Geophysical Prospecting

Geophysical Prospecting

ORIGINAL ARTICLE OPEN ACCESS

Adaptation of Iterative Supervirtual Interferometry for Enhancing Higher Modes From Surface Waves

Sverre Hassing | Deyan Draganov | Eric Verschuur

Civil Engineering and Geosciences, TU Delft, Delft, The Netherlands

Correspondence: Sverre Hassing (s.h.w.hassing@tudelft.nl)

Received: 27 November 2025 | **Revised:** 11 May 2026 | **Accepted:** 23 June 2026

Keywords: near-surface | seismics | surface waves

ABSTRACT

For geotechnical applications, surface-wave methods are one of the most important techniques for gathering information about the strength of the near surface, but these methods are sensitive to noise. Often only the fundamental mode is considered, but higher modes can also provide more information on deeper parts of the subsurface. We investigate the use of iterative supervirtual interferometry (SVI) to increase the signal-to-noise ratio and to extract more information about the higher modes. A big advantage is that the method results in data in the same domain as the original data, such that similar multichannel analysis of surface waves–based procedures can be used afterwards. SVI utilises seismic interferometry to extract the extra information in consecutive folds. Generally, the strongest event in the data is boosted the most. Although surface waves are usually the strongest, they consist of sub-events (the modes) with different amplitudes. Previously, iterative SVI has been applied to further boost the result of SVI. For surface waves, this would further diminish the presence of weaker modes. Instead, we subtract the results of SVI from the original data and use the remainder as input for the next iteration, which gives potential for highlighting different modes. We apply our methodology to three different datasets to show different properties of the iterative SVI. When the level of random noise is high, further iterations retrieve mostly the dominant event and no extra information on different modes is gained. When the level of noise is relatively low, we show that higher modes can be found in a wider frequency band in the results of each separate iteration. In both cases, the combination of all iterations provides a more complete description of the surface waves. The presence of strong random noise or spikes in the data can have negative effects on the methodology in the form of spurious events originating from the spikes.

1 | Introduction

For geotechnical investigations of the near surface on land, seismic surface-wave methods are one of the most commonly used methods. Surface waves are mostly sensitive to the shear-wave velocity of the subsurface, which is directly related to the shear modulus (Socco and Strobbia, 2004). This is the parameter of interest as it relates to the design of loading on the subsurface (Ghose and Goudswaard, 2004).

The primary aim of seismic surface-wave methods is to extract dispersion information from the data. The dispersion manifests as phase differences, recorded at different frequencies along ray paths. The surface-wave methods, however, are sensitive to noise and lateral variations of the medium in the subsurface. In early geotechnical measurements of surface waves, a method called spectral analysis of surface waves (Nazarian and Stokoe II, 1985) was used. In this method, two geophones are placed a certain distance apart with a source placed at one side. Phase differences are extracted from the spectra of the recordings from the two geophones. Due to significant problems with noise, this was repeated often with different receiver distances, source offsets and source sides. All results were added together to obtain an average dispersion curve assigned to the midpoint of the two receivers. The result is a method that is labour-intensive and still has problems with accuracy.

The development of multichannel analysis of surface waves (MASW) can be seen as a response to this method (Park et al., 1999). Linear receiver arrays are used instead of only two stations. Via a domain transformation, the seismic data can be converted to the phase velocity–frequency domain (the dispersion spectrum). In its most simple form, this transformation can be seen as an fk-transform with a subsequent exchange of the wavenumber axis for phase velocity. Park et al. (1999) suggested using the phase-shift method, which also includes a further normalisation to remove amplitude effects (Park et al., 1998). With a constant receiver spacing, longer arrays result in a higher spectral resolution and a lower wavenumber uncertainty, but a lower signal-to-noise ratio (Socco and Strobbia, 2004). The signal-to-noise ratio can be boosted by using the overlapping fold of different shots to compute the dispersion spectrum for the same stretch of geophones multiple times and stacking these.

In the phase velocity–frequency domain, dispersion curves will show up as high-energy bands. Often, it is assumed that the fundamental mode is the most energetic; this means that by picking the maxima of energy for each frequency, the fundamental-mode dispersion curve could be extracted. A second assumption is that the medium is one-dimensional. When lateral variations in the medium are present below the receiver line, the dispersion spectrum exhibits anomalous effects (Boiero and Socco, 2011). These effects result in ambiguities in the picking of dispersion curves. In the presence of low or smooth lateral variation, the recommendation is to use short receiver lines. The resulting dispersion curves then capture local characteristics of the medium. The inversion results from different local dispersion curves can then be smoothed to obtain a two-dimensional model of the medium.

As mentioned above, higher modes are often ignored. In the presence of strong velocity contrasts and inversions, the balance of energy between the different modes changes and at certain frequency bands, the higher modes can become dominant. This can cause misidentification of a high-energy band consisting of different surface-wave modes as a single fundamental mode (Socco et al., 2010). A low spectral resolution can also make it difficult to identify different peaks in the dispersion spectrum. However, the higher modes themselves carry useful information on the deeper subsurface, as they have a longer wavelength and are sensitive to greater depths than the fundamental mode at the same frequency. Different inversion schemes take higher modes into account to further constrain the velocity model (Xia et al., 2003; Luo et al., 2007; Khosro Anjom et al., 2021).

In this work, we focus on improving the signal-to-noise ratio of the data with supervirtual interferometry (SVI) and on how this technique relates to the higher modes. The technique utilises seismic interferometry to extract additional information included in multi-fold data and the output is in the same domain as the input data. This means that it can be used as a preprocessing step for any method that would ordinarily be applied. SVI was originally proposed for boosting refracted arrivals (Mallinson et al., 2011; Bharadwaj et al., 2012). For this to work, stronger arrivals had to be removed or the refracted arrivals isolated. However, the idea could be applied to all events originating from the same stationary-phase region, meaning it can also be applied to boost the surface-wave content of the seismic data (An and Hu, 2016; Xu et al., 2017). The advantage of surface waves is that they are generally the highest-energy events recorded.

However, the multi-modal characteristics of surface waves mean that the standard workflow for SVI is not applicable. After applying SVI, generally, only the fundamental mode is boosted. In order to further boost the signal-to-noise ratio, Al-Hagan et al. (2014) suggested iteratively applying SVI to the output of each iteration of SVI. This would only serve to further reduce the effects of the higher modes in the data. We propose to adjust the workflow of iterative SVI for surface waves by taking the data remainder of each iteration instead of the direct output of SVI as input for the next iteration. This remainder is calculated by subtracting the output of SVI from the input data using adaptive subtraction. The remaining part would consist of noise from sources out of the stationary-phase region and the weaker events that were boosted less by SVI.

A complication is that seismic interferometry gives rise to cross-terms between the different (higher) modes. With only sources at the surface – the case in most active surveys – these cross-terms can even overwhelm the true higher modes (Halliday and Curtis, 2008; Kimman and Trampert, 2010). These cross-terms diminish when using sources in line with the receivers (Halliday and Curtis, 2008) and when the distance to the source is large compared to the interreceiver distance. The aim of the adaptive subtraction is in part to correct for this. On the other hand, lateral variations between the receivers of interest are accounted for, as long as the other assumptions behind seismic interferometry hold.

In the Theory section, we will explain the theoretical background of seismic interferometry and SVI and use this to explain the adapted workflow for iterative SVI. Then, in the Methods section, we will discuss the implementation and introduce the three datasets that are used to investigate the effects of applying iterative SVI. In the Results section, we will show how the data is affected by iterative SVI both in the time domain and in the dispersion spectrum. In the Discussion section, we will discuss the theory and compare it to observed results. Finally, we will combine all observations in the Conclusion.

2 | Theory

SVI is based on the theory of seismic interferometry, which states that the Green's function between two receivers can be retrieved by, for example, cross-correlating two traces that recorded the same event and are located along the same travel path (Wapenaar and Fokkema, 2006). Effectively, this removes the shared travel path from the second trace and results in a trace recorded at the second receiver as if the source were located at the first receiver position, called the virtual source, as illustrated in Figure 1.

