

Suppression of Periodic Disturbances in Seismic Aftershock Recordings for Better Localisation of Underground Explosions

FELIX GORSCHLÜTER¹ and JÜRGEN ALTMANN¹

Abstract—For precise localisation of a potential underground nuclear explosion, the Comprehensive Nuclear-Test-Ban Treaty Organization, during an on-site inspection, can set up seismic sensors to find the very small signals from aftershocks. These signals can be masked by periodic disturbances from, for example, helicopters. We present a new method to characterise every such disturbance by the amplitude, frequency and phase of the underlying sine in the time domain using a mathematical expression for its Hann-windowed discrete Fourier transform. The contributions of these sines are computed and subtracted from the complex spectrum sequentially. Two examples show the performance of the procedure: (1) synthetic sines superposed to a coal-mine induced event, orders of magnitude stronger than the latter, can be removed successfully, (2) removal of periodic content from the signals of a helicopter overflight reduces the amplitude by a factor 3.3 when the frequencies are approximately constant. The procedure cannot yet cope with peaks that change frequency too fast, for example by the Doppler effect when passing, and with peaks that lie too close to each other. Improvement to solve these problems seems possible.

Key words: Seismic, localisation of underground explosions, spectral analysis, periodic noise reduction.

1. Introduction

The Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) builds up and maintains a global network of sensors to detect every underground explosion with a yield of 1 kT TNT-equivalent or better. Teleseismic detections result in localisation uncertainties on the order of 10 km. For a more precise determination of the hypocentre of the explosion and to find further indicators as to whether the explosion was a nuclear one, the CTBTO can carry out an on-site inspection (OSI) in the area of

interest if the country is a Treaty party. A seismic aftershock monitoring system (SAMS) can be placed at the surface to detect the very small vibrations produced by relaxations in the rock around the cavity. However, helicopters and vehicles used by the inspectors, noise from existing infrastructure in the country, or even intended disturbance attempts can generate seismic signals which can mask the weak aftershock signals.

Many man-made noise sources (engines, etc.) are of a periodic nature, and airborne sound can couple into the ground. Periodic signals show up as peaks in the frequency spectrum. The weak signals of aftershocks, on the other hand, are of a pulsed shape, and their spectrum is broadband. With the Fourier transform (which converts a certain interval of time-domain data into its spectrum) these properties give the opportunity to distinguish between disturbing periodic noise and impulse-type aftershock events (Fig. 1). We investigate whether the disturbing peaks can be characterised and subtracted from the superposed spectrum, so that the broadband content of the impulsive event remains.

Removing periodic noise can be done by traditional methods. If the noise peaks lie above or below the frequency range of interest, a simple low- or high-pass filter can remove them effectively. If the disturbing spectral peaks overlap with the signal spectrum—which is often the case if the latter is broadband—one can use notch filters individually tuned to the frequencies contained in the disturbance. This requires finding these frequencies and is not adapted to the different strengths of the peaks. Different methods for noise reduction in seismic data are described in ROBINSON and TREITEL (1980), BUTTKUS (2000), PETER BORMANN (2009). To better remove the peaks, we are investigating a new method that exactly

¹ TU Dortmund, Dortmund, Germany. E-mail: gorschlueter@e3.physik.tu-dortmund.de; altmann@e3.physik.tu-dortmund.de

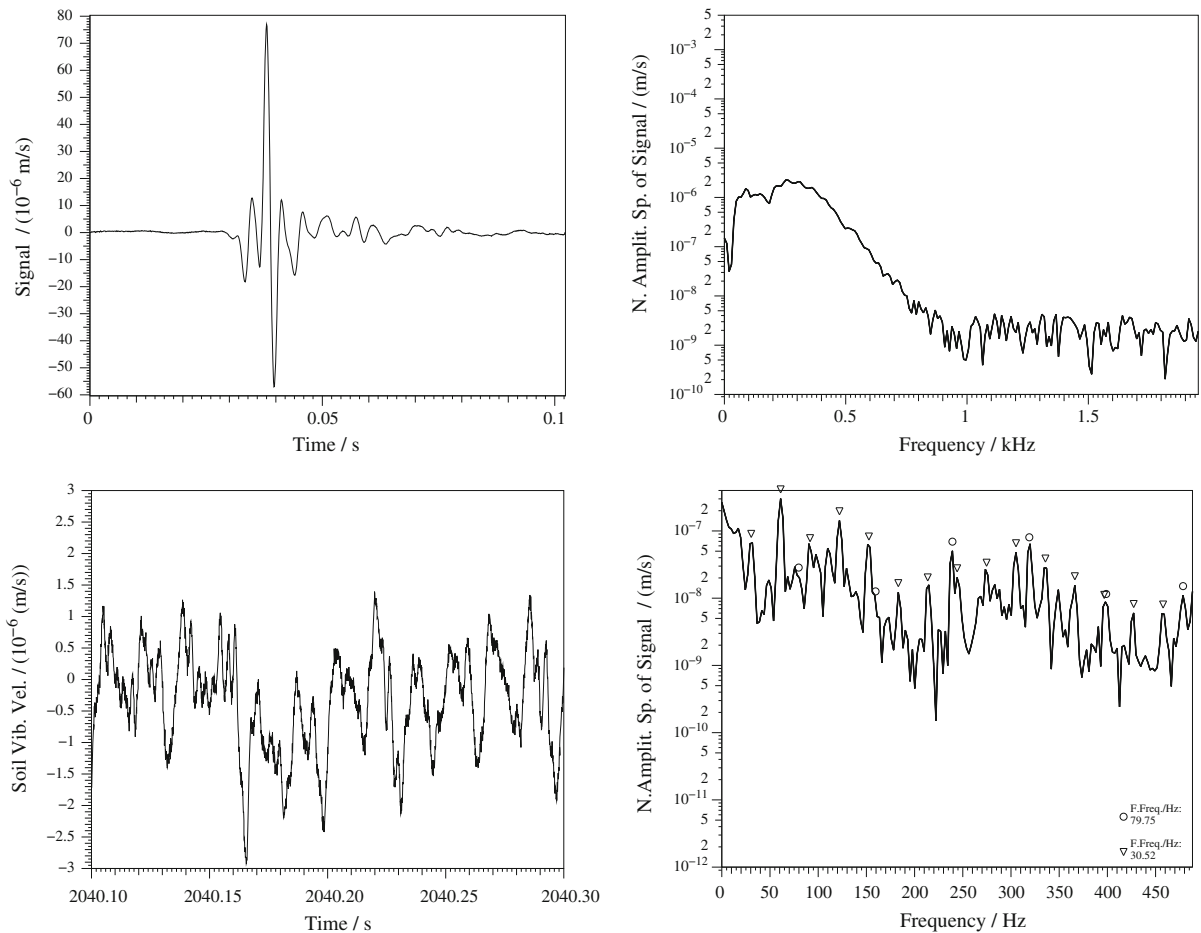


Figure 1

Demonstration of periodic and impulse events: seismic signal (*left*) and spectrum (*right*) of a firecracker at about 100 m distance from a geophone (*top*) and the seismic signal of a flying helicopter at about 1 km distance (*bottom*). In the spectra of the helicopter two harmonic series can be found stemming from the main (*triangles*) and the tail (*circles*) rotor, respectively

