Geophysical Journal International

Geophys. J. Int. (2014) **198,** 1572–1584 doi: 10.1093/gji/ggu164
GJI Seismology

Retrieval of reflections from ambient noise using illumination diagnosis

C. Almagro Vidal, D. Draganov, J. van der Neut, G. Drijkoningen and K. Wapenaar

Department of Geotechnology, Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, the Netherlands. E-mail: c.almagrovidal@tudelft.nl

Accepted 2014 May 2. Received 2014 May 1; in original form 2013 October 24

SUMMARY

Seismic interferometry (SI) enables the retrieval of virtual sources at the location of receivers. In the case of passive SI, no active sources are used for the retrieval of the reflection response of the subsurface, but ambient-noise recordings only. The resulting retrieved response is determined by the illumination characteristics of the recorded ambient noise.

Characteristics like geometrical distribution and signature of the noise sources, together with the complexity of the medium and the length of the noise records, determine the quality of the retrieved virtual-shot events. To retrieve body wave reflections, one needs to correlate body-wave noise. A source of such noise might be regional seismicity. In regions with notable human presence, the dominant noise sources are generally located at or close to the surface. In the latter case, the noise will be dominated by surface waves and consequently also the retrieved virtual common-source panels will contain dominant retrieved surface waves, drowning out possible retrieved reflections. In order to retrieve reflection events, suppression of the surface waves becomes the most important pre-processing goal.

Because of the reasons mentioned above, we propose a fast method to evaluate the illumination characteristics of ambient noise using the correlation results from ambient-noise records. The method is based on the analysis of the so-called source function of the retrieved virtual-shot panel, and evaluates the apparent slowness of arrivals in the correlation results that pass through the position of the virtual source and at zero time. The results of the diagnosis are used to suppress the retrieval of surface waves and therefore to improve the quality of the retrieved reflection response. We explain the approach using modelled data from transient and continuous noise sources and an example from a passive field data set recorded at Annerveen, Northern Netherlands.

Key words: Interferometry; Body waves; Europe.

1 INTRODUCTION

Passive seismic reflection surveys intend to use ambient-noise sources to retrieve reflection information of the subsurface. The application of seismic interferometry (SI) enables the retrieval of responses at the receiver locations as if there were a source at a chosen receiver location. Explanations and examples of how SI can be used to retrieve the subsurface response are given by Curtis *et al.* (2006), Wapenaar *et al.* (2008), Schuster (2009) and Xu *et al.* (2012). This can be achieved with either correlation, convolution or deconvolution processes.

The aim of SI with passive sources, or Passive SI, is to retrieve the Green's function from ambient-noise records. Depending on the type of noise and/or pre-processing steps, surface waves could be retrieved (e.g. Shapiro & Campillo 2004), diving body waves (e.g. Roux *et al.* 2005), but also body wave reflections (e.g. Draganov *et al.* 2009; Poli *et al.* 2012).

The quality of the retrieved response depends on the time/frequency characteristics of the recorded noise, the distribution of the noise sources, the complexity of the medium and the recording time length. Unfortunately, the preponderance in time or space of some sources with respect to others in the records limit the retrieval of the complete Green's function. When using SI by crosscorrelation, sources located near the surface would contribute predominantly to the retrieval of surface waves, whereas sources located relatively deeper would contribute mostly to the retrieval of body wave reflections. Noise recordings in regions with relatively high local and regional seismicity would facilitate the retrieval of body wave reflections, although in general the presence of anthropogenic noise would mean predominance of sources at the surface and therefore ambient noise dominated by surface waves. When this is the case, the results retrieved by SI by crosscorrelation exhibit surface waves that drown out the possible retrieved reflections, as the latter are much weaker.

In passive seismic surveys at lithospheric scale, body waves can be identified using frequency–wavenumber spectral analysis (Nishida 2013), and in cases of post-critical reflections, their estimated amplitude has been comparable to the amplitudes of surface waves (Zhan *et al.* 2010). At exploration scale, Nakata *et al.* (2011) showed that by equalizing or whitening the frequency spectrum during the retrieval process (that is applying crosscoherence instead of crosscorrelation), one could retrieve reflections even when the noise is dominated by surface waves. The drawback of using all the noise is that surface waves are still retrieved, which would require their removal after the retrieval. Forghani & Snieder (2010) show the balance between retrieved surface and body waves using SI by crosscorrelation, which opens the possibility of adaptive surface wave removal by reconstruction of waveforms with isolated surface waves (Van Wijk *et al.* 2010).

An alternative approach is to suppress the retrieval of surface waves by not using the parts of the noise dominated by surface-wave noise (Draganov *et al.* 2010). The selection of noise parts dominated by body waves can be carried out using beam-forming (Draganov *et al.* 2013) or additionally splitting the record in frequency bands for which the body-wave noise is dominant (Ruigrok *et al.* 2011). However, even when reflections are retrieved, they might not be obtained correctly due to preferential illumination of the recording array with body-wave noise from certain directions. In such a case, one needs to compensate for overillumination from dominant noise-source locations.

We propose here an efficient technique to estimate the illumination from the ambient noise. The structure of the paper starts showing how the method is based on the properties of the virtual-source function, which is on the properties of the retrieved events that pass through the position of the virtual-shot trace at time zero. Based on the diagnosis of the illumination from the noise, we carry out the discrimination of noise sections seeking body wave characteristics. The diagnosis also provides additional support for compensating overillumination. In this paper, we demonstrate its application on synthetic data set with transient sources and simultaneous noise sources. Finally, we applied this method to ambient-noise recordings from the north of the Netherlands, and compare the retrieved response from the correlated and summed noise before and after its application.

2 ILLUMINATION DIAGNOSIS WITH TRANSIENT SOURCES

In SI with transient sources, one can employ the recordings at two receivers $\mathbf{x}_A$ and $\mathbf{x}_B$ in order to retrieve the Green's function $G(\mathbf{x}_B, \mathbf{x}_A, t)$ between the receiver stations as if a source were located at receiver $\mathbf{x}_A$ (Wapenaar & Fokkema 2006). For sources acting separately in time, and having an equal autocorrelation function $S_0(t)$, located at positions $\mathbf{x}_S$ along an enclosing source boundary $\partial\mathbb{D}$, one may use the SI crosscorrelation expression:

$$\Re\{G(\mathbf{x}_B, \mathbf{x}_A, t)\} * S_0(t) \approx \frac{1}{\rho c} \oint_{\partial\mathbb{D}} u^{\mathrm{obs}}(\mathbf{x}_A, \mathbf{x}_S, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, \mathbf{x}_S, t)\, \mathrm{d}\mathbf{x}_S, \tag{1}$$

where $\Re$ stands for real part, ρ and c are the constant mass density and velocity of the medium at and outside $\partial\mathbb{D}$, respectively; $*$ denotes time convolution and $u^{\mathrm{obs}}(\mathbf{x}_A, \mathbf{x}_S, -t)$ is the time-reversed wavefield observed at $\mathbf{x}_A$ due to a transient source at $\mathbf{x}_S$.

Eq. (1) states that to retrieve the desired Green's function, one would need to integrate the correlation results from all sources of the boundary. The correlated common-source panel C^S is the correlation result from each boundary source that makes its individual contribution to the interferometric integration. For a single transient source, it is represented at a fixed receiver position $\mathbf{x}_A$ and a variable receiver position $\mathbf{x}_B$ as follows:

$$C^S(\mathbf{x}_B, \mathbf{x}_A, t) = \frac{1}{\rho c}\left(u^{\mathrm{obs}}(\mathbf{x}_A, \mathbf{x}_S, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, \mathbf{x}_S, t)\right), \tag{2}$$

as if a source were located in $\mathbf{x}_A$ that emits energy within a limited window of angles to multiple receivers $\mathbf{x}_B$. The events in this panel that pass through $t = 0$ s and the position of the virtual source are informative of the illumination characteristics of the specific source $\mathbf{x}_S$ at the boundary $\partial\mathbb{D}$. We call the collection of these events the virtual-source function (van der Neut 2013).

There are several ways to study the illumination characteristics using the virtual-source function. One approach is to make use of a slant-stack transform of field v: $\tilde{v}(p, \tau) = \int v(x, \tau + px)\, \mathrm{d}x$. In our case, we evaluate this transform at correlation time $\tau = 0$ s. Therefore, we can study the illumination contribution from $\mathbf{x}_S$ to the virtual source at $\mathbf{x}_A$ using the simplified transformation at each virtual-source location $\mathbf{x}_A$:

$$\widetilde{C}^S(\mathbf{x}_A, \mathbf{p}) \equiv \int C^S[\mathbf{x}_B, \mathbf{x}_A, \mathbf{p}\cdot(\mathbf{x}_B - \mathbf{x}_A)]\mathrm{d}\mathbf{x}_B. \tag{3}$$

$\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})$ is the ray-parameter description of the virtual-source function of the transient source S sensed at the virtual-source position $\mathbf{x}_A$. The illumination diagnosis could in principle also be done in the frequency–wavenumber domain, making use of slant-stacks of the virtual-source function over different velocity values. This procedure, though, would require that first the virtual-source function is isolated in the time domain by means of muting. The design of the muting window around the virtual-source function might become highly user-dependent. The reason to study the virtual-source function in the $\tau - p$ domain is that the analysis takes place right at its location, around $\tau = 0$ s. In this way, we avoid having to isolate the virtual-source function with a time window or to include reflections in the analysis. In addition, it has the advantage of being faster since it does not require any additional Fourier transform. The virtual-source function is analysed at every correlated common-source panel. For each slant-stack result at $t = 0$ s, we search for the ray-parameter $\mathbf{p}^S_{\mathbf{x}_A}$ at the virtual-source location $\mathbf{x}_A$, for which the source function is maximum:

$$\widetilde{C}^S(\mathbf{x}_A, \mathbf{p}^S_{\mathbf{x}_A}) = \left\|\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})\right\|_{\max}. \tag{4}$$