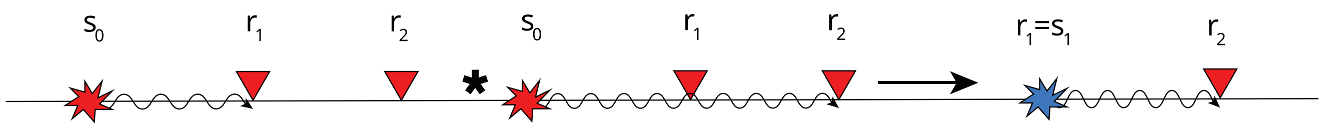

FIGURE 1 | The basic concept of seismic interferometry by cross-correlation. The trace from the original source position s_0 recorded at receiver r_1 is cross-correlated ($*$) with the trace recorded at receiver r_2. The result is the trace that would have been recorded at receiver r_2 if there were a (virtual) source at receiver r_1, which we can call s_1. The figure is based on Wapenaar et al. (2010).

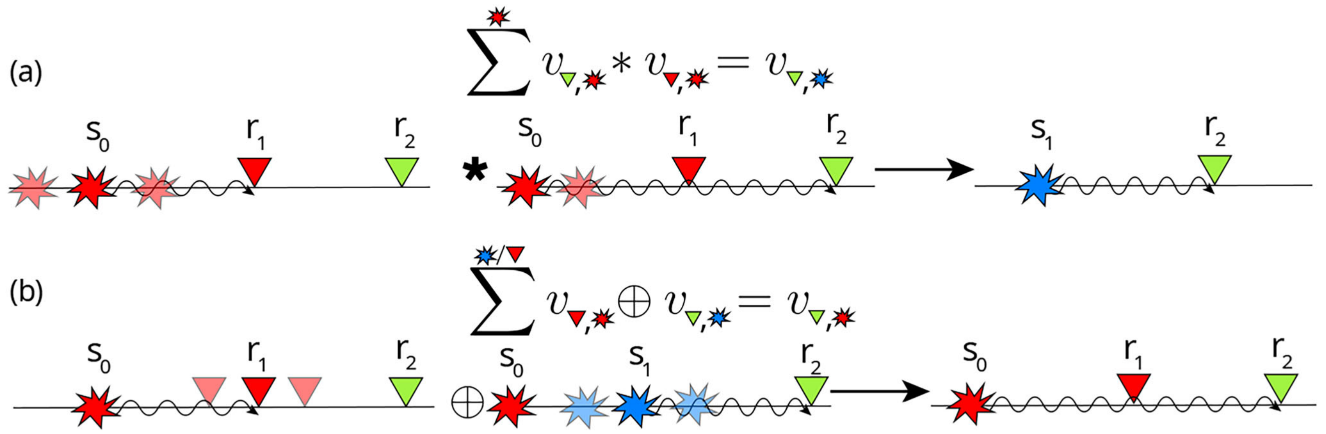

FIGURE 2 | The basic concept of SVI. In the first step shown in (a), the trace between virtual source position s_1 and receiver r_2 is retrieved from all valid source positions and then summed to boost the signal-to-noise ratio. This is effectively the same workflow as used for seismic interferometry. In the second step shown in (b), the original trace from source position s_0 to receiver r_2 is retrieved by convolving ($\oplus$) the boosted trace and the original trace from source position s_0 to receiver r_1. Again, this can be repeated for multiple different virtual source positions and summed together to boost the signal-to-noise ratio. The result is a trace characterised by the same travel path as the original one, but with a higher signal-to-noise ratio. The figure is adapted from Bharadwaj et al. (2012).

In order to retrieve the full Green's function, it is required to integrate this cross-correlation over a (smooth) source boundary effectively surrounding the medium of interest. The exact source locations do not need to be known, but the sources should be spread such that an isotropic illumination is ensured. This requirement can be simplified somewhat if we assume that the lateral variations are smooth and minor in the area between the source and receiver. The result for surface waves is that only the sources inline with the receivers (in a stationary-phase sense) will contribute to the surface-wave retrieval. These assumptions are also made during the processing for MASW, so they do not seem unreasonable.

During an active-source survey, we can exploit this by placing multiple sources in line with the two receivers. Cross-correlating the two traces for every source at a valid position (on the left side of receiver r_1 in Figure 1) results in what is kinematically the same trace each time: a trace with a virtual source at receiver r_1 and recorded at receiver r_2. As such, they can be summed so that the signal adds constructively and random noise destructively. This is the most popular seismic interferometry application and constitutes the first step of SVI, as illustrated in Figure 2a.

The first step of SVI can be repeated for all valid combinations of sources and receivers according to their geometry, and as a result the virtual sources have been shifted to the receiver positions. This is corrected in the second step of SVI. To perform the opposite operation, the original trace recorded at the first receiver is convolved with the boosted trace originating from the first receiver. The effect of this step, which is seismic interferometry by convolution (Slob et al., 2007), shown in Figure 2b, is that the travel paths characterising the two traces are 'stitched' together and the resulting trace reflects the full travel path from the (real) source s_0 to the second receiver r_2. When there are multiple receivers between the original source position and the second receiver, this process can be repeated for each of them, each time resulting in the same trace from source position s_0 recorded at receiver r_2. This means that, again, we can sum the results to boost the signal and reduce the noise. The resulting trace reflects the travel path of the original one recorded at receiver r_2, but the signal-to-noise ratio has been boosted twice proportionally to the square root of the number of shots available (Mallinson et al., 2011).

Previous investigations into surface-wave SVI have shown that the effect of SVI can be complicated and is not always positive (Hassing et al., 2024). The results of SVI can be boosted further by iteratively applying the method (Al-Hagan et al., 2014). For this, the output of an iteration of SVI is fed into the next iteration, which is repeated until the signal-to-noise ratio is considered to have improved sufficiently. As mentioned above, SVI primarily boosts the strongest event found in the data. If the desired event is not the strongest in the recording, the iterative application will only remove the desired signal. Similarly, if SVI leads to unexpected effects for surface waves, an iterative application does not resolve these issues. To address these complications,

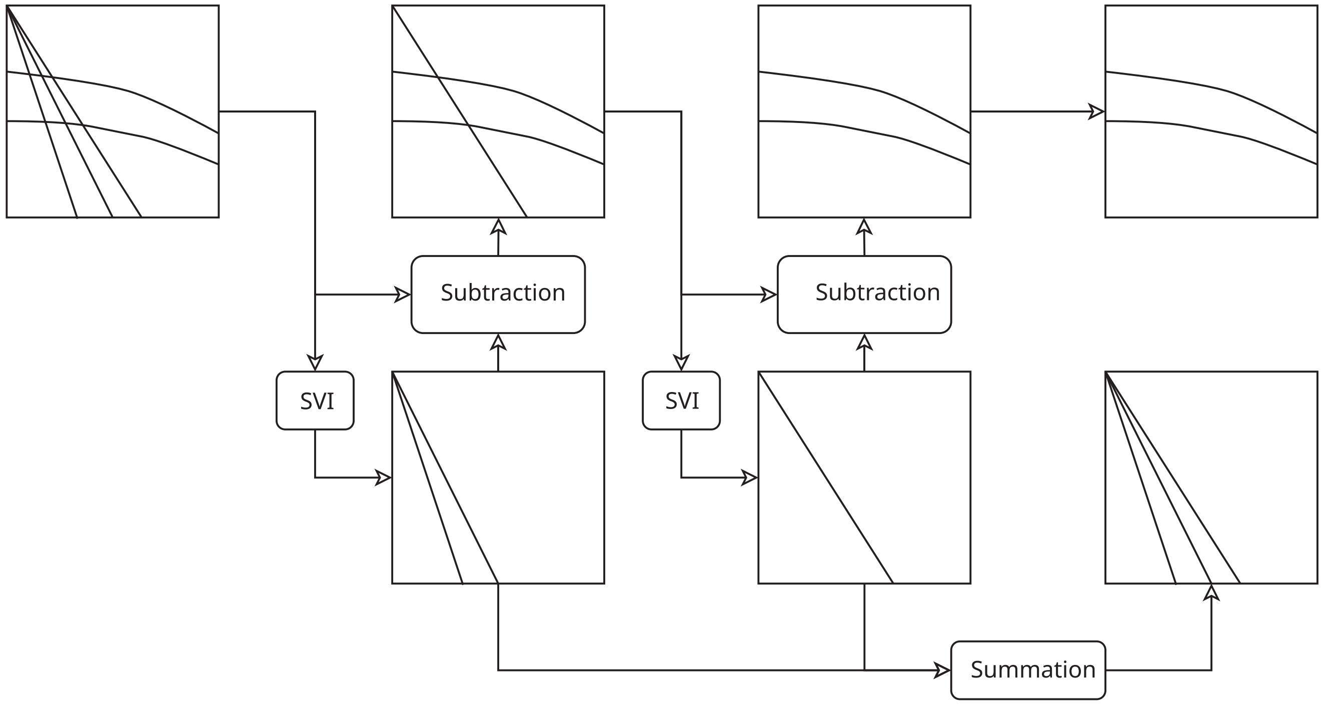

FIGURE 3 | A sketch describing the 'ideal' scenario of applying iterative supervirtual interferometry (SVI). The top-left panel shows the raw data containing both surface-wave arrivals (sketched as straight lines) and reflected arrivals (sketched as hyperbolae). After applying SVI, the fundamental mode is boosted. The boosted data are subtracted from the raw data. The resulting panel is used as input again for SVI. This can be repeated multiple times to extract different strongest events from the data. Finally, these panels can be summed together to obtain a more complete description of the surface waves with a boosted signal-to-noise ratio.

we propose to adjust the technique of iterative SVI when using surface waves.