characterises the peaks by amplitude, frequency and phase and subtracts their single spectra one after the other from the complex spectrum of the given time series.¹ The method will be explained in mathematical detail elsewhere. It promises to complement existing methods of spectral estimation, in particular for finding frequencies contained in a signal, in the time domain, such as the autoregressive moving average (ARMA) model, or in the spectral domain, such as the averaged periodogram or minimum variance spectral estimation (e.g. KAY 1988; OPPENHEIM and SCHAFFER 1999). Its first step can be seen as an

extension of the periodogram technique by finding the peaks of spectral power, but for accurate determination of the phase then it works in the complex domain where averaging becomes meaningless. We assume that the sines in the time series have constant frequency during the time interval for one spectrum, but that the amplitude, frequency and phase can change from interval to interval, so that again averaging techniques cannot be applied. Here we report on the application of our peak-identification and -subtraction method to the problem of periodic disturbances in seismic aftershock measurements as they are stipulated by the CTBT for OSIs.

The next section describes the analytical expression for the discrete spectrum of a single sine and the

¹ First results have been published in a conference poster: JÜRGEN ALTMANN and FELIX GORSCHLÜTER (2009).

algorithm used to fit it to the data. Section 3 shows examples of an application of the peak fitting and subtraction algorithm, and Sect. 4 discusses the results.

2. Background and Theory

2.1. Analytical Expression of the Complex Spectrum of a Monofrequent Sine

We assume that in each spectrum the periodic content does not change in time. This is often at least approximately fulfilled over the time interval used for one discrete spectrum. A signal consisting of periodic contributions can be expressed by a superposition of sine functions. Each such sine has an amplitude A_0 , a frequency ν_0 and a phase ϕ_0 , its continuous time course is

$$s(t) = A_0 \sin(2\pi\nu_0 t + \phi_0) \quad (1)$$

Three facts need to be taken into account (BRIGHAM 1988):

- Real signals can only be handled for a finite duration, and they can change over time, so only a short interval of data is transformed into its spectrum. Mathematical this means multiplication with a rectangle function
- Because of the first two items mentioned the discrete spectrum of a sine does not consist of two δ functions at the frequency of the sine and its negative [as the Fourier transform of the continuous $s(t)$ would be] but depends on the position of the frequency ν_0 with respect to the equidistant comb of discrete frequencies of the spectrum. To reduce the ensuing spectral leakage $s(t)$ is multiplied with a window function (Hann window in our case)

Multiplications in the time domain are equivalent to convolutions in the frequency domain causing a multiplicity of terms. After the convolution of the terms described in Table 1 the analytical expression of the spectrum of a single, monofrequent sine becomes:²

$$G(\nu) = i \frac{A_0}{4\sqrt{3}N} \left[\sin(\pi T(\nu - \nu_0)) \left(2 \cot(\pi \Delta t(\nu - \nu_0)) - \cot\left(\pi\left(\Delta t(\nu - \nu_0) - \frac{1}{N}\right)\right) - \cot\left(\pi\left(\Delta t(\nu - \nu_0) + \frac{1}{N}\right)\right) \right) \cdot e^{i(\pi T(\nu - \nu_0) - \phi_0)} - \sin(\pi T(\nu + \nu_0)) \left(2 \cot(\pi \Delta t(\nu + \nu_0)) - \cot\left(\pi\left(\Delta t(\nu + \nu_0) - \frac{1}{N}\right)\right) - \cot\left(\pi\left(\Delta t(\nu + \nu_0) + \frac{1}{N}\right)\right) \right) \cdot e^{i(\pi T(\nu + \nu_0) + \phi_0)} \right] \quad (2)$$

- Real data are gained by analogue-digital-converters (ADC) by sampling the continuous signal $s(t)$ with a certain rate (e.g., the CTBTO uses 500 Hz when performing OSI exercises); thus the data are a sequence of discrete values. This is equivalent to multiplying $s(t)$ with an equidistant Dirac comb

The numerical factor arises from the normalisation (periodogram and additionally Hann window). In the discrete Fourier transform (DFT), in order to create a discrete spectrum the result of the three steps in Table 1 is sampled in the spectral domain by

² The derivation will be published elsewhere.

Table 1

List of the components used to get the analytical expression of a sine with the properties described above

	Time domain	Frequency domain
Sine	$s(t) = A_0 \sin(2\pi\nu_0 t + \Phi_0)$	$S(\nu) = i\frac{A_0}{2} [\delta(\nu + \nu_0) - \delta(\nu - \nu_0)] e^{i\Phi_0 \mp \nu_0}$
Rectangle	$\Pi\left(\frac{t}{T} - \frac{1}{2}\right)$	$T e^{-i\pi T\nu} \text{sinc}(T\nu)$
Hann window	$\frac{1}{2} - \frac{1}{2} \cos\left(\frac{2\pi t}{T}\right)$	$\frac{1}{2} \delta(\nu) - \frac{1}{4} \delta\left(\nu + \frac{1}{T}\right) - \frac{1}{4} \delta\left(\nu - \frac{1}{T}\right)$
Dirac comb	$\frac{1}{\Delta t} \text{III}\left(\frac{t}{\Delta t}\right) = \frac{1}{\Delta t} \sum_{k=-\infty}^{\infty} \delta\left(\frac{t}{\Delta t} - k\right)$	$\text{III}(\Delta t \cdot \nu) = \sum_{k=-\infty}^{\infty} \delta(\Delta t \cdot \nu - k)$

By Fourier theory, the terms on the right have to be convolved. Note that only the expression for the rectangle does not consist of δ functions, so the overall result is a sum of infinite repetitions of the sinc functions. The δ functions of the Fourier transform of the Hann window are so close to each other ($\pm\Delta\nu = \pm\frac{1}{T}$) that summing up these three sinc functions (by convolution) results in a broadened peak at the frequencies $\pm\nu_0$

