	targDep1
Target seismic zone lower depth in km	targDep2
Signal to noise window length ratio (integer value)	sig2NoisWinLenRatio
Narrow-bands signal to noise power threshold	monoBanSig2NoisTher
Band width search limits (signal classes)	sigClas
Decomposing narrow-band list	monoBanList
Max acceptable number of events per day	maxEvPerDay
Min. threshold for searching coherencies in sigClasPower	sigClas2NoisTher
Variation coefficient search vicinities (in eq. 6) are controlled by:	variCoefViciEffectFactor
Local search and event declaration condition: minimum number of stations with coherent signal anomaly maximum number of stations involved in local event search	 minLocSearCohStNum maxLocSearStNum
Regarding the noise level or station position relative to the target: More important stations (optional) Less important stations (optional)	 morImpoSt lesImpoSt

Table S2. Control file (Parameters) used in this study

Parameter Name	Parameter Value
stInfo	[DP02 50.418 7.318],
	[DP04 50.150 7.055],
	[DP07 50.155 7.531],
	[DP10 50.420 7.586],
	[DP12 50.544 7.433],
	[ABH 49.882 7.548],
	[FACH 50.355 7.993],
	[OCHT 50.339 7.357],
	[LAGB 50.360 7.102],
	[FSH 50.076 7.109],
	[AHRW 50.541 7.076],
	[BHE 50.353 7.180],
	[TDN 50.575 6.945],
	[BIW 50.731 7.837]
velP velS Depth	[5.53 3.24 -0.9],
	[5.53 3.25 1],
	[5.65 3.32 2],
	[5.75 3.38 3],
	[5.75 3.55 5],
	[6.09 3.73 8],
	[6.10 3.77 10],
	[6.20 3.77 12],
	[6.32 3.80 15],
	[6.43 3.83 19],
	[8.00 4.76 70]
minSor2SorDis	3
targCen	[50.38 7.31]
targRadi	25
targDep1 (zone1, zone2, zone3)	3, 30, 45
targDep2 (zone1, zone2, zone3)	30, 50, 70
sig2NoisWinLenRatio	3
monoBanSig2NoisTher	0.7
sigClas	[1 5], [2 7], [3 9], [4 11], [6 14], [8 17], [10 20], [12 23], [13 25], [15 28], [16 30]
monoBanList	[1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30]
maxEvPerDay	600
sigClas2NoisTher (zone1, zone2, zone3)	2.5, 2, 2
variCoefViciEffectFactor (zone1, zone2, zone3)	2.7, 2.9, 3.1

minLocSearCohStNum	3
maxLocSearStNum	5
morImpoSt	DP12, AHRW, OCHT, BHE, LAGB
lesImpoSt	DP04, FSH