The suitability of the correlated common-source panel for reflection retrieval is now given by means of a comparison of the dominant ray-parameter in its absolute value $\|\mathbf{p}^S_{\mathbf{x}_A}\|$ with a pre-defined threshold value p_{limit} between the characteristic slowness for body waves and for surface waves. Although this discrimination test is based on the magnitude of $\mathbf{p}^S_{\mathbf{x}_A}$, the directional information can be employed in directional balancing. This discrimination test can be described as follows:

$$C^S(\mathbf{x}_B, \mathbf{x}_A, t) = \begin{cases} 0 & \text{if } \dfrac{\max\left(\|\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\|\le p_{\mathrm{limit}}}}{\max\left(\|\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\|> p_{\mathrm{limit}}}} \le R \\ \frac{1}{\rho c}\left(u^{\mathrm{obs}}(\mathbf{x}_A, \mathbf{x}_S, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, \mathbf{x}_S, t)\right) & \\ & \text{if } \dfrac{\max\left(\|\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\|\le p_{\mathrm{limit}}}}{\max\left(\|\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\|> p_{\mathrm{limit}}}} > R, \end{cases} \tag{5}$$

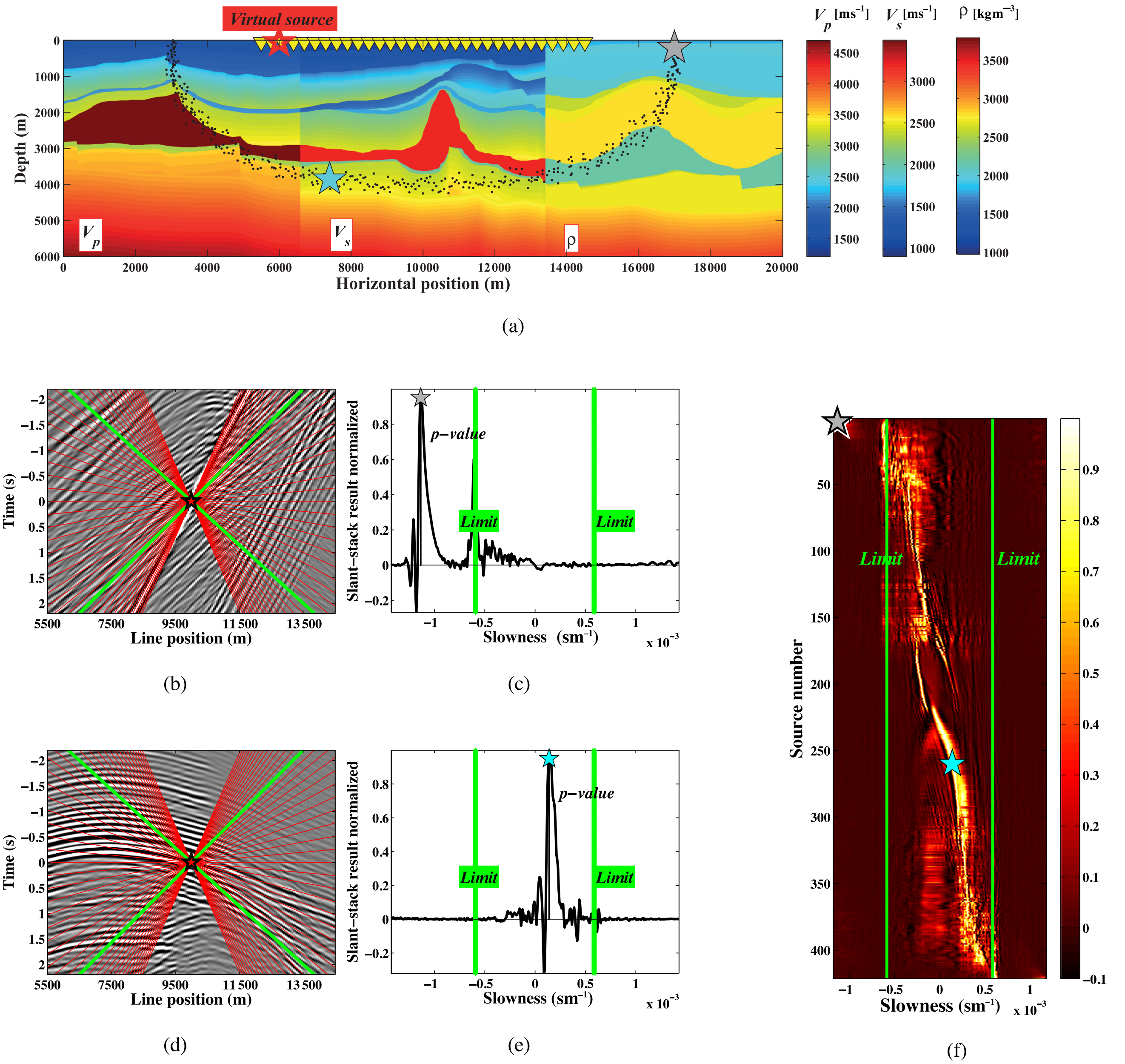


Figure 1. Slant-stack illumination diagnosis. (a) Elastic subsurface model. 181 receivers (yellow triangles) are placed at 50-m depth with 50-m spacing. 421 sources (black dots) represent the ambient-source boundary region. Two examples are shown for a shallow (grey star) and a deep source (cyan star). (b) Correlated common-source panel C^S for a virtual source at $\mathbf{x}_A = 10\,000$ m, resulting from the grey-star source in (a). Red lines indicate slownesses. Green lines are pre-defined limits between body wave and surface wave slownesses in the virtual-source function. (c) Slowness representation $\widetilde{C}^S$ of the virtual-source function from (b). Green lines indicate the forementioned limits between body wave and surface wave slownesses. (d) As in (b), but for the cyan-star source in (a). (e) As in (c), but from (d). (f) Illumination diagnosis, consisting of the results for the correlated common-source panels from sources in (a) (black dots), with the panel from (b) (grey star) discarded, and the panel from (d) (cyan star) included.

where R is some control ratio. Application of the discrimination test (eq. 5), for a certain virtual-source location $\mathbf{x}_A$, takes care that a source at $\mathbf{x}_S$ with ray-parameter $\|\mathbf{p}^S_{\mathbf{x}_A}\|$ larger than the pre-defined threshold value p_{limit} is not contributing to the final reflection-response retrieval. The value given to p_{limit} will depend on the elastic properties of the medium where the receiver array is located. One must first estimate expected values for body and surface wave velocities at the receiver location and define the threshold value with respect to these. However, it may happen that certain sources contribute in the retrieval of body and surface waves in a similar proportion. For such cases, the comparison ought to consider also the ratio R of the maxima of the source function $\widetilde{C}^S$ inside and outside the limit interval p_{limit}. A large R ratio ensures only body wave contribution by the correlation panel. This allows the test to obtain a conservative character for avoiding surface wave retrieval. (During the experiments shown in this paper, we used $R = 2$. Choosing a smaller ratio $R < 2$ resulted in including noise panels containing surface waves. That resulted in surface wave retrieval in the final result.)

Fig. 1 shows an example of the application of the slowness evaluation and discrimination procedure from eqs (3) and (5): Fig. 1(b) shows the correlated common-source panel, with the virtual-source function in the middle, for a virtual source located at $x_A = 10\,000$ m; the common-source panel before correlation is dominated by surface waves from a source close to the surface (see grey star in Fig. 1a). Fig. 1(c) gives the representation $\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})$, result of the integration over different slownesses shown with red lines in Fig. 1(b). The dominant ray-parameter $\mathbf{p}^S_{\mathbf{x}_A}$ (p-value) is outstanding, marked with a grey star. The green lines in both (b) and (c) represent the pre-defined slowness limits, which serve as a threshold slowness in the correlated common-source panel. Fig. 1(d) displays another correlated common-source panel for the same virtual-source location, but resulting from a relatively deep source (see cyan star in Fig. 1a). The respective slowness spectrum of the virtual-source function is shown in Fig. 1(e), in which the maximum ray-parameter is marked with a cyan star. The succession of such slowness distributions from correlated common-source panels for all boundary-source positions $\mathbf{x}_S$ (the black dots in Fig. 1 a) produces the Illumination diagnosis in Fig. 1(f). The contributions to the reflection-response retrieval from each of the sources can be studied using the dominant ray-parameters $\mathbf{p}^S_{\mathbf{x}_A}$ which, after the discrimination test, are between the slowness threshold values.

As a result from the application of the discrimination test in eq. (5), sources contributing to surface wave retrieval are disregarded and only sources contributing to body wave retrieval are kept. Finally, the desired retrieved response after illumination diagnosis and discrimination is obtained using

$$\Re\{G(\mathbf{x}_B, \mathbf{x}_A, t)\} * S_0(t) \approx \sum_S C^S(\mathbf{x}_B, \mathbf{x}_A, t). \tag{6}$$

For further applications of the virtual-source function and its relation to the point-spread function, the readers are referred to van der Neut (2013).