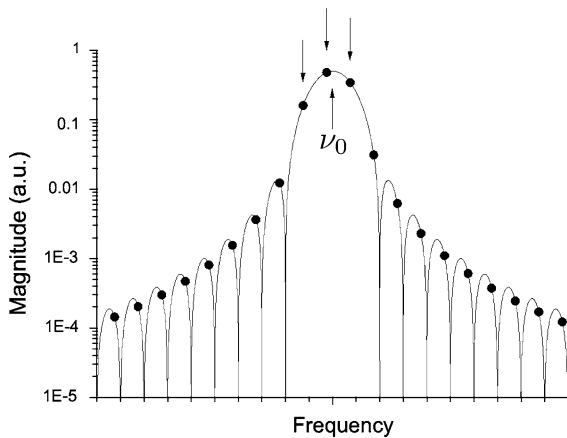


Figure 2

Plot of the absolute magnitude of $G(\nu)$ (Eq. 2) which is the theoretical spectrum of one sine, region around the peak. For a given sampling rate and number of samples the position of the discrete frequency values (dots) is fixed. The centre frequency ν_0 in function $G(\nu)$ determines the values at the discrete frequencies. One example: If ν_0 is exactly equal to one of the discrete frequencies, all magnitude values except for the highest three become zero

multiplication with a Dirac comb of spacing $\Delta\nu$ where $\Delta\nu = \frac{1}{T}$ is the inverse of the duration $T = N \cdot \Delta t$ with N the number of samples used as input, and Δt is the sampling period. This corresponds to convolving the time-domain result with the inverse transform, a Dirac comb of spacing T , producing infinite repetition that can also be seen as cyclic. If one takes Eq. (2) at the discrete frequencies $n \cdot \Delta\nu$, one gets the values of the discrete spectrum (Fig. 2). In a real recording several such contributions with different frequency, amplitude and phase are superposed, plus possibly additional contributions, for example from broadband noise. The additional sines need not be harmonics of a fundamental frequency.

2.2. Fitting Peak Parameters to a Given Spectrum and Peak Subtraction

First the discrete Fourier transform is applied to a section of the data because the algorithm works in the spectral domain. In order to find the relevant peaks, the next step is to find all the local maxima in the power spectrum and reject from them all the candidates that do not fulfil certain criteria. This is done to avoid false-positive fittings (especially in this experimental state of the algorithm where the thresholds are not yet optimised) and to save computing time by using the knowledge that the fitting will not be satisfactory.

- The line width must not exceed a certain value. It is defined as the sum of the magnitudes (absolute values)³ of the highest value of a peak and its two neighbours divided by the magnitude of the highest value. The result depends on the position of the frequency relative to the grid of discrete frequencies and ranges from 2.0 to 2.2 for a monofrequent sine. For real signals we use an empirical threshold of 2.6 in order to tolerate noise and small frequency shifts which of course lead to line broadening
- Fitting candidates must be stronger than the general frequency-dependent background of the respective sensor and the specific background of the actual spectrum in the vicinity of the peak by certain factors (here 4.0 and 20 in power, respectively). The specific background is modelled by a

³ Here and in the following we use “magnitude” in the mathematical, not the seismological sense.

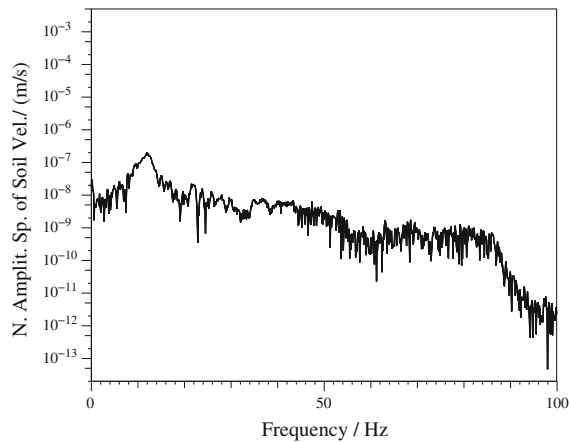


Figure 3

Magnitude spectrum of a seismic event in a coal mine

piecewise linear curve representing the valleys of the spectrum (here consisting of 20 parts)

- At each side, two discrete frequency values closest to the zero as well as to the Nyquist frequency are skipped to avoid complications with the local-maximum search and by mirrored peak components

After sorting the remaining candidates by amplitude, the algorithm is fed one by one with peaks of decreasing amplitude. Each single peak is mainly formed by three to four samples (Fig. 2). These samples have most of the power and thus are relatively least influenced by flanks of peaks at other frequencies or by the underlying signal. For each peak the three samples with highest spectral amplitudes are used. The expression $G(v)$ is fitted to their complex values by simultaneously varying the peak parameters amplitude A_0 , frequency ν_0 and phase ϕ_0 by the non-linear Levenberg–Marquardt method (PRESS *et al.* 1989). Start values for the three parameters are gained from the three samples:

- The start amplitude is the sum of the highest three magnitudes of a peak times $\sqrt{2}$
- The start frequency is the weighted mean frequency of the peak. The three frequencies with the highest magnitude are multiplied with their corresponding magnitude values, summed up and divided by the sum of the three magnitude values
- The start phase is calculated as: $\frac{\pi}{2}$ plus the phase of the value with highest magnitude (the 2π -arctangent of the complex number) minus $1.35 \cdot \pi \cdot a$

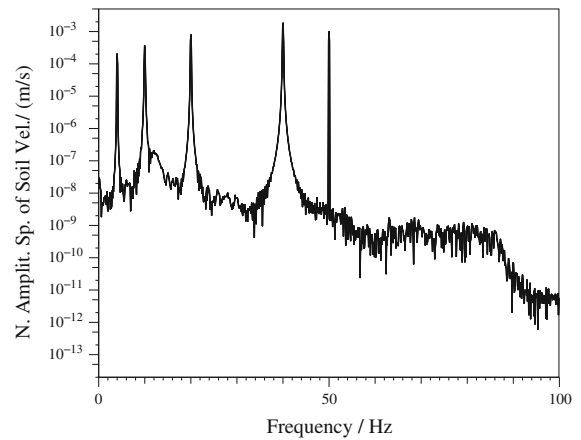


Figure 4

Magnitude of sum spectrum after superposing the complex spectra of Fig. 3 with artificial sines with amplitudes 10^3 – 10^4 times the amplitude of the seismic event. The peak at 50 Hz is narrow because the sine frequency nearly coincides with a discrete frequency

where a is the decimal fraction (non-integer part) of the start frequency in units of $\Delta\nu$.

These expressions have been derived from consideration of Eq. 2. The one for the phase was adapted according to the experiences made with many different spectra. The factor 1.35 was set after numerous tests; this is not the ideal solution, but it suffices for finding an adequate start value. After the fitting process the main criterion to decide whether or not a line has been fitted successfully is the normalised sum χ^2 of the squared deviations between the three complex samples of the discrete spectrum and the analytical expression. If a peak is to be subtracted, the complex spectrum of a peak with the gained parameters, computed with $G(v)$, is subtracted from the given spectrum, and the procedure is repeated with the next-highest peak, etc., until all candidates have been processed. Because the $G(v)$ spreads over the whole spectrum, peaks influence each other so criteria of neighbouring peaks might change after subtraction of a peak. This could result in peak candidates becoming valid which had been rejected before. To avoid this, one can redo all the steps to this point after a successful fitting and subtraction of a peak. Of course this greatly increases the processing time. For both examples in this work (the coal-mine-induced event with artificial lines and the helicopter overflight) the algorithm was restarted.

