Solid Earth Discuss., https://doi.org/10.5194/se-2019-29
Manuscript under review for journal Solid Earth
Discussion started: 11 February 2019

Improving quality of empirical Greens functions, obtained by cross-correlation of high-frequency ambient seismic noise

Nikita Afonin[1,4], Elena Kozlovskaya[1,2], Jouni Nevalainen[1,3], Janne Narkilahti[3]

[1]Oulu Mining School, POB-3000, FIN-90014, University of Oulu, Finland
[2]Geological Survey of Finland, P.O. Box 96, FI-02151, Espoo, Finland
[3]Sodankylä Geophysical Observatory POB 3000, FIN-90014, University of Oulu, Finland
[4]N. Laverov Federal Centre for the Integrated Arctic Research RAS, Arkhangelsk, Russia

Correspondence to: Nikita Afonin (nikita.afonin@oulu.fi)

Abstract. Studying the uppermost structure of the subsurface is a necessary part for solving many practical problems (exploration of minerals, groundwater studies, geoengineering, etc.). Practical application of active seismic methods is not always possible because of different reasons, such as logistical difficulties, high cost of work, high level of seismic and acoustic noise, etc. That is why developing and improving of passive seismic methods for these purposes is one of the important problems in applied geophysics. In our study, we describe the way of improving quality of Empirical Green's Functions (EGFs), evaluated from high-frequency ambient seismic noise, by using of advanced technique of cross-correlation functions stacking in the time domain (in this paper we use term “high-frequency” for the frequencies higher than 1 Hz). In compare to existing techniques, based on weight-stacking, our proposed technique makes it possible to more significantly increase the signal-to-noise ratio and, therefore quality of the EGF. The technique is based on both iterative and global optimization algorithms, where the optimized parameter is a signal-to-noise ratio of an EGF, retrieved for each iteration. The technique has been tested with the field data acquired in an area with high level of industrial noise (Pyhäsalmi Mine, Finland) and in an area with low level of anthropogenic noise (Kuusamo Greenstone Belt, Finland). The results show that the our proposed technique can be used for extraction of EGFs from high-frequency seismic noise in practical problems of mapping of the shallow subsurface in areas with high and low level of high-frequency seismic noise.

1 Introduction

Seismic methods as tools for studying of shallow subsurface structures in exploration geophysics have been developed during many years. Traditionally, seismic surveys (reflection and refraction) have been carried out using active sources. The reflection and refraction controlled-source seismic sounding methods are widely applied in exploration for oil and gas, but less commonly in mineral exploration in crystalline bedrock areas. The reasons for this have been the traditionally high cost of seismic surveys and logistical difficulties (Malehmir et al., 2012). Seismic methods as a mineral exploration tool are very good for delineation of the boundaries of certain types of mineral deposits as well as for estimating their ore potential (Kukkonen et al., 2009; Malehmir et al., 2012). There are, however, challenges in exploration of new deep targets in the vicinity of active mines, that is, in brownfield exploration. In our paper, brownfield means exploration near active mines or at the previously studied area with the purpose of getting new mineral reserves, while greenfield means exploration of new mineral deposits. Due to the large amount of heavy machinery, existing mines themselves produce strong seismic and acoustic noise. This continuous noise is overlapping in frequencies with the signals of the controlled seismic sources, creating a problem for the high-resolution active-source seismic experiments in a brownfield exploration (Place et al., 2015).

In our paper, we describe results of investigating the possibility to use high-frequency passive seismic noise interferometry with ambient noise with frequencies higher than 1 Hz (in the paper we use term "high-frequency" for this seismic noise) for extracting information about deeply seated structures in greenfield and brownfield exploration projects. For this, we develop a new method of increasing the quality of empirical Green's function (EGFs) evaluated from high-frequency industrial, anthropogenic or natural seismic noise. We partly use algorithms described in Campillo (2006), Bensen et al. (2007), Poli et al. (2012a, 2012b, 2013), Afonin et al. (2017) and implement a new algorithm of stacking of cross-correlation functions in the time domain. At present there are the several advanced algorithms of the data pre-processing (Bensen et al., 2007; Groos et al., 2012) and cross-correlation functions stacking, such as "weight stacking in time domain" (Cheng et al., 2015), "root mean square stacking" (Shirzad et al., 2014), phase cross-correlation with weight stacking (Shimmel et al., 2011), "time-frequency domain phase-weighted stacking" (Li et al., 2017). These algorithms allow increasing the signal-to-noise ratio of the EGFs extracted from high-frequency seismic noise. Nevertheless, there are some difficulties with application of these algorithms in areas with non-evenly located noise sources. The main problem is using of all cross-correlation functions for building up a final EGF. The methods based on weight stacking are partly solving the problem (Shimmel et al., 2011; Cheng et al., 2015; Li et al., 2017). Nevertheless, the incoherent cross-correlation functions are not totally excluded from building up process and can decrease quality of evaluated EGF. In our paper the term 'coherent' is used for describing cross-correlation functions with the same time lags of signal maxima and the same dominant frequency. To solve this problem, we develop a new algorithm that makes it possible to exclude incoherent cross-correlation functions from EGFs building up process and, therefore, significantly increase the signal-to-noise ratio and quality of extracted EGFs. We do not use this method in the frequency domain because it is important to stack cross-correlation functions with the same time lags and the dominant frequencies, in other words, functions that are coherent to each other.

In our paper, we are presenting details of this algorithm and illustrate its performance using passive seismic data acquired in two areas of Fennoscandia: Pyhäsalmi mine (as an example of area with high level of industrial noise) and Kuusamo Greenstone Belt area (quiet area prospective for new mining projects (Wiehed et al., 2005; Lehtonen et al., 2009).

2 Advanced technique of cross-correlation functions stacking

For solving the problems described in the introduction, we suggest our method of time-domain stacking of cross-correlation functions calculated for different time windows. We call this method signal-to-noise ratio (SNR) stacking. The general purpose of this method is to select for stacking only those cross-correlation functions that are coherent to each other. Coherence between cross-correlation functions and the EGF already retrieved at previous iteration step is estimated by comparison of the signal-to-noise ratio of EGFs before and after stacking with cross-correlation function of a next time window.

Assume that ambient seismic noise is recorded simultaneously at two different points with Cartesian coordinates $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$, $\boldsymbol{r} = [x, y, z]$ and continuous recordings are split into n time windows with the same durations. Let $a_i(\boldsymbol{r_1}, \boldsymbol{r_2}, t)$ be the cross-correlation function of these seismic records for the time window $i,\ i = 1 \dots n$, where t is time lag of the seismic records.