Instead of considering the surface waves as a single event originating from the source position, it would be more accurate to consider the different modes separately. The different modes propagate at different velocities, and the energy is distributed between them depending on the propagation medium. Often, it is assumed that most of the energy resides in the fundamental mode, but this is not necessarily the case. SVI, then, boosts the strongest event, that is, a higher mode.

The basic workflow of the iterative SVI we propose is as follows and also illustrated in Figure 3. SVI extracts the strongest event. Subtracting this result from the original data provides the remainder of the events with the noise. The amplitudes of the events are not the same after SVI and to correct for this, we apply the subtraction in a least-squares manner, similar to how multiples are removed after their estimation (Verschuur and Berkhout, 1997).

The data without the resulting strongest event can serve as input for SVI, with the new strongest event, for example, a higher mode, being boosted. This can be repeated until the data are separated into different panels containing either the desired events or a panel containing noise and other events. Finally, the panels with desired events can be summed to obtain the output of the process. Effectively, our proposed technique presents a similar workflow to that in An and Hu (2016), but the target is extracting and improving surface waves rather than suppressing them.

Only a certain range of offsets can be included for the SVI steps. Four considerations play a role. Firstly, the signal-to-noise ratio should be sufficient. If the amplitude of the signal has decayed too much, the cross-correlation result will also mostly include noise. At a certain point, stacking these data will not be beneficial. Secondly, the full arrival of the targeted wave should be recorded at both receivers. Thirdly, the source point should be in the stationary-phase region for the selected receiver pair. This can be relevant when the source is not perfectly in line with the receiver. Lastly, a minimum offset should be used to prevent near-field effects from playing a role.

The question is whether the hypothetical separation of modes of iterative SVI will bear out in reality. SVI, generally, does not fully isolate single events after boosting them. This would mean that the residual energy of the dominant event remains in the data that serve as new input for SVI. Thus, it is more likely that iterative SVI will lead to a more complete version of the strongest event remaining in the final data. The intermediate panels should provide information on what parts of the wavefield are boosted each iteration. They could also provide more information about the higher modes. To test this, we will apply iterative SVI to three different datasets. First, we use a synthetic dataset to investigate the general effects of SVI. Then, we use two different field datasets to investigate the effects of the technique we propose in the presence of a stronger second higher mode and cases where the fundamental mode is not the strongest.

3 | Methods

The general structure of the proposed iterative SVI has already been presented above. The actual implementation of SVI is adapted slightly for application to data compared to previous applications of SVI for surface waves. The first adaptation we introduce is that data outside of a certain minimum and maximum offset is not used during the processing. The minimum ensures that no near-field effects are included in the data. The maximum offset cuts off traces where the signal has attenuated too much or the first arrivals are not recorded so that the results are not contaminated with purely random noise or secondary sources.

The second adaptation we introduce is that, instead of using cross-correlation in the first step of SVI, the cross-coherence is used. The cross-coherence reduces the influence of the source signature and generally stabilises the result of SVI (Place et al., 2019). The cross-coherence can be calculated by first dividing each signal by its amplitude in the frequency domain and then performing the cross-correlation as usual.

After obtaining the results of SVI, they are subtracted from the raw data by least-squares subtraction (like adaptive noise or multiple subtraction; Robinson and Treitel, 1980; Verschuur and Berkhout, 1997). This is controlled by two factors. The first is the filter length. The most significant parts of the filter are contained at short lags, while the larger lags generally do not change the result appreciably. We keep the filter length short, at 15 samples, as, based on empirical tests, larger lengths do not improve the result. The second factor is the size of the window in time and space of the input data, within which the filter is kept constant. Smaller windows remove local effects found in the subtraction data more aggressively. The surface waves consist of longer wavelengths and wave trains of arrivals. Smaller windows also capture more noise in the resulting data. For these reasons, it is more advantageous to use no windowing and subtract the global characteristics of the SVI data from the raw data. After subtraction, the result of the subtraction and the boosted panel balanced by the subtraction filter are saved to disk. In this way, the amplitudes of both results are balanced again after they are distorted by SVI.

Then, as explained above, the remainder of the data from the previous step is used as input for the next iteration. The results from each iteration are stored separately so that they can finally be summed together. This is the crucial adaptation for surface waves in our work.

We use two different ways to evaluate the effects of SVI. Firstly, in order to estimate the changes in signal-to-noise ratio, we compare the root-mean-square (RMS) value of different sections of data. We select an offset and time range for windows where noise is found. Then, we also select an offset range for windows with a signal. Because the timing of the surface waves can vary depending on phase velocity and shot positioning, we take the time of maximum amplitude along the trace at the centre of the offset range and set the time range around this time. The ratio between the signal and noise RMS values can then be used to quantitatively assess changes in the signal-to-noise ratio. This is repeated for all shots, providing a distribution of values for the dataset. Because there is a correlation between the signal-to-noise ratio and the recording time, as we will see below, we take a sample of noise both early during the recording and late. The late noise section should be used with the caveat that it is harder to guarantee that it does not include any useful signal.

In order to compare changes in the domain more relevant for surface waves, we also compute dispersion spectra. For this, the total receiver line is subdivided with a moving window. At each window position, the dispersion spectrum is computed for the five closest shots, taking a minimum offset into account, and stacked. To compute the dispersion spectrum, first a (Hamming) windowing function is applied over the spatial dimension and then the frequency-domain beamforming method is used (Zywicki and Rix, 1999, 2005). Dispersion curves can then be picked as the local maxima on these spectra. For the picking of higher modes, it is important to keep the possibility of cross-terms of higher modes into account. For this, the original data are leading when interpreting a dispersive event as a higher mode. Information from laterally close windows can also provide extra information.

Our proposed technique is applied to three different datasets. The first is synthetic and obtained by finite-difference modelling using *fdelmod* (Thorbecke and Draganov, 2011), which uses fourth-order differential operators and a staggered grid to solve, in our case, the elastic wave equation. The velocity model is two-dimensional and consists of two different stacks of layers welded together at the centre, as shown in Figure 4. A constant Poisson's ratio of 0.1 is used, and the density is calculated using Gardner's relation (Gardner et al., 1974). A free surface is set at the top of the model to excite surface waves, while perfectly matched layers with 500 positions are used on all other edges.

The left side of the model contains a high-velocity layer at the surface, so that strong higher modes are excited. In order to predict the dispersion curves and the relative strength of each mode, we model one-dimensional models representing both layer stacks with Thomson–Haskell forward modelling (Thomson, 1950; Haskell, 1953; Buchen and Ben-Hador, 1996). Using the dispersion curves, the eigenfunctions are computed and with that, the spectral excitation of each mode is calculated (Harkrider, 1964, 1970). The results of the theoretical modelling can be seen in Figure 5. On the left side of the model, the energy jumps between different modes: below 15 Hz the fundamental mode is dominant, around 20 Hz, the first higher mode is dominant and around 25 Hz the second higher mode is dominant. On the right side of the model, the excitation of higher modes is very weak. In order to target the frequency range where different modes are dominant, we use a Ricker wavelet with a central frequency of 20 Hz for the modelling. The modelling is done with a model spacing of 0.3 m, a time step of 0.1 ms during 2 s for a total of 51 shots spaced every 10 m. The wavefield is recorded by receivers placed every 2 m between 0 and 500 m with a sampling frequency of 500 Hz.