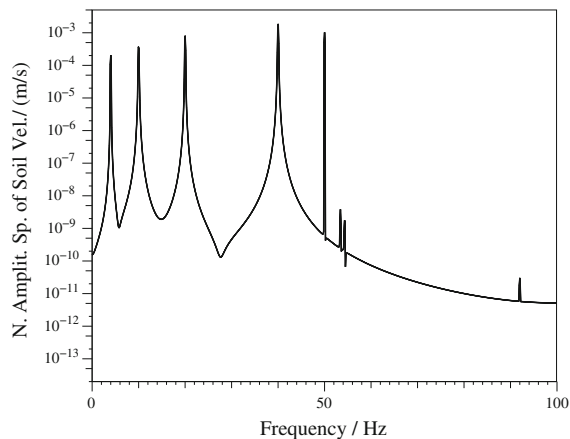


Figure 5

Magnitude spectrum of fitted and subtracted sines. Note that three sines were subtracted that had not been added before (53.4, 54.3, 92.0 Hz)

The last step is to inversely transform the remaining spectrum back to the time domain and divide the result by the window function. If a longer set of time-domain data is to be cleaned of peaks, the process is carried out for several time intervals with some overlap; $\frac{1}{4}$ overlap proved to be appropriate to reduce margin errors from small window values. After spectrum processing and inverse Fourier transform the overlapping parts of neighbouring intervals are averaged using a sliding linear weighting scheme.

3. Results of the Peak Fitting and Subtraction Algorithm

3.1. Real Seismic Event Superposed with Artificial Sine Waves

As a first demonstration we take a coal-mine-induced event in the area of Hamm-Herringen, Germany (BISCHOFF *et al.* 2010)⁴ as a model for a substantially non-periodic aftershock signal (spectrum Fig. 3, signal Fig. 7) and superpose it with artificial sine functions with different frequencies,

⁴ Data kindly provided by Monika Bischoff and Sebastian Wehling-Benatelli (Institut für Geologie, Mineralogie und Geophysik; Ruhr-Universität Bochum). Measured: 1 February 2007, 19:18:28 UTC during the “HAMNET” acquisition period. Location depth is 1,056 m, distance between source and sensor 1,441 m, Mercalli intensity: -0.48, vertical component of a 3D-velocity sensor.

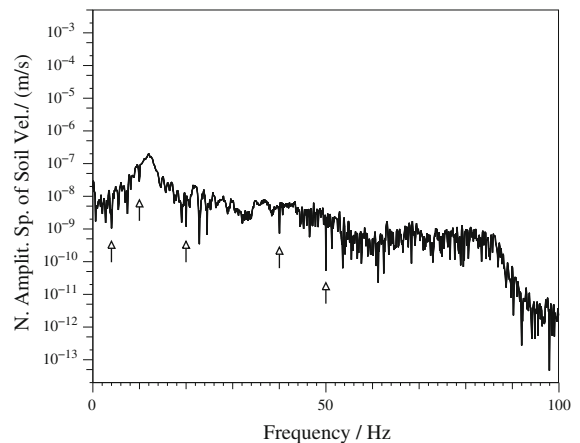


Figure 6

Magnitude spectrum remaining after peak fitting and subtraction from the superposed spectrum of Fig. 4. The frequencies of the artificial lines (respectively of the notches) are marked with arrows

amplitudes and phases as an example of periodic disturbances. The sampling rate for these data is 200 Hz and the segment length is only 15 s. To get a good spectral resolution we take a single spectrum with 2,048 samples, corresponding to 10.24 s of the time domain. By adding these pure sine functions (spectrum in Fig. 5) to the seismic event (spectrum in Fig. 3) one gets the signal and spectrum shown in Figs. 8 and 4, respectively. In this example, the sine amplitudes are 10^3 – 10^4 times the event amplitude (Table 2).

After addition of extremely strong artificial peaks, the seismic event is completely masked (see magnitude scales in Figs. 7, 8). The peak fitting results in sine parameters that are impressively close to the original ones (Table 2). As a consequence, their spectral contributions are subtracted nearly fully. The event again becomes the strongest component in the data, and its shape is reproduced well (Fig. 9). The remaining (periodic) differences from the original data (Fig. 7) are mainly caused by the spectral power which is lost in the notches at the peak positions (see Figs. 6, 10). The reasons for this are non-ideal peak parameters due to small contributions from other peaks, and of course by the underlying broadband signal. Lower magnitudes at certain frequencies in the broadband spectrum of an impulse event cause periodic disturbances similarly to higher ones. As the algorithm is fed with the superposed data, it has no

Table 2
Parameters of all lines (Fig. 5, left: input, right: found lines)

Input (exact)			Found		
Frequency (Hz)	Amplitude (mm/s)	Phase (rad)	Frequency (Hz)	Amplitude (mm/s)	Phase (rad)
4	0.5	0	3.9999997529	0.5000212055	-0.00001572
10	1	1	10.000016047	1.0001643250	0.99938260
20	2	2	20.000000164	2.0000225443	1.99999110
40	5	3	40.000000293	5.0000028117	2.99998852
50	2.5	4	49.999999850	2.4999975712	4.00000328
			53.404826315	9.0516943955e-006	4.38969166
			54.323507210	4.9913225312e-006	0.59328760
			91.996010077	7.7939926521e-008	0.35863423

The three lines at the bottom are categorised as valid lines and subtracted by mistake—these have not been added before and thus their subtraction from the spectrum leads to a manipulation of the data which should be avoided

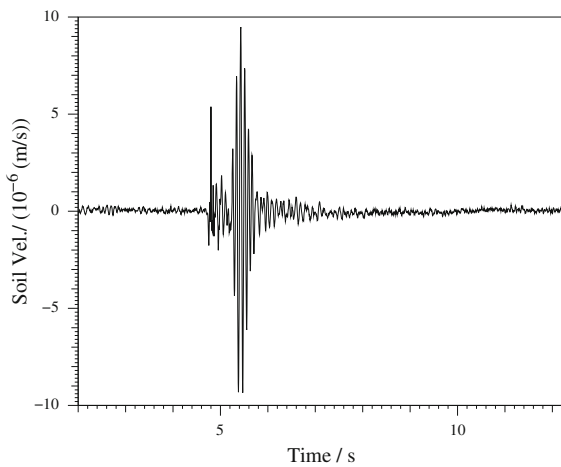


Figure 7

Seismic signal in the time domain (associated with the spectrum in Fig. 3). Note that the peak-to-peak value is $1.9 \cdot 10^{-5} \frac{\text{m}}{\text{s}}$

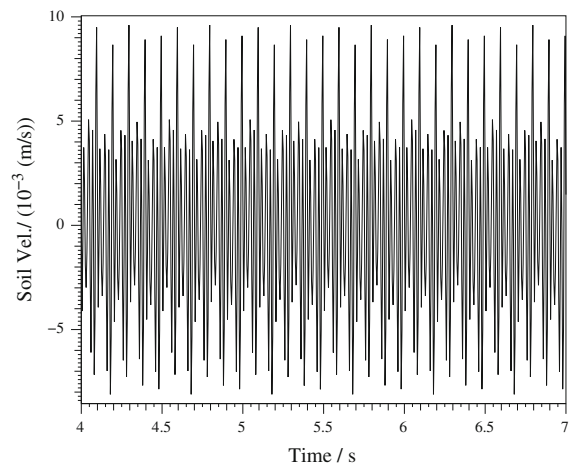


Figure 8

Time-domain signal after addition of artificial peaks (associated to spectrum of Fig. 4). Note that the peak-to-peak value is $1.8 \cdot 10^{-2} \frac{\text{m}}{\text{s}}$

ability to distinguish between the contribution of the signal and the peak. It fits the theoretical function to the complex sum of all contributing partial signals. Just “filling” the notches in Fig. 6 to arrive at a smoother spectrum is difficult because the correct phase is not known. At this stage of research the resulting notches are tolerated.