Let t_m is the maximum time lag in a cross-correlation function (length of cross-correlation); t_{ds} is a maximum time of wave propagation between the two points; $|t_m| \gg |t_{ds}|$ and $-t_m \le t \le t_m$. Let $-t_{ds} \le \triangle t_e \le t_{ds}$ is the time lag interval on cross-correlation function corresponding to the expected seismic phase (body or surface wave). In this case, selection of t_{ds} and $\triangle t_e$ is based upon *a priori* information about seismic velocities in the studied area. The value of $\triangle t_e$ is at least two periods of the expected signal dominant frequency. In the case of evaluation of surface wave parts of EGFs, this frequency usually corresponds to the frequency of noise with the largest amplitude that can be estimated by time-frequency analysis of noise seismic records. Let $a_i^{max}(\boldsymbol{r_1},\boldsymbol{r_2},\triangle t_e)$ is the maximum value of cross-correlation function in the time interval $\triangle t_e$. Then, the signal-to-noise ratio of a cross-correlation function calculated for the $i\ th -$ time window ($SNR(a_i(\boldsymbol{r_1},\boldsymbol{r_2},t))$) is:

$$SNR(a_i(\boldsymbol{r_1},\boldsymbol{r_2},t)) = \frac{a_i^{max}(\boldsymbol{r_1},\boldsymbol{r_2},\triangle t_e)}{\frac{1}{2|t_m-t_{ds}|}\left(\int_{t_{ds}}^{t_m} a_i^2(\boldsymbol{r_1},\boldsymbol{r_2},t)dt+\int_{-t_m}^{-t_{ds}} a_i^2(\boldsymbol{r_1},\boldsymbol{r_2},t)dt\right)} \tag{1}$$

Let $a_i(\boldsymbol{r_1},\boldsymbol{r_2},t)$ and $a_j(\boldsymbol{r_1},\boldsymbol{r_2},t)$ are cross-correlation functions calculated for different time windows $i \in (1..n)$ and $j \in (1..n)$ and $c(\boldsymbol{r_1},\boldsymbol{r_2},t) = a_i(\boldsymbol{r_1},\boldsymbol{r_2},t) + a_j(\boldsymbol{r_1},\boldsymbol{r_2},t)$ is EGF retrieved from these two cross-correlation functions. If $a_i(\boldsymbol{r_1},\boldsymbol{r_2},t)$ and $a_j(\boldsymbol{r_1},\boldsymbol{r_2},t)$ are coherent to each other and $i \ne j$, then expressions $SNR(a_i(\boldsymbol{r_1},\boldsymbol{r_2},t)) < SNR(c(\boldsymbol{r_1},\boldsymbol{r_2},t))$ and $SNR(a_j(\boldsymbol{r_1},\boldsymbol{r_2},t)) < SNR(c(\boldsymbol{r_1},\boldsymbol{r_2},t))$ have to be true, according to the principle of interference. Condition $i \ne j$ is necessary in order to avoid stacking of functions with itself. Therefore, increasing SNR of the retrieved EGF after stacking with cross-correlation function can be used as a criterion for excluding incoherent functions from the stack and building up the EGF with high signal-to-noise ratio.

Based on the criteria described above, an expression for calculation of EGF for k-th iteration may be written as

$$G^k(\boldsymbol{r_1},\boldsymbol{r_2},t) = \sum_{\substack{i=1 \\ i\ne k}}^{n}\left(G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t) + a_i(\boldsymbol{r_1},\boldsymbol{r_2},t) * \delta(G_i^k, a_i)\right), \tag{2}$$

where $k = 1 \dots n$ is number of initial function; n is number of time windows; $i = 1, \dots, n$; $G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t)$ is EGF, which corresponds to k-th – initial function and evaluated in previous iterations:

$$G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t) = \begin{cases} a_k(\boldsymbol{r_1},\boldsymbol{r_2},t), & i = 1 \\ G_{i-1}^k(\boldsymbol{r_1},\boldsymbol{r_2},t), & i \ne 1 \end{cases}. \tag{3}$$

The operator of selection may be written as

$$\delta\left(G_i^k, a_i\right) = \begin{cases} 0, SNR\left(G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t) + a_i(\boldsymbol{r_1},\boldsymbol{r_2},t)\right) < SNR\left(G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t)\right); \\ 1, SNR\left(G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t) + a_i(\boldsymbol{r_1},\boldsymbol{r_2},t)\right) \ge SNR\left(G_i^k(\boldsymbol{r_1},\boldsymbol{r_2},t)\right); \end{cases} \tag{4}$$

The result of this algorithm are n candidates for EGF. Let us denote signal-to-noise ratio as function of initial function index $SNR(G^k(\boldsymbol{r_1}, \boldsymbol{r_2}, t)) = f(k), k = 1, \dots, n$. Then, the condition for the final EGF selection can be written as $m = argmax(f(k))$, where m denotes the index of selected EGF. According to this condition, the EGF with maximum signal-to-noise ratio will be selected as the final one. As the function $f(k)$ may have several local maxima in the parameter space k, the condition for the final EGF selection ensures selection of the global maximum of this function in the parameter space k.

In the proposed algorithm, maximizing the signal-to-noise ratio of the retrieved EGF is ensured by stacking of only cross-correlation functions coherent to each other and selection of EGF with the maximum signal-to-noise ratio from all calculated candidate EGFs. In other words, the proposed algorithm is analogous to the direct search methods of global optimization. It is necessary to remember, however, that EGF with maximum signal-to-noise ratio does not correspond to a true EGF if the dominant noise sources are located outside the stationary phase area. Therefore, it is important to use the system of observations that allows estimating azimuthal distribution of noise sources. Moreover, the method is based on assumption that sources of ambient seismic noise produce a signal with relatively broad bandwidth and cannot produce an ideal harmonic signal of single frequency.

The method also can be used for calculation of azimuthal distribution of strongest noise sources. For this a 2-D array of seismic recording station is necessary. In this case, the time lags, corresponding to expected signal Δt_e in Eq. 1 have to be a function of expected seismic velocity and azimuth of expected wave approach $\Delta t_e = f(v, \varphi)$. Then signal-to-noise ratio for each pair of stations of the array is the function of initial function index, velocity and azimuth of approach $SNR(G^k(\boldsymbol{r_1}, \boldsymbol{r_2}, t)) = f(k, v, \varphi), k = 1, \dots, n,\ v_{min} \le v \le v_{max},\ 0 \le \varphi \le 360$. Limits of velocity have to be calculated according to *a priory* information about seismic velocities in the studied area. Global maximum of the function correspond to the strongest or the most coherent wavefiled. Therefore, the method allows estimating azimuths to the strongest source of noise wavefield.

We suggest that this method can be used for extraction of EGFs from high-frequency industrial, anthropogenic, or natural seismic noise. Moreover, this method does not require that a diffuse field is used for calculating EGFs. Therefore, application of this method to the data of optimally selected seismic recording array might decrease significantly the time necessary for registration of ambient seismic noise, which is very important for practical applications of passive seismic interferometry. For studying the possibilities of using this method for extraction of EGFs from high-frequency seismic noise, we use the data from two passive seismic experiments carried out in areas with different seismic noise characteristics. The first area is characterized by high level of industrial noise (Pyhäsalmi underground mine site) that is usually observed in brownfield exploration areas, while the second area is seismically very quiet and characterized by a limited amount of local anthropogenic (roads) and natural (rivers) high-frequency seismic noise sources. Such noise characteristics are typical for greenfield exploration areas.

3 Experimental data

3.1 Pyhäsalmi mine area

As an example of using high level industrial seismic noise for estimation of EGFs, we used the seismic noise at the site of Pyhäsalmi mine, Finland. For this purpose we installed 24 3-component DSU-SA MEMS seismic sensors with the autonomous RAUD eX data acquisition units manufactured by Sercel Ltd. along a 10-km-long line crossing the mine area with interstation distances of about 100 m (for PLB03-PLB13 and PLB14-PLB22) and 2 km (PLB01, PLB02, PLB23, PLB24) (Figure 1). The seismic stations recorded continuous seismic data from 1.11.2013 to 5.11.2013 with a sampling frequency of 500 Hz.

The profile configuration was selected based on results of previous passive seismic measurements in Pyhäsalmi. These studies showed that the mine is the main source of seismic high-frequency noise.

The profile crossing the mine area consists of two parts, and each of these consists of 12 sensors: the western part has direction from the mine to the west (PLB01-PLB13), and the eastern part has direction from the mine to the east (PLB14-PLB24). Each part of the profile includes one sensor installed closest to the mine (PLB13 and PLB14). The horizontal components were oriented to the true north and east (NS and EW-components, respectively). Thus, rotation of the horizontal components before seismic noise analysis was not necessary.