2.1 Retrieval of reflections from passive transient sources

In the previous section, we introduced the process of the illumination diagnosis. Here, we are going to apply it on a synthetic model. Fig. 2 presents the analysis and results of the illumination diagnosis and discrimination from transient sources used in a synthetic 2-D elastic model, based on the geology of the north of the Netherlands

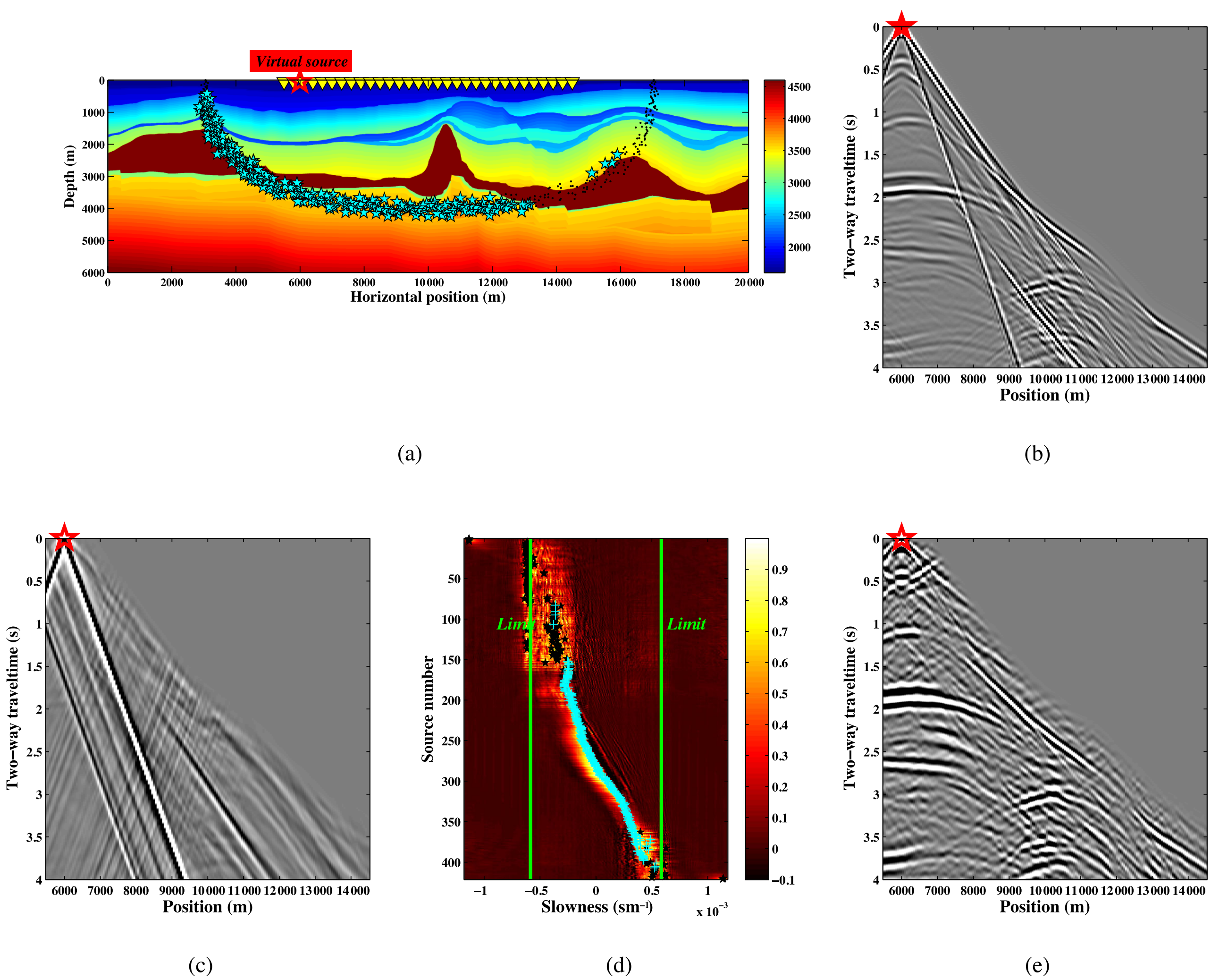


Figure 2. Illumination diagnosis for transient sources. (a) Same P-wave velocity model from Fig. 1(a), but now with the 421 subsurface sources represented differently whether they are dominated by surface wave arrivals (black dots) or by body waves (cyan stars). (b) Directly modelled reflection response for an active source at $\mathbf{x}_A = 6000$ m (the red open star in a). (c) Retrieved virtual common-source panel at the same location, obtained using all correlated common-source panels. (d) Illumination diagnosis with $\mathbf{p}^S_{\mathbf{x}_A}$ values of all correlated source panels. (e) As in (c), but after application of slowness evaluation and discrimination for suppression of surface wave retrieval.

(Duin *et al.* 2006). Fig. 2(a) shows the *P*-velocity model employed (*S*-velocity and density models use the same subsurface distribution with non-constant values, see Fig. 1a), with the location of the ambient sources marked with black points and cyan stars, the receiver array with yellow triangles and the location of the virtual source at $x_A = 6000$ m indicated by the open red star. The sources enclose the receiver array from below, providing full illumination from the subsurface to the receiver array. The field employed in the crosscorrelation is the pressure field, and all sources employed are monopoles since the distance between the sources and the acquisition array assures that the recordings are in the far-field regime. If this was not the case, also recordings from dipole sources would have been required (Wapenaar & Fokkema 2006). The result of integrating the contribution of each correlated common-source panel is displayed in Fig. 2(c). The retrieved response is so dominated by surface waves, that reflection arrivals are hardly visible. Reflections, though, must be retrieved since the sources enclose the receivers. Fig. 2(d) shows the illumination diagnosis for all correlated common-source panels. The normalization of the slant-stack results scales the amplitudes but preserves the sign. The minimum and maximum slowness values are dependent on the expected surface wave velocity. In this modelling, the surface wave velocity is $860\,\mathrm{m\,s^{-1}}$ ($1.16 \times 10^{-3}\,\mathrm{s\,m^{-1}}$) . We chose values of $\pm 1.2 \times 10^{-3}\,\mathrm{s\,m^{-1}}$ in order to ensure all expected surface waves were scanned. The rows in the diagram represent the ray-parameter distribution of the source function $\widetilde{C}^S(\mathbf{x}_A, \mathbf{p})$ from one of the boundary sources. The black and cyan stars in the illumination diagnosis indicate dominant ray-parameter $\mathbf{p}^S_{\mathbf{x}_A}$ for surface wave slowness and body wave slowness, respectively. Panels with maximum slowness below the threshold slowness value but not fulfilling the control ratio *R* condition are also represented with black stars. Correlated common-source panels with dominant ray-parameter in cyan are kept for the following step of integration in the SI reflection-response retrieval procedure. The transient sources that lead to kept panels are indicated with the cyan stars in Fig. 2(a). The correlated common-source panels with dominant ray-parameter in black are discarded from further usage in the integration for reflection retrieval; the position of the sources giving rise to them are indicated by the black dots in Fig. 2(a). Correlated panels with dispersed ray-parameter distributions around the ray-parameter limit are discarded, for the ratio of the maxima of the source function $\tilde{C}^S$ inside and outside the limit interval is relatively small. This is notable in Fig. 2(d), for source numbers between 20 and 150. Fig. 2(e) shows the result of applying eq. (6) (the summation step in the SI retrieval process) only to the kept (the cyan) correlated common-source panels from Fig. 2(d).

The retrieved virtual common-source panel now shows clearly all the expected reflection arrivals. This can be seen by comparing it with the directly modelled panel in Fig. 2(b) for an active source at the position of the virtual source. The comparison also shows that our slowness evaluation and discrimination procedure has suppressed the retrieval of surface waves. Furthermore, we can see that also the direct *P*-wave arrivals are not retrieved. This is due to the fact that in the discrimination procedure, we chose the discrimination limit (the green line in Fig. 2d) slightly smaller than the slowness of the direct wave: $p_{\mathrm{limit}} = 6 \times 10^{-4}\,\mathrm{s\,m^{-1}}$, whereas the direct wave velocity at the virtual-source location is $6.25 \times 10^{-4}\,\mathrm{s\,m^{-1}}$.

Looking at Fig. 2(a), it can also be seen that due to the complexity of the subsurface model, changing the position of the virtual source would mean changing the positions of the sources that contribute to the retrieval of reflections. For the virtual source in Fig. 2(a), the majority of the sources to the right of the steep subsurface structure do not contribute to the retrieval of reflections, even if they are relatively deep sources.

3 ILLUMINATION DIAGNOSIS FOR AMBIENT-NOISE RECORDINGS

For SI with uncorrelated noise sources, Wapenaar & Fokkema (2006) derived a relation to retrieve the Green's function $G(\mathbf{x}_B, \mathbf{x}_A, t)$ as:

$$\Re\left\{G(\mathbf{x}_B, \mathbf{x}_A, t)\right\} * S_0(t) \approx \left\langle u^{\mathrm{obs}}(\mathbf{x}_A, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, t)\right\rangle, \tag{7}$$

where the noise sources are assumed to have the same autocorrelation function $S_0(t)$, $u^{\mathrm{obs}}(\mathbf{x}_A, -t)$ stands for the time-reversed total recorded noise at $\mathbf{x}_A$ due to all the noise sources and $\langle \cdot \rangle$ denotes ensemble average. For field applications, the ensemble average is exchanged for averaging over long recording times. As the long-time recordings are stored in time windows with certain length, the time averaging is exchanged for summation over all *i* time windows:

$$\Re\left\{G(\mathbf{x}_B, \mathbf{x}_A, t)\right\} * S_0(t) \approx \sum_i \left(u^{\mathrm{obs}}(\mathbf{x}_A, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, t)\right)_i . \tag{8}$$

To apply the slowness evaluation and discrimination procedure to such recordings, we define the correlated noise panel C^i as

$$C^i(\mathbf{x}_B, \mathbf{x}_A, t) = \left(u^{\mathrm{obs}}(\mathbf{x}_A, -t) * u^{\mathrm{obs}}(\mathbf{x}_B, t)\right)_i . \tag{9}$$

From here on, we can apply the illumination-diagnosis procedure using eqs (3)–(6) in the same way as for the transient noise sources.