3.2. Comparison with Notch-Filtered Data

If periodic noise with constant frequencies disturbs the data it is common to use notch filters to suppress it and thus to enhance the signal-to-noise ratio. As the amplitude of our artificial sines (Sect.

3.1) is strong, the values from spectral leakage even in the far flanks need to be suppressed (only at the 50 Hz sine is the width narrow, Fig. 4) the notches need to have a certain width. A compromise is to be made: On the one hand an increased width leads to a larger suppressed bandwidth which influences the original data, on the other hand if the width is too small, parts of the peak remain in the signal. The Blackman filter used provides only -74dB attenuation in the stopband—after a single application the peaks are markedly reduced but still remain significantly stronger than the surrounding frequencies (compare with peaks in Fig. 4, divide peak heights by $\approx 5,000$). We carried out tests with different filter

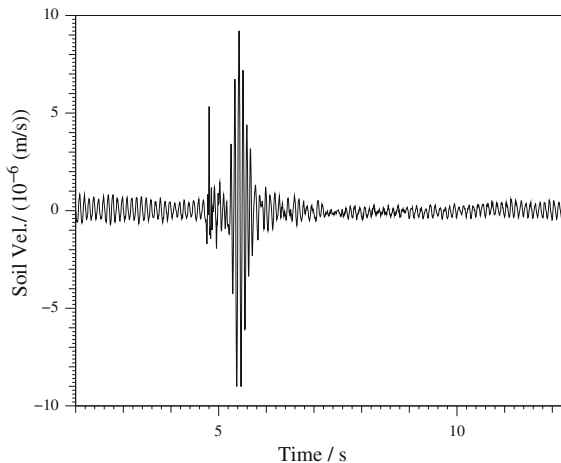


Figure 9

Time-domain signal after peak fitting and subtraction in the superposed spectrum (Fig. 6) and back transformation

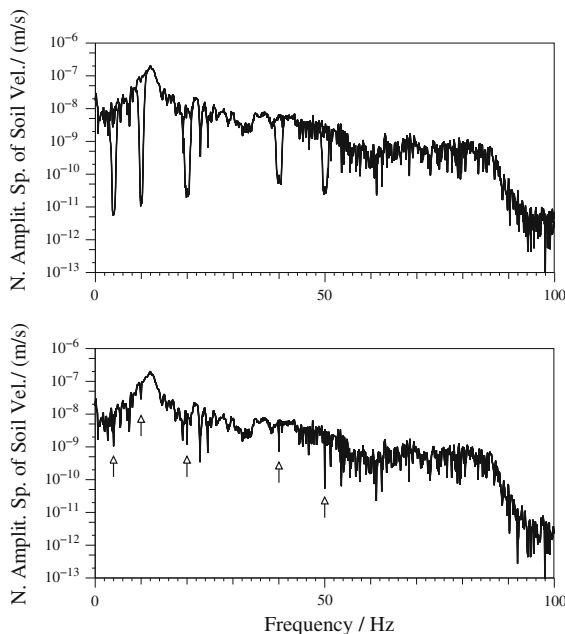


Figure 10

Spectra of the coal-mine-induced event (1,024 values from 2,048 signal samples) after addition and suppression of artificial peaks. Comparison of the results gained by application of notch filters (*top*) and the peak-subtraction algorithm (*bottom*), in both figures the *upper line* is the original spectrum (without artificial sines) which is substantially congruent except for frequencies close to the suppressed peaks. The filter notches are located exactly at the frequencies of the sines

kernels [between 1.2 and 2.4 Hz notch-filter width (between the two points where the magnitude is reduced to one half)] and single and double

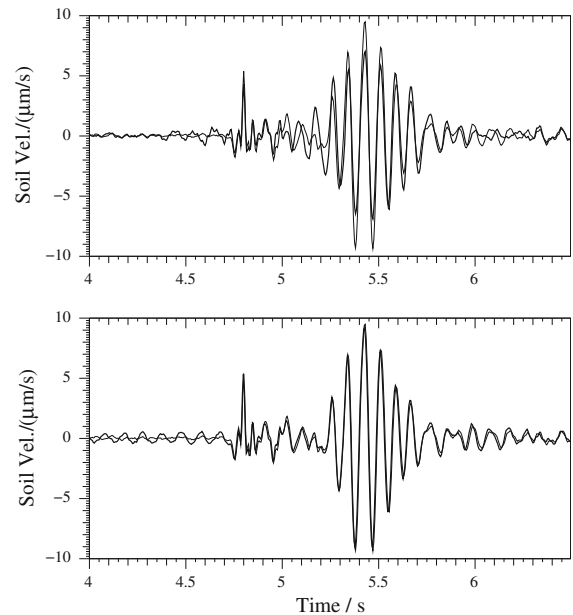


Figure 11

Time-domain signal of the coal-mine-induced event after addition and suppression of artificial peaks, section around the event. Comparison of the results gained by application of notch filters (*top*) and the peak-subtraction algorithm (*bottom*), the *thin lines* are the original signal

application. The best results were gained by double application of notch filters with a width of 1.6 Hz, that is 16.384 times the frequency step of $\Delta v \frac{1}{N\Delta t} = 0.0977$ Hz, with 801 kernel values i.e. 4 s—these are shown in Figs. 10 and 11.

It is possible that a better adjustment of the parameters of the notch filters could lead to further improvement of its performance. Nevertheless, this approach is less flexible as the frequencies of the notches have to be found before application—which again needs a line-finding algorithm (even though it could be simpler than the described one and phase and amplitude would not be important). If there were even more lines, the signal shape could change dramatically, whereas the changes produced by the line subtraction algorithm are smaller. Figure 11 shows that the P onsets are clearly reproduced by both approaches. The peak-subtraction algorithm lets the event shape largely unchanged; a small oscillation is added before the event. The notch filters change the signal shape more strongly: the surface wave starts to increase earlier, reaches lower amplitudes and decreases more slowly. In addition, a spurious signal

of increasing amplitude appears before the P onset. With the line subtraction, on the other hand, the oscillation before the event is stationary. The influence of these changes on the characterisation of the event needs to be reviewed by an aftershock analyst and hence are not commented by us.