3.2 Kuusamo Greenstone Belt area

As an example of an area with low level of anthropogenic seismic noise, we select an area located in the Kuusamo Greenstone Belt (KuGB), Finland, because of numerous previous geological and geophysical studies there (Silvennoinen, 1991; Bruneton et al., 2004; Yliniemi et al., 2004; Silvennoinen et al., 2007; Poli et al., 2012; Pedersen et al., 2013; Silvennoinen et al., 2014; Tiira et al., 2014; Vinnik et al., 2014; etc.). Moreover, according to studies by Weighed et al. (2005) and Lehtonen et al. (2009), this area is prospective for gold- and diamond deposits.

For testing of our method of cross-correlation function stacking, we use the data collected during a passive seismic experiment in KuGB area in the August of 2014. One of the targets of this experiment was to investigate the possibility of high-frequency EGFs extraction from anthropogenic or natural seismic noise in regions with low ambient noise level.

The temporary seismic array (Figure 2) consisted of five three-component velocimeters Trillium Compact produced by Nanometrics (Canada) and 24 three-component accelerometers DSU-SA MEMS with autonomous RAUD eX data acquisition units manufactured by Sercel Ltd. (France).

As we can see in Figure 2, the seismic array represents a triangle. The lengths of the sides of this triangle are about 4-6 km. The BB sensors were installed in the vertices of this triangle and collocated with MEMS accelerometers. Moreover, each of the vertices was surrounded by a circular array with small aperture (about 1400-1500 m), consisting of six accelerometers. The array recorded continuous seismic data from 28.08.2014 to 10.09.2014 with a sampling rate of 500 samples per second. Such an array configuration makes it possible to estimate the azimuthal distribution of the high-frequency noise sources and also to extract high-frequency EGFs from records of small aperture arrays.

4 Analysis of the seismic noise

4.1 Time-frequency analysis

One of the most important steps of the data preparation before extraction of EGFs is time-frequency analysis. It is necessary for selection of a frequency band with high amplitudes of the ambient noise. For this, we analyse characteristics of the seismic noise recorded at different distances from the potential noise sources. In the Pyhäsalmi experiment, the most probable noise sources are located inside the underground mine and in the open pit. For the time-frequency analysis of the seismic noise, we used records of sensors installed at different distances from the mine and open pit (PLB24 and PLB14 (Figure 1)). Figure 3 (a, b) shows the result of this analysis.

From Figure 3 (a, b), one can see two main frequency bands with high amplitudes of the seismic noise recorded closest to the mine: about 3-4 Hz and about 10-100 Hz, respectively. Moreover, the amplitudes of the noise in these frequency bands decrease with increasing distance from the mine. Therefore, we can assume that the sources of the noise for these frequency bands are located inside the underground mine and in the open pit. Based on this analysis, we selected the frequency band of 1-100 Hz for pre-filtering of the noise before calculation of the cross-correlation functions.

In the KuGB experiment, a temporary seismic network was installed in a quiet area without any significant industrial activity; therefore, we can assume that the high-frequency seismic noise might be produced by multiple natural (for example, rivers) and/or anthropogenic (for example, roads) sources. In this case, analysis of the time-frequency characteristics of the seismic noise is necessary. For this, we calculate time-frequency diagrams in the frequency band 0.1-100 Hz and examples of these diagrams for two stations are presented in Figure 3 (c, d)).

Figure 3 (c, d) shows that records from both stations have amplitude maximums in the frequency band of 0.1-1 Hz. Seismic noise recorded by the KU05 station is also characterized by periodically high amplitudes in the frequency band of 40-100 Hz (Figure 3 (c)). This noise may be caused by anthropogenic (transport) or natural (for example, wind) factors. Station KU02 is located close to the river that can be a source of continuous seismic noise with high amplitudes in the frequency band of 40-80 Hz (Figure 3 (d)). Therefore, for estimation of high-frequency EGF, we select to pre-filter the data band-pass filter of 1-100 Hz.

4.2. Analysis of azimuthal distribution of the noise sources

Classical methods of passive seismic interferometry are based on diffuse field approximation (Wapenaar et al., 2008, 2010). One of the most important conditions for using this approximation is isotropic and homogeneous azimuthal distribution of noise sources (Mulargia, 2012). Therefore, the second important procedure of data preparation before estimation of EGFs is analysis of the azimuthal distribution of the noise sources during the experiment's period. In our study, we considered two cases. In the first case, the main sources of high-frequency seismic noise in the Pyhäsalmi area are most probably located inside the mine and in the open pit. Thus, the assumption about isotropic and homogeneous azimuthal distribution is not valid. As shown in Wapenaar et al. (2010), in such cases one cannot assume diffuse field approximation. That is why the measurements of the noise were made along a profile (linear array) consisting of two parts crossing the mine site and oriented EW. However, signals from other noise sources can be also present in the wavefield during the data acquisition period. That is why we made additional analysis of azimuthal distribution of noise sources. For calculation of the azimuthal distribution of the noise sources, the well-known methods are frequency-wavenumber (f-k) analysis (Neidell et al., 1971; Douze et al., 1979) and beamforming in the time domain (Rost et al., 2002; Schweizer et al., 2012). The linear configuration of the Pyhäsalmi array does not allow application of f-k analysis and beamforming, however. For understanding of the directivity of the seismic noise wavefields for different frequency bands, we apply the horizontal-to-vertical ratio rotate method proposed in Nakamura et al. (1989), investigated in Barazza et al. (2009), and implemented into Geopsy software (http://www.geopsy.org).

In our study, we analyse records of seismic noise with duration of 10 min for each hour of records. We apply this procedure to records from stations which are the most distant from the mine, located in both parts of the profile (PLB01 and PLB24). and for. We have selected two frequency bands (2-5 Hz and 5-10 Hz) for analysis, because they characterize strong and stable seismic noise, from which it is possible to retrieve surface waves. The result is shown in Figure 4 as a percentage of record time during which the recorded wavefields approached from a certain azimuths with respect to the total time of the record. In Figure 4, azimuth of 0 degree corresponds to the true north and shadowed sectors denote the azimuths to the noise sources. Radial sizes of these sectors are proportional to the relative source-acting time calculated as a percentage of the total measurement time. Anglular sizes of the sectors denote errors of the azimuth calculation.

In Figure 4 (a, b) one can see strong directivity of the noise wavefields from the east. This proves that the main noise source for the eastern part of the profile for frequency bands of 2-5 Hz and 5-10 Hz is the mine. Considering the western part of the profile, there is no such clear directivity of the noise wavefields as revealed for the eastern part. One can see near-homogenous azimuthal distribution of the noise sources for azimuths between about 250 and 300 degrees. This could be explained by location of the profile close to the open pit that occupies a larger area than the underground mine. Because of this, the point-source approximation of noise sources is not valid. From these results we can conclude that if we simply stack all calculated cross-correlation functions for a pair of stations (in particular, in the eastern part of the profile) the final EGF will be biased. Therefore, for estimation of the EGF with minimum bias, we need to apply the advanced method of stacking described above.

In the second case, we considered the KuGB area with low level of high-frequency noise. In order to investigate spatial and azimuthal distribution of the strongest noise sources, we applied the procedure described above to the data of each of small-aperture arrays. The crosscorrelation functions were calculated between the central sensor and the other sensors of the array. Figure 5 presents results of the calculations of the azimuths of the strongest seismic noise sources.

In Figure 5, one can see that for the different small-aperture arrays there are also different azimuths to the sources in the different frequency bands and directions to the sources depend on the frequency. Taking into account the size of our temporary array, we can assume that the sources of high-frequency seismic noise are located at distances of about 0.7-3 km from the centres of the small-aperture arrays.