In continuous ambient-noise recordings, the characteristics of the virtual-source function will depend on the noise sources acting during the recording time. Evaluation of the virtual-source function for relatively short windows would diagnose the illumination characteristics of the noise sources present during that time window.

The ensemble of illumination diagnosis results over consecutive time windows produces the illumination record. This display shows the succession of the dominant illumination in time along the noise record according to the window length applied.

The choice of the time-window length before the correlation fundamentally depends on two factors: the desired deepest reflection to be retrieved and the nature of the recorded noise. The time window should be at least as long as the expected two-way traveltime down to the deepest target reflector. With such a window, the correlation process would remove the traveltime of the direct arrival from the traveltime of its multiple and would retrieve the desired reflection from the target reflector. Increasing the length of the time window would result also in the correlation of later arrivals contributing to the retrieval of the same reflection with higher signal-to-noise ratio. Concerning the nature of the noise sources, in the case of our modelled data, we are not assuming transient noise signals, so long noise panels will improve the correlation quality, enable a proper reflection retrieval from the correlated panel and avoid the retrieval of spurious or non-physical events. So, for the modelled data, both factors demand longer time windows.

With field measurements, the sources of body-wave noise would be of limited time duration and would be present at discrete time periods during the passive survey. Using longer time windows would increase the risk of more than one source of body-wave noise being captured by the window, the illumination diagnosis would produce the information only for the strongest of these sources and this would practically mean loss of useful information. Furthermore, the longer the time window, the higher the risk of capturing more surface-wave noise. The latter might drown the present body-wave

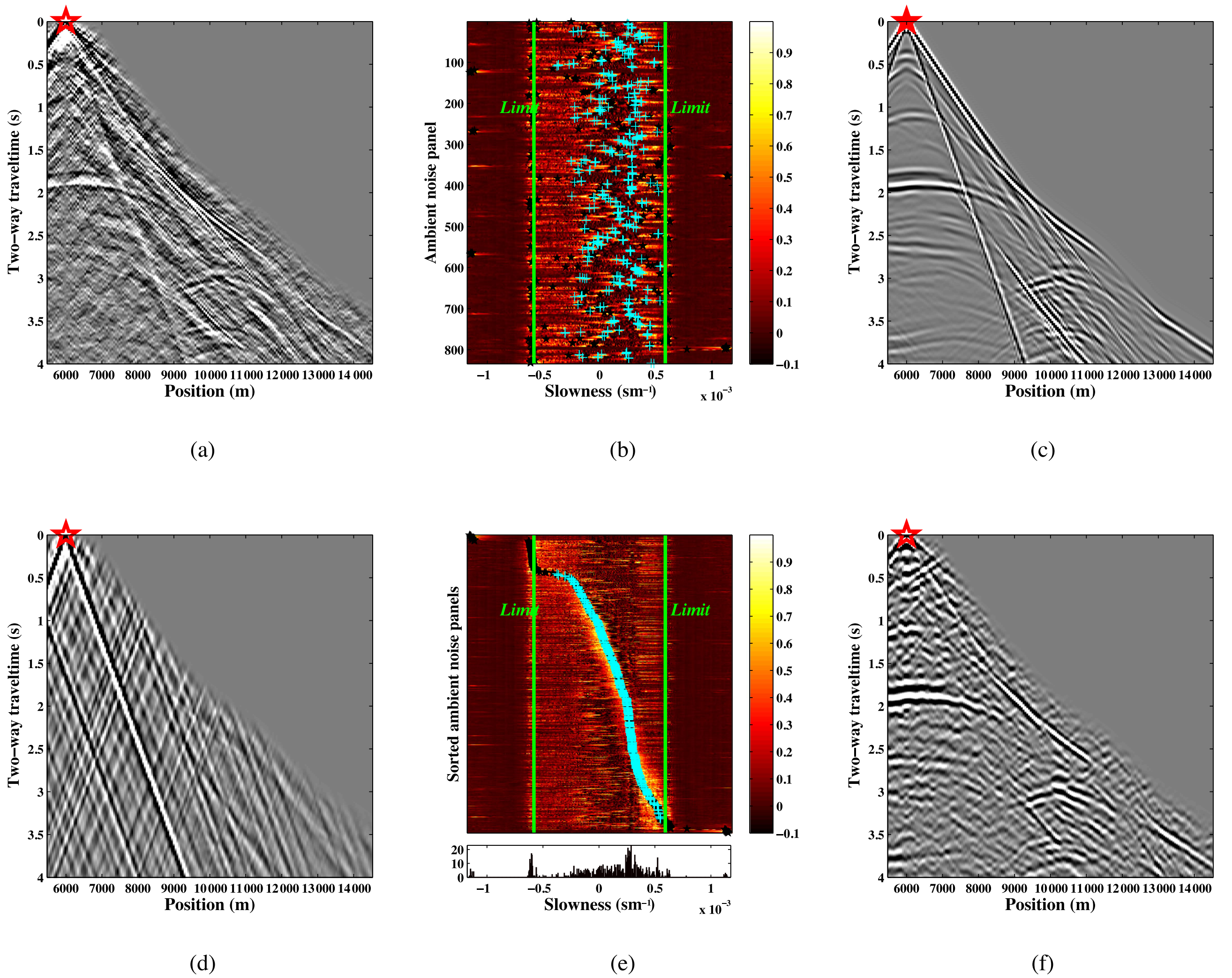


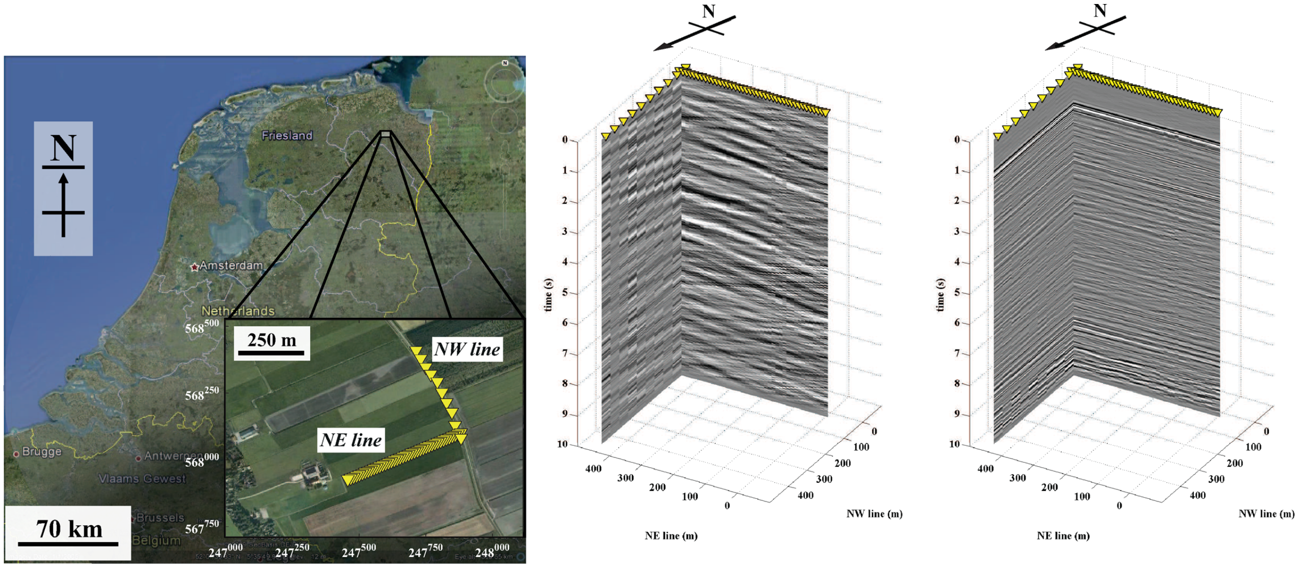
Figure 3. Illumination diagnosis for random noise sources in the model shown in Fig. 2(a). (a) Retrieved virtual common-source panel for a virtual-source position $x_A = 6000$ m obtained using crosscoherence. (b) Illumination record, constructed using 834 noise panels of 10-s length, with 5 s of record overlap. (c) Directly modelled common-source panel for an actual source at the virtual-source position. (d) Retrieved common-source panel obtained using all correlated panels. (e) Illumination diagnosis, with body wave dominated panels highlightened in cyan at their characteristic ray-parameter $\mathbf{p}^i_{\mathbf{x}_A}$; Diagram beneath shows the histogram of $\mathbf{p}^i_{\mathbf{x}_A}$. (f) Same result as in (d) after discrimination and discard of panels with dominant surface waves, that is, after summation only over the correlated panels with cyan stars, weighted according to the histogram in (e).

noise. So, for the case of field data, the two factors state opposing demands and thus a compromise should be sought.

The diagnosed illumination characteristics are then used to decide if a correlated noise panel $C^i(\mathbf{x}_B, \mathbf{x}_A, t)$ would contribute to the retrieval of mainly body waves or of mainly surface waves, and therefore be kept or discarded, respectively, for the consecutive summation.

