# 2D body-wave seismic interferometry as a tool for reconnaissance studies and optimization of passive reflection seismic surveys in hardrock environments

Chamarczuk, Michał; Malinowski, Michał; Draganov, Deyan

**DOI**
10.1016/j.jappgeo.2021.104288

**Publication date**
2021

**Document Version**
Accepted author manuscript

**Published in**
Journal of Applied Geophysics

**Citation (APA)**
Chamarczuk, M., Malinowski, M., & Draganov, D. (2021). 2D body-wave seismic interferometry as a tool for reconnaissance studies and optimization of passive reflection seismic surveys in hardrock environments. *Journal of Applied Geophysics*, *187*, Article 104288. https://doi.org/10.1016/j.jappgeo.2021.104288

# 2D body-wave seismic interferometry as a tool for reconnaissance studies and optimization of passive reflection seismic surveys in hardrock environments


ABSTRACT

Despite the unrivalled spatial resolution and depth penetration of active-source seismic methods used for mineral exploration in hardrock environment, economic and environmental restrictions (e.g., source permitting) may preclude its full-scale application. In such a case, 2D passive reflection seismics can be considered a cost-effective way to perform reconnaissance-type survey and provide body-wave structural imaging using ambient-noise seismic interferometry (ANSI). This is, however, conditional to the presence of noise sources in the subsurface, for example produced by underground mining activity. Here, we propose a 2D ANSI workflow as an intermediate step prior to a full-scale 3D ANSI survey and an affordable tool in brownfield exploration, e.g., when trying to update current geological models beyond the drilled area. We test the applicability of this approach by analysing selected receiver lines from a 3D passive dataset acquired over the Kylylahti mine in Finland. Our methodology aims at choosing the optimal processing strategy at possibly lowest acquisition (2D geometry) and computational (small amount of data) cost. We address the fundamental questions in ANSI, i.e., (i) how much AN should one record and (ii) which SI processing approach should one choose. Therefore, we test different processing steps necessary to produce virtual shot gathers (VSG): preprocessing, selection of the ambient-noise portion, and selection of the method for retrieving the impulse responses between the receivers (crosscorrelation - CC, crosscoherence - CCh, multidimensional deconvolution - MDD). We conclude that trace energy normalization and high-pass filtering are the preferred preprocessing steps, while the best imaging is obtained when VSGs are retrieved using MDD applied in the noise-volume approach or CC in the event-driven approach. An event-driven approach may significantly reduce the acquisition time: for the Kylylahti dataset, using 10 events with energetic body-wave arrivals, extracted from one

hour of data, was enough to provide results comparable to the results from the noise-volume approach using the complete one hour of noise.

# 1. INTRODUCTION

The most comprehensive method to fully resolve the structural complexity characterizing highly deformed crystalline rocks hosting mineralization (referred to as 'hardrock environment', Eaton et al., 2003) are the 3D seismic surveys (e.g., Malehmir et al., 2012a). However, such surveys are not always the method of choice (Koivisto et al., 2012). Economic and environmental restrictions, source permitting, and challenging terrain conditions might significantly reduce the feasibility of a full-scale 3D survey at a given site (Cheraghi et al., 2012). Thus, 3D surveys are mainly conducted at well-recognized sites with ongoing exploration/production (brown-field exploration) with the aim to expand the knowledge about subsurface/reserves beyond the current geological models' boundaries (Malehmir et al., 2012b; White et al., 2012; Singh et al., 2019). In such cases, a 3D passive survey based on the principles of ambient-noise seismic interferometry (ANSI) offers a cost-effective solution.

The successful applications of seismic interferometry (SI) in reflection imaging, and in particular the possibility of using ambient noise (AN) instead of active (controlled) sources (Draganov et al., 2013), eventually brought the concept of SI to the mining industry (Cheraghi et al., 2015). Recent applications demonstrated the feasibility of ANSI to support imaging in operating mine environments using surface waves (Olivier et al., 2015b; Czarny et al., 2016) and body waves (Cheraghi et al., 2015; Roots et al., 2017; Polychronopoulou et al., 2020). These experiments involved both underground (Olivier et al., 2015a) and surface measurements (Cheraghi et al., 2015).

2D seismic reflection projects aim at reconnaissance and initial exploration at a regional scale (Cheraghi et al., 2011; Calvert and Li, 1999). However, when dealing with a complex 3D medium, only limited information of the true orientation of reflectors can be obtained from 2D surveys (Malehmir et al., 2012a, White et al., 2012). Ideally, one should combine both 2D (for higher resolution and lower acquisition cost) and 3D surveys (for wider azimuthal illumination and proper reflection positioning) (e.g., Hajnal et al., 2010).

Both 2D and 3D seismic surveys face the same problems typical for hardrock environments: strong scattering of seismic waves, low velocity gradients and small impedance contrasts between the rocks (with a notable exception of the massive sulphide mineralization), resulting in the inherently low signal-to-noise ratio (SNR), which is additionally degraded by the presence of anthropogenic noise (e.g., due to the mine infrastructure) (Eaton et al., 2003). All these factors impact negatively the reflections present in the 3D active-source data and decrease the overall SNR (compare the top and middle rows in Figure 1a).

Lower fold of the 3D surveys (and hence necessity of using wider bins and resulting lower resolution) is dictated by the source cost/effort. In the conventional, orthogonal design, shot lines are spaced between 1 to 2 receiver line spacings. In this regard, SI allows to obtain virtual-shot gathers (VSGs) at every receiver position, thus the dense array of receivers theoretically suffices to obtain high-fold 3D coverage and thus reflectivity similar to the active data (see bottom row in Figure 1a). The limitation though is related to the necessity of placing additional receivers to maintain the crossline fold.

As compared to the active acquisition, when designing a 3D passive survey, one should additionally consider, e.g., recording time, array geometry and its orientation with respect to the dominant noise sources, and number of receivers and their spacing (note again that sources will be retrieved at receiver positions). In such cases, the 2D geometry provides the minimal array configuration required for evaluating the dominant AN events present in the study area. Observations derived from processing steps of 2D ANSI regarding basic AN characteristics (periodicity and location of noise-sources activity and body-to-surface wave content) can help estimate the length of the recording time and location of the array. Another advantage of 2D ANSI is that for distant sources (i.e., the distance to a source is much larger than the length of the array), the plane-wave approximation allows one to treat the arriving energy as separate plane waves with small ray parameters corresponding to body waves (Ruigrok et al., 2010). Thus, as compared to 3D ANSI, the 2D approach allows broadening the effective stationary-phase region and utilizing more AN sources, as long as their phases are consistent and in-plane with the array.

Acknowledging the aforementioned limitations of 2D imaging and differences between passive and active surveys, we evaluate the 2D ANSI method as an intermediate step prior to a full-scale 3D ANSI survey, as well as a cost-effective solution for brownfield exploration. Towards this end, we use passive seismic data acquired over the Kylylahti mine in Finland and synthetic data simulated using the geological model of the mine area. The methodology we develop in this study might be used to: (i) evaluate the acquisition parameters for a potential follow-up 3D seismic survey (both active and passive), (ii) estimate the length of the recording time and selection of SI processing steps for 3D ANSI, (iii) mapping the general structural framework in the area of interest, and (iv) constructing 3D geologic model from a network of seismic profiles.

We investigate the whole 2D ANSI processing flow, including data preprocessing, up to the VSGs retrieval (Figure 2). We put special emphasis on the choice of (i) an SI technique used to retrieve the impulse responses between the receivers and (ii) segments of recorded AN used for retrieval of VSGs, providing best-quality imaging. With the latter, we address the fundamental questions related to ANSI: (i) how much noise one should use (acquisition time), (ii) whether one should process continuous recordings (i.e., AN volumes) or noise panels containing events from separate sources (i.e., event-driven approach; Draganov et al., 2013).

Since the imaging part, i.e., selection of a migration algorithm, is not the scope of this study, we choose a conventional approach used in hard-rock data processing and mining applications (e.g., Adam et al., 2003, Malehmir et al., 2012) of dip-moveout (DMO) stack and post-stack time migration to qualitatively compare the results of various SI approaches.

Note that our processing workflow can be evaluated at three different levels of details. At the level of 'General procedure' (Figure 2a), it contains all the steps we believe must be investigated in 2D ANSI. At the level of 'Recommended approach' (Figure 2b), we gather all the tools that should be used to address the 'General procedure'. At the final level of 'Variables specific for Kylylahti' (Figure 2c), we summarize our case-specific selection of tools and processing parameters.

We first describe the methodology used in this study: SI methods, selection of AN segments using illumination diagnosis, different stacking approaches and semblance analysis. Next, we briefly introduce our 2D ANSI workflow followed by the description of the dataset used in this study (Kylylahti array). To provide the basis for verification of the 2D ANSI results, we further perform numerical tests and investigate the overall feasibility of reflection retrieval with active and passive seismic methods using a simplified geological model representative for our study area. For the passive case, we additionally investigate the role of illumination imposed by one-sided AN source localization. Subsequently, the 2D ANSI workflow is applied to the field recordings to assess the performance of every processing step. We further compare VSGs retrieved using all 9 approaches (3 VSGs retrieval techniques vs. 3 segments of AN data). We evaluate the reflectivity retrieved in VSGs by visual inspection and an automatic quantitative measure (semblance). Then, we compare migrated sections for all 9 configurations. Finally, we show 2D ANSI processing results with the preferred workflow applied to three adjacent receiver lines from the Kylylahti array. The migrated sections consistently show repeatable reflectivity patterns, which were previously identified in the synthetic data.

# 2. METHODOLOGY

Our 2D ANSI processing workflow (Figure 2) builds upon and combines experiences from previous SI experiments for both oil and gas exploration and mining-industry applications (Draganov et al., 2013; Cheraghi et al., 2015) as well as studies analyzing the performance of different SI methods: crosscorrelation (CC), crosscoherence (CCh), and multidimensional deconvolution (MDD) (Snieder et al., 2009; Nakata et al., 2011; and Wapenaar et al., 2011). We start with a brief description of the specific challenges faced when adapting ANSI to hardrock environments, as well as justification of the 2D approach in case of the available 3D passive data and complex geology.

## 2.1 Challenges of adapting ANSI to hardrock environments

The complex hardrock environment is very challenging for active-source seismics, and thus poses a big challenge for ANSI as well, as the changes in temporal and spatial stationarity of noise sources may cause destructive interference of potential reflection events during stacking (compare rows in Figure 1b). It means that results from stacking smaller amounts of data might exhibit higher SNR than those obtained from more data (see top row of Figure 1b, where stack for one day exhibits different coherent events than those obtained for more AN). This counterintuitive observation is even more evident for weak reflectivity observed in 3D passive data (see bottom row in Figure 1a), which might be very easily hindered during the stacking process (in the Kylylahti data such reflections exhibit low SNR and are observed only along 20-30 neighbouring traces out of all 994 traces). Therefore, contrary to the conventional approach used in ANSI, i.e., recording as much noise as possible and then stacking all the noise panels (e.g., Cheraghi et al., 2015; Chamarczuk et al., 2018), one needs to be more selective in the stacking process. It applies both to the noise-volume and event-driven approach. As a remedy, we propose to include novel illumination-diagnosis techniques in the ANSI workflow, allowing to (i) assess the temporal and spatial stationarity of noise sources (useful for designing the 3D survey orientation and minimum recording time) and (ii) stack the periods of data containing body-wave illumination (useful in designing the processing workflow).

### 2.2 Value of the 2D approach at initial stages of processing of 3D data

When analyzing a full 3D passive dataset, the 2D ANSI approach seems to be a necessary intermediate step allowing to test the performance of an SI processing flow (SI techniques, AN segmentation, and preferred illumination-diagnosis techniques) at much lower computational cost. For example, the results shown in the bottom row of Figure 1a were obtained using an event-driven approach combined with CC (one of our preferred approaches) in a full 3D manner and visual inspection of hundreds of VSGs. Computing this collection of VSGs took approximately 60 hours on graphical processing units (GPU), and was preceded by several preprocessing steps (requiring analysis of the whole dataset, i.e., 600 hours of AN) necessary for the event-driven approach (which was only one of the possible solutions). Without initial testing using the 2D workflow (i.e., considering a representative receiver line), choosing the optimal combination of preprocessing, noise-panel selection, evaluating methods to retrieve impulse responses and the actual responses (e.g., choosing between causal and acausal part) and finally the stacking itself, would have been a daunting task.

### 2.3 Impulse-response retrieval: transient vs diffuse fields

Ambient noise can originate from a diffuse wavefield caused by multiple scattering in a heterogeneous medium and/or energy from transient sources in a deterministic medium (Wapenaar et al., 2004). From a practical point of view, it means that the reflection response of a medium can be obtained using SI either by correlating long recordings (possibly overlapping in time) of uncorrelated noise sources (diffuse-wavefield case) or by stacking correlations from separately acting sources (deterministic-wavefield case). In this study, we refer to those two cases as noise-volume and event-driven approach, respectively. To indicate the relevance of evaluating the influence of different segments of AN, we follow Wapenaar et al. (2006) and describe the retrieval of the impulse response (i.e., Green's function) by CC for the case of transient and uncorrelated noise sources.

The transient-source case relates to the situation when a noise panel contains the wavefield resulting from a single seismic source. In such a case, we can write the wavefields $u$ recorded by receivers at $x_A$ and $x_B$ in the frequency domain as

$$u(x_A, x_S, \omega) = S(x_S, \omega) G(x_A, x_S, \omega), \quad (1)$$

$$u(x_B, x_S, \omega) = S(x_S, \omega) G(x_B, x_S, \omega), \quad (2)$$

where $G(x_A, x_S, \omega)$ and $G(x_B, x_S, \omega)$ are the Green's functions recorded by receivers at $x_A$ and $x_B$, respectively, and $S(x_S, \omega)$ denotes the source time function.

Then, the correlation of those two wavefields is given by

$$C(x_B, x_A, \omega) = \oint_{S_{src}} |S(x_S, \omega)|^2 u(x_B, x_S, \omega)\, u^*(x_A, x_S, \omega) dx_S \quad (3)$$

where superscript asterisk $*$ denotes complex conjugation. Hence, the correlation function $C(x_B, x_A, \omega)$ is proportional to the Green's function between $x_B$ and $x_A$, $G(x_B, x_A, \omega)$, convolved with the averaged transient-source wavelet.

In the case of recording uncorrelated noise sources characterized each by source time function $N(x_S, \omega)$, the responses at $x_A$ and $x_B$ are defined as

$$u(x_A, \omega) = \oint_{S_{src}} N(x_S, \omega) G(x_A, x_S, \omega)\, dx_S \quad (4)$$

and

$$u(x_B, \omega) = \oint_{S_{src}} N(x'_S, \omega) G(x_B, x_S, \omega) dx_S, \quad (5)$$

respectively.

The assumption of uncorrelated noise sources invokes processing of continuous recordings of simultaneously acting passive sources. In this case, the summation over the sources is replaced with time averaging. We assume that two noise sources $N(x_S, \omega)$ and $N(x'_S, \omega)$ are mutually uncorrelated for any $x_S \neq x'_S$ and have an equal power spectrum. The ensemble average $\langle \quad \rangle$ taken over them is equal to

$$\langle N(x'_S, \omega) N^*(x_S, \omega) \rangle = S(\omega) \delta(x_S - x'_S) , \quad (6)$$

where $S(\omega)$ is the autocorrelation of the AN source.

Then, the correlation of equations 4 and 5 is defined as

$$C(x_B, x_A, \omega) = \langle u(x_B, \omega) u^*(x_A, \omega) \rangle. \quad (7)$$

The above discrimination between AN considered as originating from transient and simultaneously acting sources is our motivation to investigate the influence of different segments of the AN recordings on the reflection retrieval in crystalline environments. Equations 3 and 7 state that the correlation function yields the Green's function between $x_B$ and $x_A$ (Wapenaar et al., 2011). Considering practical applications, the characteristics of the AN wavefield at a given site determine the eventual preponderance of either approach.

The imprint of a source signature $S(\omega)$ on the correlation result $C(x_B, x_A, \omega)$ is removed by applying a wavelet deconvolution per every correlated trace. The deconvolution operator is estimated by extracting a short segment around t=0 s from the correlation result $C(x_B, x_A, \omega)$ for $x_A$=$x_B$, i.e., the trace autocorrelation (AC).

The important parameter to establish before the actual correlation is the length of the noise panels to be used as the input for SI (Draganov et al., 2007; Almagro Vidal et al., 2014; Cheraghi et al., 2016). Regardless of transient sources or volumes of noise, the minimum length should be greater than or equal to the two-way traveltime (TWT) to the deepest reflection event of interest. To capture possible surface-related multiples of that event (as well as multiply scattered contributions), this length should be further extended to at least double that TWT. Furthermore, because our processing is specific for a situation in which seismic events are induced by underground mine activity, the record length should account for the maximum depth of mining operations (~1000 m depth). Taking these factors into account, for all the analyses shown in this study, we use a window length of 10 s for both the noise-volume and event-driven approach.

## 2.4 Impulse-response retrieval: SI methods

Further part of the comparison analyzed in this study relates to the application of the three main SI methods to retrieve impulse responses (Wapenaar et al., 2011).