After the modelling, noise is added to the data. This is done by adding a random value to each data point taken from a Gaussian

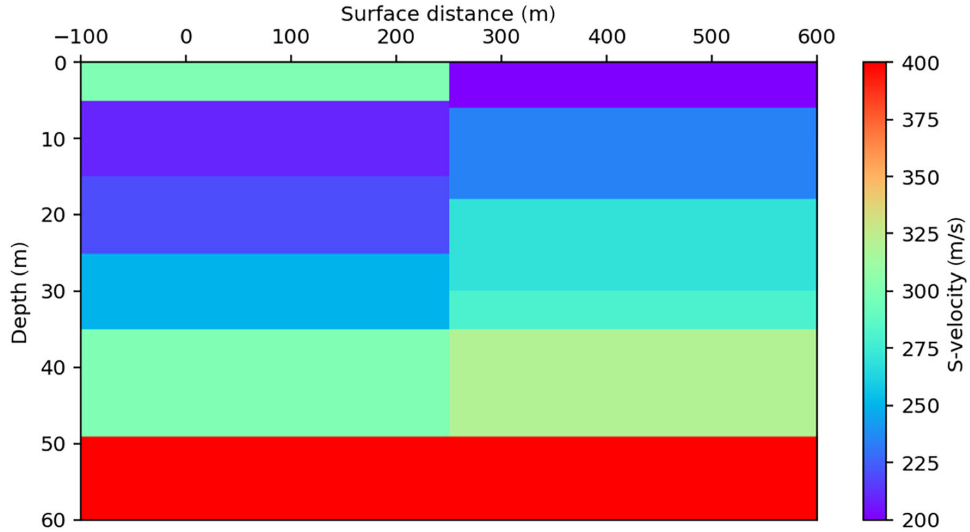

FIGURE 4 | The shear-wave velocity model used in the finite-difference modelling of the synthetic data. It consists of two horizontally layered stacks of homogeneous layers welded together at the centre of the model. The left model contains a high-velocity layer at the surface to create strong higher modes. The ‘halfspace’ extends down to 100 m, but is cut off here to show more detail in the shallow parts. The halfspace velocity is equal throughout the model.

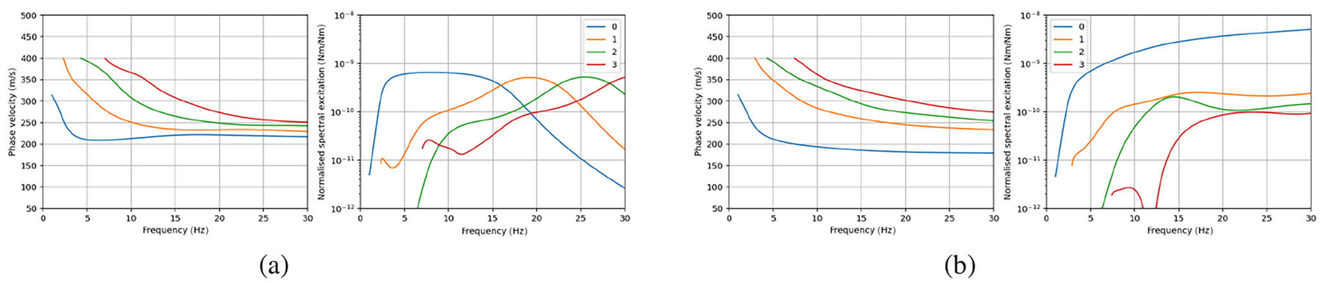

FIGURE 5 | Theoretical dispersion information for the first four modes of the (a) left and (b) right laterally homogeneous regions in the model in Figure 4. In each subfigure, the left panel shows the dispersion curves computed with the Thomson–Haskell propagator method and the right panel shows the relative spectral excitation of different modes.

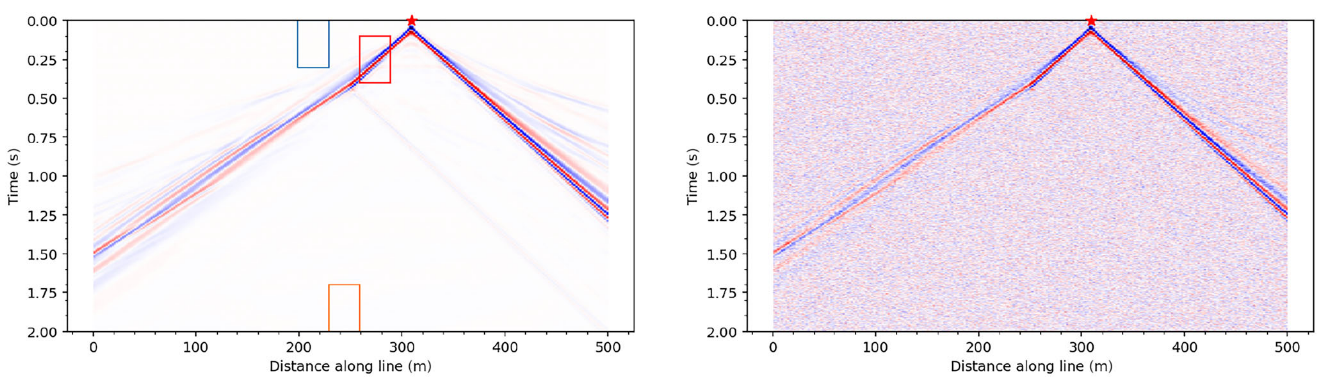

FIGURE 6 | An example of one of the common-shot gathers from the synthetic dataset with a source location at 309 m along the model without noise (left) and with noise added (right). The boxes in the left figure indicate the relative position of the sections of data used to compute the signal-to-noise ratio in Figure 10a. The blue box indicates early noise, the red box signal and the orange box late noise.

distribution with a standard deviation of 10% of the maximum amplitude in that shot. The data resulting from an example common-shot gather with and without noise added are shown in Figure 6. The sections of data used for determining the signal-to-noise ratio of 0.3 s in length and 30 m in width are also shown. The spatial windowing for the computation of dispersion spectra is done with a window length of 100 m, a minimum offset of 150 m and a step of 2 m. The farther offsets are used as higher modes are more apparent there.

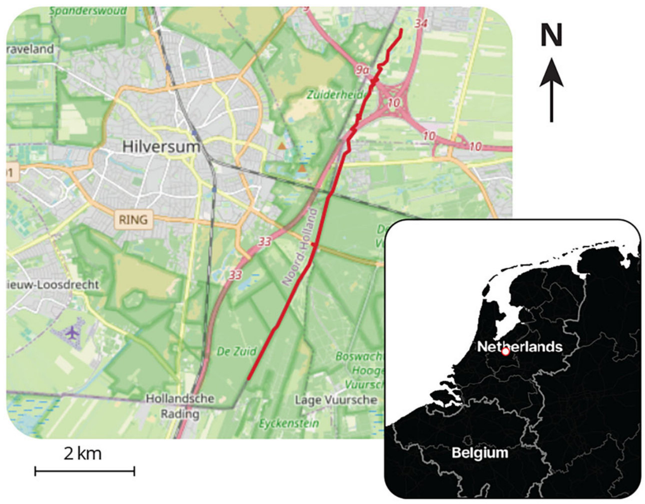

FIGURE 7 | Location of the receiver line used for the SCAN data. Map from OpenStreetMap contributors under ODbL.

The second dataset is a subset of a field dataset. A quarter of the shots from a section of line 001–011 of the so-called SCAN data (Rehling et al., 2023) are taken and downsampled to a sampling rate of 100 Hz to reduce the total amount of data. This results in 302 shots spaced around 25 m apart, recorded by 1500 receivers, whose locations are shown in Figure 7. The line mostly follows a path in the forest along the provincial boundary between North Holland and Utrecht, but crosses two busy highways in the northern part of the selected line. The SCAN project itself focuses on P-wave reflection data for the deep subsurface. As such, there is little information on the shear-wave velocity structure of the near surface. Information from boreholes close to the line shows that the top 20 m of the subsurface consists almost completely of sand with some variation in grain size (TNO - Geologische Dienst Nederland, na). As such, higher modes are not the focus for this dataset.