5 Empirical Green's functions estimation

For estimation of EGFs, it is necessary to apply a procedure for data preparation. This procedure includes several steps, such as spectral whitening, removing parts of records with earthquakes, blasts and missed data. This procedure is applied to the data of both experiments in our study. In the previous parts of our paper, we have demonstrated that the Pyhäsalmi mine is the source of continuous and strong seismic noise in the frequency band of 1-10 Hz. Therefore, we extract EGFs separately for the east and west part of profile.

Each part of the profile includes one sensor installed in the closest vicinity to the mine, and we calculated cross-correlation functions between those sensors and each of the other sensors in each corresponding part of the profile. Industrial seismic noise may consist of surface and body waves, because of different types of noise sources.

There are several methods of stacking the cross-correlation functions in the time domain, for example, the root-mean-square method of Shirzad et al. (2014) and the weighted stack by Cheng et al. (2015). We compare the signal-to-noise ratio of EGFs estimated by our method and the root-mean-square and weighted methods of cross-correlation functions stacking for the Pyhäsalmi experiment with respect to the surface wave signal seen in EGFs. Results of this comparison are presented in Figure 6.

In Figure 6, one can see that after application of SNR-stacking method we obtained the EGF with the highest signal-to-noise ratio of surface waves compared to the other two methods of stacking. This is because we used only cross-correlation functions coherent to each other in our stacks. As one can see from Figure 7, the algorithm selects only several cross-correlation functions to building up the EGF. Nevertheless, it does not mean that there are only few cross-correlation functions coherent to each other. It means that after some iterations the signal-to-noise ratio might not increase any more by stacking.

We analyse the apparent velocities obtained from the maximums of each of the cross-correlation functions and the apparent velocities from the cross-correlation functions selected by our algorithm of stacking (Figure 8). This figure shows that most of the retrieved EGFs have group velocities of about 4500 m/s. After applying simple stacking procedure to these cross-correlation functions, the group velocity of the surface wave part of the resulting EGF is about 4500 m/s. This cannot be true velocity, as it is too high for surface waves propagating in the uppermost layer. As can be seen from Figure 8, our SNR-stacking algorithm has selected only EGFs with group velocity of about 3400 m/s. This velocity is close to the minimal from all group velocities and is in agreement with group velocities of surface waves and S-wave velocities in the uppermost part of the bedrock in Fennoscandia (Kobranova, 1986; Dortman, 1993; Silvennoinen et al., 2007; Janik et al., 2009; Poli et al., 2013, etc.). Therefore, after applying our stacking method, we can analyse retrieved EGFs with true group velocity and minimal error.

Solid Earth Discuss., https://doi.org/10.5194/se-2019-29
Manuscript under review for journal Solid Earth
Discussion started: 11 February 2019

We apply our method of stacking to cross-correlation functions calculated for the east and west part of profile for the frequency bands 2-5 Hz and 5-10 Hz separately. After stacking, we analyse particle-motion diagrams of the waves retrieved from the seismic noise. Figure 9 shows result of stacking and particle motion analysis of EGFs.

In Figure 9, we presented only EGFs that probably contain also body waves, because other EGFs, namely those calculated for the western part in the band of 5-10 Hz and the eastern part in the band of 2-5 Hz include only surface-wave parts. Figure 9 (a) shows that the seismic noise recorded in the western part of the profile retrieves mainly Rayleigh waves with group velocity of about 3400 m/s. Other wave is marked on figure 9 (a) as an S-wave because the particle motion diagram corresponds to this type of wave. Nevertheless, this wave has apparent velocity of 5700 m/s, which is too high. Therefore, we speculate that this can be an artefact that cannot be used for further analysis. In the frequency band of 5-10 Hz, the EGFs calculated for the eastern part of the profile (Figure 10, b) consists of Rayleigh wave. The other arrivals could correspond to one reflected P-wave and three reflected S-waves. Apparent velocities of reflected P-, S1-, S2-, and S3-wave are about 4480 m/s, 3192 m/s, 3261 m/s and 2543 m/s, respectively. Our assumption that these phases may correspond to retrieved body waves is based upon comparison of their travel times with travel times of body waves recorded during previous active source experiment in Pyhäsalmi (Heinonen et al., 2012). Alternatively, the extracted waves may correspond to other phases, for example, to direct waves generated by sources inside the mine. In the same time, these phases can be also artefacts. Unfortunately, these assumptions cannot be proved using our data and it would be necessary to use the higher density array for precise phase identification of body waves. The group velocity of the Rayleigh wave is about 3400 m/s. In our study, the error in velocities estimation is assumed proportional to about 0.25 of the wavelength of an extracted signal. The error of the polarization calculation is about 1-3 degrees.

For the KuGB experiment, we calculate cross-correlation functions for each small-aperture array and apply the SNR-stacking algorithm for EGFs evaluation. Cross-correlation functions are calculated between the central sensor and each other sensor of the corresponding small-aperture array. In Figure 10 (b), we present result of EGFs calculation by this method for one of the small-aperture array (SK1-SK8 in Figure 2).

In Figure 10, one can see that after application of simple stacking, there are many implicit maximums in the retrieved EGFs. Due to this, it is not possible to calculate the azimuth to noise sources and seismic wave velocities. However, application of the SNR-stacking allows retrieval of the EGFs with maximums corresponding to surface wave propagating from a virtual source with apparent velocity of about 320-350 m/s. These wave could be Rayleigh wave, or acoustic wave propagating in the air. This assumption is based on the fact that velocity of 350 m/s is close to both to velocity of sound in the atmosphere and to the velocity of surface wave propagating in the shallow quaternary sediments in the uppermost subsurface. For precise determination of the wave type, it would be necessary to have more dense observation network. Thus, using our SNR-stacking algorithm we extracted surface waves from high-frequency seismic noise. As we noticed in previous sections, the noise sources were distributed stochastically, both in the space and in the time, and intensity of these noise sources was small. The body waves are not seen in the Figure 10 (b), because for their identification a higher density array is necessary.

Solid Earth Discuss., https://doi.org/10.5194/se-2019-29
Manuscript under review for journal Solid Earth
Discussion started: 11 February 2019

6 Discussion

The classical passive seismic interferometry is based on diffuse-field approximation, because of the equivalence of correlation properties of the multiple-scattering and resulting wavefields. Therefore, it is possible to evaluate EGFs from averaged cross-correlation functions (Campillo et al., 2003). In practice, one needs averaging over long time intervals (more than 1 year) because of the inhomogeneous and anisotropic distribution of ambient seismic noise sources during short time intervals (Wapenaar et al., 2010). This is a limitation for practical application of passive seismic interferometry as a method of applied geophysics, because it is not always possible to have long-term data acquisition experiments for solution of applied problems (mining exploration, microseismic zonation etc.). In such applied problems, the alternative may be to use ballistic waves, not scattered from heterogeneities, but produced by some localized sources of seismic noise (Mulargia, 2012). The major challenge in this case is retrieving body waves from seismic noise. Recently, some techniques of body-wave extraction were proposed in Almagro Vidal et al. (2014) and Panea et al. (2014). The main idea of these techniques is separating ambient seismic noise into a body-wave part and a surface-wave part. One could expect that combination of these separation methods with our technique of stacking would significantly increase the quality of retrieved body waves. Nevertheless, for this it would be necessary to have the data of dense high-resolution seismic arrays, so making new experiments would be a next step in development of our technique.

There are several methods based on weighted-stacking of cross-correlation functions, both in time and frequency domains, which allow significantly increase quality of extracted EGFs (Shimmel et al., 2011; Cheng et al., 2015; Liu et. al., 2016; Li et. al. 2017, etc.). Nevertheless, the full exclusion of incoherent cross-correlation functions from the process of building up the EFGs, based on signal-to-noise ratio increasing criteria, allows to obtain EGFs of better quality, as has been shown in previous sections. Of course, signal-to-noise ratio increasing criteria is possible to use in the frequency domain, but this would make the algorithm significantly more complicated. It is necessary to remember, however, that this algorithm can be applied because ambient noise sources are generally characterized by relatively wide frequency band. In this case one can expect that increasing of signal-to-noise ratio for dominant frequency would result in increase of signal-to-noise ratio of all other frequencies of the signal, as shown in Bensen et. al. (2007).