CCh is equivalent to the CC normalized in the frequency domain:

$$C_{ch}(x_B, x_A, \omega) = \int_{S_{src}} \frac{u(x_A, x_S, \omega) u^*(x_B, x_S, \omega)}{|u(x_A, x_S, \omega)||u(x_B, x_S, \omega)| + \varepsilon}, \quad (8)$$

where $u(x_A, x_S, \omega)$ and $u(x_B, x_S, \omega)$ are the responses at $x_A$ and $x_B$ in the frequency domain, $\omega$ denotes the angular frequency, and the asterisk denotes complex conjugate. In CCh, the nominator is equal to the CC (equation 3) and is divided by the amplitude cross-power spectrum $|u(x_A, x_S, \omega)||u(x_B, x_S, \omega)|$. The regularization parameter $\varepsilon$ is added to provide numerical stability, and could be estimated for example by taking 1% of the cross-spectrum value per each frequency component, averaged over many time windows. The spectral division in equation 8 removes any contributions related to noise-source wavelets. Thus, the estimation of $S(\omega)$ and wavelet deconvolution required by the CC approach is omitted in CCh processing.

Both CC and CCh are trace-by-trace operations and their definitions assume lossless medium, isotropic illumination from the sources and regular source distribution (assumptions not possible to meet in real data scenario).

There were many solutions proposed to account for directionally biased illumination (see Bakulin and Calvert, 2006; Snieder et al., 2006; Mehta et al., 2007). However, most of them are essentially deconvolution-based trace-by-trace operations which mostly account for source-wavelet issues and do not address the asymmetric illumination of noise sources (Wapenaar et al., 2011). Wapenaar et al. (2008) proposed a method in which deconvolution is performed on all traces simultaneously thus allowing to account for assumptions limiting the validity of the Green's functions retrieved using CC and CCh. In this multidimensional deconvolution (MDD), an improved version of the frequency-domain Green's function $G^S(x_B, x_S, \omega)$ is obtained by deconvolving the raw correlation output $C(x_B, x_A, \omega)$ by a 2D deconvolution operator $\Gamma(x, x_A, \omega)$:

$$C(x_B, x_A, \omega) = \int_{S_{rec}} G^S(x_B, x, \omega)\Gamma(x, x_A, \omega)\, dx, \quad (9)$$

where $G^S(x_B, x_S, \omega)$ is the scattered part of the Green's function (total Green's function minus the direct wave) and $\Gamma(x, x_A, \omega)$ is so-called point-spread function (PSF; van der Neut et al., 2010, 2011). Equation 9 shows that the CC function is actually a blurred variant of the Green's function $G^S(x_B, x_S, \omega)$, where the blur in time and space is quantified by the PSF and is connected to source-related factors (source time functions, source distribution, relative strength, etc.). The underlying assumption for MDD is that PSF be optimally obtained from CC (equation 3) such that it accounts for the source-related distortions. Note that the integration in equation 3 is performed along the source boundary $S_{src}$, while in equation 9 - over receivers, which removes the requirement for regular source distribution. However, this also means that regular receiver distribution is required. MDD is realized by solving for $G^S(x_B, x, \omega)$ in equation 9 by deconvolving $C(x_B, x_A, \omega)$ with $\Gamma(x, x_A, \omega)$. To avoid ill-posedness, equation 9 is solved for each source position $x_A$ and for each available source component at $x_A$, resulting in an ensemble of equations for $G^S(x_B, x, \omega)$. $\Gamma(x, x_A, \omega)$, as proposed by van der Neut et al. (2011), can be obtained by time windowing the CC output around t=0 s (summed over sources for the transient case or over time instances for uncorrelated noise sources). This time widowing yields a butterfly-shaped seismic record with its thinnest part at $x_A$=$x_B$ (Nishitsuji et al., 2016) and with slopes determined by the apparent slowness of the dominant events. In practice, MDD (equation 9) is recast in a matrix form using a least-square approach or a singular value decomposition (see e.g., Nishitsuji et al., 2016 for the details of MDD discretization).

In this study, we approximate the PSF by extracting the butterfly-shaped seismic record from body-wave events captured in individual noise panels (in the event-driven approach). For the noise-volume approach, we extract the PSF from every correlated 10-s-long panel. For both approaches, we invert for $G^S(x_B, x, \omega)$ using all noise panel simultaneously. For longer recordings, this becomes a computationally intensive task. As we are interested in the general performance of the noise-volume approach, we test MDD on an exemplary 1-hour-long recording.

The main difference between all three methods is related to the different type of deconvolution inherent in each of them. In CC, windowed AC of the master trace is used as a source-function estimate to divide each trace in the spectral domain. Deconvolution in CCh is done by dividing the CC output of two traces by the multiplication of the amplitude spectra of both traces, i.e., a cross-power normalization is actually performed. Compared to CC and CCh, the deconvolution in MDD is much more comprehensive, i.e., deconvolution is performed simultaneously for every trace in the currently analysed VSG. The theoretical improvements from applying MDD compared to CC are: (i) removing source signature(s), (ii) improved radiation characteristics of the retrieved source, (iii) relaxation of the assumptions of a closed surface of regularly sampled sources (when directional illumination is present), and (iv) works correctly in dissipative medium. A disadvantage of MDD is that it requires a regular sampling of the receiver array.

From a theoretical point of view, CC should result in the correct phase and relative amplitude of arrivals, CCh - only the phase, and MDD - the absolute amplitude, phase, and correction for the unbalanced illumination. The robustness of CCh is useful when recordings contain unwanted instrument noise or poor-coupling effects. It could be appealing in case of surveys performed in hardrock environment, where the terrain conditions can vary quickly across the survey, yet the unwanted effect of CCh is data whitening, which can be harmful for retrieving weak reflectivity.

### 2.5 Evaluation of preprocessing on virtual zero-offset data

Reflection retrieval using ANSI involves correlating the separate body-wave sources or volumes of noise with dominant presence of body-wave arrivals. In such cases, the first aim of the preprocessing is to assure that the AN segments contain body waves with higher energy than surface waves. Additionally, the routine part of ANSI preprocessing is trace-energy normalization applied for each noise panel (Draganov et al., 2007, 2013). The energy normalization aims at equalizing the contribution of each correlated panel, i.e., fulfilling the assumption of equal energy of the different AN sources (Ruigrok et al., 2010; Nishitsuji et al., 2016). The inevitable consequence of applying CC to noise panels is enhancing the strongest event present in the given data segment (Almagro Vidal et al., 2014). As a consequence, the virtual-source function is determined by the strongest event in the pre-correlated noise panel. The illumination diagnosis method (Almagro Vidal et al., 2014) can be used to scan for panels with dominant presence of body waves. However, by providing appropriate preprocessing, even those noise panels which are dominated by surface waves might be turned into useful data.

In our 2D ANSI workflow, we use AC (Clearbout, 1968; Daneshvar et al., 1995) to obtain virtual zero-offset data and asses the influence of given preprocessing on the Green's function retrieval. Analysing virtual zero-offset data allows for direct visual assessment of amplitudes and phases of waveforms in a time window of reflections indicated by simulated active-source data.

### 2.6 Selection of ambient-noise segments

The next step in the 2D ANSI processing workflow is the extraction of AN segments. The formulations of CC for transient sources (equation 3) and uncorrelated noise sources (equation 7) indicate two possible approaches in processing AN data (Draganov et al., 2013): (1) an event-driven approach, where separate noise sources can be detected and extracted from the continuous AN recordings, or (2) a noise-volume approach, where long, continuous data are automatically separated into time windows of equal length with the assumption that most of the windows contains body-wave events that after stacking will not be masked by retrieved surface waves. For data extraction in the event-driven and noise-volume approaches, we apply dedicated illumination diagnoses. For the noise-volume approach, we apply a 2D illumination diagnosis (Almagro Vidal et al., 2014; Panea et al., 2014) to choose 1 hour of the AN recordings dominated by low-slowness arrivals. Note that, when using the noise-volume approach with all the noise, there is no need for illumination diagnosis. For selecting body-wave events in the event-driven approach, we use the two-step wavefield evaluation and event detection (TWEED) method (Chamarczuk et al., 2019).

In the event-driven approach, we aim to choose sources bounding the target area. According to Wapenaar et al. (2008, and 2011) even though the MDD approach can be carried out without assumptions with respect to the regularity of the source positions. The MDD results quality depends mainly on the source density (with a rule of thumb of average horizontal distance between sources being less than half of the dominant wavelength; see Wapenaar et al., 2008). Thus, during noise-panel selection we try to find events fulfilling this condition and located approximately on the contour outlining the main target (in this case Kylylahti ultramafic body).

Assuming high-frequency body-wave events with peak frequency around 50 Hz (see Figure 3b) and background velocity of the host rock of 5000 m/s, the optimal source separation should be less than 50 m. As the distribution of sources is constrained by the location of the noise sources, the sources subset used in this study will only approximate the desired distribution.

In order to assure retrieval of reflections when applying the noise-volume approach to only 1 hour of AN, we have to make sure that this hour is characterized by a dominant body-wave energy. To estimate the general body-wave energy content we use the aforementioned 2D illumination diagnosis (Almagro Vidal et al., 2014; Panea et al., 2014). In this approach, we use the slant-stack transform (Chapman, 1981) of the wavefield $v$, $\tilde{v}(p,\tau) = \int v(x,\tau + px)dx$, where $p$ is the ray parameter, $x$ is the offset, and $\tau$ is the intercept time at $p = 0$. The slant-stack at $\tau = 0$ for each correlated noise panel $C^S$ can be described as

$$\tilde{C}^S(x_A, p, \tau) = \int C^S[x_B, x_A, \tau + p \cdot (x_B - x_A)]dx_B, \tag{10}$$

where $\tilde{C}^S$ is the representation of the virtual-source function of the transient source $S$ in the $\tau -$ p domain. Therefore, $\tilde{C}^S$ describes the dominant ray-parameter contribution from the transient source to the virtual source located at $x_A$ and recorded at $x_B$. Then, a discrimination test is performed by comparing the dominant ray-parameter value $max\left(\left\|\tilde{C}_L^s(x_A, \boldsymbol{p})\right\|\right)$ with a predefined ray-parameter threshold $p_{limit}$ characteristic for the body waves in the recording area.

For the event-driven approach, the number of used noise panels could be far lower than for the noise-volume approach (which is the case for the Kylylahti data), thus if for the latter it is sufficient to have the majority of the panels containing body-wave events, the event-driven approach demands that every noise panel contains body-wave events. To assure this, we use illumination diagnosis method extended to 3D, i.e., the TWEED method (Chamarczuk et al. 2019). TWEED was developed to overcome the insufficient crossline receiver spacing (i.e., no receiver lines in the crossline direction deployed) by simultaneous analysis of the adjacent (parallel) receiver lines.

We use the illumination diagnosis for extracting two AN segments dominated by body-wave activity and surface-wave activity. Then, we apply the same SI processing to both volumes and compare the resulting VSGs.

## 2.7 Semblance analysis

Because of the inherent ambiguity in visual comparison of reflection patterns observed in pre-stack data, we propose to use a similarity measure. Similarity measures are commonly used in comparison of multiple datasets from various sources (Cooper and Cowan, 2008) and are well-known tools for analysing active-source seismic data (Neidell and Taner, 1971). Aiming to decrease the subjectivity of visual comparison of our passive results, we incorporate the semblance method, which enables comparison of time-series data in quantitative manner.

Semblance filtering compares two datasets on the basis of their phase as a function of frequency. The semblance is calculated using the continuous wavelet transform (CWT; e.g., Sinha et al., 2005). The CWT is defined as the correlation of the given time series $h(t)$ with a scaled arbitrary wavelet $\Psi$:

$$CWT(u,s) = \int_{-\infty}^{\infty} h(t) \frac{1}{|s|^{0,5}} \Psi^* \left(\frac{t-u}{s}\right) dt, \quad (11)$$

where s denotes the scale, u is displacement, and $*$ denotes the complex conjugate. Using the wavelet approach allows to account for temporal variability in the spectral character. Comparison of the two wavelet-transformed time series can be achieved using the cross-wavelet transform according to

$$CWT_{1,2} = CWT_1 \times CWT_2^*, \quad (12)$$

with the result being a complex quantity with an amplitude $A = |CWT_{1,2}|$ and local phase $\theta = tan^{-1}(Img(CWT_{1,2})/Re(CWT_{1,2}))$. $CWT_{1,2}$ is the relation between the imaginary and real part of the cross-wavelet transform and is valued between $-\pi$ and $+\pi$. Then, the similarity between the two wavelet-transformed time series can be defined as semblance:

$$S = cos^n(\theta), \quad (13)$$

where n is a positive odd integer. S is a measure of the phase correlation between the two datasets and takes values between -1 and 1. In this study, we use the semblance value to compare the reflectivity patterns present in the VSGs retrieved with the 2D ANSI processing and in the synthetic active-source data.

### 2.8 Imaging approach

After applying our 2D ANSI workflow, we use the VSGs as input to standard time imaging to retrieve reflectivity sections. For simplicity, we use conventional constant-velocity Stolt f-k migration (Stolt and Benson, 1986) applied on DMO-corrected sections (with integral T-X DMO run on common-offset planes, Hale (1984)). Additionally, we apply a top mute to remove first arrivals and, in the case of the ANSI results, SI artefacts earlier than the first arrivals, and balance the amplitudes by dividing by the root-mean-square (RMS) value. We argue that the expected quality of the migrated sections can be already deduced from comparison of the VSGs. Since the scope of this study is limited to explaining and comparing different processing strategies for 2D ANSI in the mineral-exploration context, we focus on the SI methodology itself. Hence, selection of the imaging techniques for the recovery of the best-possible reflectivity image is outside the scope of this study. Despite the fact that considering the complexity of structures as in the Kylylahti mine area, pre-stack depth migration was deemed the most-appropriate approach (see Heinonen et al., 2019; Singh et al., 2019), we prefer to use the above post-stack time migration approach. Since the potential of 2D ANSI is to use it as a reconnaissance tool in mineral exploration, we assume that no detailed knowledge of the velocity structure is known prior to acquisition, which hampers application of pre-stack depth migration.

## 3. DATASET

The 2D passive data used in this study comprises a single receiver line (line 7) of the Kylylahti array (Chamarczuk et al., 2019). Additionally, lines 8 and 9 are used to show the consistency of the 2D ANSI processing results. Figure 3a shows the layout of the Kylylahti array, highlighting the lines used in this study and their relation to the known extent of the Kylylahti mineralization. The Kylyahti array was deployed as a part of the COGITO-MIN project tackling the cost-effectiveness of various novel seismic exploration technologies targeting high-resolution resource delineation (Riedel et al., 2018). The primary purpose of the Kylylahti array deployment was to advance the development of ANSI imaging techniques for mineral exploration and provide a baseline for testing novel array-processing techniques (see Chamarczuk et al., 2020).

The array was deployed over the active Kylylahti polymetallic mine (Outokumpu mineral belt, Eastern Finland) in the direct vicinity of the town of Polvijärvi. The array consisted of 994 receiver stations distributed regularly over a 3.5 x 3 km area with 200 m line spacing and 50 m receiver spacing. Each receiver station was equipped with a bunched string of six 10-Hz vertical-component geophones and a wireless data logger, recording AN at 2 ms for 20 hours per day during 30 days, resulting in ~600 hours of passive seismic data. The Kylylahti mine was active during the whole recording period. Routine mining activities included, among others, drillings (surface and underground), transporting ore and waste rock (surface and underground), scaling (underground), mine ventilation (surface). Another source generating strong energy are the mine blasts which occurred daily at depths ranging from a few hundred meters down to approximately 800 meters below the surface. We expect all of these activities to significantly contribute to the AN wavefield in the Kylylahti area and provide us the opportunity to record body-wave arrivals. In Figure 3b, we show the power-spectral-density (PSD) estimate for the whole array averaged over one day of recordings. The PSD analysis indicates a broad frequency content of AN in the Kylylahti area, with highest energy peaks between 10-15 Hz and 30-40 Hz. Areas in the direct vicinity of the mine (denoted with red dashed line in Figure 3b) exhibit PSD peaks also in the 65-80 Hz range.

## 4. NUMERICAL TESTS

To investigate the feasibility of 2D ANSI in a setting dominated by operating-mine activity, we perform 2D numerical tests including: (i) synthetic active-source data to provide a benchmark of the imaging quality expected from surface-seismic data (see Figure 4a); (ii) passive seismic simulation with regular source distribution to show the maximum achievable performance (in the case of the array deployed directly over the mine and assuming the AN sources to be generated by mine-related activities) of 2D ANSI (see Figure 5), and (iii) supporting test to evaluate the influence of the directional AN sources illumination breaking the omnidirectional AN distribution condition, which is a situation commonly encountered in field experiments (see Figure 6). The synthetic active-source data are our benchmark in this study for verifying the fidelity of reflections visible in the 2D ANSI. Furthermore, the synthetic active-source data juxtaposed with the 2D ANSI results from the field and synthetic data should indicate the potential deviations from desired imaging results, i.e., misplaced and/or flattened reflections, artifacts (near-surface noise, non-physical reflections), and general hints in terms of SNR. The velocity model used for modelling includes the ore body and is representative for one of the receiver lines of the Kylylahti array (receiver line 7 in Fig. 3).