One of the shots is shown in Figure 8a. The signal-to-noise ratio of the data is generally high and, although some reflections are visible, the surface waves are dominant. Some shots contain strong secondary events. Because some shots are not placed exactly on the receiver line, one of the assumptions made for SVI to work, we set the minimum offset to 500 m. The maximum offset is placed at 2000 m, where the slower, main package of surface-wave energy is no longer recorded. Again, the sections of the data used for determining the signal-to-noise ratio are shown. The spatial windowing is done using windows of 300 m in length with a step of one third of the window length throughout the offset range used for SVI.

The third dataset is more challenging. We shall call it the Farm dataset, recorded onshore in South Dakota, the United States. The site is located inside a glacial valley, and the near surface consists of a layer of glacial sediments overlying limestones. This velocity contrast is likely to excite higher modes, and different dispersive events can be observed in the data. A channel is found between 150 and 400 m along the line. The dataset contains 517 shots interleaved between the geophones, whose spacing is 3 m. A (zoomed-in) example of a shot is shown in Figure 8b. The dataset is noisy, and many shots contain events – refracted and direct waves – that dominate over the surface waves. As discussed above, the SVI boosts the strongest event. In order to ensure that the surface waves are boosted, we apply a top and bottom mute to isolate the surface waves. The minimum offset is set to 12 m, while the maximum offset is set to 250 m, as no signal is visible at further offsets. Due to the muting, comparisons of the signal-to-noise ratio are difficult, as sections with no signal are hard to extract. Therefore, the signal-to-noise ratio will not be computed for the dataset.

4 | Results

An example of results of the different iterations of SVI on a shot from the synthetic data can be seen in Figure 9, with the associated changes in signal-to-noise ratio in Figure 10a. As expected, the first iteration results in a large increase in the signal-to-noise ratio. The random noise is reduced, and the shape of the dispersion curve has become more consistent. An interesting feature to note is that the random noise has obtained a gradient of intensity, increasing with longer traveltimes. As a result, the noise is stronger for late traveltimes than for early traveltimes.

This effect becomes more pronounced in the following iterations, as more of the signal has been extracted. The signal-to-noise ratio decreases significantly with longer traveltimes and is lower than the original signal-to-noise ratio for each separate iteration. The decrease in signal-to-noise ratio of the separate higher iterations is expected, as each consecutive iteration separates increasingly weaker events. The signal-to-noise ratio of the cumulative iterations does show a slight decrease over the iterations, but the signal-to-noise ratio of all iterations combined is still significantly higher than in the original data. As a result of the second iteration, the first higher mode is visible between 12 and 20 Hz and part of the second higher mode is visible between 23 and 30 Hz. The dominant mode is mostly removed after the first iteration, and the second-strongest mode becomes visible. At low frequencies, the fundamental mode is consistently still visible. As more signal is extracted after each iteration and the noise becomes relatively stronger, it becomes harder to accurately extract the available information.

The dispersion curves picked based on the synthetic dataset are shown in Figure 11. After the first iteration, the spread of the picks has decreased. In subsequent iterations, the frequency band in which higher modes are found becomes broader in both parts of the model, but as the signal-to-noise ratio decreases again, the spread of the picks also increases. In the left part of the model (Figure 11a), the parts of the dispersion spectrum where the modes come close to each other become more difficult to separate in the later iterations of SVI, which can be seen as many of the picks for the first higher mode around 17 Hz – where the amplitude of the fundamental mode and the first higher mode is nearly equal in Figure 5a – are located between the analytical fundamental and first higher modes. This could be caused by correlation cross-terms between the two modes.

Now, we turn to the second dataset. The signal-to-noise ratio over different iterations is shown in Figure 10b. Larger samples of noise are taken to average out the secondary events that are active along the recording line. The pattern of the signal-to-noise

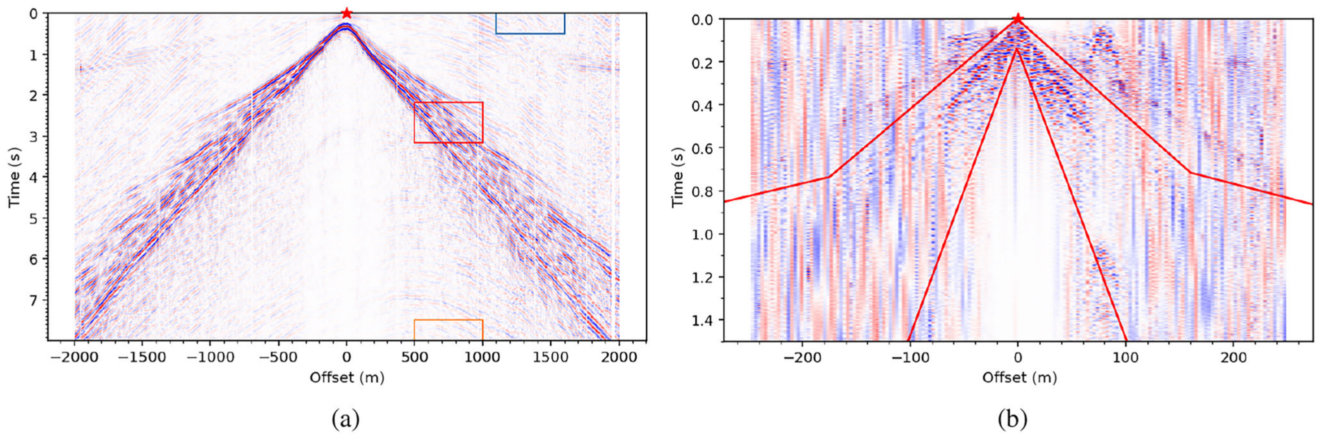

FIGURE 8 | Example common-shot gathers of the two field datasets showing the (a) SCAN data and (b) the Farm data. For visualisation purposes, all traces are normalised by their root-mean-square value and the Farm data are limited to 1.5 s of recording. Like in Figure 6, three boxes indicate the relative positions of the data sections used for computing the signal-to-noise ratio. Each dataset shows different types of noise and different recorded events. The SCAN data is dominated by surface waves, but reflections and secondary events are visible. The Farm data shows more noise in general and both a strong air wave and a refracted wave below the noise. To isolate the surface waves, a top and bottom mute are applied during processing; the boundaries of the top and bottom mute are indicated by the red lines.

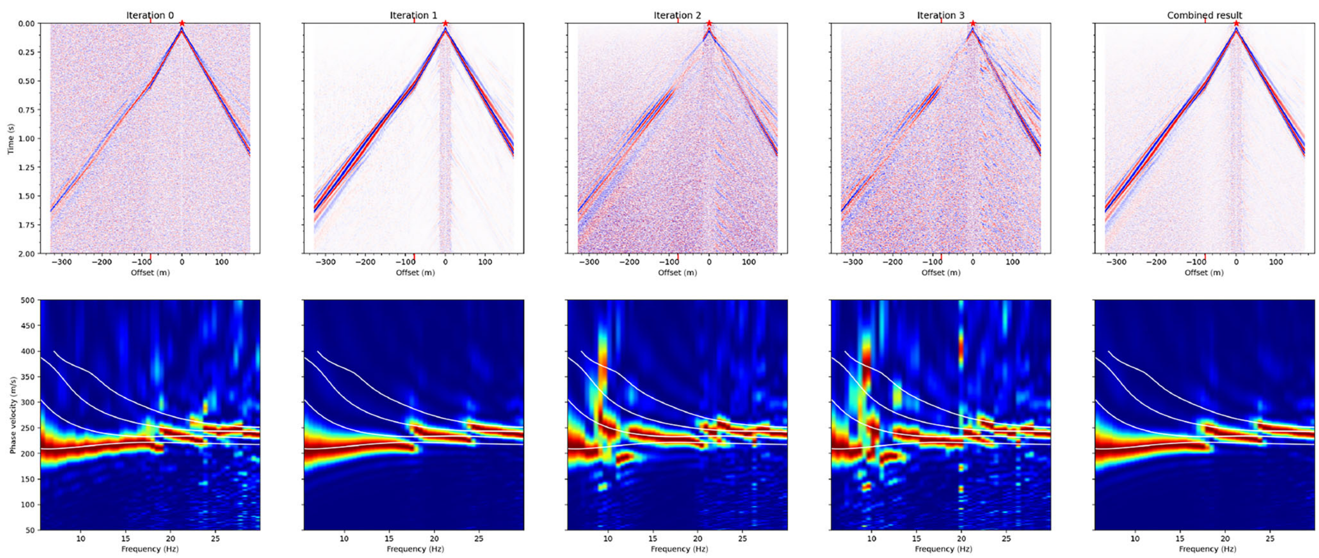

FIGURE 9 | Results of iterative SVI applied to a common-shot gather in the synthetic dataset, where the top row shows the time-domain data. The traces below the minimum offset of 12 m are not included in the processing, but copies of the original data are inserted there to show contrast and consistency throughout the iterations. The red bars above and below the images indicate where the lateral variation is found in the data. The dispersion spectra in the bottom row are computed by taking all traces in the offset range −280 and −180, which corresponds to distances between 50 and 150 m along the receiver line. Before computation of the dispersion spectrum, a Hamming window is applied over the spatial axis and the spectra are normalised per frequency. The white lines show the theoretical dispersion curves from Figure 5a. The first column shows the original data, and the following columns show the results of the iterative SVI after each iteration. Then, the final column shows the sum of the results from each iteration.

ratio over the iterations is similar to the one for the synthetic data, where the first iteration shows a large improvement in signal-to-noise ratio and this decreases over further iterations.