The algoritm proposed in this paper has several limitations and drawbacks. One of the most important limitations is relationship between the time of seismic wave propagation between neighbouring sensors and the dominant period of the retrieved EGFs. If the time of surface-wave propagation is about one or two periods of this wave, then it would not be possible to separate body and surface waves, similarly as in seismic experiments with active sources. Moreover, the increase of signal-to-noise ratio of one event, for example, a retrieved surface wave, might lead to the decrease of the signal-to-noise ratio for a retrieved body wave. The results obtained for KuGB area demonstrate, however, that using the SNR stacking method might be useful for building up an EGF by stacking ballistic surface wave signals retrieved from the ambient noise. The high quality of surface waves in EGF was achieved both for brownfield and greenfield exploration areas. Experimental data used in our study is insufficient to make detailed evaluation how this technique is working with body wave signals, and it will be the subject for future research.

7 Conclusion

Results of our study suggest that classical approaches for EGFs evaluation from ambient seismic noise (Campillo, 2003) cannot be considered as a universal tool for extracting high-frequency EGFs. In particularly, in quiet areas with low level of antropogenic and industrial noise the method would require long registration time because sources of high-frequency wavefield are weak and stochastically distributed both in space and time. One of the ways to treat the problem is to use ballistic waves and develop and improve methods of selection of the coherent parts of the wavefield. Study of azimuthal distribution of ambient noise sources using array techniques is necessary to do prior to making passive seismic experiments, both in greenfield and brownfield exploration areas.

The presented algorithm of cross-correlation functions stacking in a time domain allows to increase significantly signal-to-noise ratio of retrieved EGFs. In our study we demonstrated that under certain conditions the body waves could be extracted from high-frequency industrial seismic noise using the proposed algoritm of stacking in the time domain. This was illustrated with the data collected during passive seismic experiment near the Pyhäsalmi underground mine. Nevertheless, for more detailed testing of possibility of extracting body waves, it would be necessary to analyse the data collected with a higher-density seismic array near mine.

The presented algorithm of stacking makes it possible to extract EGFs from ambient seismic noise with frequencies higher than 1 Hz recorded in the quiet area without strong sources of industrial noise. This has been demonstrated by application of our new technique to the data collected in the Kuusamo Greenstone belt area that is characterized by low level of anthropogenic seismic noise and has no industrial sites located nearby.

8 Acknowledges

The study is a part of SEISLAB project funded by the European Regional Development Fund (ERDF), Council of Oulu region (Finland) and Pyhäsalmi Mine Oy.

References

Afonin, N., Kozlovskaya, E., Kukkonen, I., and DAFNE/FINLAND Working Group: Structure of the Suasselkä postglacial fault in northern Finland obtained by analysis of local events and ambient seismic noise, Solid Earth, 8, 531-544, doi:10.5194/se-8-531-2017, 2017.

Barazza, F., Malisan, P., and Carniel, R.: Improvement of H/V technique by rotation of the coordinate system: Communications in Nonlinear Science and Numerical Simulation, 14(1), 182-193, 2009.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., and Yang, Y.: Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysical Journal International, 169(3), 1239-1260, 2007.

Bruneton, M., Pedersen, H. A., Farra, V., Arndt, N. T., Vacher, P., Achauer, U., and Grad, M.: Complex lithospheric structure under the central Baltic Shield from surface wave tomography: Journal of Geophysical Research: Solid Earth, 109(B10), 2004.

Campillo, M.: Phase and correlation of ‘random’ seismic fields and the reconstruction of the Green function: Pure Applied Geophysics, 163, 475–502, doi:10.1007/s00024-005-0032-8, 2006.

Cheng, F., Xia, J., Xu, Y., Xu, Z., and Pan, Y.: A new passive seismic method based on seismic interferometry and multichannel analysis of surface waves: Journal of Applied Geophysics, 117, 126-135, 2015.

Dortman, N. B.: Petrophysics. A Handbook. Book 1. Rocks and Minerals, 1992.

Douze, E. J. and Laster, S. J.: Statistics of semblance: Geophysics, 44, 1999–2003, doi:10.1190/1.1440953, 1979

Groos, J. C., S. Bussat, and J. R. R. Ritter.: Performance of different processing schemes in seismic noise cross-correlations: Geophysical Journal International 188.2, 498-512, 2012.

Heinonen, S., Imaña, M., Snyder, D.B., Kukkonen, I.T. and Heikkinen, P.J.: Seismic reflection profiling of the Pyhäsalmi VHMS-deposit: A complementary approach to the deep base metal exploration in Finland: Geophysics, 77, WC15-WC23, 2012.

Janik, T., Kozlovskaya, E., Heikkinen, P., Yliniemi, J., and Silvennoinen, H.: Evidence for preservation of crustal root beneath the Proterozoic Lapland-Kola orogen (northern Fennoscandian shield) derived from P and S wave velocity models of POLAR and HUKKA wide-angle reflection and refraction profiles and FIRE4 reflection transect: Journal of Geophysical Research: Solid Earth, 114(B6), 2009.

Kobranova, W. N.: Petrophysics: Nedra, Moscow, 1986.

Kukkonen, I., Heikkinen, P., Elo, S., Heinonen, S.,Laitinen, J., and HIRE Working Group of the Geological Survey of Finland: HIRE Seismic Reflection Survey in the Kemi Cr mining area, northern Finland: Report Q23/2009/64, 2009.

Lehtonen, M., O’Brien, H.: Mantle transect of the Karelian Craton from margin to core based on PT data from garnet and clinopyroxene xenocrysts in kimberlites: Bulletin of Geological Society of Finland, 81, 79-102, 2009.

Li, G., Niu, F., Yang, Y., and Xie, J.: An investigation of time–frequency domain phase-weighted stacking and its application to phase-velocity extraction from ambient noise's empirical Green's functions. Geophysical Journal International, 212(2), 1143-1156, 2017.

Liu, X., Ben-Zion, Y., and Zigone, D.: Frequency domain analysis of errors in cross-correlations of ambient seismic noise. Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society, 207(3), 1630-1652, 2016.

Malehmir, A., Durrheim, R., Bellefleur, G., Urosevic, M., Juhlin, C., White, D.J.,Milkereit, B.,Campbell, G.: Seismic methods in mineral exploration and mine planning: A general overview of past and present case histories and a look into the future: GEOPHYSICS 77:5, WC173-WC190, 2012.

Mulargia, F.: The seismic noise wavefield is not diffuse: The Journal of the Acoustical Society of America, 131(4), 2853-2858, 2012.

Nakamura, Y.: A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface: QR Railway Tech. Res. Inst., 30(1), 25-33, 1989.

Neidell, N., and Turhan Taner, M.: Semblance and other coherency measures for multichannel data: Geophysics, 36, 482–497, 1971.

Panea, I., Draganov, D., Almagro Vidal, C., and Mocanu, V.: Retrieval of reflections from ambient noise recorded in the Mizil area, Romania: Geophysics, 79(3), Q31-Q42, 2014.

Pedersen, H. A., Debayle, E., Maupin, V., and POLENET/LAPNET Working Group: Strong lateral variations of lithospheric mantle beneath cratons–example from the Baltic Shield: Earth and Planetary Science Letters, 383, 164-172, 2013.