### 4.1 Synthetic model

For all synthetic tests, we use a 3D seismic impedance model based on a simplified geological model of the Kylylahti area (Riedel et al., 2018). The model is based on comprehensive drilling (i.e., the geology at this location is well known) and consists of the following four main rock units: (1) the sulphide-bearing schist (SULBS), (2) Outokumpu ultramafic rocks (OUM) – Kylylahti body, (3) Outokumpu altered ultramafic rocks (OME), and (4) massive to semi-massive sulphide (S/MS) mineralisation. The ore body is located at approximately 300 m depth (indicated by a yellow inclusion in Figures 4a and 5a). The petrophysical characterization of the targets indicates that S/MS should cause a strong reflected signal when in contact with any of the hosting rocks, mainly due to the notably higher densities of ore compared to those of the other rock types (Luhta, 2019). In Table 1, we provide the P-wave velocities, densities, and impedances of the units building the input model used for the acoustic modelling. The white dashed lines shown in Figure 4a indicate areas of expected reflectivity.

For the 2D synthetic modelling, we used a receiver spread mimicking the field acquisition geometry, i.e., a line of 29 sensors placed on the top of the model with 50 m spacing giving a total length of the line of 1400 m. For such a configuration, the theoretical maximum unaliased frequency is equal to $f_{un} = V/(4Bsin(\theta))$, where $V$ is the average medium velocity, $B$ is the bin size, and $\theta$ is the geological dip. Assuming a dominant dip of the target of 60°, a common-depth-point (CDP) bin size of 25 m, and average P-wave velocity in the medium of 6000 m/s, the maximum unaliased frequency is approximately 58 Hz. The dominant frequency of the AN sources in the Kylylahti area is not higher than 60 Hz (see Figure 3b), thus we do not expect frequency aliasing for the passive results due to the steep dips.

The synthetic modelling was done using a 2D finite-difference acoustic modelling scheme (Thorbecke and Draganov, 2011). First, we test the performance of the 2D active-source seismic method using a linear array of sources deployed on the surface. To facilitate comparison with data retrieved using SI, the synthetic shots are collocated with the receivers. We use a pressure source with a Ricker wavelet with centre frequency of 60 Hz and 40 Hz for the active and passive case, respectively. The synthetic active-shot gather is shown in Figure 4b. In Figure 4c, we show the migrated section obtained from all synthetic active-shot gathers. The reflectivity related to the main geological units (shown with dashed black lines) is visible both on the pre-stack (Figure 4b) and post-stack data (Figure 4c). We can distinguish three reflection packages (RPs) related to the velocity contrast between the host rock and the background (RP1), the S/MS mineralization (RP2), and the bottom of the Kylylahti body (RP3).

Next, we simulate the passive seismic survey. We use idealized regular noise distribution of underground sources. The rectangular polygon of sources together with the free surface form a surface enclosing the area of potential reflectivity. Theoretically, integrating over this surface (summing over separate sources) should provide a reliable estimate of the subsurface reflectivity (Wapenaar and Fokkema, 2006). The exemplary procedure to obtain Green's functions for the synthetic data by CC for a central receiver acting as a virtual source is as follows: for a fixed source position, we crosscorrelate the trace at the central receiver with the traces at all other receivers; we repeat this for all sources along the 'box'; the result is then summed per receiver over all sources. The erroneous amplitudes visible in the VSG obtained for the synthetic passive case (Figure 5b) are related to deviating from the far-field approximation of source boundary from the receivers and the assumption of smooth impedance contrast across the source boundary (see the 'Discussion' section for more detailed explanation). The migrated section in Figure 5c exhibits reflections in all the areas of expected reflectivity. We note, that it is very unlikely that AN sources in the actual field situation would appear with such a regular distribution and that serious deviations from this preferred illumination could be expected in actual field conditions; however, we want to examine the best theoretically achievable performance of SI assuming preferable alignment of mine-induced seismic sources in the Kylylahti geological setting.

#### 4.1.1 Directional illumination test

To investigate the influence of directional distribution of AN sources, we investigate three scenarios with sources distributed along one of the three sides of the target area (distributions of sources shown as insets in Figure 6). We compare migrated images obtained using VSGs produced from pressure sources illuminating the target area from the left (Figure 6a), bottom (Figure 6b), and right (Figure 6c) side of the rectangle bounding the target area.

The migrated sections obtained from directionally biased source distributions are generally dominated by artifacts, but it is still possible to track the reflectivity in the expected areas (indicated with dashed black contours). Sources distributed underneath the target (Figure 6b) provide the clearest image of the three cases, in which each RP can be visually separated. The sources illuminating the target from the left side (Figure 6a) provide an image similar to the one from the bottom distribution, yet the presence of a strong dipping artifact stretching from the depth of 800 m until ~1000 m distorts the reflection related to the bottom of the OUM formation. The relatively worst image is provided using sources distributed along the right side of the target (Figure 6c), with a prominent horizontal artifact stretching for the whole section at the depth of 400 m and masking the RP related to part of the target with high-impedance inclusion (see RP2 in Figure 6c). On the other hand, the section shown in Figure 6c exhibits the highest level of SNR in the area between the RPs. Overall, the simulation results shown in Figure 6 aid the interpretation of the migrated field data by explaining artifacts related to directional source distributions.

### 4.2 Validity of the 2D approach: 3D synthetic modelling

One may argue that the qualitatively good results of ANSI imaging applied to 2D synthetics might be misleading as we are ultimately aiming at imaging complex 3D structures. In order to support the reliability of the 2D ANSI approach (Figures 5-6) we additionally performed 3D finite-difference acoustic modelling. We used SOFI3D open-source modelling code (Bohlen, 2002) to simulate 648 separately acting AN sources and the full 3D model for the Kylylahti area (Riedel et al. 2018). The locations of those sources were obtained from the result of the InterLoc procedure (Dales et al., 2017a) applied to the events detected with TWEED (Chamarczuk et al. 2019, 2020) using the Kylylahti array data. In such a way, we used realistic 3D distribution of passive sources. The sections shown in Figure 7 were obtained along the same receiver line as in the 2D synthetic case. Similar to the test of the directional illumination in 2D discussed above (Figure 6), we selectively stack sources on the left, bottom, and right of the target (Figure 7b, 7c, 7d, respectively). Additionally, we produce VSGs with all the sources included (Figure 7e) and a subset of sources mimicking the event-driven stacking (Figure 7f). When comparing the 2D results with those from the 3D approach, we can note that albeit the individual reflection packages are slightly shifted, structures inferred from the purely 2D approach agrees well with the synthetic model. Therefore, we conclude that the 2D ANSI can provide relatively robust imaging of 3D structures in the Kylylahti area.

# 5 FIELD DATA APPLICATION

## 5.1 Auto-correlations of traces

In this section, we evaluate the influence of the different AN preprocessing techniques by comparing virtual zero-offset traces. We obtain the zero-offset data by stacking the ACs of arbitrarily chosen one-hour-long AN segment. We focus on comparing the SNR in the time window of expected reflection arrivals and the general resemblance of the virtual data to its active counterpart. We compare zero-offset data for the 17th trace in the active-shot synthetic gather shown in Figure 4b and the corresponding trace extracted from line 7 of the Kylylahti array. Figure 8 shows the comparison of stacked zero-offset virtual traces retrieved using the various preprocessing schemes applied prior to AC. To facilitate the comparison between the synthetic active-shot trace and virtual ACs, we show the zero-offset active-shot trace after concatenating its time-reversed variant (Figure 8a). In order to demonstrate the influence of each processing step, we show the raw AC result (Figure 8b). Applying time windowing (Figure 8c) and filtering (Figure 8f and 8g) enhances the peaks in the time window not related to reflections (outside the grey shaded area shown in Figure 8) and the traces exhibit high-amplitude ringing noise. The high-amplitude event appearing between the pulse at t=0 s and the grey shaded area (the expected arrival time of target reflections) in the AC traces suggests possible problems with the near-surface noise (caused by destructive interference of reflection and spurious events) retrieved in VSGs. On the other hand, peaks visible around the time window related to target reflections suggest the possibility of retrieving such reflections in the VSGs. The 'ringing' appearance of traces indicates possible problems due to overall low SNR in the retrieved VSGs. The auto-coherence (Figure 8e) exhibits a single positive peak, which is due to the spectral whitening performed intrinsically with this process (it is the AC normalized similarly to CCh). Energy normalization does not significantly affect the shape of AC (Figure 8h). However, it has to be applied to assure equal contribution from separate stacks of correlated AN panels (Draganov et al., 2009). In general, applying one-bit and sliding-window energy normalization (Figures 8d and 8i, respectively) provides AC traces exhibiting highest amplitudes near the area of the expected reflections, while the ringing-amplitude effect visible in the raw CC (Figure 8b) is highly reduced. Based on those results, for the final processing of the field data we choose the routine time-windowing and energy normalization followed by high-pass filtering (Figure 8g) to further enhance the expected body-wave content.

## 5.2 Noise-volume quality control and selection using illumination diagnosis

In order to assure selection of high-quality AN segments for the 2D ANSI noise-volume approach, we apply illumination diagnosis to determine periods with the desired body-wave illumination. Subsequently, to show the relevance of illumination quality control (QC), we investigate the consequence of applying CC to two different hours of noise. Towards this end, we extract two AN segments dominated by body-wave activity and surface-wave activity, respectively. We apply the same SI processing to both volumes and compare the resulting VSGs.

Figure 9 shows the illumination diagnosis panels for three adjacent receiver lines 7, 8, and 9 (highlighted in Figure 3a). We obtain this plot by employing equation 10 for the whole day of recording (divided into 10-s-long noise panels) from those three lines and automatically picking the slowness characterizing the strongest event in each noise panel. We denote picks with green and black crosses for low- and high-slowness event, respectively. Note that line 7 contains the highest number of the low-slowness events, as it is located in the direct vicinity of the mine and the mine is expected to produce body-wave events. We select 1-hour-long recordings based on their illumination characteristics. Collating observations from all 3 receiver lines, we select the first 1-hour-long segment by choosing a period when at least several events with dominant slowness values fall into the limit of body-wave slownesses (< 0.2 s/km) for all lines (see hour A in Figure 9). We choose a typical AN recording dominated by surface-wave energy as the second data segment. For that, we select an hour when zero low-slowness events occurred simultaneously on adjacent receiver lines (see Hour B in Figure 9). As indicated in the ACs of the zero-offset virtual traces (Figure 8), the most optimal preprocessing sequence is RMS energy normalization followed by a high-pass filtering. We apply this sequence to both selected hours and then retrieve VSGs using equation 7. The VSG retrieved using hour A (Figure 10b) has higher SNR compared to the VSG obtained using hour B (Figure 10c). Both VSGs exhibit the same reflection events, but the gather obtained from the low-slowness hour is characterized by less artifacts (see the events inside the blue rectangles in Figures 10b, c). To assure minimum number of artifacts, for further comparisons we select VSGs obtained using hour A.

## 5.3 Event-driven 2D ANSI

We evaluate the performance of the event-driven approach of 2D ANSI using body-wave events detected with the TWEED method. In order to construct the contour enclosing the target (see subsection 4.1), we use the InterLoc method (Dales et al., 2017a) to compute the location of every event captured with TWEED. Those locations were already used to calculate 3D synthetics (subsection 4.2). From these hypocenters, we select an ensemble of events mimicking the synthetic regular noise-sources distribution shown in Figure 5a. In Figure 11a, we indicate ten body-wave events selected for the event-driven approach considering the orientation of line 7 and the geological section shown in Figure 4a. In Figure 11b, we show the seismograms of those 10 events. Note that in order to detect and locate those events, we needed to scan the AN data over the whole recording period. Some of the selected events overlap with the low-slowness events from hour A (marked by dashed white circles in Figure 11a). Compared to the event-driven approach, body waves from the single hour are distributed directionally and illuminate the target area mainly from the right side. This suggests that the imaging using the event-driven approach should produce less artifacts related to directional illumination compared to the noise-volume approach.

## 5.4 2D ANSI methods applied to various segments of ambient noise

### 5.4.1 Visual inspection

In this section, we apply CC, CCh, and MDD techniques to the AN segments consisting of (i) a single event, (ii) 10 events, and (iii) AN volume of 1-hour recording with the preferred illumination characteristics (hour A). In the subjective, visual comparison of the results described here we focus on: (i) resemblance to the synthetic active-source data, (ii) near-surface effects (up to 0.1 s TWT), (iii) general reflectivity content, and (iv) random noise on traces.

The VSG obtained with MDD (Figure 12 b, e, and h) exhibits the most prominent reflectivity for 1 hour (see the blue rectangle in Figure 12h). This result is also resembling the synthetic active-shot gather best. Generally, MDD exhibits the highest SNR of traces of all three 2D

ANSI techniques, and the first 0.1 seconds of the VSG exhibit distinguishable reflectivity. We also note that increasing the AN volume seems to have constructive influence on the quality of VSG retrieved with MDD.

The VSGs obtained using the CCh approach (Figures 12 c, f, and i) exhibit the lowest quality of all three applied SI techniques. The traces are noisy, almost no reflectivity is visible, and the near-surface artifacts seem to either dominate the whole gathers as in the case of 10 events and 1 hour (see Figure 12f and 12i, respectively) or the whole gather is dominated by random noise (Figure 12c). The best result from the CCh appears to be achieved for the case of 1-hour-long recording (Figure 12i), in which the reflectivity is partially similar to the one obtained with MDD for 1-hour-long recording (Figure 12g). Note the lack of any coherent events in Figure 12c.

The VSGs retrieved using CC (Figures 12d, g, and j) exhibit higher quality compared to the CCh results. The near-surface noise visible in the CCh results is not retrieved in all VSGs obtained using CC (note that some reflectivity in the first 0.1 s can be clearly tracked). The reflections in the green area expected from the synthetic active shot are best retrieved in case of 10 events (Figure 12g). Surprisingly, even the single-event CC (Figure 12c) allows to retrieve some reflectivity, yet clearly stacking over higher number of noise panels increases the SNR of most reflections and retrieves the new events. Generally, all VSGs (except the CCh for 1 event) exhibit more prominent reflectivity in the shallow parts of the data (first 0.1 s of TWT).

Another reflective feature retrieved with 2D ANSI and visible in the synthetic active-source data are the two events denoted with the shaded green colour in Figure 12. These are retrieved with all MDD approaches and CC for 10 events.

Apart from reflections expected from the synthetic active-source data, the VSGs contain some more coherent events. However, it is difficult to interpret them because their origin is uncertain.

Obviously, they are not predicted by our simplified geological model, but we should stress that it is hard to obtain shallow reflectivity from the real active-source data due to the shot-generated noise, muting, and low fold at shallow depths. Thus, they might be related to true geological features in the subsurface. The reflection events in the area of interest retrieved with 2D ANSI are shifted towards earlier times and exhibit less steeper dips compared to the synthetic active-source data.

### 5.4.2 Semblance analysis

In order to compare reflectivity patterns retrieved with the synthetic active and field passive data in a more objective manner, we calculate semblances employing equation 11. CWT is calculated as both a function of scale and time, and, therefore, allows measuring the temporal change of the phase. We use part of the traces falling into the spatio-temporal window denoted with the blue rectangle in Figure 12 as input data. We calculate the semblance between every part of the trace falling in the analysed window and its corresponding trace in the synthetic active shot. As a result, we obtain a 2D matrix with phase and amplitude correlation coefficient for every trace. For comparison purposes, we average all results over amplitudes and obtain the mean phase-correlation value for every sample per each trace. In Figure 13, we show these average semblances calculated for VSGs obtained with the nine different processing approaches. To indicate the benchmark value, Figure 13a shows the part of the synthetic active-source used as base input for semblance calculation and Figure 13b shows its auto-semblance exhibiting maximum correlation represented with a red colour. The semblance plots in Figure 13c to 13k are presented in the same layout as VSGs in Figure 12.

The red patches visible in the semblance plots indicate areas of high correlation, while the blue colour denote high anti-correlated part of data (all values fall in the range between -1 to 1). Considering that the 2D ANSI results contain significant amount of noise (see Figure 12), we expect that areas outside the targeted reflection would be strongly uncorrelated, because the synthetic active-source data do not contain noise. Thus, if the VSG contains a coherent reflection similar to the one observed in the synthetic active-source data, the semblance plot should display a broad continuous red patch extending over the whole plot with generally similar curvature as the reflection visible in Figure 13a. To some extent this feature can be observed in the MDD (Figures 13f, and i) and CC results (Figure 13h) and is highlighted with white arrows. A semblance anomaly appearing as scattered remnant of the above feature appears in Figures 13j, and 13k, possibly indicating the faded imprint of expected reflectivity. Another potentially significant feature is the red area visible in the top-right segment of Figures 13 c, e, f, j, k, and g, possibly related to partial correlation with the direct wave shown in Figure 13a. To facilitate distinguishing between semblance anomalies related to the direct wave and reflection event, we indicate the line separating both type of arrivals with the black dashed line.