Nevertheless, the application of the procedure as defined for the transient sources might not be optimal for the situation with ambient-noise recordings. The results shown in the previous section assumed a regular spatial distribution of the sources in the subsurface. In practice, body-wave noise might illuminate the receiver array more frequently from some directions than from others. This would affect the retrieval process adversely by distorting the retrieved reflection response. The illumination diagnosis provides an easy remedy for such situations. Upon sorting the noise panels by their dominant ray-parameter, the subsurface illumination distribution is observed, and enables statistical estimations for illumination balancing. The frequency of occurrence of illumination from a certain direction can be used to define weights W_i for the summation of the correlated noise panels. If the panels are individually amplitude-normalized, the weights W_i are set to be inversely proportional to the occurrence frequency of the ray-parameter value $\mathbf{p}^i_{\mathbf{x}_A}$. Illumination balancing with respect to ray-parameter can also be found in Ruigrok *et al.* (2010). Thus, the application of the illumination diagnosis and discrimination test to the ambient-noise recordings can be defined as follows:

$$C^i(\mathbf{x}_B, \mathbf{x}_A, t) = \begin{cases} 0 & \text{if} \quad \dfrac{\max\left(\|\widetilde{C}^i(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\| \le p_{\text{limit}}}}{\max\left(\|\widetilde{C}^i(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\| > p_{\text{limit}}}} \le R \\ W_i(\mathbf{p}^i_{\mathbf{x}_A})\,\left(u^{\text{obs}}(\mathbf{x}_A, -t) * u^{\text{obs}}(\mathbf{x}_B, t)\right)_i & \\ & \text{if} \quad \dfrac{\max\left(\|\widetilde{C}^i(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\| \le p_{\text{limit}}}}{\max\left(\|\widetilde{C}^i(\mathbf{x}_A, \mathbf{p})\|\right)_{\|\mathbf{p}\| > p_{\text{limit}}}} > R. \end{cases} \qquad (10)$$



(a) (b) (c)

Figure 4. (a) Geographical location of Anneveen, Northern Netherlands, where the noise recordings are taken. Close-up: Two perpendicular receiver lines are indicated by yellow triangles. Note: Number of receivers and spacing are different in both lines: NE line has 40 receivers, while NW line has only 10; Space sampling is 12 and 48 m, respectively. (b) 3-D display of an ambient-noise panel dominated by surface-wave noise. (c) 3-D display of an ambient-noise panel dominated by body-wave noise.

To minimize the possibility of not selecting body-wave noise, we allow certain time overlap between consecutive windows during our illumination diagnosis. Furthermore, this enables a smoother analysis in time of the illumination record. The choice for an optimal time overlap between noise panels must compensate for a precise detection of surface wave presence, without extending the computational time costs of scanning larger amount of noise panels.

3.1 Retrieval of reflections from synthetic ambient noise

We apply the above-described method to a synthetic continuous noise recording of 12 min, generated using the model and source distribution from Fig. 2(a). During the noise modelling, each ambient-noise source is activated randomly in time for 10 s. For the application of the illumination diagnosis, we divide the continuous recordings into 10-s-long noise panels with 5 s overlap. Figs 3(a), (d) and (f) show the respective results for retrieval of reflections in the form of virtual common-source panels for a virtual source at $x = 6000$ m after application of the illumination diagnosis.

Fig. 3(a) shows the retrieved virtual common-source panel after application of SI by crosscoherence as in Nakata *et al.* (2011), using all noise panels. Reflections show up with a whitened frequency spectrum, but under the presence of surface waves; see for comparison the directly modelled reflection response in Fig. 3(c). Fig. 3(b) shows the illumination record of the synthetic noise. By resorting the noise panels in Fig. 3(b) according to their maximum in the slowness spectrum, we obtain the illumination diagnosis displayed in Fig. 3(e). Under it, we show the histogram of $\mathbf{p}^i_{\mathbf{x}_A}$, upon which the weights (W_i) in eq. (10) are estimated. The retrieved common-source panels using SI by crosscorrelation before and after slowness evaluation, discrimination and weighting (expression 10) are shown in Figs 3(d) and (f), respectively. In contrast to the crosscoherence result from Fig. 3(a), in Fig. 3(f) the illumination diagnosis has not only succeeded to retrieve the reflections, but has successfully suppressed the surface and direct waves.

3.2 Retrieval of reflections from field data

In the previous sections, we showed how the illumination diagnosis should be applied to transient or ambient-noise sources considering a line of receivers above a 2-D medium. Field applications for retrieval of reflections from ambient noise using a line of receivers can lead to misleading results due to the lack of the 3-D character of the wavefield. Surface-wave noise coming at the receiver line from the crossline direction might be recorded with apparent slowness characteristic of body waves. Such arrivals will be inherited also by the SI retrieved results and be misinterpreted as retrieved reflections. To avoid such erroneous interpretations, ambient-noise recordings in the field should be carried out using areal arrays. For the application of illumination diagnosis, the minimum optimal geometry is to use crossing lines.

We apply the illumination diagnosis for retrieval of reflections to ambient noise recorded near the town of Annerveen in the north of the Netherlands. During the recording, an Earth tremor was detected by the array. We will use this event in the analysis.

For the retrieval of reflections in virtual common-source panels, we work with two perpendicular lines of receivers as displayed in Fig. 4(a). The first line has an NE orientation and is composed of 40 receivers equally spaced at 12 m. The second line follows an NW orientation and has 10 receivers with 48 m spacing. Both arrays are buried at 50 m depth in the subsurface. The sampling frequency is 250 Hz. A total of 23 hr and 56 min of ambient noise has been processed for this work, split into 34 434 noise panels of 10 s length with 7.5 s overlap between them.

Fig. 4(b) shows an ambient-noise panel along both lines dominated by surface waves coming from one side of the lines (road noise). We can see that along the NE line, the noise appears to be

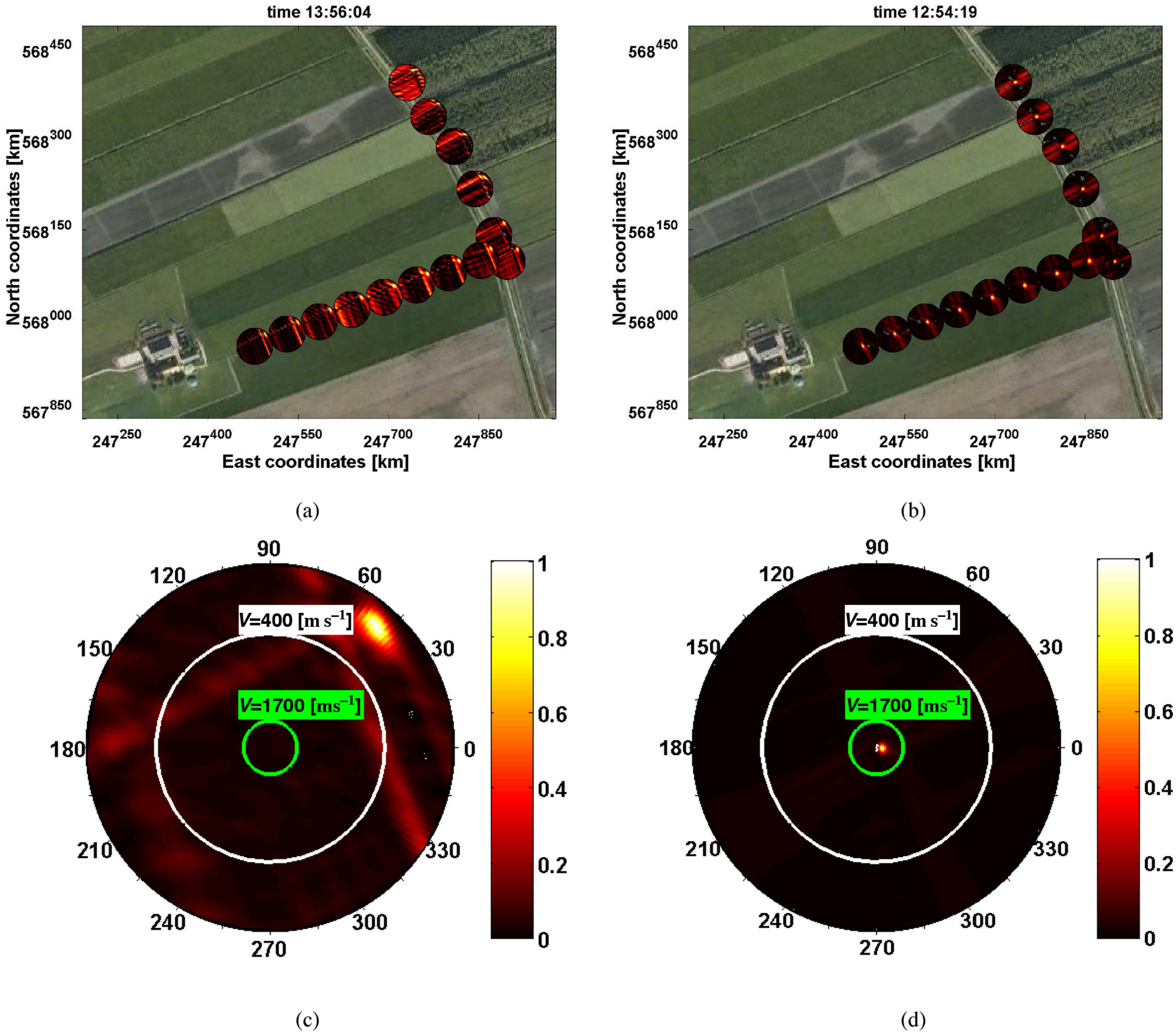


Figure 5. Illumination diagnosis at different time sections. (a) Virtual-source illumination diagnosis at some of the receiver locations in Fig. 4(a), due to surface-wave noise. (b) Illumination diagnosis of an Earth tremor. (c) Integration of the illumination diagnosis from (a). (d) Same as in (c), from (b).

characterized by a low ray-parameter, that in a 2-D setting might cause it to be interpreted as body waves; however, along the NW line, though, the arrivals are characterized by a ray-parameter typical for surface waves. Fig. 4(c) is an example of another noise panel with arrivals from a deep source (Earth tremor), characterized by a low ray-parameter in both perpendicular lines.