3.3. Real Seismic Helicopter Data

Helicopter data are good examples for showing how the algorithm behaves if real data are processed. The engine of a helicopter typically is a turbine which produces turbulent airflow, resulting in non-periodic signals. This turbine typically powers one main and one tail rotor, having a certain ratio of revolution rates and consisting of blades cutting the air which produces strong periodic signals.

Signals used in this section were recorded by us during a test measurement of the OSI Division of the CTBTO near Varpalota, Hungary, in September 2011.⁵ Ground velocity was measured by 4.5 Hz geophones; the sample rate was 10 kHz. For this analysis we use $N = 8,192$ samples (that is, 0.8192 s for each spectrum). This is a tradeoff: on the one hand, large values mean good spectral resolution and thus good separation of neighbouring peaks; on the other hand, peak broadening increases if there is a frequency shift in this time interval which yet cannot be handled by the fitting algorithm. The spectrogram of 90 s of a helicopter overflight is shown in Fig. 12. Note that there is significant power (far) below 250 Hz, the Nyquist frequency applicable at the SAMS during OSI exercises of the CTBTO. In the beginning of the time interval shown, the helicopter approaches the sensor so the frequencies received (white horizontal lines) appear to be higher than corresponding to the revolution rates of the rotors and their harmonics (Doppler shift). The velocity component towards the sensor decreases and becomes negative after the helicopter passed the point of closest approach to the sensor, resulting in lower

frequencies from the departing helicopter. At long distances the velocity component towards the sensor is almost constant so that the lines produced by the helicopter become horizontal again. Because of the long duration, an overlap between adjacent time intervals of $\frac{1}{4}$, that is, 2,048 samples, is used. Figure 13 shows the frequencies and amplitudes of the successfully subtracted peaks for each spectrum. The resulting spectrogram is shown in Fig. 14. Figure 15 contains quantitative information about the spectral power that was reduced by peak subtraction. Shown is the relative difference in the spectral power sums before and after subtraction, normalised to the power sum of the original spectrum. As the power sum of a spectrum approximately equals the mean-square value of the signal in the time domain, this is a relative measure of the remaining signal strength in that time interval. Small values mean small differences in the total spectral power. As there are no circumstances known that the algorithm ever increases the power sum (by peak subtraction with wrong phase or whatever) this occurs if no peaks are subtracted. Except for the centre region, the mean-square value in the time domain is reduced by 70–90 %, corresponding to multiplication by a factor 0.3 to 0.1—that is, the root-mean-square value is reduced with a factor 0.5 to 0.3. Because this example has strong periodic content, there is much to subtract; remaining are mainly broadband noise and sines with smaller amplitude, except at the centre where the higher-frequency sines change frequency too rapidly to be represented by Eq. 2 and thus are not subtracted. Two figures are given to show extreme cases. Figure 16 corresponds to a time with very constant lines and shows good results. The rms value is reduced by a factor about 2.5, the peak-to-peak value by a factor 3.3. As the algorithm fits a monofrequent sine to the data, results get worse if frequencies change in time during the analysed interval, as shown in Fig. 17. Here the rms value is reduced with a factor 0.88, the peak-to-peak one with 0.93.

In comparison to the example with few artificial sines in Sect. 3.1, in the good example the peak-to-peak value has been reduced by only a factor 3.3, and the remaining amplitude is still relatively high. If only a simple amplitude criterion were used to detect

⁵ The measurements were done for the Preparatory Commission for the Comprehensive-Nuclear-Test Ban Treaty Organization (CTBTO), Vienna, under Contract No. 2011-1260. We thank the On-Site Inspection Division of the CTBTO for the good co-operation.

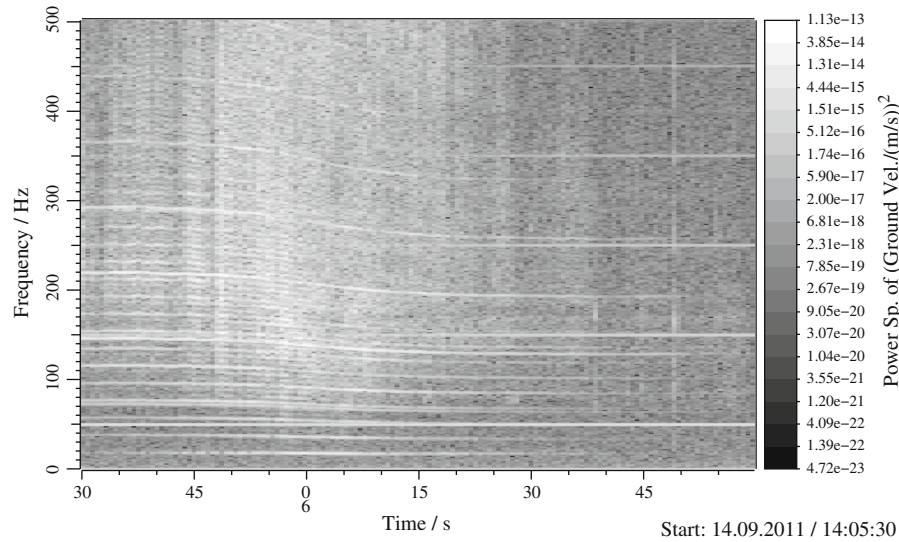


Figure 12

Sequence of 145 power spectra with 8,192 samples of a helicopter overflight. The sampling rate is 10 kHz, here the lowest 500 Hz are shown. There are three harmonic series of frequencies, one is mains-line hum probably produced by a nearby power generator (constant 50, 150, 250 Hz, etc.), one stems from the main rotor (multiples of around 6.0 Hz, every third is stronger) and one is from the tail rotor (fundamental frequency around 68 Hz). The variation of the Doppler shift, as the source moved with a velocity component towards the sensor, passed the point of closest approach (at about 14:06:00), and moved away from it again, is easy to recognise