Place, J. A. P., A. Malehmir, K. Högdahl, C. Juhlin, and K. Persson Nilsson: Seismic characterization of the Grängesberg iron deposit and its mining-induced structures, central Sweden: Interpretation, 3, SY41–SY56, doi: 10.1190/INT-2014-0212.1, 2015.

Poli, P., Pedersen, H. A., Campillo, M. and the POLENET/LAPNET Working Group: Emergence of body waves from cross-correlation of short period seismic noise: Geophysical Journal International, 188(2), 549-588, 2012a

Poli, P., Pedersen, H., Campillo, M., LAPNET Working Group: Body-wave imaging of Earth's mantle discontinuities from ambient seismic noise: Science, 338, 1063-1065. DOI: 10.1126/science.1228194, 2012b.

Poli, P., Campillo, M., Pedersen, H. and the POLENET/LAPNET Working Group: Noise directivity and group velocity tomography in a region with small velocity contrasts: the northern Baltic shield application to the northern Baltic Shield: Geophysical Journal International, 192(1), 413-424, DOI: 10.1093/gji/ggs034, 2013.

Rost, S. and Thomas, C.: Array seismology: methods and applications: Rev. Geophys., 40, 2-1–2-27, doi:10.1029/2000RG000100, 2002.

Sambridge, M.: Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space: Geophysical Journal International, 138(2), 479–494, 1999.

Schweitzer, J., Fyen, J., Mykkeltveit, S., Gibbons, S. J., Pirli, M., Kühn, D., and Kværna, T.: Seismic Arrays in New Manual of Seismological Observatory Practice(NMSOP-2): German Research Centre for Geosciences, ISBN 3-9808780-0-7.9., 1–80, 2012.

Shapiro, N. M. and Campillo, M.: Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise: Geophysical Research Letter, 31, 2004.

Shapiro, N.M, Campillo, M., Stehly, L., and Ritzwoller, M. H.: High-Resolution Surface-Wave tomography from ambient seismic noise: Science, 307, 1615 –1618, 2005.

Schimmel, M., Stitsmann, E., and Gallart, J.: Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale: Geophysical Journal International, 184, 494-506, 2011.

Shirzad, T., and Shomali, Z. H.: Extracting seismic body and Rayleigh waves from the ambient seismic noise using the rms-stacking method: Seismological Research Letters, 86(1), 173-180, 2015.

Silvennoinen A.: Kuusamon ja Rukatunturin kartta-alueiden kalliopera. – Espoo : Geological survey of Finland, 1991.

Silvennoinen, H., and Kozlovskaya, E.: 3D structure and physical properties of the Kuhmo Greenstone Belt (eastern Finland): Constraints from gravity modelling and seismic data and implications for the tectonic setting: Journal of Geodynamics, 43(3), 358-373, 2007.

Silvennoinen, H., Kozlovskaya, E., Kissling, E., Kosarev, G., and POLENET/LAPNET working group: A new Moho boundary map for the northern Fennoscandian Shield based on combined controlled-source seismic and receiver function data: GeoResJ, 1, 19-32, 2014.

Tiira, T., Janik, T., Kozlovskaya, E., Grad, M., Korja, A., Komminaho, K., and BrŰckl, E.: Crustal Architecture of the Inverted Central Lapland Rift along the HUKKA 2007 Profile: Pure and Applied Geophysics, 171(7), 1129-1152, 2014.

Vidal, C. A., Draganov, D., Van der Neut, J., Drijkoningen, G., and Wapenaar, K.: Retrieval of reflections from ambient noise using illumination diagnosis. Geophysical Journal International, 198(3), 1572-1584, 2014.

Vinnik, L., Oreshin, S., Makeyeva, L., Peregoudov, D., Kozlovskaya, E., Pedersen, H., and Jämsen, T.: Anisotropic lithosphere under the Fennoscandian shield from P receiver functions and SKS waveforms of the POLENET/LAPNET array: Tectonophysics, 628, 45-54, 2014.

Wapenaar, K., Draganov, D., and Robertsson, J., (Editors): Seismic interferometry: History and present status: Society of Exploration Geophysicists, Geophysics Reprint Series No. 26, ISBN 978-1-56080-150-4, 2008.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A.: Tutorial on seismic interferometry: Part 1 — Basic principles and applications: Geophysics, 75(5), 75A195, 2010.

Weihed, P., N. Arndt, K. Billström, J-C. Duchesne, P. Eilu, O. Martinsson, H. Papunen, and R. Lahtinen: Precambrian geodynamics and ore formation: The Fennoscandian Shield: Ore Geology Reviews, 27(1–4), 273–322, doi: 10.1016/j.oregeorev.2005.07.008, 2005.

Yliniemi, J., Kozlovskaya, E., Hjelt, S. E., Komminaho, K., Ushakov, A., and SVEKALAPKO Seismic Tomography Working Group: Structure of the crust and uppermost mantle beneath southern Finland revealed by analysis of local events registered by the SVEKALAPKO seismic array: Tectonophysics, 394(1-2), 41-67, 2004

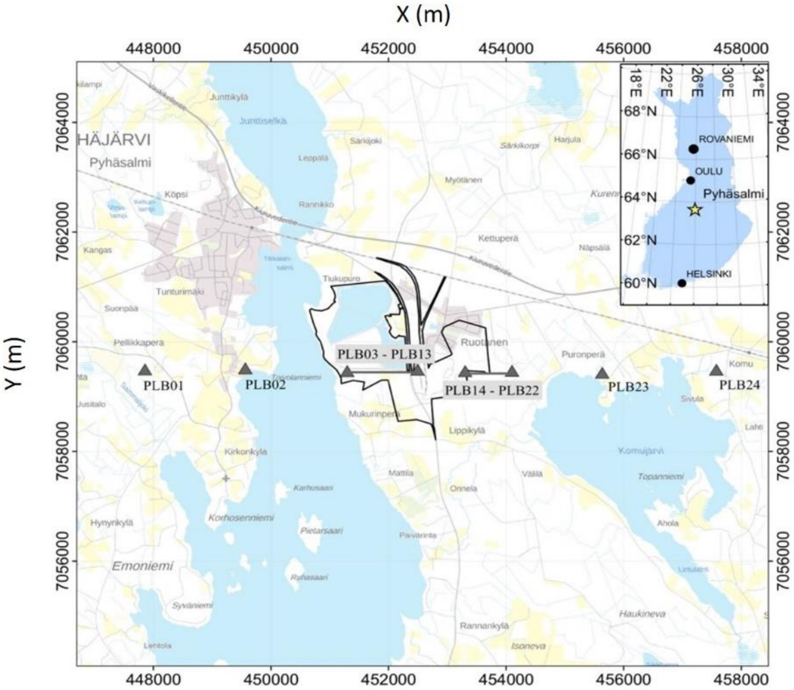

Figure 1: Map of the experiment in Universal Transverse Mercator coordinate system with the two parts of the profile (PLB01-PLB13 – west part of profile; PLB14-PLB24 – east part of profile). Black lines are the boarders of the mine and open-pit territories.

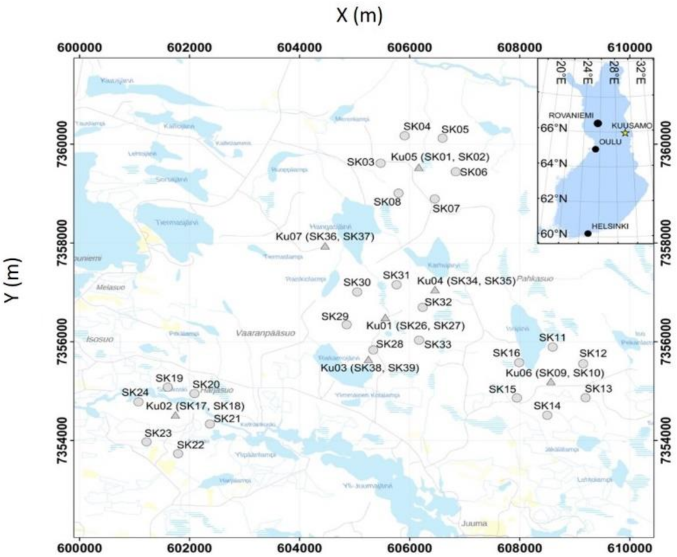

Figure 2: Configuration of the temporary seismic array in Universal Transverse Mercator coordinate system: white triangles – positions of geophones; white dots – positions of accelerometers.