In Fig. 5, we compare the use of the illumination diagnosis in the same area, with separate noise panels from different time sections. Figs 5(a) and (b) display the illumination diagnosis from some of the receiver locations ($\mathbf{x}_A$) with respect to the rest of the array, due to the ground tremor and to surface wave ambient noise, respectively. Figs 5(c) and (d) show the integration from all the individual illumination diagnosis from Figs 5(a) and (b), respectively. Note that the normalization of the slant-stack results here scales the amplitudes and takes the absolute value. The array required directional slowness-balancing for the illumination diagnosis not to suffer from spatial aliasing caused by the array design. Although one gets to identify the dominant ray-parameter $\mathbf{p}^i_{\mathbf{x}_A}$ for the Earth tremor or the ambient sources located at the surface, the results still show the spatial aliasing imprint in the perpendicular direction of the respective lines.

The illumination diagnosis is closely related to the beam-forming method (Lacoss *et al.* 1969), since it also analyses crosscorrelations of wavefields. However, our approach is different in the sense that we directly interpret the correlated incident field at the receiver location, as the source function of the reconstructed virtual source. For similar reasons, we apply our method in the $\tau - p$ domain at $\tau = 0$ s only, to restrict ourselves to the incident field only (without having to apply a time window). Moreover, the illumination diagnosis is independent between stations because one could use the source function at any virtual-source position, therefore one gets as many diagnosis results as receivers there are available.

Besides the results in Fig. 5, due to the aliasing and differences in space sampling of the two crossing lines, NW and NE, we decided not to carry out the illumination diagnosis using both lines' receivers together, but instead using each line's independently: At each correlated noise panel, the illumination conditions are analysed by detecting the dominant $\mathbf{p}^i_{\mathbf{x}_A}$ at each of the receiver lines.

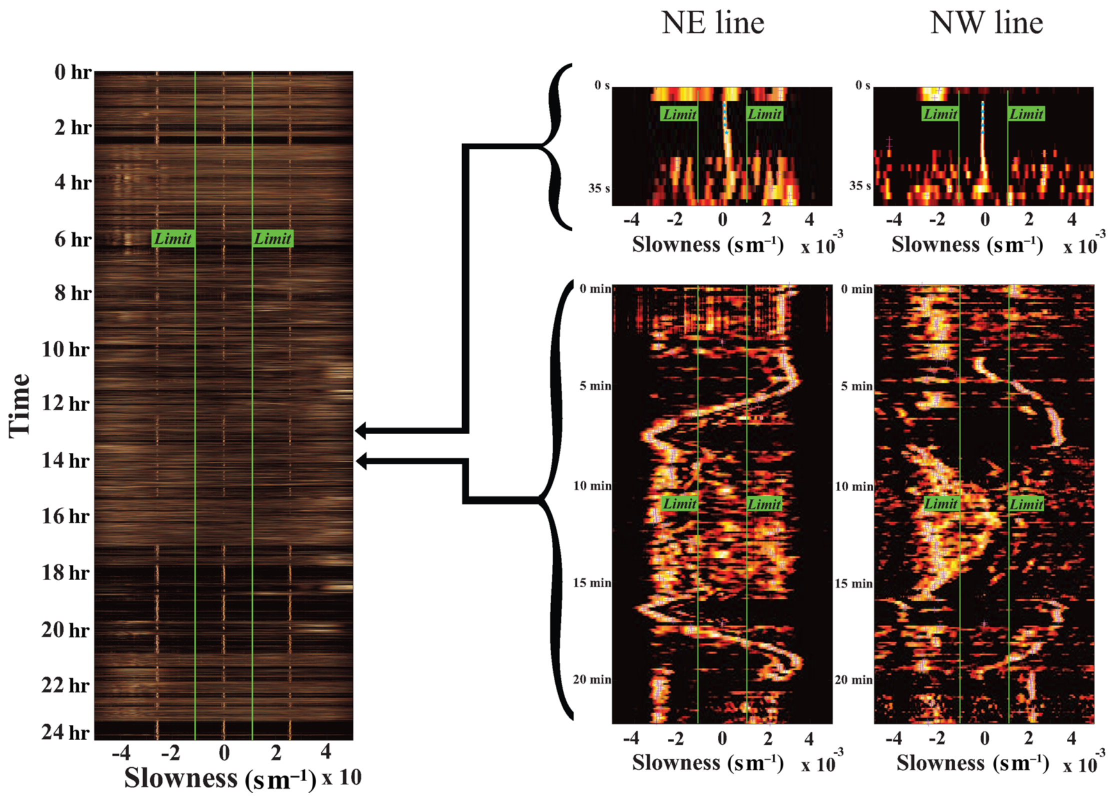


Figure 6. Illumination diagnosis on nearly 24 hr of noise along line NE (left-hand side) with two different features: a weak Earth tremor (middle top); and a surface source in motion identified by a pseudo-helix feature in the illumination records (middle bottom). The magenta dots indicate the dominant $\mathbf{p}^i_{\mathbf{x}_A}$ for each of the correlated panels. The 3-D nature of the noise being from body wave or surface wave sources can be judged when the illumination diagnosis along the NE line is complemented by the illumination diagnosis along the NW line (right).

The estimated surface wave velocity at the site is $370\,\mathrm{m\,s^{-1}}$ ($2.7\times10^{-3}\,\mathrm{s\,m^{-1}}$). Therefore, the illumination diagnosis had a minimum velocity to start scanning with of $200\,\mathrm{m\,s^{-1}}$ ($5\times10^{-3}\,\mathrm{s\,m^{-1}}$).

In Fig. 6, we can see that there are numerous correlated noise panels along the NE line that are dominated by arrivals with low $\mathbf{p}^i_{\mathbf{x}_A}$ values, which fall inside the limits for being characteristic of body-wave noise. However, to decide in a 3-D sense whether a noise panel is characterized by body or surface-wave noise, we have to take a closer look at the illumination characteristics of the noise along the NE line (Fig. 6, middle) and compare them to the illumination characteristics of the corresponding noise panels along the NW line (Fig. 6, right). Then, we can see that some of the low $\mathbf{p}^i_{\mathbf{x}_A}$ values along the NW line correspond to low $\mathbf{p}^i_{\mathbf{x}_A}$ values along the NW line (Fig. 6, middle top and right top, correspond to the Earth tremor from Fig. 4c). On the other hand, the illumination diagnosis of both lines may inconsistently identify surface-wave noise as well when only one line exhibits low $\mathbf{p}^i_{\mathbf{x}_A}$ values (Fig. 6, middle bottom and right bottom, corresponding to road noise or farming activities). Therefore, only correlated noise panels from time windows that are dominated by low $\mathbf{p}^i_{\mathbf{x}_A}$ values on the diagrams for both the NE and NW lines are being selected for the subsequent summation of the correlated noise panels.

In Fig. 7, the illumination diagnosis is applied to 90 min of noise. Figs 7(a) and (b) show the dominating $\mathbf{p}^i_{\mathbf{x}_A}$ along the record in both NW and NE lines. For that short amount of data, very few panels were detected to be suitable for body wave retrieval (blue stars), while the great majority was dominated by surface waves (magenta dots). The resorting of the diagnoses from the two lines are shown in Figs 7(c) and (d). The weights, estimated from the ray-parameter histograms, is in this case not necessary because of the small illumination available.

Because of the difference in resolution for the analysis of the virtual-source function $\widetilde{C}^i$ along the two lines due to different amounts of receivers (NW line: 10, NE line: 40) and sampling (NW line: 48 m, NE line: 12 m), for this data set an additional test is applied to the correlated panels with desired low $\mathbf{p}^i_{\mathbf{x}_A}$ values in both lines: We carried out a polarization analysis on each panel that was diagnosed to be dominated by body waves before correlation, profiting from multicomponent receivers (Vx, Vy and Vz). Fig. 8 shows two panels with dominant body waves (a) and for comparison also dominated by surface waves (b). Hodogram pairs consist of $Vx - Vz$ and $Vy - Vz$, in blue and red colour, respectively. These are obtained from 0.5-s sections at the same time of the respective components, every 0.9 s. Hodograms are displayed in every six traces of the NE line. In Fig. 8(b), surface-wave noise produce elliptical features in both hodograms. Fig. 8(a) shows in its hodograms a rather more polarized behaviour, confirming body wave particle vibration. The first section (until 6 s) corresponds to body wave arrivals from an Earth tremor. At time 7 s appear the shear wave arrivals.