Results of separate common-shot gathers are shown in Figure 12. For these data, it is also worth considering the remaining parts of the data that form the input for the next iteration and are discarded at the end. The noise gradient that was visible in the synthetic data is not prominent here. Instead, spurious (acausal) events are visible, angling away from the source position. These are the strongest after the first iteration. When there are large, random spikes in the original data, these spurious events can have a large amplitude.

In two examples of dispersion spectra for two different spatial windows in Figure 13, we can see that the fundamental mode is primarily boosted after the first iteration of SVI. Further iterations do still include energy from the fundamental mode.

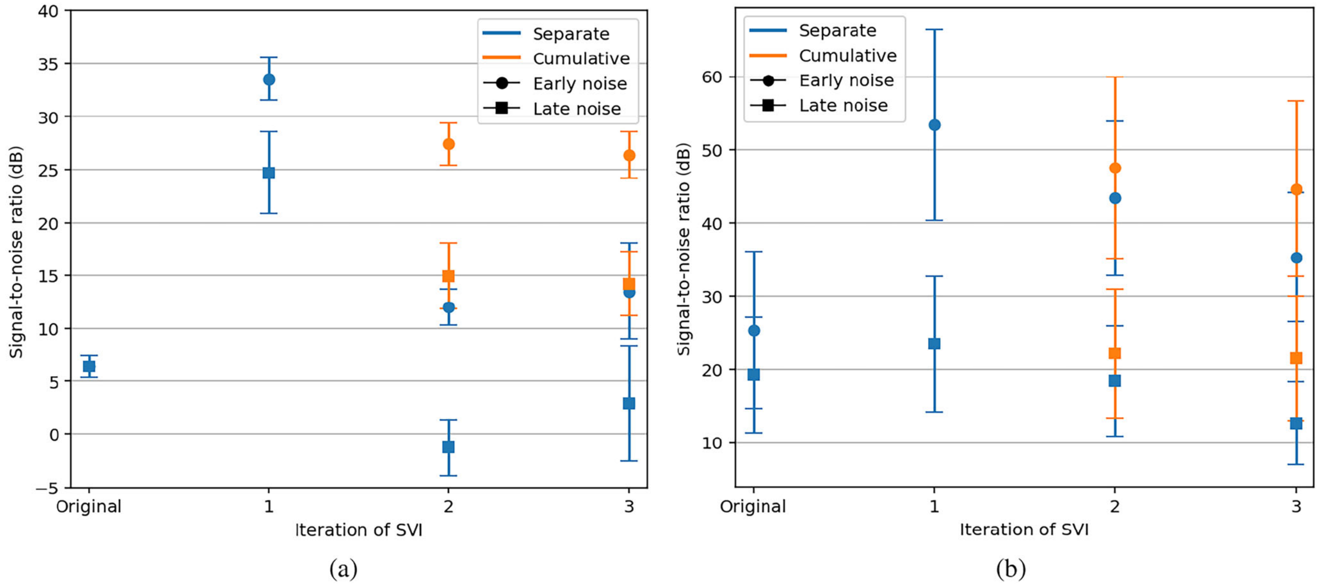

FIGURE 10 | Signal-to-noise ratios of (a) the synthetic dataset and (b) the SCAN dataset throughout each iteration of SVI computed by taking the ratio between the root-mean-square values of small sections of data. Representations of these sections are shown in Figures 6 and 8a. Each time, the red box indicates the signal, the blue box indicates early noise and the orange box indicates late noise. This is computed for every shot, with the location based on the offset. The resulting mean and standard deviation of all signal-to-noise ratios per iteration are plotted here for early (circles) and late (squares) noise. The effect of stacking all previous iterations is also shown in orange.

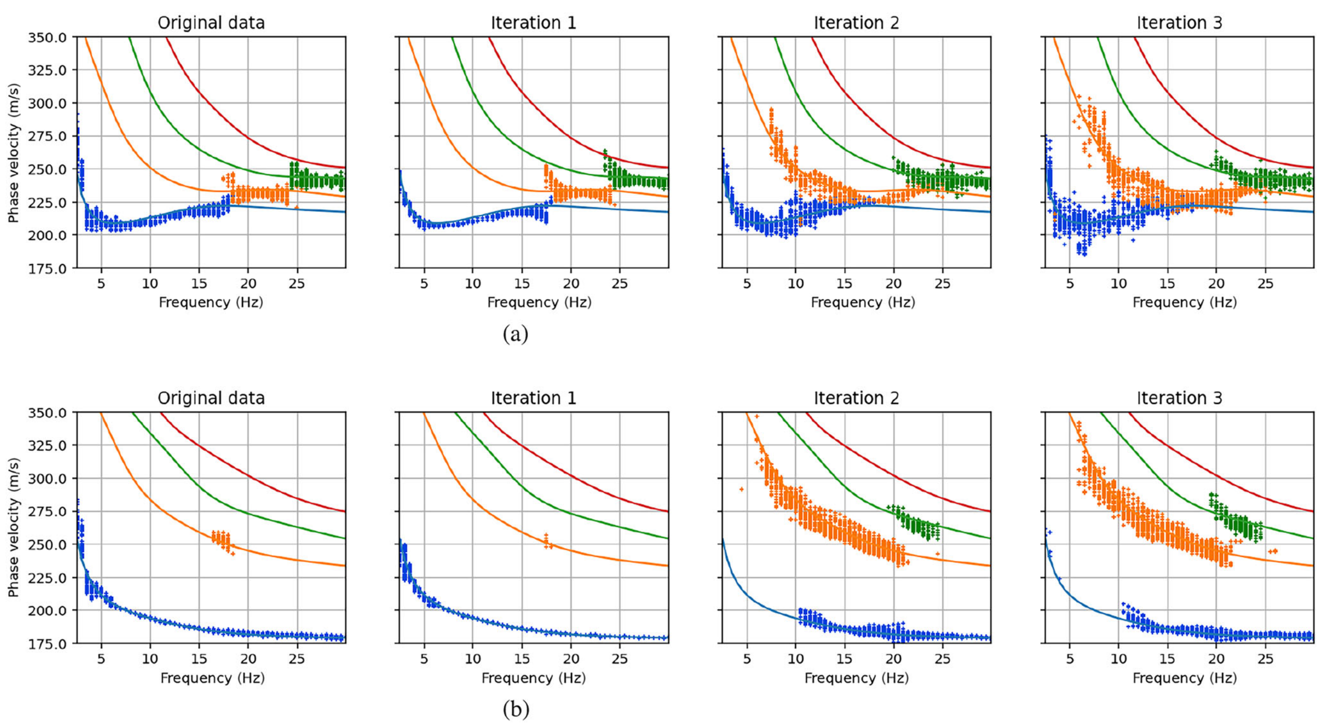

FIGURE 11 | Manually picked dispersion curves for the original data and the results of three iterations of SVI based on the synthetic dataset. The lines indicate the analytical solution to the dispersion curves from Figure 5. The data are split into the two different laterally homogeneous parts of the model with (a) all windows whose centres are below 200 m along the model and (b) all windows above 300 m along the model.