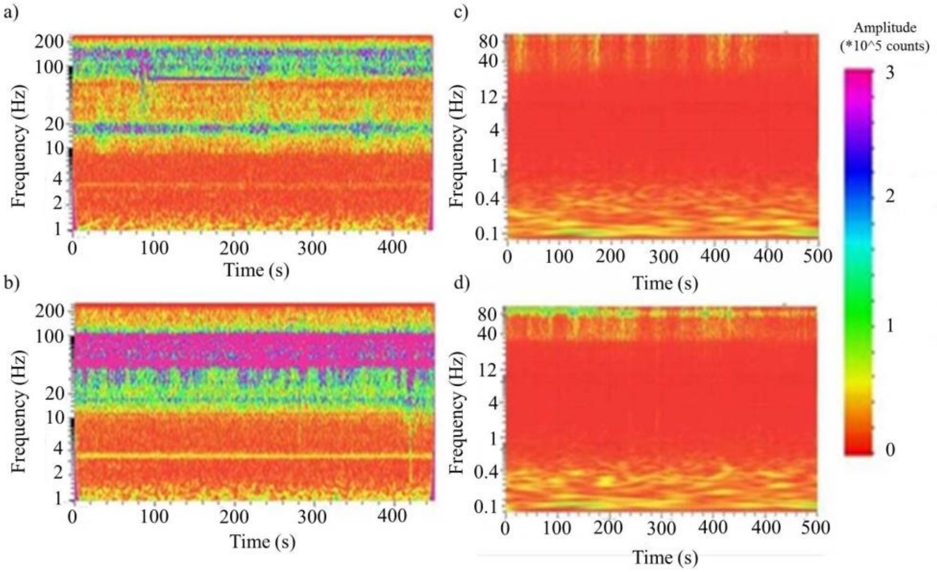

Figure 3: Result of time-frequency analysis of seismic noise recorded by the sensor a) the most distant from the mine in the Pyhäsalmi experiment, b) closest to the mine in the Pyhäsalmi experiment, c) most distant from a noise source (river) (KU05) in the large-aperture array in Kuusamo experiment, d) KU02 which is closest to the river in the large-aperture array in Kuusamo experiment.

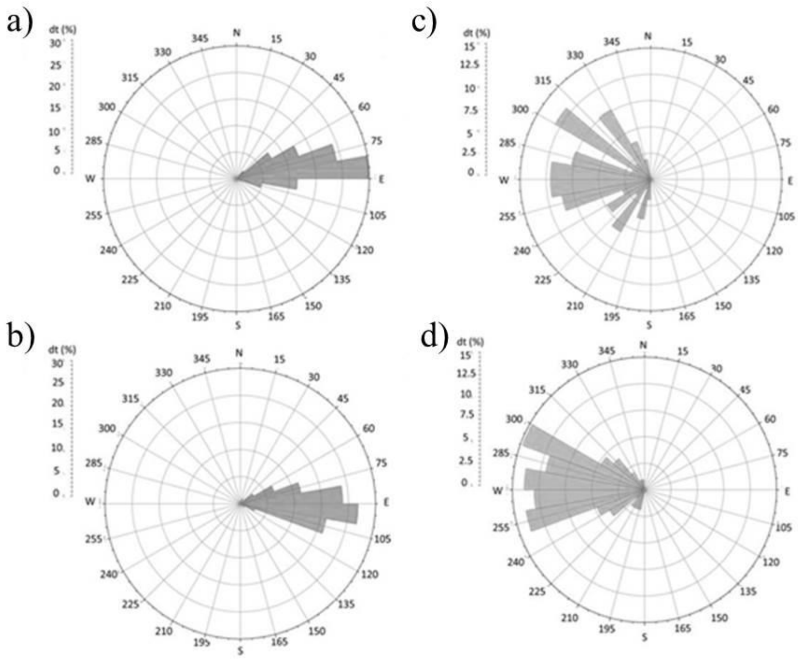

Figure 4: Result of azimuthal distribution calculation for different frequency bands for the Pyhäsalmi experiment: a) west part of profile, band of 2-5 Hz; b) west part of profile, band of 5-10 Hz; c) east part of profile, band of 2-5 Hz; d) east part of profile, band of 5-10 Hz.

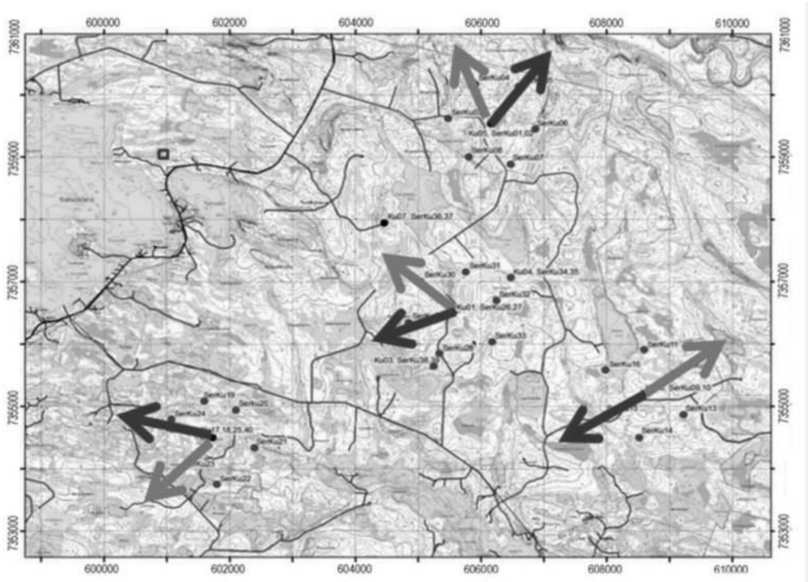

Figure 5: Azimuths to main noise sources: dots – stations of the temporary seismic array; black arrows show azimuths to noise sources in the frequency band 10-50 Hz; grey arrows show azimuths to noise sources in the frequency band 5-10 Hz.

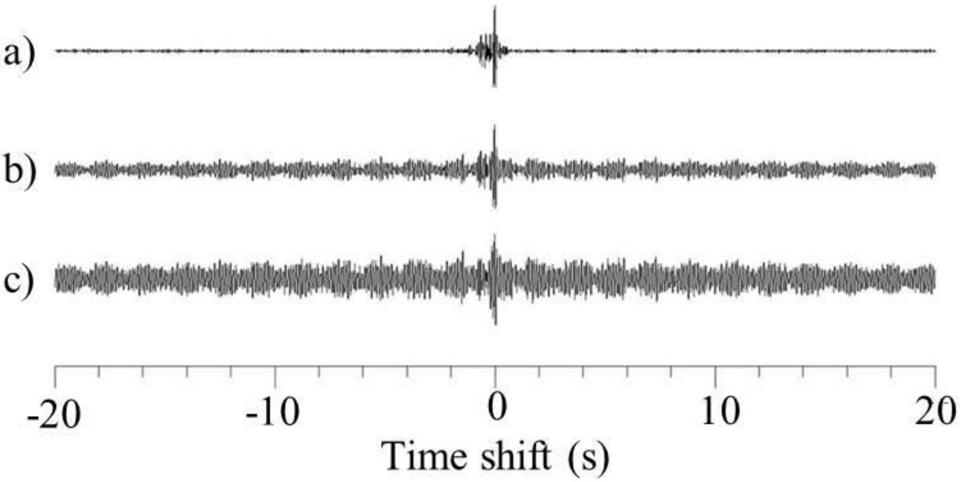

Figure 6: Final EGF's (vertical components) calculated by different methods of stacking in the time domain for the frequency band 5-10 Hz: a) SNR-stacking (SNR=71); b) Weight-stacking (SNR=10.4); c) RMS-staking (SNR=15.6).