Figs 9(a) and (b) display the retrieved common-source panels from crosscoherence at $x_A = 0$ m in both lines NW and NE, respectively. The summation of all correlated noise panels is displayed in Figs 9(c) and (d). Both methods have succeeded in retrieving arrivals, but these are surface waves, that dominate the SI results and therefore retrieved reflections are not observable. Figs 9(e) and

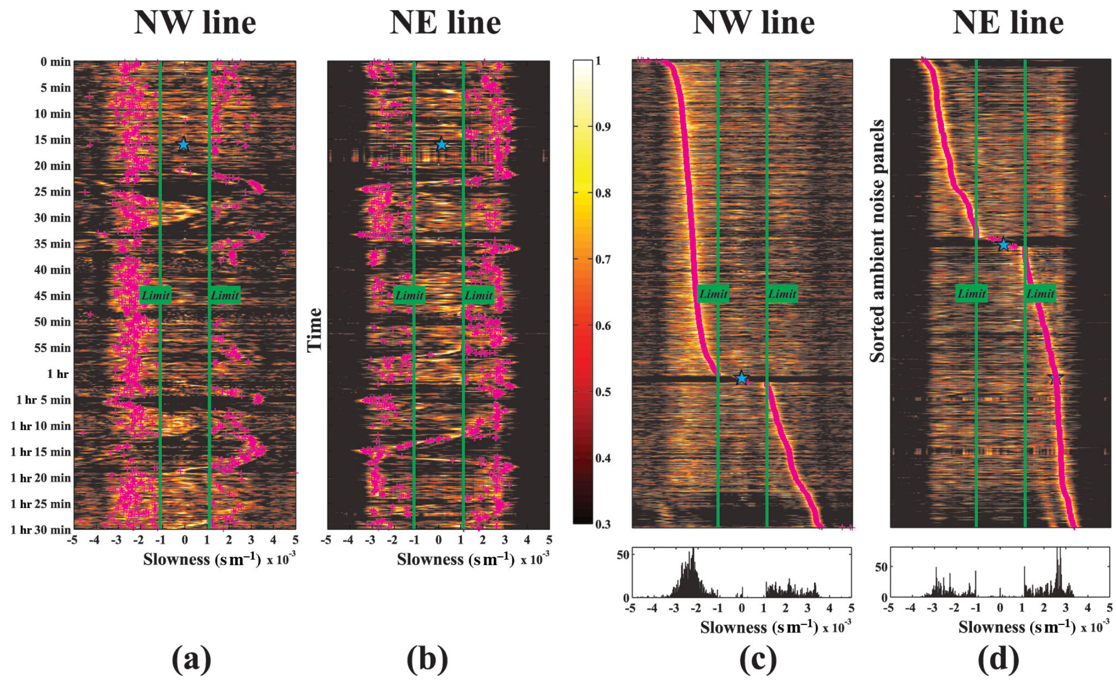


Figure 7. Illumination diagnosis with 90 min of the total noise record from Fig. 6. (a) Illumination record for the NW line at position $x_A = 0$ m. Ray-parameter $\mathbf{p}^i_{\mathbf{x}_A}$ is highlightened with a cyan star (dominating body waves) or a magenta dot (dominating surface waves). (b) Illumination record for the NE line at position $x_A = 0$ m. (c) Illumination diagnosis for all correlated panels in (a), with the dominant ray-parameters $\mathbf{p}^i_{NW,\mathbf{x}_A}$ highlightened; Diagram below shows the histogram of $\mathbf{p}^i_{NW,\mathbf{x}_A}$. (d) Illumination diagnosis for all correlated panels in (b). Diagram below shows the histogram of $\mathbf{p}^i_{NE,\mathbf{x}_A}$.

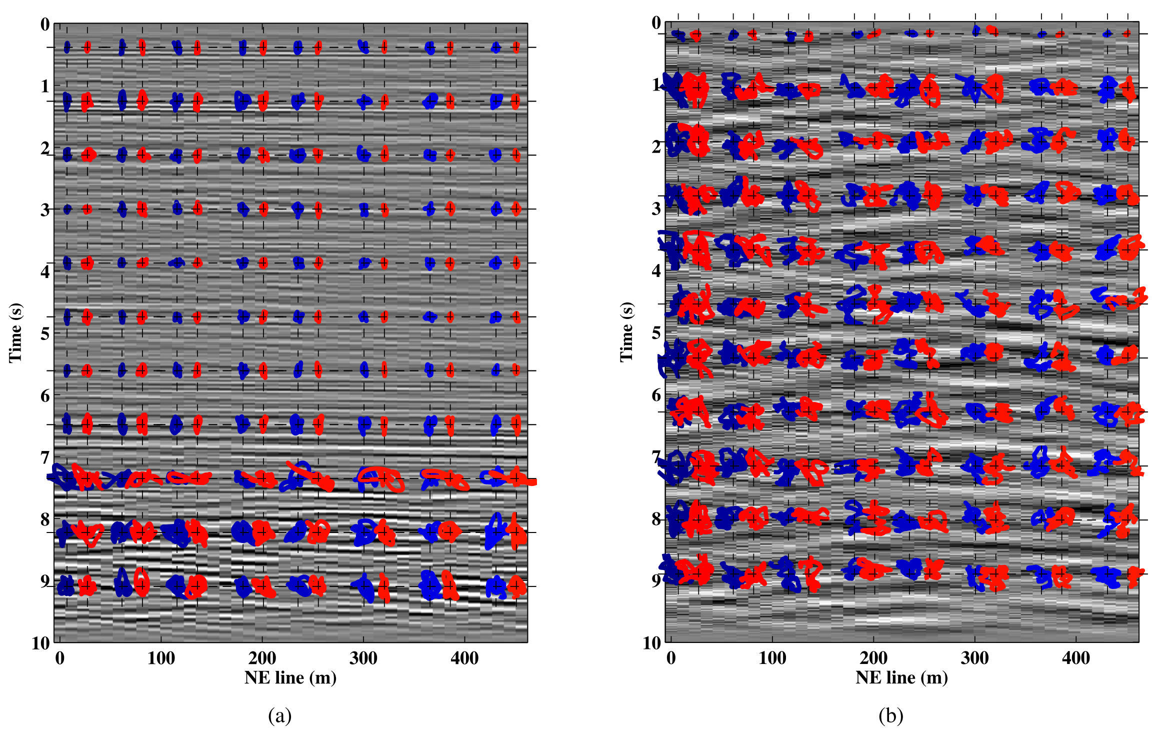


Figure 8. (a) The Vz component of an NE line ambient-noise record, dominated by body waves, with the respective hodograms every 0.8 s in every sixth trace, between components $Vx - Vz$ (blue) and $Vy - Vz$ (red). (b) As in (a), with ambient noise dominated by surface waves.

(f), present the respective results from (c) and (d), which are using crosscorrelation, but after illumination diagnosis. Among the 34 541 noise panels analysed, only five passed the test along both NE and NW lines. Of these positive cases, only four were suitable for reflection retrieval as observed at the polarization analysis, to be sure that they indeed are dominated by body-wave noise. We did this, because the NW line is sampled only by 10 geophones, which makes the illumination diagnosis, in this case, difficult. These numbers show the small amount of data available for the retrieval of reflections in the ambient noise recorded at the location of the

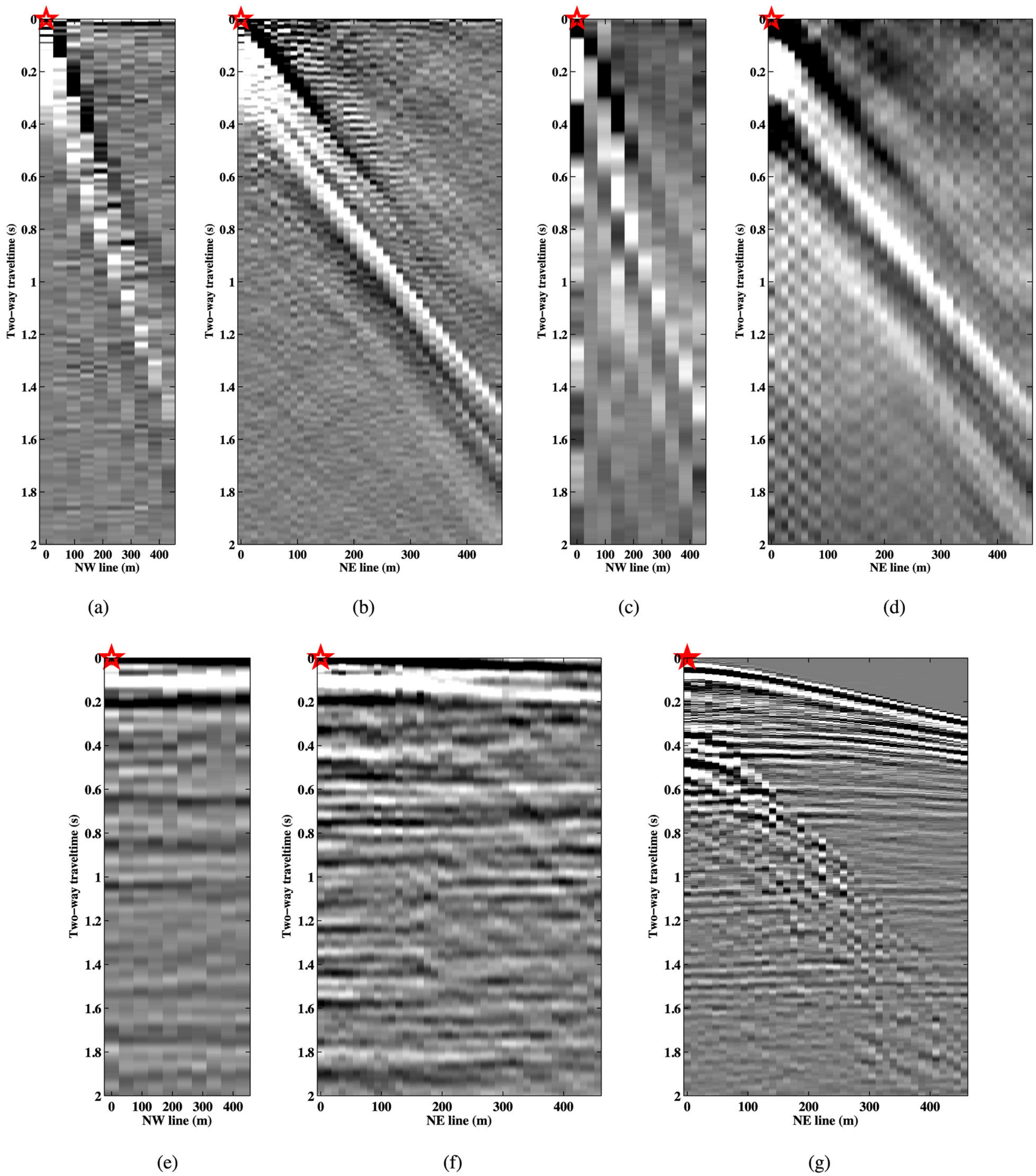


Figure 9. Retrieved Vz-component common-source panels for virtual source at $x_A = 0$ m for the NE line and $x_A = 0$ m for the NW line. (a) Crosscoherence result for the NW line. (b) Crosscoherence result for the NE line. (c) Crosscorrelation result using all panels for the NW line. (d) Crosscorrelation result using all panels for the NE line. (e) Same result as in (c) after discrimination and discard of panels with dominant surface waves. (f) Same as in (e) for the NE line. (g) Reference response for an active source located at the surface above the virtual-source location.

acquisition array. This is related to the continuous anthropogenic activities, which result in continuous generation of strong surface waves, and the seismicity of the area, which is very low.