The general orientation of positive anomalies in Figure 13 c-k is horizontal, implying that the coherent features are stretching across the receivers, yet the anomalies have a narrow temporal extent (usually up to several time samples). The semblance plots shown in Figures 13 c, d, e, and g exhibit relatively broad, scattered red patterns indicating similarity which is likely random. Thus, they are not related to credible reflectivity content in the synthetic active-source data and we qualify them as not-resembling the expected result. Overall, we interpret the semblance anomalies denoted with white arrows in Figure 13f, 13h, and 13i, as features related to part of the reflection shown in Figure 13a. This means that the VSG obtained with MDD on 10 events and 1 hour (see Figure 12e and 12h, respectively) as well as the VSG obtained with CC applied on 10 events (see Figure 12g) exhibit a reflection event similar to the one in the synthetic active-source data.

### 5.4.3 Imaging results

For all nine 2D ANSI configurations presented in Figure 12, we retrieve VSGs for every receiver position. Subsequently, we apply top-mute and amplitude scaling, common-offset

DMO with a constant velocity V=6000 m/s, and normal CDP stack. The CDP stack is migrated using constant-velocity Stolt migration and time-to-depth-converted with a constant velocity of 6000 m/s.

The resulting depth sections are shown in Figure 14 using the same layout as in Figure 12. As the quality of the imaged reflectivity differs significantly, we focus only on the general SNR of the retrieved images and quality of reflections retrieved with 2D ANSI in the areas of the reflectivity predicted by the synthetic model (see dashed black lines in Figure 14). The migrated sections obtained using the single-event approach (Figure 14 b-d) exhibit a similar, low-frequency blurred pattern of reflectivity for every tested method, with low SNR, where the target RPs are hardly distinguishable from the image noise. For the single-event approach, RP1 is best retrieved using MDD (Figure 14b) and is to some extent visible in the CCh result (Figure 14c). RP3 is best visible in the CC panel (Figure 14c). In the single-event case, all three sections (Figure 14b-d) contain reflectivity in the expected areas, yet they are difficult to interpret as they are masked by the reflection artifacts of similar order of amplitude. The RP2 retrieval is of the poorest quality.

The images obtained using 10 events (Figure 14e-g) are much clearer than those obtained from the single event. The lowest number of artifacts is obtained with the MDD approach (Figure 14e), yet the reflectivity packages expected from the synthetic data are best visible in the CC section (Figure 14g). Again, RP2 is poorly constrained. The x-shaped reflection visible in the CC result (see Figure 14g, in the proximity of the RP3 area) is discernible also in the CCh result (Figure 14f), yet it is shifted towards shallower depths. The CCh section obtained for 10 events (Figure 14f) exhibits almost no coherent reflections in the shallower part (up to 750 meters) and the only recognizable feature is the x-shaped reflection related to RP3.

The reflectivity images obtained using the noise-volume approach (Figure 14 h-j) bring the highest quality image for the MDD and CCh case. Especially, the MDD with 1 hour of AN

(Figure 14h) enables retrieval of reflections related to all RPs expected from the synthetic data (Figure 14a). On the other hand, the CC result for 1 hour (Figure 14j) brings a relatively poor image where no expected reflections can be tracked, with the exception of an 'x'-shaped reflection similar to the observed in Figures 14f and g, barely visible again in the RP3 area. The CCh results for 1 hour (Figure 14i) show RP1 and RP2; however, the ringing noise in the RP2 area and broad, horizontal artifacts, visible at approximately 750 m depth, are most likely not related to any geological features.

From all the images shown in Figure 14, the image resembling best the synthetic migrated section is obtained for the MDD 1-hour approach (Figure 14h). The second-best image resembling the synthetic active-source data is obtained using the event-driven CC approach (Figure 14g). Based on the visual inspection and semblance analysis of VSGs, we select the most optimal 2D ANSI approach, which is MDD applied to 1 hour of AN (see Figure 12h and 13h for the VSG and migrated section, respectively) and we use this approach to process the adjacent receiver lines. In Figure 15, we show migrated images for receiver lines 7, 8, and 9. Persistence of the imaged features across the receiver lines corroborates our findings.

# 6 DISCUSSION

## 6.1 Optimizing the array – inferences from synthetic modelling

The synthetic passive data obtained with a preferable, i.e., regular distribution of sources around the target, provide ANSI results similar to the ones from the synthetic active-source data, but also contain artifacts. The artifacts in the virtual-shot domain (Figure 5b) are represented with arrivals visible before the line of the first breaks. Furthermore, in the synthetic active shot, the reflection related to the target at 0.1 s at ~20-29th trace is stronger than the reflection on the opposite side of the gather, while the synthetic passive case exhibits a reversed tendency. These amplitude errors are related to the imperfect distribution of simulated sources, i.e., instead of a sphere with a large radius and/or sources in the far-field (Wapenaar et al., 2010), we used the rectangular polygon of sources located in the direct vicinity of the target. Furthermore, because the location of the mine imposes inducing seismic events mostly in the direct vicinity of the Kylylahti body, the contour of sources is crossing a sharp contrast in impedances (see the left flank of sources in Figure 5a). Not complying with the required source assumptions results in Green's functions with a correct phase of the arrivals, but with distorted amplitudes, which is clearly visible in Figure 5b. Fulfilling these assumptions in the field conditions would require moving the recording array away from the underground mine infrastructure, to approach the far-field approximation, but at the expense of a one-sided illumination. As shown in Olivier et al. (2016), the mine tunnels can act as scatterers, hence approximating the inhomogeneous medium, where seismic energy is scattered back to the receivers. In such case, one-sided illumination might be sufficient (Wapenaar et al., 2006a).

The image obtained from migration of the passive synthetic source data (Figure 5c) contains RPs similar to those in the synthetic active-source data, yet we could see a strong horizontal artifact hindering the clear outline of RP2; this artefact is mainly arising due to the right flank of the subsurface sources (see Figure 6c).

The illumination test shown in Figure 6 allows us to investigate the consequence of directionally biased illumination, which is a common issue in field measurements. The important conclusion from the reflection patterns visible in Figure 6 is the possibility to image the target even using an irregular sources distribution. The relatively best image is retrieved with sources underneath the target (see Figure 6b); however, such distribution is difficult to achieve in field conditions. Depending on the noise-sources location, we could obtain a response of the same structures but represented with different reflectivity patterns. For this reason, if it is possible to estimate the AN sources distribution prior to deployment (e.g., from the locations of the dominant noise sources in the area), one could estimate what part of the target would be illuminated best and how to layout the recording array with respect to the dominant AN source locations.

In the active mining camps, most of the AN sources would be related to routine mining activities concentrated in one place. Thus, the array should be oriented in accordance to the mutual orientation of the target and the mine area, with the general requirement to obtain a recording geometry allowing to capture the sources which emit wavefronts with ray paths connecting the traces being crosscorrelated and the point to be imaged. In the case of an operating mine where most activities are vertically aligned under the surface, the spatial distribution of ambient-noise sources may be approximated by a situation, where the sources are distributed in the vertical flanks (as shown in Figures 6a and 6c). For instance, if the target of interest is a dipping reflector, then the recording array could be deployed: (i) at some distance from the mine area (such that the far offsets for sources located toward the dipping direction are obtained, or (ii) directly above the mine for the sources located in direction opposite to the dip of the target reflector (see e.g., Roots et al., 2017 for details of imaging the dipping reflectors with SI).

### 6.2 Influence of data preprocessing

We incorporated two well-established SI QC tools: virtual zero-offset traces (Claerbout, 1968) and illumination diagnosis (Almagro Vidal et al., 2014) as parts of our ANSI workflow. Analysing virtual zero-offset traces allows for computationally effective evaluation of the preprocessing at the initial stage of data analysis. The pitfall related to assessment of amplitudes in a reflection time window from AC traces relates to the estimation of the deconvolution operator in the CC case. The side lobes visible in the grey shaded areas in Figure 8 are possibly related to the reflectivity targets; time-windowing of ACs around t=0 s might remove such events during the source-function deconvolution usually performed after stacking of all correlated panels (Draganov et al., 2009). This is a consequence of having relatively low frequency (causing broadening of the side lobes) in our data. Therefore, we additionally applied high-pass filtering. Out of all compared preprocessing techniques, the one-bit normalization is particularly effective solution for extracting coherent information from AN. By removing the amplitude information, it could potentially retrieve all coherent events travelling between the two receivers (Väkevä, 2019). However, for body-wave retrieval, it requires preferential illumination from body-wave sources (like having a receiver line oriented inline towards a railway, see e.g., Quiros et al., 2016). Otherwise, the body-wave events may be hindered by interfering, and usually much stronger. surface waves. In the Kylylahti case, the mine area is located approximately perpendicular to the line orientation, and thus applying one-bit normalization could result in degrading the body-wave arrivals. However, as demonstrated by Väkevä (2019), one-bit normalization in conjunction with bandpass filtering, and f-k filtering can be effectively used for suppressing the dominant surface-wave content and reveal the reflectivity content in the Kylylahti area. The recent developments in autocorrelation studies using AN recordings (Clayton, 2020) indicate the potential to further improve the performance of the preprocessing step in the 2D ANSI workflow.

### 6.3 Body-wave- vs surface-wave-dominated recordings

We used illumination diagnosis to identify periods of AN dominated by body waves. Noise-volume selection (one-hour-long recording in this study) is an ambivalent choice and has implication in the resulting VSGs (see Figure 10 b, and c). We argue that even when stacking continuous data (noise-volume approach), it is beneficial to perform illumination diagnosis and scan for the noise panels dominated by low-slowness events. The VSGs shown in Figure 10b and 9c suggest that stacking over volumes of noise recorded during different periods (see illumination characteristics of these periods in Figure 9) bring generally similar results, yet varying in terms of SNR of the retrieved reflections and number of artifacts. Practically, it means that acquiring longer recordings does not necessarily bring better results, as improvement mainly depends on the eventual body-to-surface-wave content ratio. The potential pitfall of stacking an hour dominated by body waves is that despite capturing events with low slownesses, their distribution might be asymmetric as shown in Figure 11a. An event-driven approach allows directly to choose which sources we want to stack and, hence, overcomes the directional-illumination issue. However, the need for scanning more data and computing the illumination direction of recorded sources makes this approach more computationally expensive. This process can be automated though with machine-learning tools (Chamarczuk et al. 2019, 2020). Furthermore, even scanning the whole available data volume does not guarantee proper illumination, i.e., the subset of the available sources to choose from is determined by the location of the sources comprising the AN at a given recording site, which in the case of the Kylylahti data is mostly limited to the extent of the mine infrastructure. The practical implication of choosing an event-driven approach over a noise-volume approach is in the required recording time, since theoretically a few body-wave events should bring equal results to stacking over long periods of noise, thus possibly reducing the necessary acquisition time if a number of events, deemed sufficient, is already detected.

### 6.4 Quality of the virtual shot gathers

The main goal of applying 2D ANSI is to produce VSGs, which will allow to obtain structural imaging comparable to the one from the active-source surveys, but using ambient-noise sources (Draganov and Ruigrok, 2015). We rated the performance of SI processing strategies by comparing VSGs obtained using three different SI methods and benchmarking the results with the synthetic active-source data. In the case of the Kylylahti data, the CCh method yields the noisiest results, with relatively better performance for 1 hour of noise (see Figure 12 i). We argue that possibly many reflections are retrieved with CCh, yet they are buried in the noisy traces of the retrieved VSGs. A reflection in the target area is present (see blue rectangles in Figures 12f and i), yet it is hard to track as it exhibits low SNR. The poor performance of CCh is mostly related to the relatively low SNR of any reflection events in the noise panels, which further gets undermined in the correlated gathers. The consequence of applying CCh is bringing all recorded events to the same amplitude level. As a result, when the raw noise panels contain surface waves, the virtual shots after CCh will contain reflection events with amplitudes of the same magnitude as surface waves. The CC produces generally higher-quality results than CCh (Figure 12 d, g, and j). The best result is obtained for the event-driven approach (see Figure

12g). The CC result for 1 event (Figure 12d) is comparable for the result retrieved using 1 hour of noise (Figure 12j). Furthermore, the result for 10 events (Figure 12g) resembles to some extent the MDD result for 1 hour of noise (Figure 12h), exhibiting both reflectivity in the target area (denoted with blue rectangles in Figure 12) and also deeper reflections (denoted with green dashed colour in Figure 12) visible in the MDD result. The better performance of CC over CCh is related to the fact that the deconvolution operator in CC is derived from the data itself, and allows suppressing the ringing-amplitude pattern visible in the CCh results. Theoretically, MDD should bring better results than CC, as deconvolving by PSF should correct for varying noise-sources signatures, intrinsic attenuation, and irregular noise-source illumination. Accordingly, the MDD technique seems to produce VSGs resembling most the synthetic active-source data (see Figure 12b-h). Next to the reflections expected from the synthetic data, deeper reflectivity is also retrieved. The best result for the MDD case is achieved using 1 hour of AN (Figure 12h). Since MDD relies on deblurring the correlation output with PSF (equation 9), the main reason for differences observed in the virtual shots compared to the other two techniques is related to the PSF estimation. Deblurring the correlation function with PSF should eliminate the crosstalk from the correlation function and give the deblended virtual-source response. The potential distortions of the MDD result might be related to crosstalk contributions contained in PSF itself (Wapenaar et al., 2011), as well as incorrect extraction of the PSF. The exact influence of the PSF estimation on the reflection retrieval deserves a separate study, but is also thoroughly discussed, e.g., in Nishitsuji et al. (2016). On top of the shallow reflectivity, the MDD results brought also very clear reflections at ~1 s (not shown here); however, their fidelity is yet to be verified and is outside the target depths for exploration (but such reflectors were present in the active-source imaging of Heinonen et al. 2019 and Singh et al. 2019).

Generally, all VSGs exhibit very prominent reflectivity at shallower depths. The differences between the three techniques are mostly related to the specific type of deconvolution implied by them. The event-driven approach for MDD and CCh (Figure 12e and f) performs worse than the noise-volume approach. Nevertheless, these gathers still exhibit some reflectivity, which is promising in terms of similar SI applications with refined processing. For the MDD method, when adding more AN, both the deconvolution operator and the scattered field become updated, while for the CC only the raw correlation is stacked and the deconvolution operator is derived from the stacked correlogram, hence, theoretically, it cannot account for the whole complexity of the wavefield. For the CC method, the event-driven approach (see Figure 12g) performs best, and it could be potentially further enhanced by applying deconvolution per-event (Ruigrok et al., 2010).

### 6.5 Imaging

The general diversity of the retrieved reflectivity makes it hard to directly asses the quality of the processing approaches for selecting an optimal one using these migrated images. For this reason. and from the point of view of computational efficiency, the assessment of the effectiveness of the processing strategies should be carried out before migration, at the level of noise panels and VSGs (including e.g., the introduced semblance evaluation method). We leave it up to the reader to review the presented images and draw their own conclusions. Yet, based on both the visual similarity to the synthetic shots and semblance analysis, we think that the results from MDD with 1 hour of AN (Figure 14h) and CC with 10 events (Figure 14g) exhibit reflectivity resembling most the image obtained from the synthetic data. The results of imaging for adjacent receiver lines (Figure 15) suggest the redundancy of 2D ANSI imaging (provided the same processing sequence is applied to the data from every line). The reflectivity packages in the synthetic section are visible in line 7 collocated with the synthetic model, and the reflectivity patterns visible along the two adjacent lines deployed to the south of line 7 (Figures 15 c and d) are persistent.

The results shown in Figure 15 were obtained with only 1 hour of AN. The hour used for imaging was selected after thorough illumination diagnosis (see Figure 9 for illumination-diagnosis QC panel). Note that even if the amount of data we used is not significant, we still needed to process the bigger dataset to increase the probability of capturing AN with satisfying illumination characteristics. Some of the RPs visible in the migrated sections might still be related to out-of-plane targets and the full-3D ANSI processing could address their geometry properly.

### 6.6 Future developments and recommendations

The core of the 2D ANSI methodology is the comparison of different SI processing techniques to determine the most optimal sequence of processing steps for a given case study. Practically, it means that it requires repeating the complete processing flows starting from extraction of the recorded passive data up to the generation of the VSGs. This idea is illustrated in Figure 16, where the 2D ANSI methodology is represented with parallel processing flowcharts allowing for the practical implementation of the comparison between various tools advocated in this study.

We believe that 2D ANSI is capable of imaging targets in a hardrock environment. However, the quality of the imaging, apart from the acquisition geometry and strength of the impedance contrasts and the complexity of the medium (e.g., dip angles), depends on the selected SI method and the selection of AN segments. The results obtained for the real-case scenario of the complex structure at the Kylylahti site are generally quite noisy, and most of the compared approaches differ significantly in the imaging quality. Therefore, when applying 2D ANSI, a comparison of different approaches should be an essential part of the processing workflow.