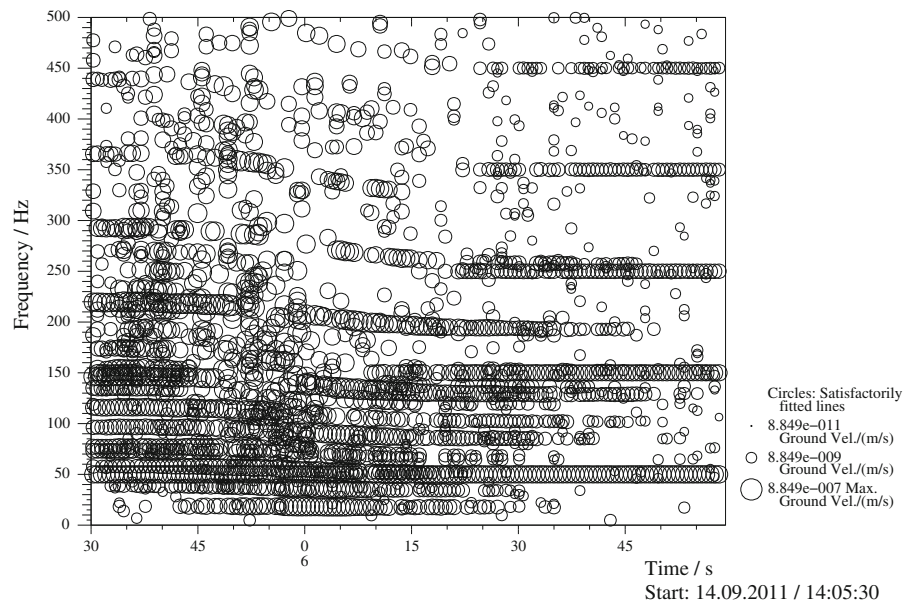


Figure 13

Sequence of results of the peak fitting and subtraction algorithm. Every *circle* stands for a successfully subtracted peak (with an amplitude of at least 10^{-4} times the amplitude of the strongest peak). Because the algorithm fits the theoretical function of a pure sine (which has a frequency that does not change in time) results get worse in times of changing frequencies. When the helicopter reaches its closest point to the sensor the Doppler shift passes through zero quickly and changes to lower frequencies afterwards. The line starting at 220 Hz shows a gap from 5:50 to 6:00. During these 10 s the frequency shifts by 9 Hz from 217 to 208 Hz. The line with the highest frequency that is reliably subtracted starts at 97 Hz and is shifted down to 93 Hz

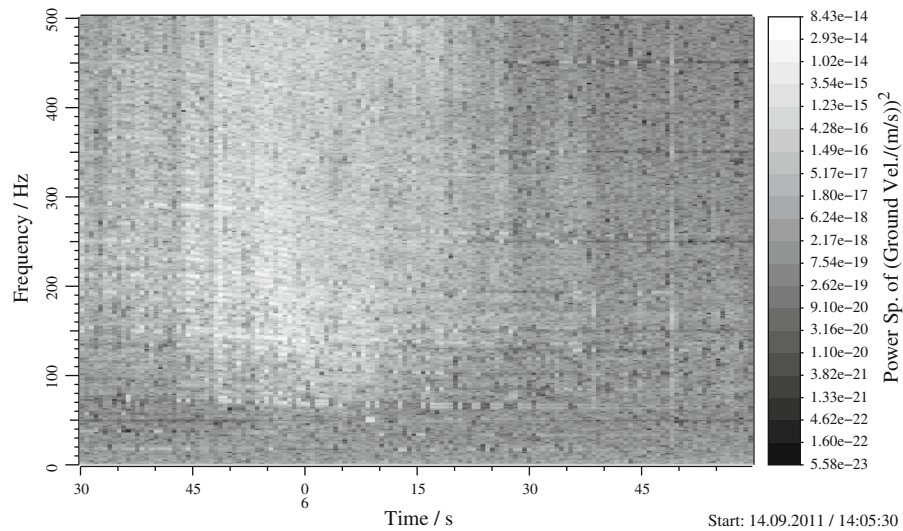


Figure 14

Same sequence as Fig. 12 but after subtraction of the peaks shown in Fig. 13. Most of the peaks are reduced successfully, at some regions one can even recognise that the *bright lines* changed to ones *darker* than the surrounding; these are the notches mentioned in Sect. 3.1 (see Fig. 6)

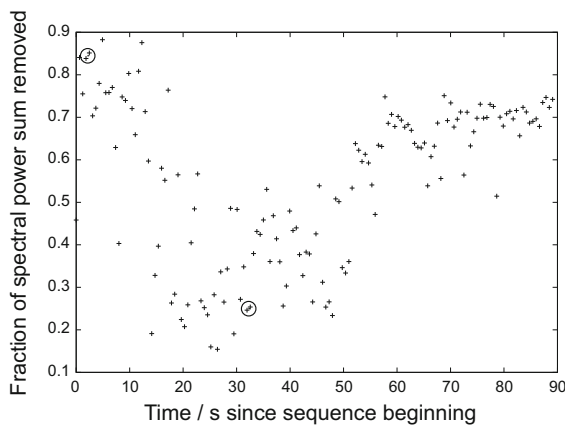


Figure 15

Degree of peak removal, same time interval as in Figs. 12–14. The differences in the spectral power sums before and after subtraction are normalised to the power sum of the original spectrum. Values of 0.7–0.9 mean that ≈ 70 –90 % of the mean-square value in the time domain is reduced, corresponding to a reduction by 0.5–0.3 in root-mean-square value. The time-domain data of the marked intervals are shown in the Figs. 16 and 17

a pulse masked in the original signal, the pulse would have to be relatively strong. Further research will look into improved procedures for peak removal. This example allows estimation of the tolerated frequency rate of change. As the frequency shift is proportional to the frequency, in times of changing frequencies the fitting gets worse for higher-order

harmonics. Considering the line starting at 220 Hz, peak fitting is interrupted during 10 s. During that time its frequency shifts by 9 Hz which means (linear shift assumed) that a frequency shift in the order of $0.9 \frac{\text{Hz}}{\text{s}}$ in this case (or generalised: $0.6 \frac{\Delta\nu}{T}$) is too strong to be handled. On the other hand, the line with the highest frequency that is reliably subtracted starts at 97 Hz and is reduced in 10 s by 4 Hz, which means that the algorithm is able to handle frequency shifts of at least $0.4 \frac{\text{Hz}}{\text{s}}$, respectively $0.3 \frac{\Delta\nu}{T}$.

4. Discussion and Outlook

The peak-finding and peak-subtraction procedure described is a new method of cleaning signals and spectra from unwanted periodic noise. The algorithm fitting the analytical expression of a monofrequent peak to the maxima of a spectrum works well if the frequencies are not too close and do not change in time. If the peaks do not interfere with each other strongly, artificial sine functions added to real data and to pure noise can be subtracted successfully even if the sine amplitudes are higher by orders of magnitude. This is the case when the frequencies differ by more than about $10\Delta\nu$. In real helicopter signals, 70–90 % reduction of the mean-square value can be