The comparison of the results from the three methods in Fig. 9 shows that the application of illumination diagnosis has successfully suppressed the retrieval of surface waves, while the retrieval of body wave events has been enhanced. Comparing these events with reflection arrivals recorded using an active source (Fig. 9g), we can conclude that at least some of the retrieved events are reflections. Note that the active data were shot with a source at the surface, while the virtual source in the retrieved data is at 50 m depth.

Due to the small-slowness values in the used correlated panels, together with the short aperture of the receiver array, the retrieved results exhibit predominantly nearly horizontal events.

No weights were required to balance the illumination, since the dominant ray-parameters were nearly homogeneously distributed between 8×10^{-5} and $12 \times 10^{-5}\,\mathrm{s\,m^{-1}}$ from approximately the same azimuthal direction. The moveout is not recovered due to the very limited illumination angles in the noise panels selected for retrieval. Earlier reflections are not properly estimated due to the lower frequency content of the retrieved response. At early arrival times, there is more moveout in the reflection response, making it harder to reconstruct with limited illumination. Moreover, there is the imprint of the virtual-source function overlying the early reflections.

Note that some of the retrieved horizontal events might be non-physical. Increased illumination of the recording lines by body-wave noise could contribute, in addition to the potential retrieval of reflection hyperbolae, to the suppression of non-physical arrivals by destructive interference.

4 DISCUSSION

The frequency characteristics of the retrieved body waves would depend and be inherited from the frequency characteristics of the recorded body-wave noise. These frequency characteristics are often dominant at low frequencies (Ruigrok *et al.* 2011). Some studies have accomplished retrieval of reflections with frequency content comparable to that of active surveys (Draganov & Panea 2011), even with human-induced noise (Nakata *et al.* 2011). Except for the case of the earthquake, however, body wave ambient noise in the study area is more dominant at lower frequencies due to oceanic waves. The reflection response obtained from passive SI will bear this low-frequency characteristics, and might be an incentive to combine the reconstructed result with active survey results which usually lack low-frequency content. By complementing passive and active results, the merging of both (virtual-) panels may produce a broad-band reflection response.

The attempt to balance the subsurface illumination could be further improved as well. In addition to the weights, other applications, such as directional balancing (Curtis & Halliday 2010) and multidimensional deconvolution (Wapenaar *et al.* 2008) could be used.

The size of the correlation windows to be used for illumination diagnosis is very important. In the above results, we used windows of 10-s length because of a limited amount of sources acting inside the panel in the synthetic experiment. The use of shorter time windows might improve the illumination analysis over time by discarding less portions of the noise dominated by surface waves, but would consequently result in retrieval of reflections from shallower parts of the subsurface. This is because the window length dictates the maximum two-way traveltime that could be retrieved. The optimal time windows to be used will depend on the ambient-noise source length and coincidence.

Complementary, besides avoiding the retrieval of surface waves, the proposed method also allows gathering of only the useful data for reflection retrieval, discarding unnecessary data and reducing later processing and storage costs.

5 CONCLUSIONS

We proposed a method to analyse the illumination characteristics of recorded ambient noise to be used for passive SI. To apply the method, the recorded noise panels are correlated to obtain correlated panels and from them we used the events that pass through the virtual-source position at zero time. We transformed these events to the slowness domain and analysed them to classify the different noise panels as being dominated by surface wave or body-wave noise. The illumination diagnosis from the correlated panels enables defining balancing weights in order to compensate overillumination from certain directions.

We applied the analysis to field data acquired in the north of the Netherlands. We dealt with the multiazimuth radiation recognition using two orthogonal receiver lines as acquisition setup. We showed that the retrieved results when using all recorded noise exhibit mainly surface waves. The illumination analysis of the ambient noise allowed to identify and discard parts of the recording as dominated by surface-wave noise. The illumination diagnosis detected and isolated several noise panels dominated by body-wave noise. Polarization analysis of the selected noise panels confirmed the dominant ambient noise to be due to body waves. The amount of body-wave noise panels and their respective illumination was less than desired for a complete reflection retrieval. Nevertheless, comparison of the results retrieved after illumination diagnosis with active-source seismic data confirmed that some of the retrieved events are reflection arrivals.

ACKNOWLEDGEMENTS

This work is supported by the Netherlands Research Centre for Integrated Solid Earth Science (ISES).

We are thankful to the anonymous reviewers of this paper for their constructive comments and valuable discussions.

REFERENCES

Curtis, A. & Halliday, D., 2010. Directional balancing for seismic and general wavefield interferometry, *Geophysics,* **75,** SA1–SA14.

Curtis, A., Gerstoft, P., Sato, H., Snieder, R. & Wapenaar, K., 2006. Seismic interferometry—turning noise into signal, *Leading Edge,* **25,** 1082–1092.

Draganov, D. & Panea, I., 2011. Seismic interferometry with ambient noise from Mizil area, Romania, in *Proceedings of 73th EAGE Conference and Exhibition,* Vienna, Austria.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A. & Wapenaar, K., 2009. Reflection images from ambient seismic noise, *Geophysics,* **74,** A63–A67.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A. & Wapenaar, K., 2010. Event-driven seismic interferometry with ambient seismic noise, in *Proceedings of 72nd EAGE Conference and Exhibition,* Barcelona, Spain.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A. & Wapenaar, K., 2013. Seismic exploration-scale velocities and structure from ambient seismic noise (>1 Hz), *J. geophys. Res.,* **118,** 4345–4360.

Duin, E., Doornenbal, J., Rijkers, R., Verbeek, J. & Wong, T., 2006. Subsurface structure of the Netherlands—results of recent onshore and offshore mapping, *Neth. J. Geosci.,* **85**(4), 245–276.

Forghani, F. & Snieder, R., 2010. Underestimation of body waves and feasibility of surface-wave reconstruction by seismic interferometry, *Leading Edge,* **29,** 790–794.

Lacoss, R., Kelly, E. & Toksöz, M., 1969. Estimation of seismic noise structure using arrays, *Geophysics,* **34,** 21–38.

Nakata, N., Snieder, R., Tsuji, T., Larner, K. & Matsuoka, T., 2011. Shear-wave imaging from traffic noise using seismic interferometry by cross-coherence, *Geophysics,* **76**(6), SA97–SA106.

Nishida, K., 2013. Global propagation of body waves revealed by cross-correlation analysis of seismic hum, *Geophys. Res. Lett.,* **40,** 1691–1696.

Poli, P., Campillo, M. & Pedersen, H., 2012. Body-wave imaging of Earth's mantle discontinuities from ambient seismic noise, *Science,* **338,** 1063–1065.

Roux, P., Sabra, K.G., Gerstoft, P., Kuperman, W.A. & Fehler, M.C., 2005. *P*-waves from crosscorrelation of seismic noise, *Geophys. Res. Lett.,* **32,** L19303, doi:10.1029/2005GL023803.

Ruigrok, E., Campman, X., Draganov, D. & Wapenaar, K., 2010. High-resolution lithospheric imaging with seismic interferometry, *Geophys. J. Int.,* **183,** 339–357.

Ruigrok, E., Campman, X. & Wapenaar, K., 2011. Extraction of *P*-wave reflections from microseisms, *C. R. Geosci.,* **343,** 512–525.

Schuster, G.T., 2009. *Seismic Interferometry,* Cambridge Univ. Press.

Shapiro, N. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31,** L07614, doi:10.1029/2004GL019491.

van der Neut, J., 2013. Downhole interferometric illumination diagnosis and balancing, *Geophys. Prospect.,* **61,** 352–367.

Van Wijk, K., Mikesell, D., Blum, T., Haney, M. & Calvert, A., 2010. Surface wave isolation with the interferometric Green tensor, *SEG Expanded Abstr.,* **29,** 3996–4000.

Wapenaar, K. & Fokkema, J., 2006. Green's function representations for seismic interferometry, *Geophysics,* **71,** SI33–SI46.

Wapenaar, K., van der Neut, J. & Ruigrok, E., 2008. Passive seismic interferometry by multidimensional deconvolution, *Geophysics,* **73,** A51–A56.

Xu, Z., Juhlin, C., Gudmundsson, O., Zhang, F., Yang, C., Kashubin, A. & Lüth, S., 2012. Reconstruction of subsurface structure from ambient seismic noise: an example from Ketzin, Germany, *Geophys. J. Int.,* **189,** 1085–1102.

Zhan, Z., Ni, S., Helmberger, D. & Clayton, R., 2010. Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise, *Geophys. J. Int.,* **182,** 408–420.