We recommend using 2D ANSI as a reconnaissance tool prior to the massive 3D deployments (either active or passive) as it allows to determine the cycle of body-wave event activity (related to the mine operations). As we show in Figure 9, some periods of AN exhibit higher density of body-wave events. For instance, receiver line 7, i.e., the line which is closest to the mine area, could serve as a good indicator of potential body-wave content in AN in the Kylylahti area. Knowing that the mine operations produce body-wave events in repeatable cycles could allow reducing the recording time with a full array to periods when mine-induced sources with preferable illumination are likely to be recorded. Our results can be used to derive general recommendations in terms of planning future SI experiments for mineral exploration purposes: (i) orienting the acquisition array considering the location of the dominant noise sources and targets, and (ii) considering the possibility to reduce the continuous recording time to preselected time period, i.e., it is safe to assume that scheduled drilling and mine blasting occurring at several hundred meters below the surface would produce body-wave events. Each passive dataset would be recorded in a different AN setting, yet because we address active mine camps, the general characteristic of the AN wavefield would exhibit similar dominant features due to the dominance of mine-induced noise (Cheraghi et al., 2015; Oliver et al., 2016a,b; Dales et al., 2017a,b; Roots et al., 2017) and particularly unusually high ratio of body-to-surface wave energy with highly asymmetric distribution. In the Kylylahti case, for a single 3D virtual shot gather, reflection events are observed only along 20-30 traces out of all 994 receivers (top row in Fig. 1a) and are much less prominent than in the active 3D data (bottom row in Fig. 1a). The processing method aiming for automatic detection of such sparse, coherent events deserves dedicated tailored approach including scanning of hundreds of receivers for hundreds of virtual shots, which is not established yet and computationally not feasible. At this point we argue that prior to developing the 3D ANSI methodology, it is beneficial to know which combination of SI techniques and AN segmentation offers the highest probability for reflection retrieval and as such we need to develop 2D ANSI methodology first.

# 7 CONCLUSIONS

We introduced a 2D ambient-noise seismic interferometry (ANSI) processing workflow, which can be applied to passive seismic data acquired along a test profile to serve as a reconnaissance tool for AN evaluation and future more detailed seismic acquisition (passive or active) in active-mine environments. Using synthetic and field data from the Kylylahti mine (Finland), we indicated the relevance of 2D ANSI in general structural delineation and optimization of both the acquisition design and AN recording parameters. The synthetic modelling indicated that the passive data can be used to reproduce similar details of the complex geological model as the active-source data. The differences are attributed mainly to AN sources illuminating the target from different angles than the active shots. The key point of our workflow was the comparison of the performance of SI by multidimensional deconvolution (MDD), crosscoherence (CCh), and crosscorrelation (CC) on various AN segments: single body-wave event, event-driven approach using 10 body-wave events, and a noise-volume approach using 1 hour of AN recordings. The primary general conclusion of the comparison is the necessity to recognize the spatial and temporal distribution of the AN sources in the recording area. Based on this information, synthetic tests should be performed and the 2D receiver line for passive acquisition should be subsequently oriented with respect to the expected dominant AN sources and the imaging target. After acquisition of 2D passive data, different processing schemes should be evaluated using the methodology we proposed in this study. We showed that the final outcome of the 2D ANSI workflow provides initial target delineation, which facilitates the decision about conducting follow-up 3D surveys (active or passive) or using a network of 2D lines. These future experiments should be designed with the acquisition parameters, length of recording time, and SI processing workflow indicated by the initial 2D ANSI assessment.

The application of the full processing workflow on 2D receiver lines extracted from a passive dataset recorded at the Kylylahti mine led to the following conclusions specific for this case study. (1) The effectiveness of the AN preprocessing could be evaluated on zero-offset data. For the Kylylahti dataset, the sequence of trace-energy normalization and high-pass filtering provided the highest amplitudes in the retrieved body-wave arrivals and minimized the artifact contribution. (2) The 2D illumination diagnosis applied to AN for the noise-volume approach increased the signal-to-noise ratio of the reflection events in the retrieved VSGs, and thus we recommend it as a routine SI processing step. Illumination diagnosis applied for the event-driven approach provided results bearing similar quality, but obtained using significantly smaller amount of AN data. For the Kylylahti dataset, using 10 body-wave events, extracted from one hour of AN, was enough to provide results comparable to the results from the noise-volume approach using the complete one hour. (3) VSGs retrieval using the MDD method applied using the noise-volume approach and CC using the event-driven approach provided the highest quality data with reflection events resembling the active-source data the most. (4) Semblance analysis is an effective tool to aid the visual comparison of the passive and active-source data in selected spatio-temporal windows. (5) For the optimal selection of an SI technique and AN segment, the subsurface image can be obtained using a simple post-stack migration scheme, which requires only little knowledge on the velocity model. (6) The 2D ANSI processing workflow applied to the Kylylahti data provided images of the subsurface acceptable in terms of the general delineation of the target structures, as verified by comparison of results along adjacent receiver lines.

## ACKNOWLEDGEMENTS

The COGITO-MIN project was funded under the ERA-MIN initiative and received funding in Poland from the National Center for Research and Development (NCBR). We thank numerous people from the University of Helsinki, the Geological Survey of Finland, IG PAS, Boliden, and NovaSeis for arranging, deploying, and maintaining the Kylylahti array. This paper received comments from 12 reviewers, which are all thanked for their constructive comments. Despite the almost impossible job of satisfying all of them simultaneously, they helped us to make the paper more comprehensible.

## REFERENCES

Adam, E., G. Perron, G. Arnold, L. Matthews, and B. Milkeriet, 2003, 3D seismic imaging for VMS deposit exploration, Matagami, Quebec, in D. W. Eaton, B. Milkereit, and M. H. Salisbury, eds., Hard rock seismic exploration: SEG.

Almagro Vidal, C., D. Draganov, J. van der Neut, G. Drijkoningen, and K. Wapenaar, 2014, Retrieval of reflections from ambient noise using illumination diagnosis: Geophysical Journal International, 198, 1572-1584, doi: 10.1093/gji/ggu164.

Bakulin, A., and R. Calvert, 2006, The virtual source method: theory and case study: Geophysics, 71, no. 4, SI139 SI150, doi: 10.1190/1.2216190.

Balestrini, F., D. Draganov, A. Malehmir, P. Marsden, and R. Ghose, 2020, Improved target illumination at Ludvika mines of Sweden through seismic-interferometric surface-wave suppression: Geophysical Prospecting, 68, 200-213. doi:10.1111/1365-2478.12890.

Bohlen, T., 2002, Parallel 3-D viscoelastic finite difference seismic modelling. Comput. Geosci. 28, 887–899, doi:10.1016/S0098-3004(02)00006-7

Calvert, A. J., and Y. Li, 1999, Seismic reflection imaging over a massive sulfide deposit at the Matagami mining camp, Quebec: Geophysics, 64, 24 32, doi: 10.1190/1.1444521.

Chamarczuk, M., M. Malinowski, Y. Nishitsuji, J. Thorbecke, E. Koivisto, S. Heinonen, S. Juurela, M. Mężyk, and D. Draganov, 2019, Automatic 3D illumination diagnosis method for large-N arrays: robust data scanner and machine learning feature provider: Geophysics, 84, no. 3, Q13 Q25, doi: 10.1190/geo2018-0504.1.

Chamarczuk, M., Y. Nishitsuji, M. Malinowski, M. and D. Draganov, 2020, Unsupervised Learning Used in Automatic Detection and Classification of Ambient-Noise Recordings from a Large- Array: Seismol. Res. Lett., 91, 370-389, doi: 10.1785/0220190063.

Chapman, H., 1981, Generalized Radon transforms and slant stack: Geophysical Journal International, 66, 445-453, doi: 10.1111/j.1365-246X.1981.tb05966.x.

Cheraghi, S., A. Malehmir, and G. Bellefleur, 2011, Crustal-scale reflection seismic investigations in the Bathurst Mining Camp, New Brunswick, Canada: Tectonophysics, 506, 55 72, doi: 10.1016/j.tecto.2011.04.011.

Cheraghi, S., A. Malehmir, and G. Bellefleur, 2012, 3D imaging challenges in steeply dipping mining environment: New lights on acquisition geometry and processing from the Brunswick no. 6 seismic data, Canada: Geophysics, 77, no. 5, WC109, doi: 10.1190/geo2011-0475.1.

Cheraghi, S., J. A. Craven, and G., Bellefleur, 2015, Feasibility of virtual source reflection seismology using interferometry for mineral exploration: A test study in the Lalor Lake volcanogenic massive sulphide mining area, Manitoba, Canada: Geophysical Prospecting, 63, 4, 833-848, doi: 10.1002/2015JB011870.

Cheraghi, S., D. J. White, D. Draganov, G. Bellefleur, J. A. Craven, and B. Roberts, 2016, Passive seismic reflection interferometry: A case study from the Aquistore CO2 storage site, Saskatchewan, Canada, Geophysics, 82, no. 3, doi: 10.1190/geo2016-0370.1.

Clayton, R., 2020, Imaging the Subsurface with Ambient Noise Autocorrelations, Seismological Research Letters, 91, 2A, 930-935, doi: 10.1785/0220190272.

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, 33, 264 269, doi: 10.1190/1.1439927.

Cooper, G.R.J., and D.R. Cowan, Comparing time series using wavelet-based semblance analysis: Computers and Geosciences, 34, no. 2, 2008, 95-102, doi: 10.1016/j.cageo.2007.03.009.

Czarny, R., H. Marcak, N. Nakata, Z. Pilecki, Z. Isakow, 2016, Monitoring velocity changes caused by underground coal mining using seismic noise: Pure and Applied Geophysics, 173, no. 6, 1907-1916, doi:10.1007/s00024-015-1234-3.

Dales P., P. Audet, G. Olivier, and J.P. Mercier, 2017a, Interferometric methods for spatio-temporal seismic monitoring in underground mines: Geophysical Journal International, 210, 731 742, no. 2, doi: 10.1093/gji/ggx189.

Dales, P., P. Audet, and G. Olivier, 2017b, Seismic interferometry using persistent noise sources for temporal subsurface monitoring: Geophysical Research Letters, 44, 10, 863-870, doi: 10.1002/2017GL075342.

Daneshvar, M., R., C. S. Clay, and M. K. Savage, 1995, Passive seismic imaging using microearthquakes, Geophysics, 60, 1178-1186, doi: 10.1190/1.1443846.

Draganov, D., K. Wapenaar, W. Mulder, J. Singer, and A. Verdel, 2007, Retrieval of reflections from seismic background-noise measurements: Geophysical Research Letters, 34, L04305, doi: 10.1029/2006GL028735.

Draganov, D., X. Campman, J. W. Thorbecke, A. Verdel, and K. Wapenaar, 2009, Reflection images from ambient seismic noise: Geophysics, 74, no. 5, A63-A67, doi: 10.1190/1.3193529.

Draganov, D., X. Campman, J. W. Thorbecke, A. Verdel, and K. Wapenaar, 2013, Seismic exploration-scale velocities and structure from ambient seismic noise (>1 Hz): Journal of Geophysical Research: Solid Earth, 118, no. 8, 4345-4360, doi: 10.1002/jgrb.50339.

Draganov, D. S., and E. Ruigrok, 2015, Passive seismic interferometry for subsurface imaging. In M. Beer, E. Patelli, I. Kougioumtzoglou, and I. Siu-Kui Au, eds., Encyclopedia of earthquake engineering,1-13, doi:10.1007/978-3-642-36197-5_378-1.

Eaton, D, B. Mikereit, and M. Salisbury, 2003, Hardrock Seismic Exploration, SEG.

Hajnal, Z., D. White, E. Takacs, S. Gyorfi, I. R. Annesley, G. Wood, C. O'Dowd, and G. Nimeck, 2010, Application of modern 2D and 3D seismic reflection techniques for uranium exploration in the Athabasca Basin, in Lithoprobe: Parameters, processes and the evolution of a continent: Canadian Journal of Earth Sciences special issue, 47, 761 782, doi: 10.1139/E10-026.

Heinonen, S., M. Malinowski, F. Hloušek, G. Gislason, S. Buske, E. Koivisto, and M. Wojdyla, 2019, Cost-effective seismic exploration: 2D reflection imaging of the mineral system of Kylylahti, Easter Finland: Minerals, 9, no. 5, 263, doi: 10.3390/min9050263.

Koivisto, E., A. Malehmir, P. Heikkinen, S. Heinonen, and I. Kukkonen, 2012, 2D reflection seismic investigations at the Kevitsa Ni-Cu-PGE deposit, northern Finland: Geophysics, 77, no. 5, WC149 WC162, doi: 10.1190/GEO2011-0496.1.

Luhta, T., 2019, Petrophysical properties of the Kylylahti Cu-Au-Zn sulphide mineralization and its host rocks: M.S. thesis, University of Helsinki.

Malehmir, A., R. Durrheim, G. Bellefleur, M. Urosevic, Ch. Juhlin, D.J. White, B. Milkereit, and G. Campbell, 2012a, 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, no. 5, WC173 WC190, doi: 10.1190/GEO2012-0028.1.

Malehmir, A., C. Juhlin, C. Wijns, M. Urosevic, P. Valasti, and E. Koivisto, 2012b, 3D reflection seismic investigation for open-pit mine planning and exploration in the Kevitsa Ni-Cu-PGE deposit, northern Finland: Geophysics, 77, no. 5, WC95-WC108, doi: 10.1190/geo2011-0468.1

Mehta, K., A. Bakulin, J. Sheiman, R. Calvert and, R. Snieder, 2007, Improving the virtual source method by wavefield separation: Geophysics, 72, no. 4, V79 V86, doi: 10.1190/1.2733020.

Nakata, N., R. Snieder, T. Tsuji, K. Larner, and T. Matsuoka, 2011, Shear wave imaging from traffic noise using seismic interferometry by cross-coherence: Geophysics, 76, no. 6, SA97 SA106, doi: 10.1190/geo2010-0188.1.

Neidell, N. S., and M. Turhan Taner, 1971, Semblance and other coherency measures for multichannel data: Geophysics, 36, no. 3, 482-497, doi: 10.1190/1.1440186.

Nishitsuji, Y., S. Minato, B. Boullenger, M. Gòmez, K. Wapenaar, and D. Draganov, 2016, Crustal-scale reflection imaging and interpretation by passive seismic interferometry using local earthquakes: Interpretation, 4, no. 3, SJ29-SJ53, doi:10.1190/INT-2015-0226.1.

Olivier, G., F. Brenguier, M. Campillo, P. Roux, N. Shapiro, and R. Lynch, 2015a, Investigation of coseismic and postseismic processes using in-situ measurements of seismic velocity variations in an underground mine: Geophysical Research Letters, 42, 9261-9269, doi:10.1002/2015GL065975.

Olivier, G., F. Brenguier, M. Campillo, R. Lynch and P. Roux, 2015b, Body-wave reconstruction from ambient noise seismic noise correlations in an underground mine: Geophysics, 80, no. 3, KS11-KS25, doi: 10.1190/geo2014-0299.1.

Panea, I., D. Draganov, C. A. Vidal, and V. Mocanu, 2014, Retrieval of reflections from ambient noise recorded in the Mizil area, Romania: Geophysics, 79, Q31-Q42, doi: 10.1190/geo2013-0292.1.

Polychronopoulou, K., A. Lois, and D. Draganov, 2020, Body-wave passive seismic interferometry revisited: mining exploration using the body waves of local microearthquakes: Geophysical Prospecting, 68, 232-253, doi:10.1111/1365-2478.12884.

Quiros, A. D., L. D. Brown, and D. Kim, 2016, Seismic interferometry of railroad induced ground motions: body and surface wave imaging: Geophysical Journal International, 205, 1, 301-313, doi: 10.1093/gji/ggw033.

Riedel, M., C. Cosma, N. Enescu, E. Koivisto, K. Komminaho.; K. Vaittinen, and M. Malinowski, 2018, Underground Vertical Seismic Profiling with Conventional and Fiber-Optic Systems for Exploration in the Kylylahti Polymetallic Mine, Eastern Finland: Minerals, 8, no.11, 538, doi: 10.3390/min8110538.

Roots, E., A. Calvert, and J. Craven, 2017, Interferometric seismic imaging around the active Lalor mine in the Flin Flon greenstone belt, Canada: Tectonophysics, 718, 92-104, doi: 10.1016/j.tecto.2017.04.024.

Ruigrok, E., X. Campman, D. Draganov, and K. Wapenaar , 2010, High-resolution lithospheric imaging with seismic interferometry: Geophysical Journal International, 183, 339-357, doi: 10.1111/j.1365-246X.2010.04724.x.

Singh, B., M. Malinowski, F. Hloušek, E. Koivisto, S. Heinonen, O. Hellwig, S Buske, M. Chamarczuk, and S. Juurela, 2019, Sparse 3D Seismic Imaging in the Kylylahti Mine Area, Eastern Finland: Comparison of Time vs Depth Approach: Minerals, 9, no. 5, 305, doi: 10.3390/min9050305.

Sinha, S., P. S. Routh, P.D. Anno, and J. P. Castagna, 2005, Spectral decomposition of seismic data with continuous-wavelet transform: Geophysics, 70, no. 6, P19 P25, doi: 10.1190/1.2127113.

Snieder, R., J. Sheiman, and R. Calvert, 2006, Equivalence of the virtual source method and wave-field deconvolution in seismic interferometry: Physical Review E, 73, 066620, doi:10.1103/PhysRevE.73.066620.