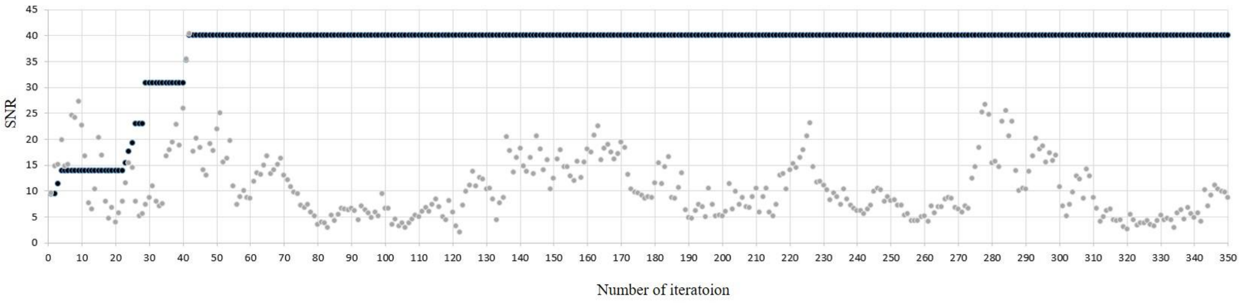

Figure 7: Build up process of EGF from cross-correlation functions by different methods: black dots – by SNR-stacking, grey dots – by simple stacking.

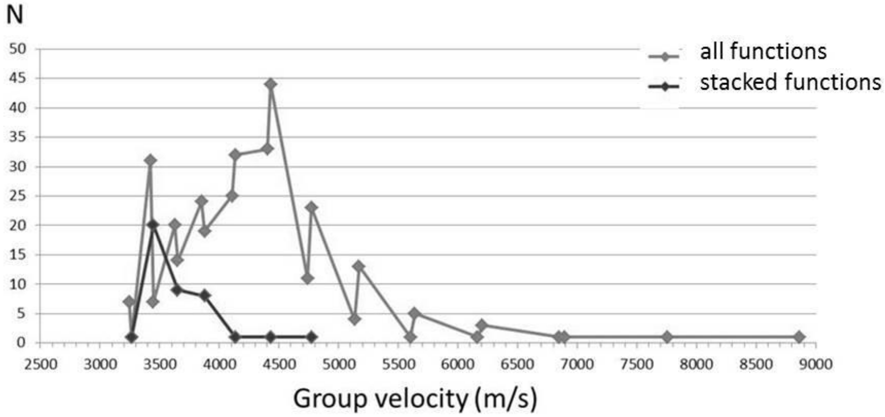

Figure 8: Distribution of EGF by group velocities for frequencies of 5-10 Hz.

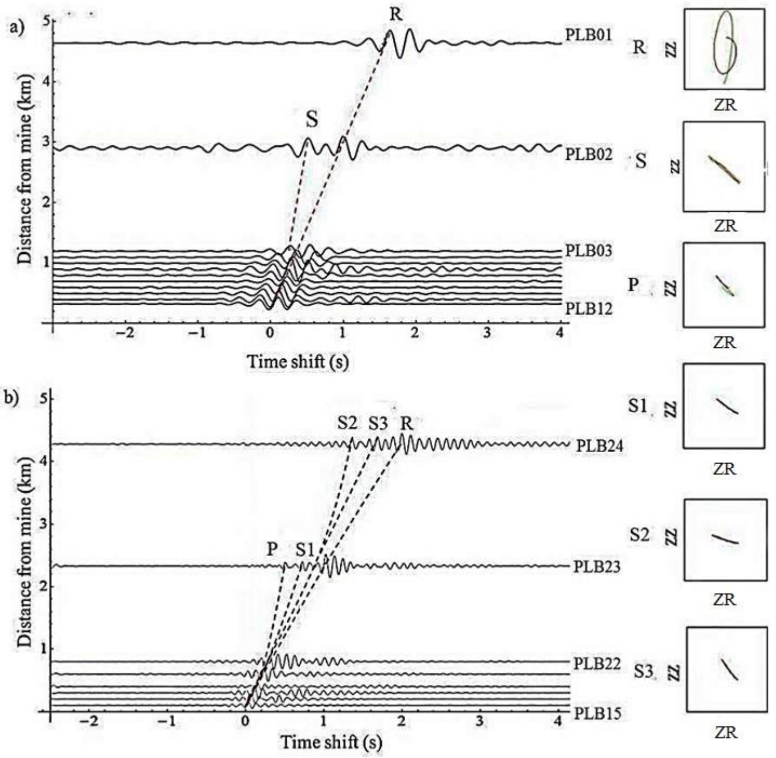

Figure 9: Result of stacking and particle analysis of EGF: a) western part of the profile in the frequency band 2-5Hz; b) eastern part of the profile in the frequency band 5-10Hz. 5

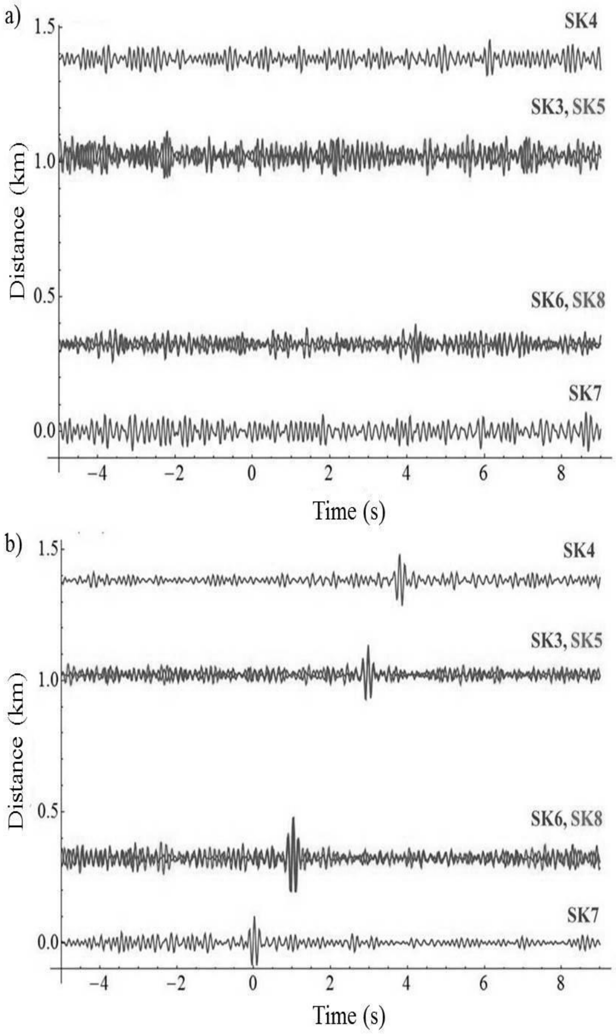

Figure 10: Empirical Green's functions calculated from records of small-aperture array in the frequency band of 5-10 Hz and stacked (vertical components): a) by simple stacking method; b) by SNR-stacking method. The EGFs in subplots a) and b) are sorted according to distance from sensor SK7.