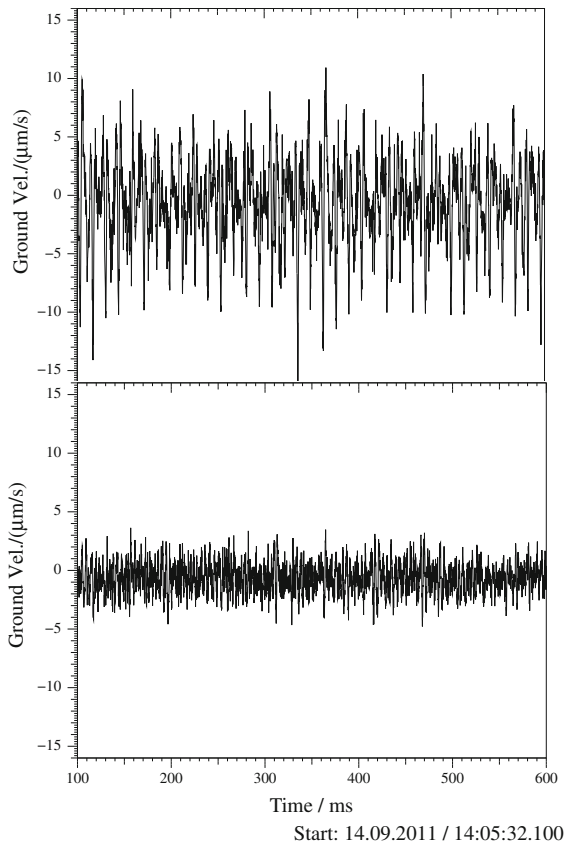


Figure 16

0.5 s of the time-domain signal of the helicopter overflight when the algorithm gives good results (computed by inverse FFT of the 4th and 5th spectrum with spectral-power-sum reduction factors 84 and 85 %, see Fig. 15, 2.1 s, corresponding to rms-value reduction by about 60 %). *Top* original signal, *bottom* after subtraction of peaks

achieved. The remaining signal mainly consists of non-periodic components (probably produced by turbulent airflow) and of peaks strongly varying in time. If the frequency change stays below about 0.3 times the spectral frequency step during the time used for one spectrum, the theoretical expression represents the peak shape well enough and subtraction can succeed. On-going research is devoted to three problems:

- Including the small contributions of distant peaks by trying to optimise, in a second stage, the parameters of all found peaks simultaneously
- In case of close peaks, when the expression for a single one is no longer usable, optimise the parameters of two in parallel

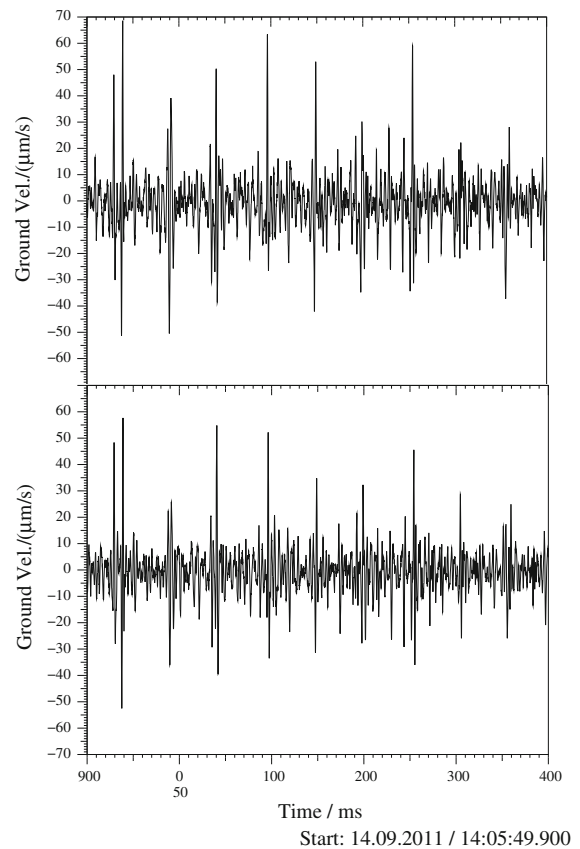


Figure 17

0.5 s of the time-domain signal of the helicopter overflight when the algorithm gives bad results (computed by inverse FFT of the 32th and 33th spectrum with spectral-power-sum reduction factors 22 and 21 %, see Fig. 15, 19.9 s, corresponding to rms-value reduction by about 12 %). *Top* original signal, *bottom* after subtraction of peaks

- Investigate how cases can be handled when the frequency changes faster than the limit above, e.g., by time-varying Doppler shift

The procedure promises to become a flexible tool for removing unwanted periodic disturbances from seismic records, including the possibility to select which peaks to remove if some are of interest but they are disturbed by others.

At the moment our procedure works on a single time series. It is expected that expanding it to array signals will improve the results. The least that one expects is better differentiation between spectral peaks that occur consistently at several sites from those that are due to random fluctuations. Beyond that, signals cleaned from unwanted periodic

disturbances should give better results in all forms of array processing.

Acknowledgments

We thank the OSI Division of the CTBTO for their good cooperation. Further thanks go to Sebastian Wehling-Benatelli (Institut für Geologie, Mineralogie und Geophysik; Ruhr-Universität Bochum, Germany) and Monika Bischoff (formerly Institut für Geologie, Mineralogie und Geophysik, Ruhr-Universität Bochum, Germany, now Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Germany) for providing the data and the interpretation of the HAMNET project which was part of the Collaborative Research Centre 526 “Rheology of the Earth” at Ruhr-University Bochum and funded by the German Research Foundation (DFG). Finally we are grateful for the comments from two anonymous reviewers that contributed much to improving the article.

REFERENCES

- JÜRGEN ALTMANN and FELIX GORSCHLÜTER. *Removing Periodic Noise to Detect Weak Impulse Events*. Poster presented at International Scientific Studies (ISS 09) Conference, Vienna, 10–12 June, 2009. Available from http://www.ctbto.org/fileadmin/user_upload/ISS_2009/Poster/SEISMO-08J%20%28Germany%29%20-%20Jurgen_Altmann%20and%20Felix_Gorschluter.pdf (16. Dec. 2011).
- MONIKA BISCHOFF, ALPAN CETE, RALF FRITSCHEN and THOMAS MEIER. *Coal mining induced seismicity in the Ruhr Area, Germany*. *Pure Appl. Geophys.*, 167:63–75, 2010.
- PETER BORMANN. *Seismic signals and noise*. pages 1 – 34. Deutsches GeoForschungsZentrum GFZ, Potsdam, 2009.
- E. ORAN BRIGHAM. *Fast Fourier Transform and Applications*. Prentice-Hall, Englewood Cliffs NJ, 2nd edition, 1988.
- BURKHARD BUTTKUS, editor. *Spectral Analysis and Filter Theory in Applied Geophysics*. Springer, 2000.
- STEVEN M. KAY. *Modern Spectral Estimation: Theory and Applications*. Prentice-Hall, Englewood Cliffs NJ, 1988.
- ALAN V. OPPENHEIM and RONALD W. SCHAFFER. *Discrete Time Signal Processing*. Prentice-Hall, Englewood Cliffs NJ, 2nd edition, 1999.
- WILLIAM H. PRESS, SAUL A. TEUKOLSKY, WILLIAM T. VETTERLING and BRIAN P. FLANNERY. *Numerical recipes in Pascal: the art of scientific computing*. Cambridge University Press, 1st edition, 1989.
- ENDERS A. ROBINSON and SVEN TREITEL, editors. *Geophysical Signal Analysis*. Prentice-Hall, Englewood Cliffs NJ, 1980.

(Received January 5, 2012, revised October 8, 2012, accepted October 10, 2012, Published online January 1, 2013)