Snieder, R., M. Miyazawa, E. Slob, I. Vasconcelos, and K. Wapenaar, 2009, A comparison of strategies for seismic interferometry: Surveys in Geophysics, 30, 503 523, doi: 10.1007/s10712-009-9069-z.

Stolt, R. H., and A. K. Benson, 1986, Seismic migration: Theory and practice, handbook of geophysical exploration, 5, Geophysical Press.

Thorbecke, J. W., and D. Draganov, 2011, Finite-difference modeling experiments for seismic interferometry: Geophysics, 76, no. 6, H1 H18, doi: 10.1190/geo2010-0039.1.

Väkevä, S., 2019, Using Three-Component Data for Seismic Interferometry Studies at the Kylylahti Mine, Eastern Finland: M.S. thesis, Univeristiy of Helsinki.

van der Neut, J., E. Ruigrok, and K. Wapenaar, 2010, Retrieving the earth’s reflection response by multidimensional deconvolution of ambient noise: 72nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, P406, doi: 10.3997/2214-4609.201401259.

van der Neut, J., J. Thorbecke, K. Mehta, E. Slob, and K. Wapenaar, 2011, Controlled-source interferometric redatuming by crosscorrelation and multidimensional deconvolution in elastic media: Geophysics, 76, no. 4, SA63, doi: 10.1190/1.3580633.

Wapenaar, K., J.W. Thorbecke, D. Draganov, and J. T. Fokkema, 2002, Theory of acoustic daylight imaging revisited: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2269 2272, doi: 10.1190/1.1817165.

Wapenaar, K., J.W. Thorbecke, and D. Draganov, 2004, Relations between reflection and transmission responses of 3-D inhomogeneous media: Geophysical Journal International, 156, 179 194, doi: 10.1111/j.1365-246X.2003.02152.x.

Wapenaar, K., and J. Fokkema, 2006, Green's functions representations for seismic interferometry: Geophysics, 71, no. 4, 1JA-Z75, doi: 10.1190/1.2213955.

Wapenaar, K., J. van der Neut, and E. Ruigrok, 2008, Passive seismic interferometry by multidimensional deconvolution: Geophysics, 73, A51-A56, doi: 10.1190/1.2976118.

Wapenaar, K., E. Slob, R. Snieder, and Andrew Curtis, 2010, Tutorial on seismic interferometry: Part 2 — Underlying theory and new advances: Geophysics, 75, no. 5, doi: 75A211-75A227, doi: 10.1190/1.3463440.

Wapenaar, K., J. van der Neut, E. Ruigrok, D. Draganov, J. Hunziker, E. Slob, J. Thorbecke, and R. Snieder, 2011, Seismic interferometry by crosscorrelation and by multidimensional deconvolution: a systematic comparison: Geophysical Journal International, 185, no. 3, 1335 1364, doi: 10.1111/j.1365-246X.2011.05007.x.

White, D. J., D. Secord, and M. Malinowski, 2012, 3D seismic imaging of volcanogenic massive sulfide deposits in the Flin Flon mining camp, Canada: Part 1 — Seismic results: Geophysics, 77, no. 5, WC47, doi: 10.1190/geo2011-0487.1.

# FIGURES

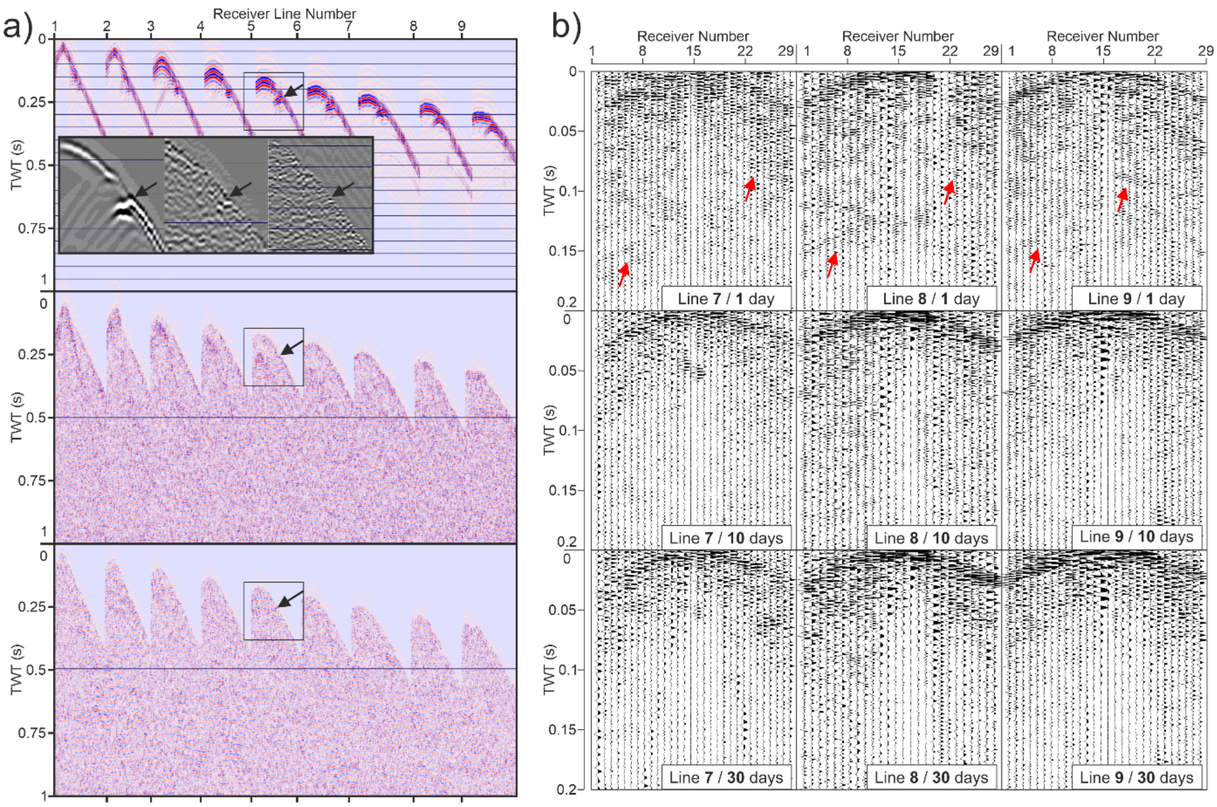


Figure 1 (a) Comparison of exemplary collocated shot gathers illustrating performance of 3D imaging in the Kylylahti mine area, (top row) synthetic active-shot gather, (middle row) field active-shot gather, and (bottom row) field virtual shot gather. Insets show zoomed part of shot gathers indicated with black polygons. (b) Influence of using more ambient-noise (AN) data for retrieval of virtual shot gathers (VSGs) without accounting for temporal and spatial stationarity of noise sources (blind stacking). Each row represents three VSGs obtained for three adjacent receiver lines. Each column represent VSGs obtained at the same master-trace position for increasing AN volumes. TWT stands for two-way traveltime.

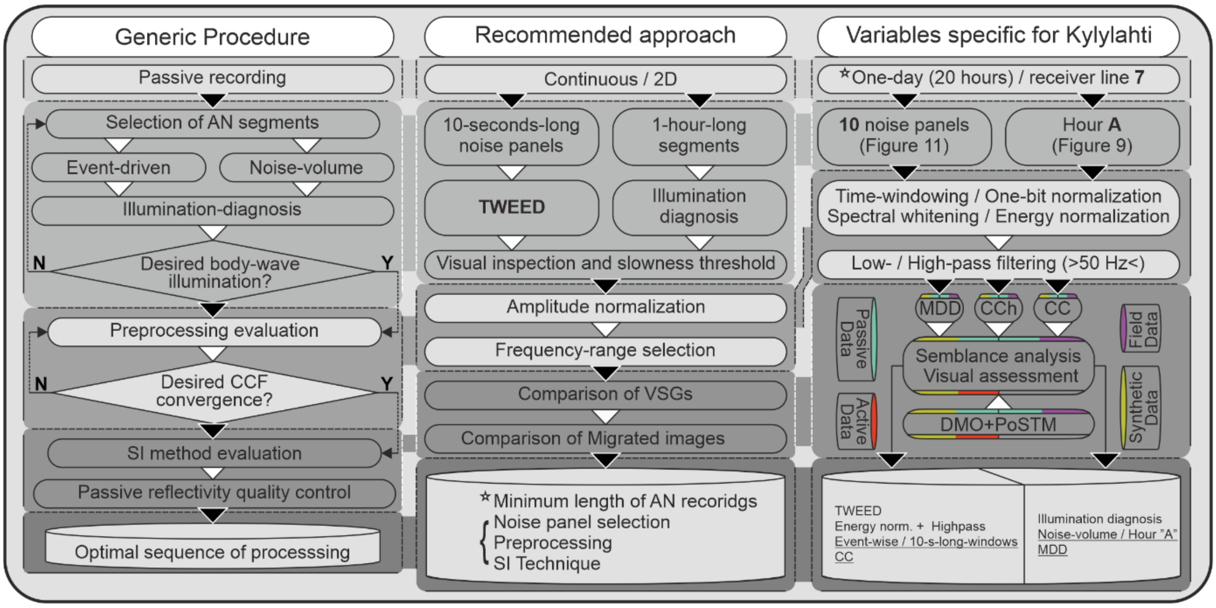


Figure 2. 2D ANSI processing workflow developed in our study. At the 'Generic procedure' level, it contains all the important steps (grey-scale coloured) and their ingredients to be investigated. Under the 'Recommended approach', we list all the tools/procedures that can be used at each step. The parameters specific for the case of the Kylylahti data we investigate are listed in the 'Variables specific for Kylylahti' flowchart.

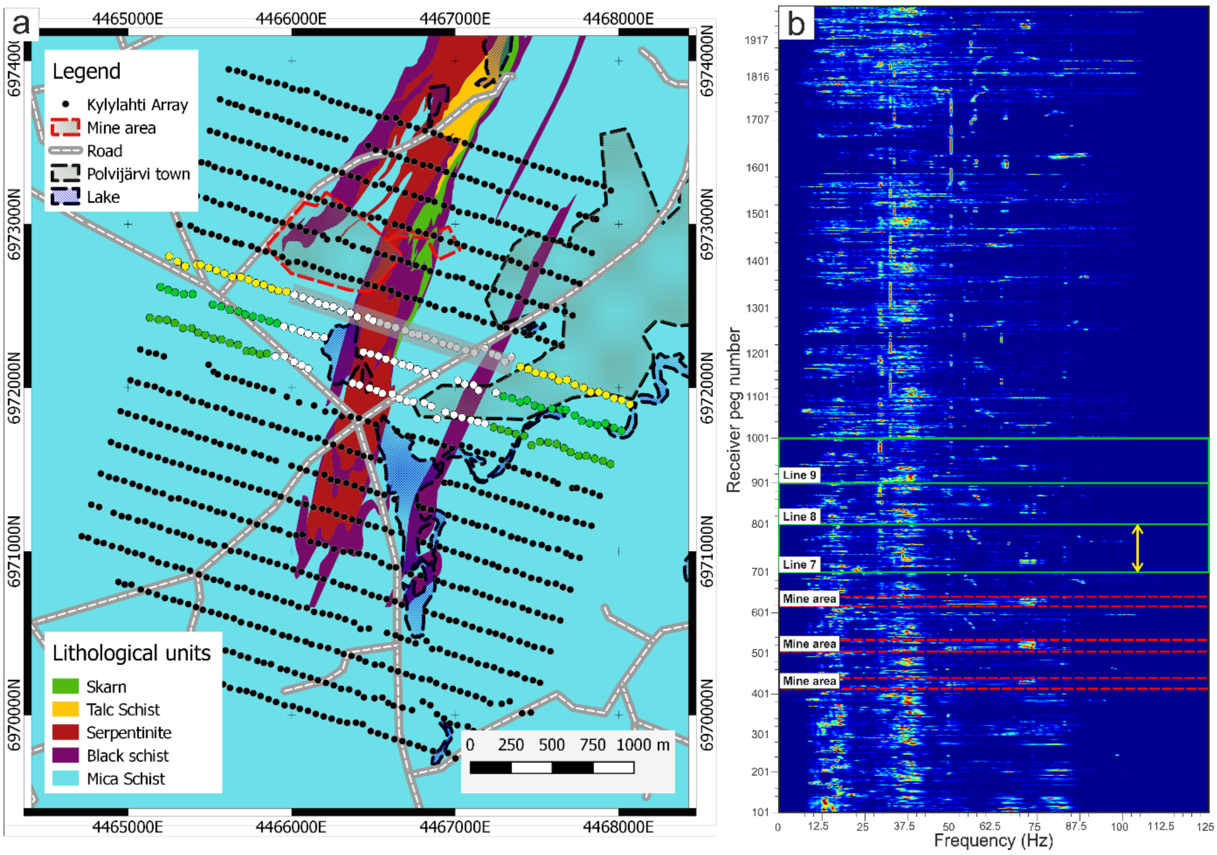


Figure 3. Layout of the Kylylahti array. Receiver line 7, selected for evaluation of the 2D ANSI processing strategy, is shown with yellow dots. Receiver lines used for illumination diagnosis and testing of results redundancy are denoted with green dots. Grey transparent stripe denotes the horizontal extent of the velocity model used for modelling; white dots indicate the part of the selected lines we use for 2D ANSI data processing. (b) Power Spectral Density averaged over one day of recording for the Kylylahti array. The green solid lines denote the receiver lines selected for analysis (as in (a)), the yellow arrow shows the spatial extent of receiver line 7, and the red dashed lines indicate the receivers located in the operating mine site.

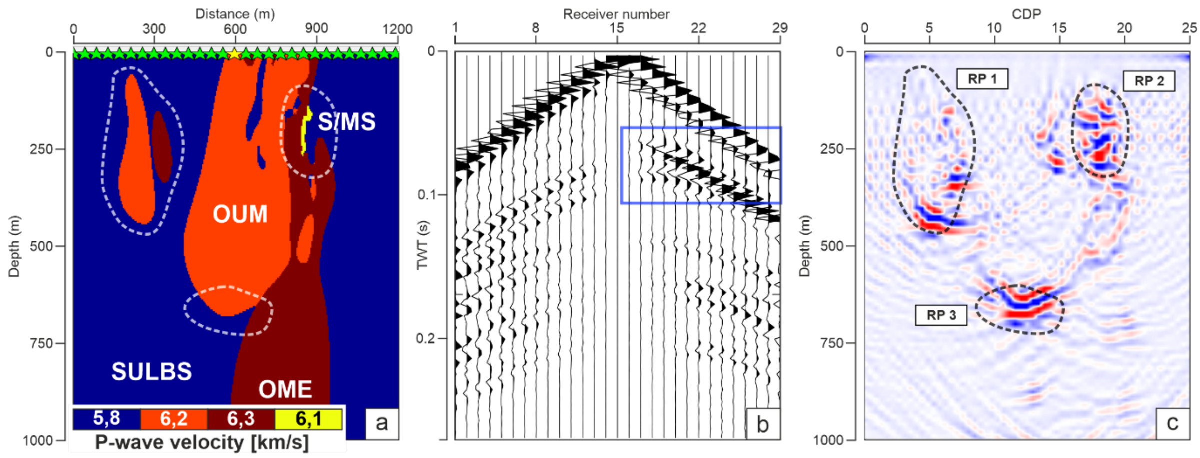


Figure 4. Synthetic active-source data tests. (a) Velocity model used as an input for the forward-modelling study, green stars denote the source distribution, while the yellow star denotes the location of the active shot used to record the shot gather shown in (b). The blue rectangle marks the part of data with a reflection arrival from the massive to semi-massive sulphide (S/MS) mineralisation; this reflection is also used in the semblance analysis. (c) Post-DMO (post dip moveout) migration of the synthetic active-source data. RP1-RP3 are the group of reflectors discussed in the text. CDP stands for common depth point.

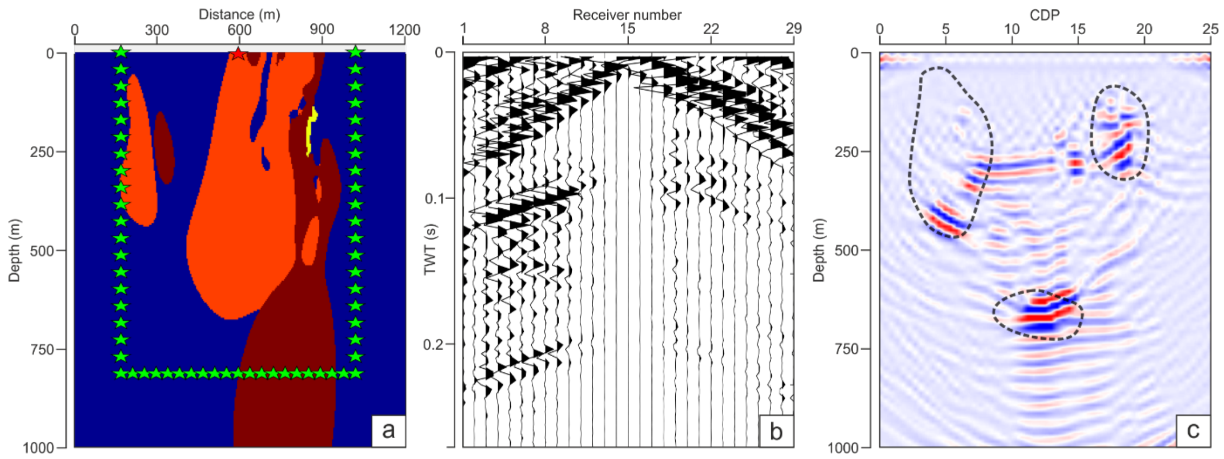


Figure 5. Synthetic passive data tests. (a) Same velocity as in Figure 3a used as an input for the forward-modelling study, green stars denote the regular passive-source distribution, while the red star denotes the location of the retrieved virtual shot used to record the VSG shown in (b). (c) Post-DMO migration of the synthetic VSGs.