When a second dispersive event is visible, the combination of all iterations shows an increased consistency in the fundamental mode but also retains information on the second dispersive event.

The fundamental mode picked over all spatial windows is shown in Figure 14. The first iteration causes a large increase in the frequency band in which the fundamental mode can be identified. Data become more and more scarce in the higher

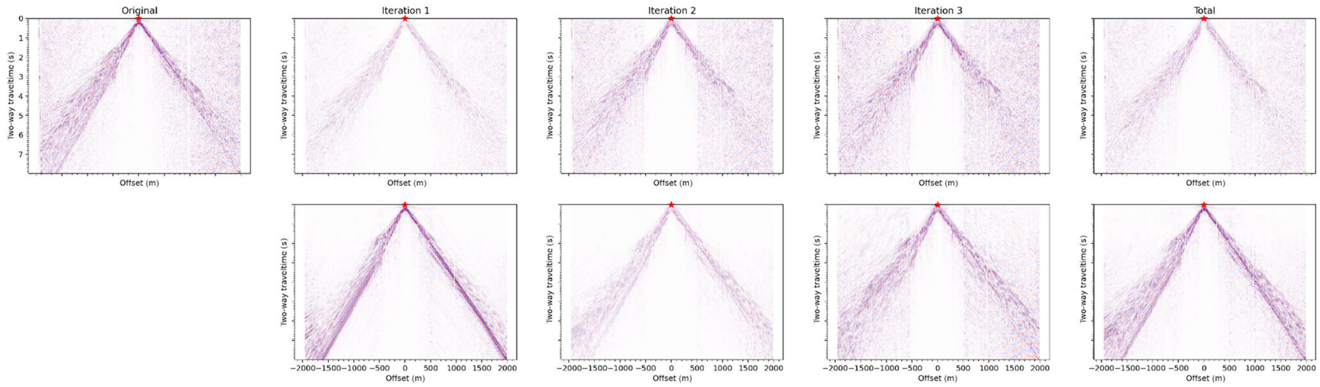

FIGURE 12 | An example of a common-shot gather from the SCAN data throughout the iterative SVI process. Following the illustrated flow in Figure 3, the leftmost panel shows the original gather and each subsequent panel shows the result of SVI after each iteration, with the bottom row containing the result of SVI and the top row the data remainder (the input for the following iteration). The bottom right panel shows the sum of the results from the three iterations, while the top right shows the final remainder. Again, like in Figure 9, the original traces are inserted below the minimum offset of 500 m.

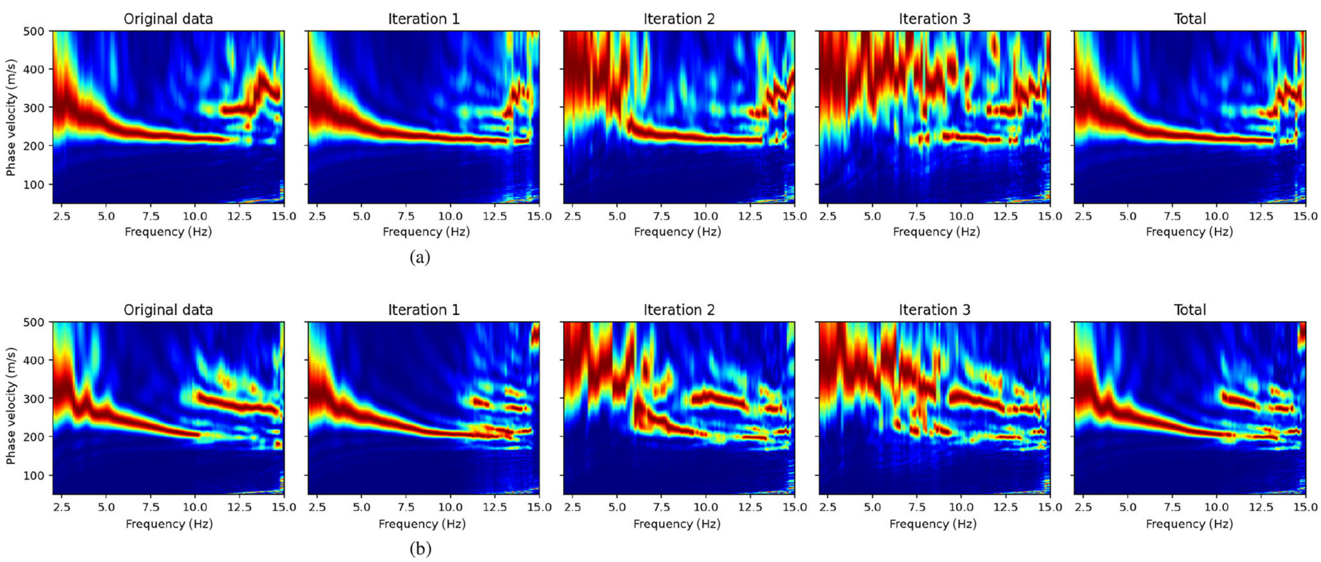

FIGURE 13 | The dispersion spectra computed from stacked spectra from different spatial windows along the receiver line. In (a), the fundamental mode is boosted after the first iteration of SVI. Later iterations primarily retrieve more of the fundamental mode. In (b), a second dispersive event that appears to be a higher mode is visible. The later iterations do recover more of this event.

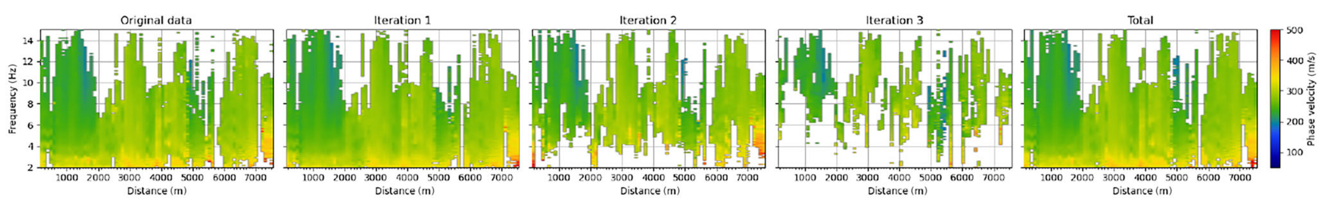

FIGURE 14 | The fundamental mode picked for each iteration of SVI for the SCAN dataset.

iterations, but the higher iterations still do include information on the fundamental mode. As such, the combined result includes more information on the fundamental mode, but also more noise.

Figure 15 shows two examples of the effect of iterative SVI on the Farm dataset. Generally, as before, the dispersion curve is more consistent after application of iterative SVI, but in some cases the higher-frequency tail of the fundamental mode disappears in the final result, as can be seen in Figure 15b. However, other shots that were unusable due to the presence of secondary events became usable shots in the final result, such as in Figure 15a.

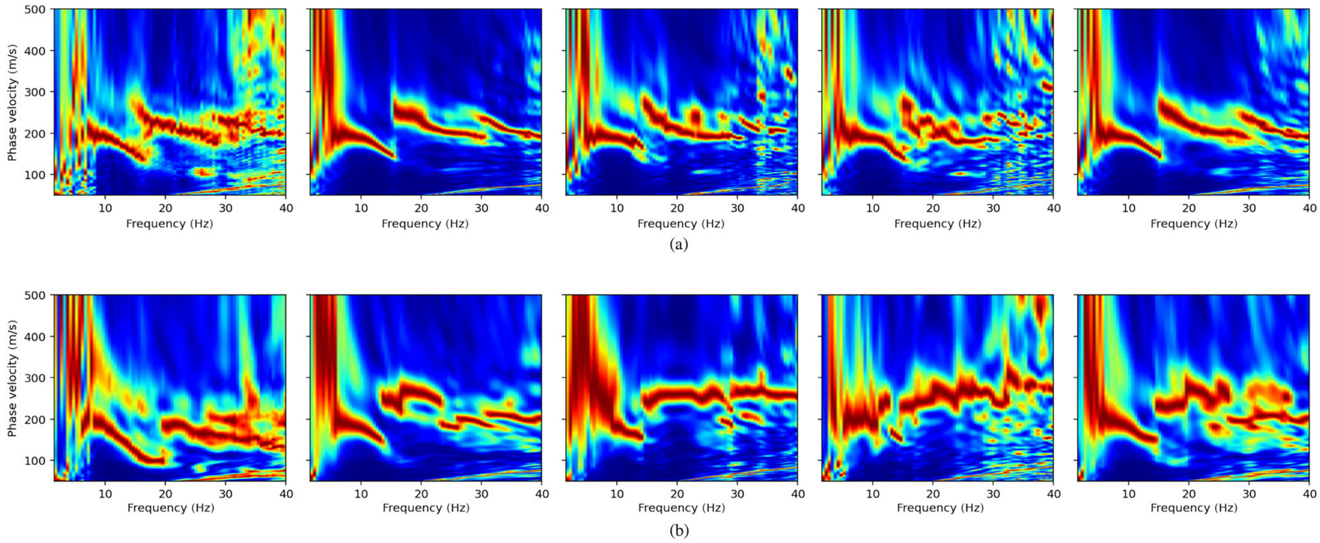

FIGURE 15 | The stacked dispersion spectra resulting from iterative SVI from two spatial windows of the Farm dataset. (a) At 650 m along the line, a shot where a secondary event was active distorts the initial spectrum, but this is restored after SVI is applied. (b) The fundamental mode is harder to distinguish above 15 Hz after the application of iterative SVI.