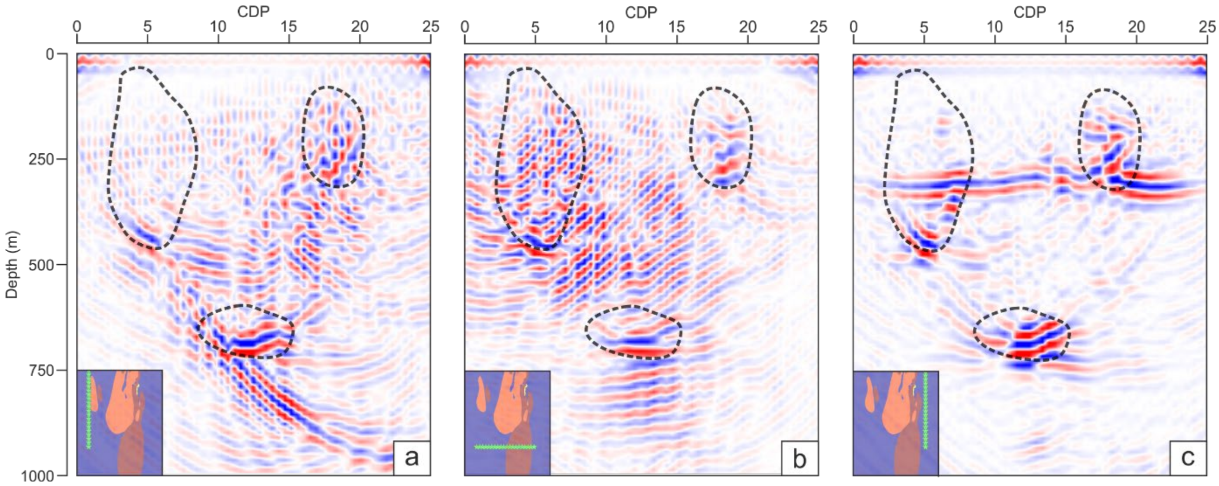


Figure 6. Post-DMO migration of the synthetic passive data obtained from simulating directional illumination with passive sources. Depth images obtained using sources distributed along a line (a) to the left of the geological target, (b) underneath the geological target, and (c) to the right of the geological target.

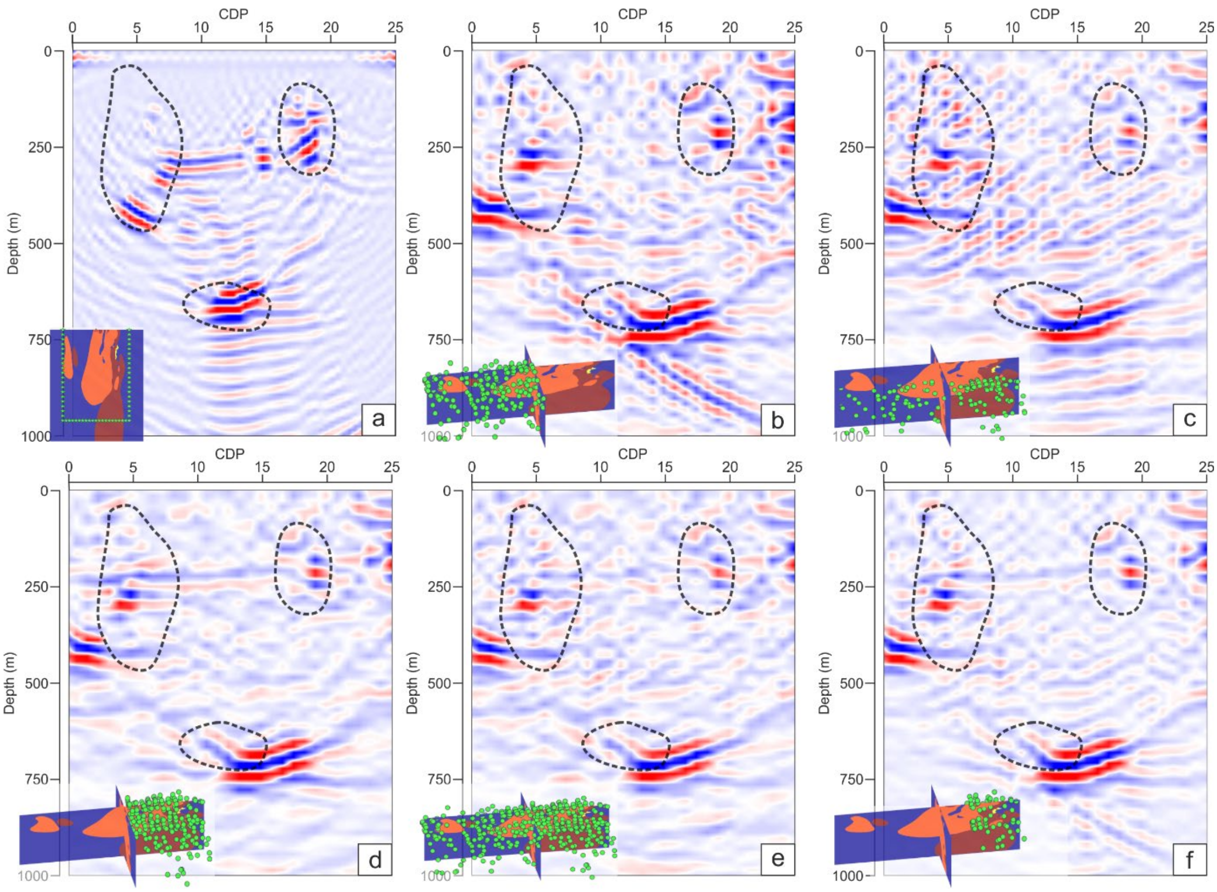


Figure 7. Comparison of Post-DMO stacks of the 2D synthetic passive data (a), and 3D synthetic passive data using various approaches of stacking the passive sources (b-f). The migrated images for the 3D scenario are obtained using VSGs produced from sources illuminating the target area from the left (b), bottom (c), right (d), all 648 modelled sources (e), and subset of sources mimicking the event-driven approach (f).

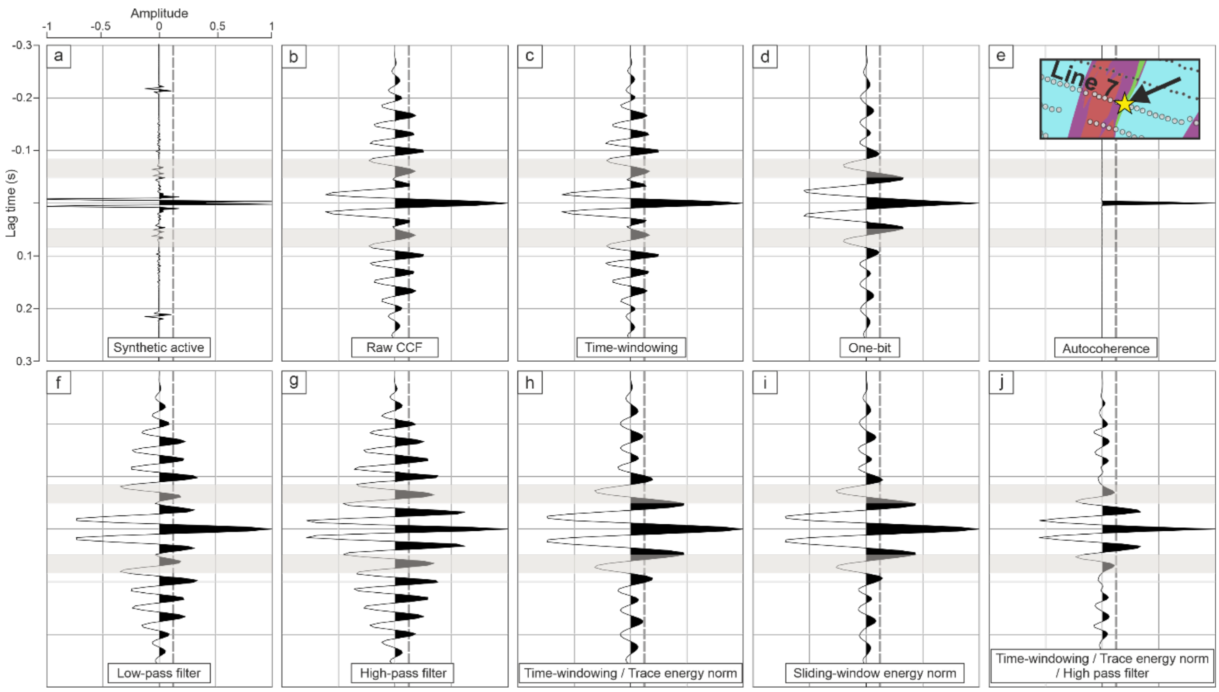


Figure 8. Influence of basic preprocessing procedures on the retrieval of zero-offset virtual traces. Synthetic active-shot trace concatenated with its time-reversed version (a) and (b) zero-offset trace without any preprocessing are show for reference. Virtual zero-offset traces are retrieved using: (c) windowing in the time domain, (d) one-bit normalization, (e) auto-coherence, (f) low-pass filtering, (g) high-pass filtering, (h) time windowing followed by trace energy normalization, (i) sliding-window energy normalization, and (j) sequence of time windowing, trace energy normalization, and high-pass filtering. The inset in (e) indicates the location of the zero-offset trace with respect to the synthetic velocity model and receiver line 7.

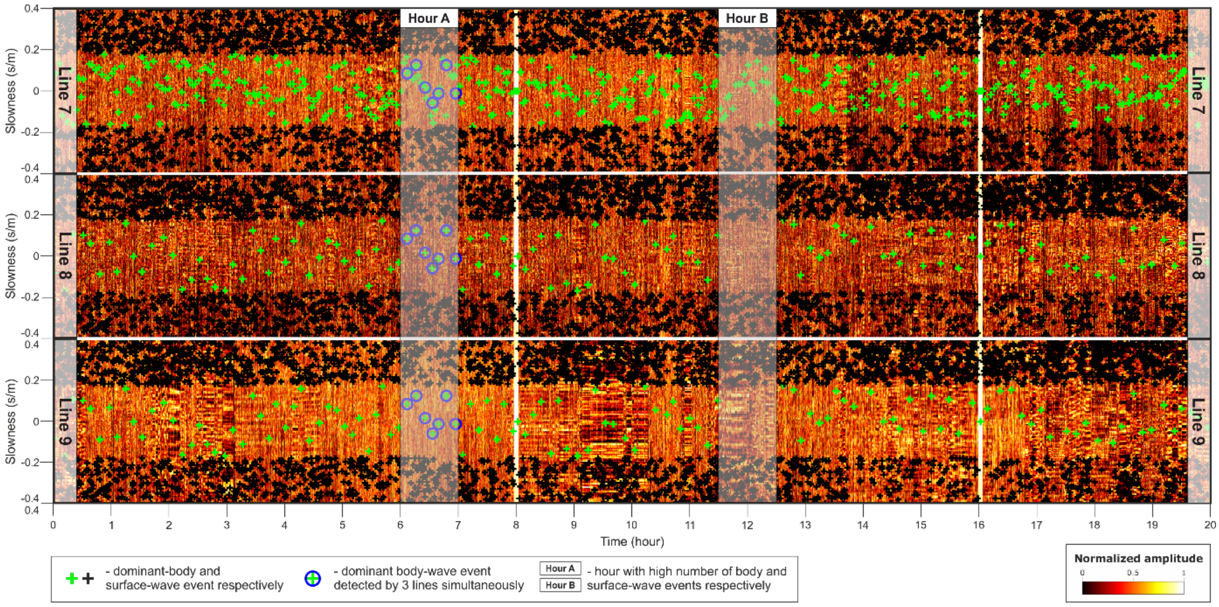


Figure 9. Illumination diagnosis panel obtained from the TWEED method (Chamarczuk et al. 2019) showing a distribution of normalized slant-stack values for every noise panel recorded by three receiver lines during a whole day of passive acquisition.

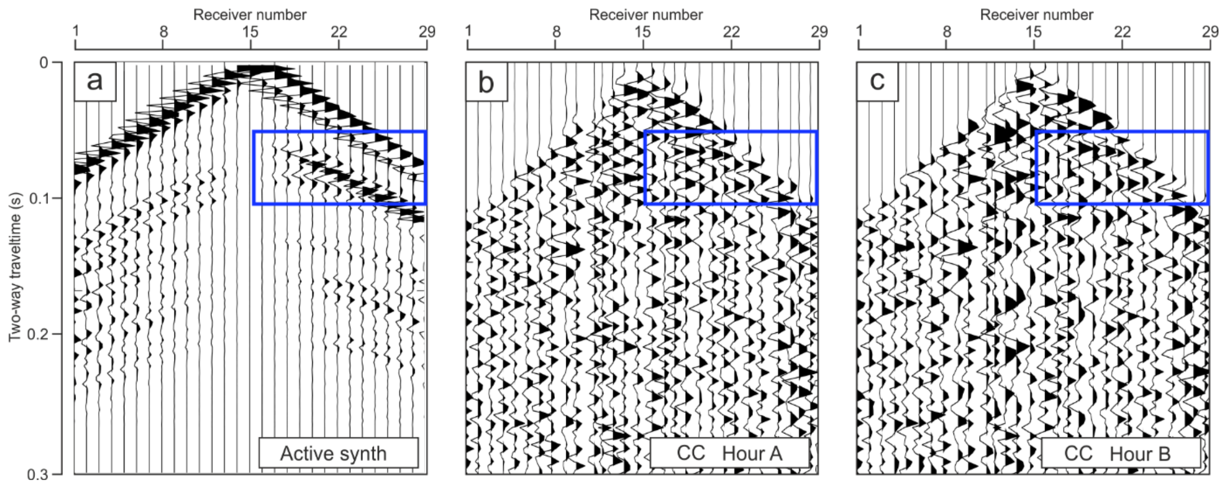


Figure 10. Comparison of (a) an active-shot gather and VSGs retrieved with CC using one hour of AN dominated by (b) body-wave events and (c) surface-wave events. Blue rectangles mark the part of data selected to compare the reflectivity content and used in the semblance analysis.

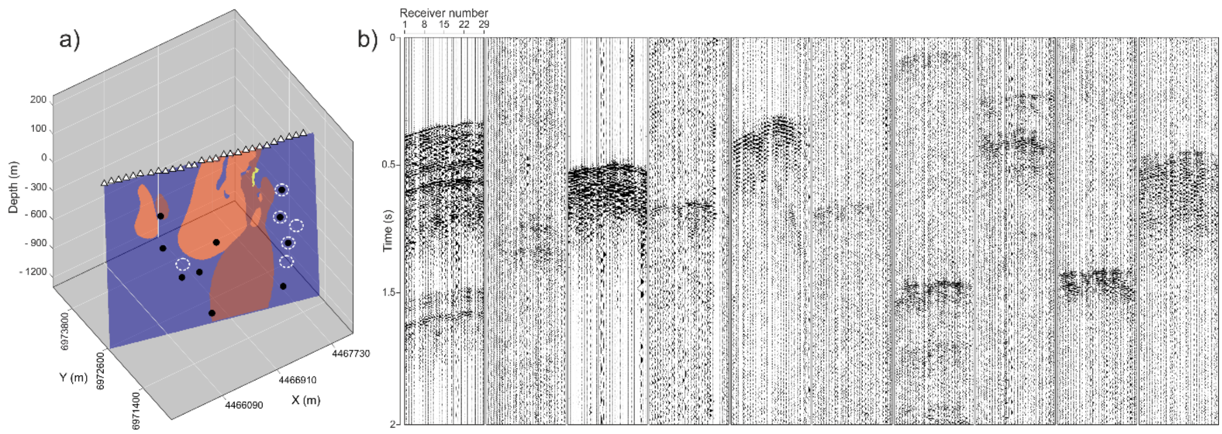


Figure 11. (a) Location of recording line 7 with respect to the velocity-model section used for the forward-modelling study. Location of the body-wave events detected during hour A indicated in Figure 7 is denoted with white circles, body-wave events selected for the evaluation of the event-driven 2D processing and enclosing the target area are shown with black dots. (b) Seismograms of body-wave events recorded by receiver line 7, and selected for the evaluation of the event-driven 2D processing.

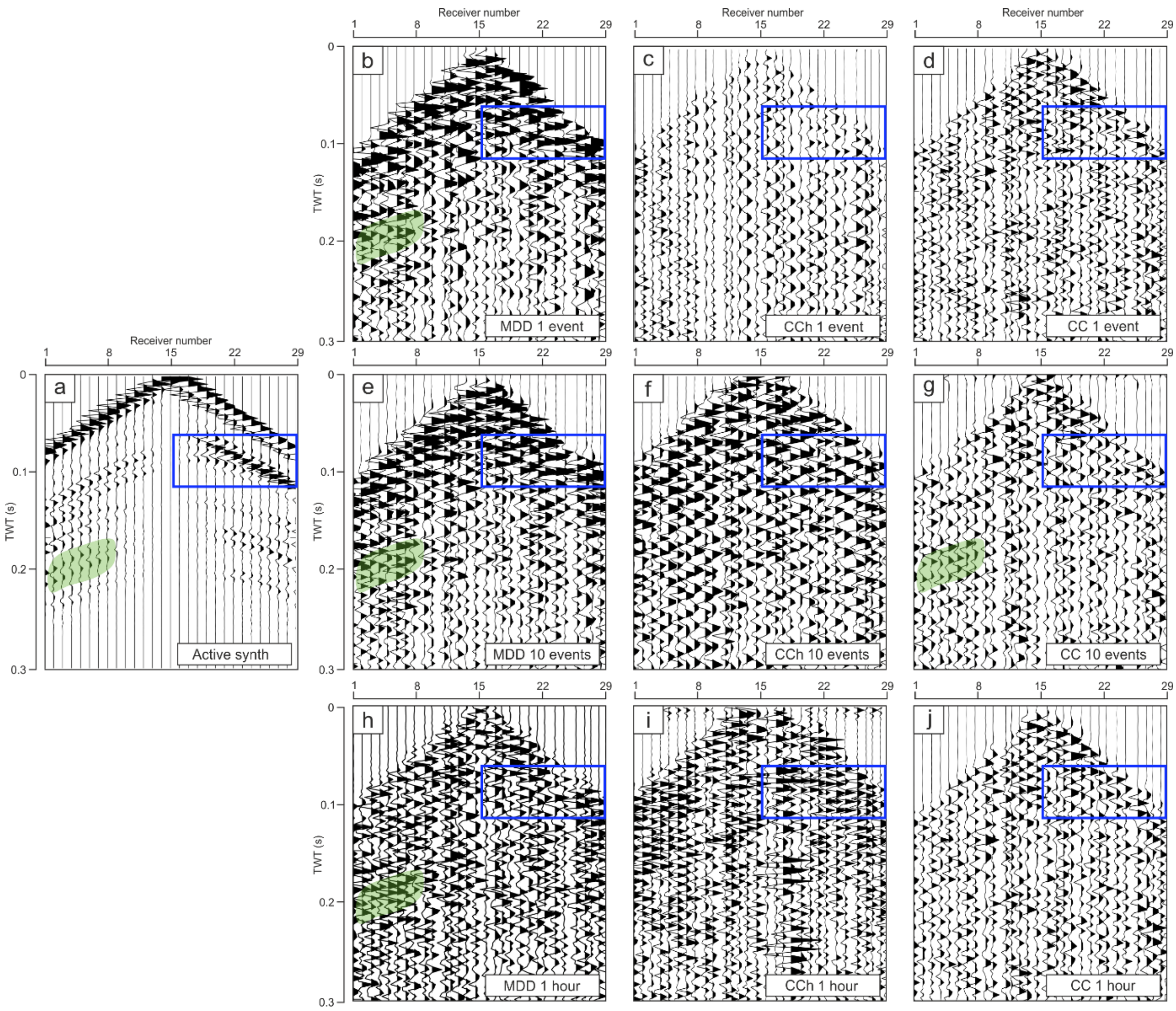


Figure 12. VSGs obtained from applying 2D ANSI on different segments of AN recordings selected on the basis of illumination diagnosis characteristics. (a) Synthetic active shot; (b, e, and h) VSGs obtained using MDD on a single event, 10 body-wave events, and 1 hour of AN, respectively; (c, f, and i) VSGs obtained with CCh using a single event, 10 body-wave events and 1 hour of AN, respectively; (d, g, and j) VSGs obtained with CC using a single event, 10 body-wave events, and 1 hour of AN, respectively. Blue rectangles mark the part of the data selected to compare the reflectivity content and used in semblance analysis. Green shaded areas indicate part of the data, where deeper reflectivity appears on both active-shot gathers and VSGs, and are discussed in the text.

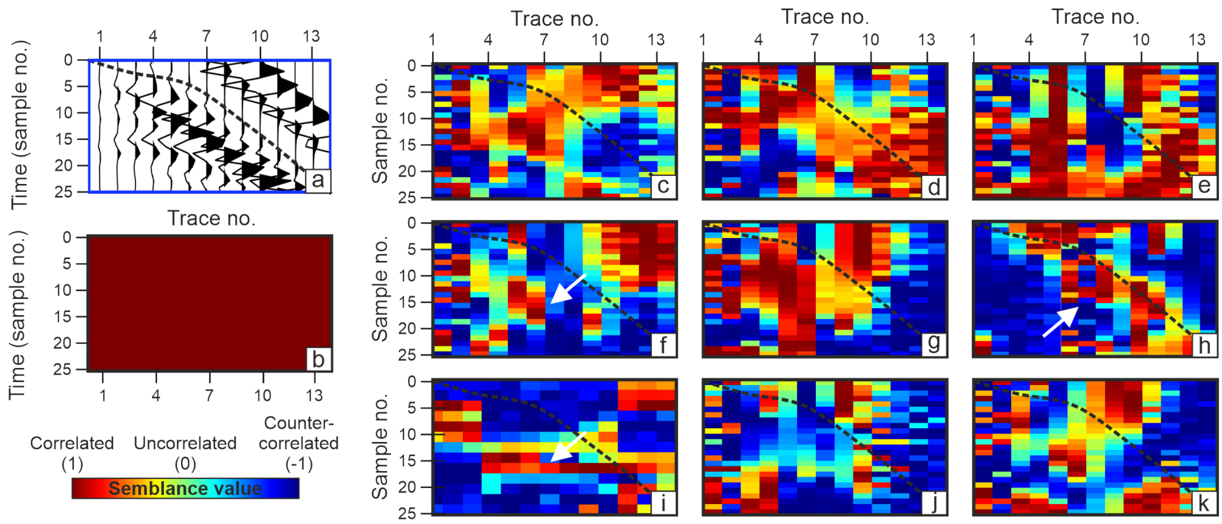


Figure 13. Semblance analysis of the virtual shots retrieved using 2D ANSI. The results are calculated in the spatio-temporal windows denoted with the blue rectangles shown in Figure 10. (a) Extracted part of the synthetic active-source data used for comparison with the passive data; (b) auto-semblance output calculated for the data shown in (a); (c, f, and i) semblance results for the reflection event obtained using MDD on a single event, 10 body-wave events, and 1 hour of AN, respectively; (d, g, and j) semblance results for the reflection event obtained using CCh on a single event, 10 body-wave events and 1 hour of AN, respectively; (e, h, and k) semblance results for the reflection event obtained using CC on a single event, 10 body-wave events, and 1 hour of AN, respectively. The black dashed line indicates the separation between the direct wave and the reflection event.

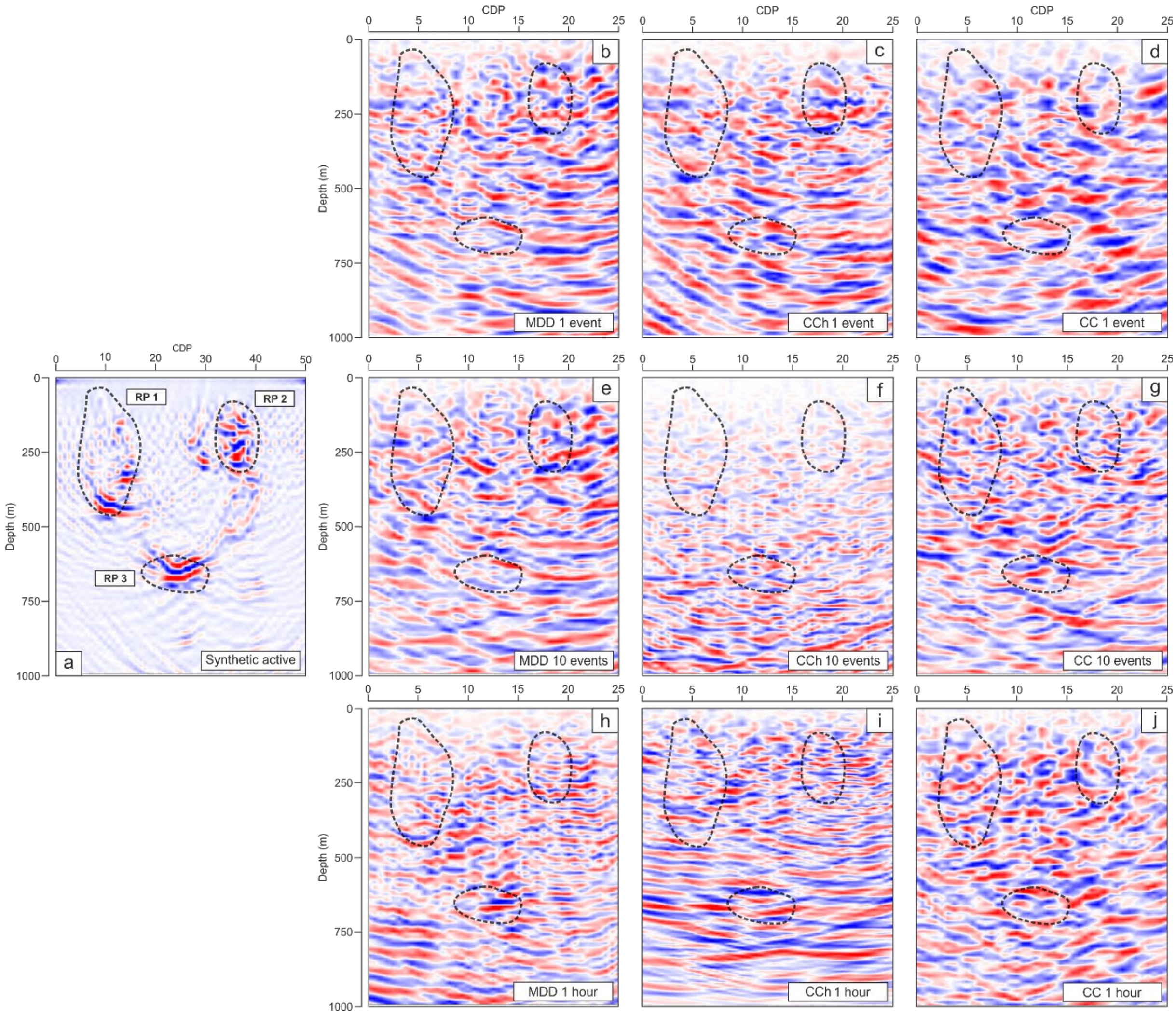


Figure 14. Comparison of migrated depth sections obtained from the nine different 2D ANSI processing strategies. The sections in a) to j) correspond to the order of the VSGs in Figure 12.

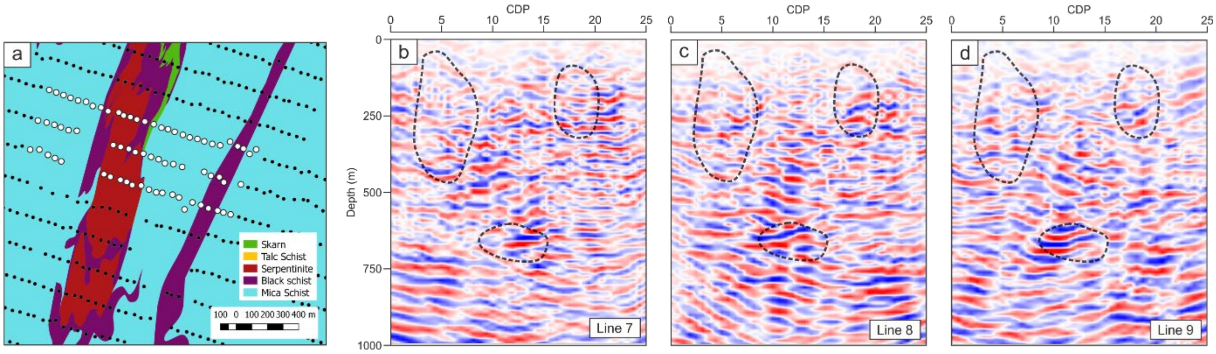


Figure 15. Images using the MDD results obtained from 1 hour of AN recordings for 3 adjacent recording lines denoted with white circles in (a). Depth sections are shown for the following receiver lines: (b) line 7, (c) line 8, and (d) line 9. Consistent reflectivity can be observed in areas where reflections in the synthetic active (Figure 4c) and synthetic passive (Figure 5c) migrated sections are visible.

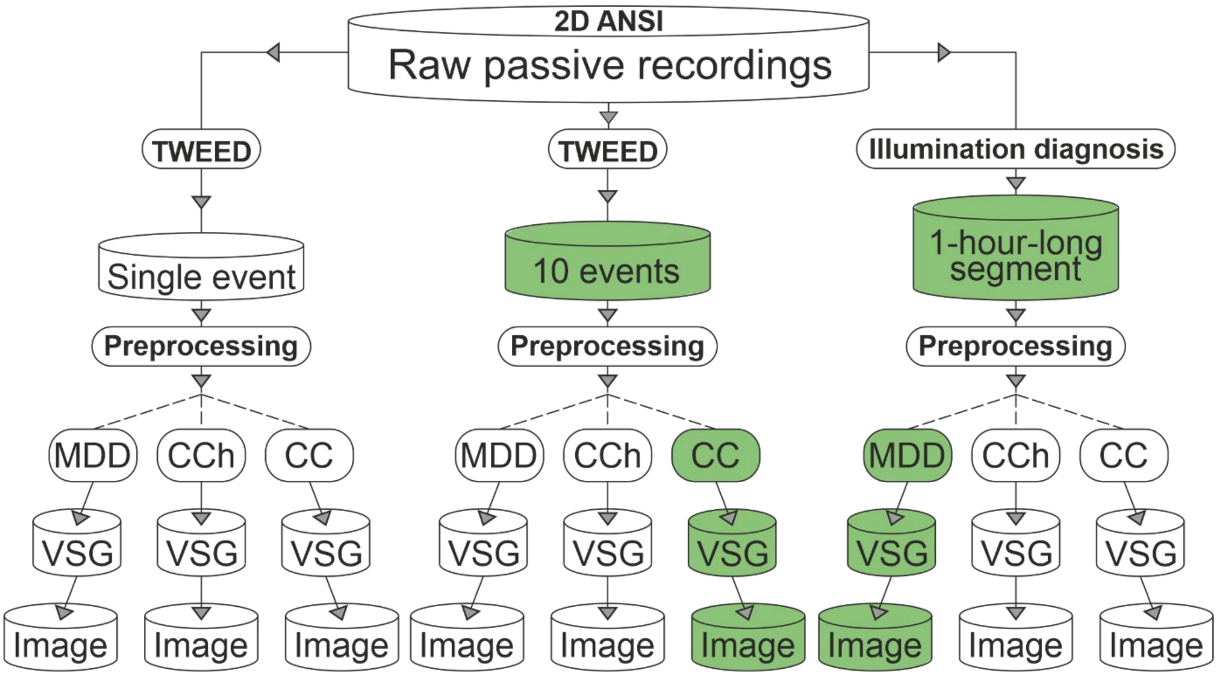


Figure 16. Summary of the 2D ANSI methodology and comparison strategy. The core of the comparison is represented by parallel flow diagrams. The optimal SI processing sequences selected for the Kylylahti data are denoted with green colour. Parts of the workflow in bold denote processing steps implicit for a given processing route.

| | SULBS | OUM | OME | S/MS |
|---|---|---|---|---|
| P-wave velocity (km/s) | 5,8 | 6,2 | 6,3 | 6,1 |
| Density (g/cm$^3$) | 2,9 | 2,9 | 3,1 | 3,8 |
| Impedance (km/s g/cm$^3$) | 16,82 | 17,98 | 19,53 | 23,18 |

Table 1. Average elastic rock properties of the geological units in the synthetic model

## LIST OF ACRONYMS

| | |
|---|---|
| AC | autocorrelation |
| AN | ambient-noise |
| ANSI | ambient-noise seismic interferometry |
| CC | crosscorrelation |
| CCh | crosscoherence |
| CDP | common-depth-point |
| CWT | continuous wavelet transform |
| DMO | dip-moveout |
| GPU | graphical processing units |
| MDD | multidimensional deconvolution |
| OME | Outokumpu altered ultramafic rocks |
| OUM | Outokumpu ultramafic rocks |
| PSD | power-spectral-density |
| PSF | point-spread function |
| QC | quality control |
| RMS | root-mean-square |
| RP | reflection packages |
| S/MS | massive to semi-massive sulphide |
| SI | seismic interferometry |
| SNR | signal-to-noise ratio |
| SULBS | sulphide-bearing schist |
| TWEED | two-step wavefield evaluation and event detection |
| TWT | two-way traveltime |
| VSG | virtual shot gathers |