The dispersion curves picked based on spatial windowing of the dataset are shown in Figure 16. Two zones with low phase velocities are visible between 200 and 250 m, which coincides with the known location of a channel, and between 900 and 1050 m along the line. Where before, the higher modes were able to be picked in a larger frequency band in later iterations of SVI, this is not the case here. The amount of data that could be picked decreases over each iteration. Instead, the combination of all iterations yields the largest amount of data.

5 | Discussion

Iterative SVI shows clear improvements over a single iteration of SVI when the data contains multiple different events. For surface waves, it is important to consider the different modes separately. When the fundamental mode is already dominant, a single iteration improves the signal-to-noise ratio and generally increases the frequency band where the mode can be found. Further iterations still retrieve energy from the fundamental mode. There is a balance where more energy of the event can be retrieved from further iterations but also when more noise is included.

When there is a strong higher mode, the higher mode can be recovered better when the dominant parts of the fundamental mode have been separated in the first iteration. Generally, the first higher mode shows up more clearly in the result of the second iteration, but remnants of the fundamental mode are still found because the SVI cannot fully isolate single events. Including further iterations in the final result gives a more complete representation of the surface waves and increases the signal-to-noise ratio. Treating the results of the iterations separately in some cases makes it possible to obtain a more complete representation of the higher mode. This could be used for picking the dispersion curve of a higher mode. When more random noise is present, SVI struggles to accurately retrieve the dominant event. Then, it is not accurately removed from the data, and it is not possible to extract other events in further iterations.

Even though the iterative method can handle data containing multiple events better. It is still advisable, if possible, to try to isolate the desired parts of the wavefield. This prevents crosstalk and makes sure that the event that is boosted is actually the one that is targeted. A disadvantage of seismic interferometry is that this crosstalk can also happen between the higher modes. Identification of a true mode can be made more secure by comparison with the original data and laterally close data.

When the data contain large, random spikes, these will generate spurious events sloping away from the source position even after a single iteration. These can be prevented by cleaning the data before processing. They also provide a clue to the presence of the 'noise gradient' found in the synthetic data. The Gaussian noise acts as a large number of random spikes in the data. Each of them generates a small spurious event that extends towards positive traveltimes. The random amplitude of the noise means that the summation of all the events looks like random noise, but the density of events is a function of the number of data points at earlier traveltimes. Of course, the data are sampled equidistantly in time, meaning that the level of noise increases linearly with time. It is possible that by including shots that would generate a signal at acausal times, this effect would be decreased.

There are some apparent events in the original data that seem to disappear after the application of SVI, for example, Figure 15b. According to the theory, all events originating from the same stationary-phase region should stack constructively. As these events are also what we are interested in, any event that disappears should be looked at with a critical eye. The original event could simply be random noise or it could be constructive noise from an undesired source. Generally, it would not be advisable to use these events to pick dispersion curves.

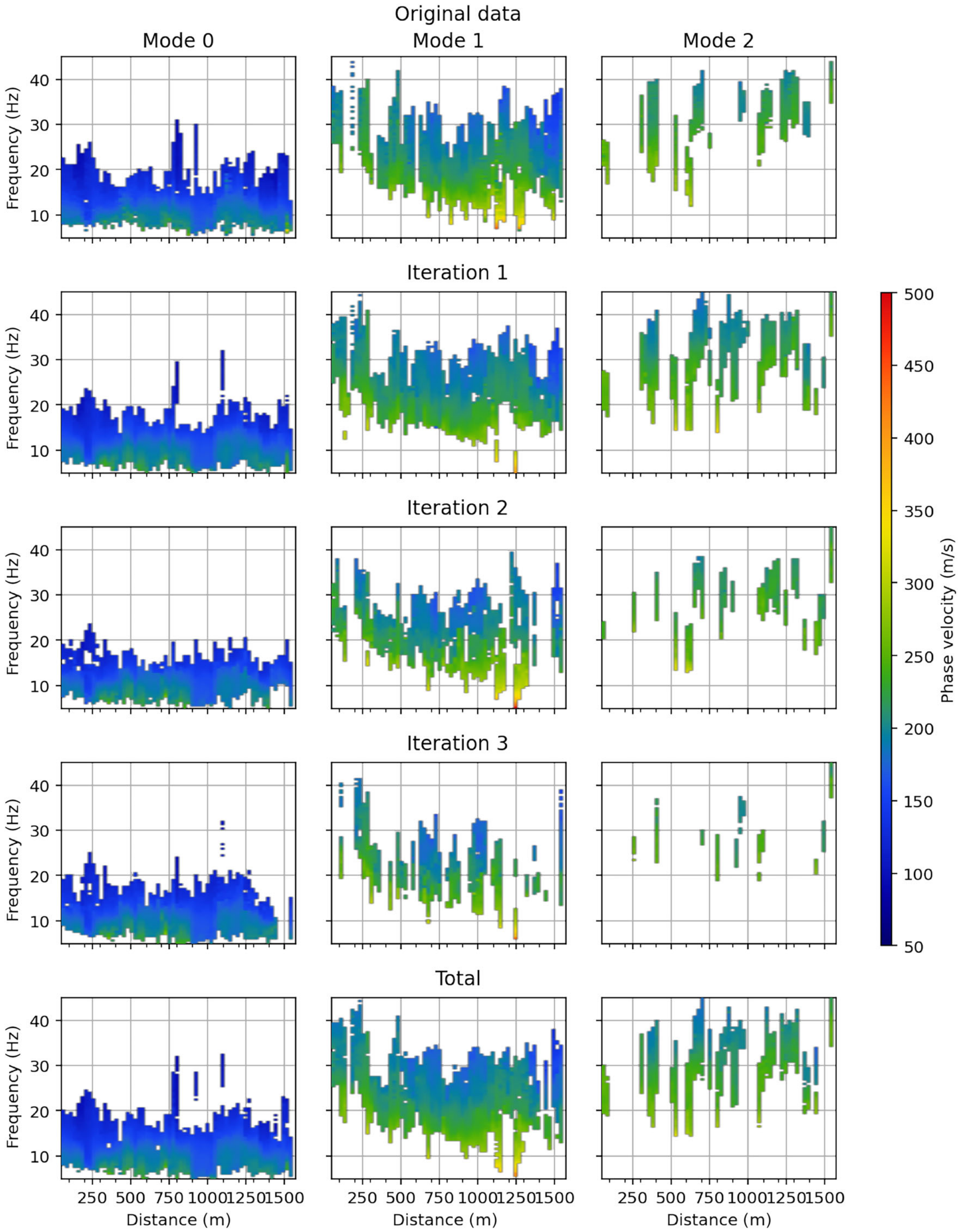

FIGURE 16 | The first three modes manually picked for all spatial windows of the Farm dataset and for each iteration. Here, the later iterations show little improvement in identifying the higher modes compared to the data of the combined iterations.

Our results show that a single iteration does not extract a complete single mode. However, the different iterations do show potential in highlighting different events, making it possible to extract more information from the different events. As it is generally difficult to get information on higher modes from surface-wave data, this could prove to be a valuable addition to processing workflows. In nearly all cases, the signal-to-noise ratio and the consistency of the dispersion curves improved after application of the iterative SVI we proposed. In some cases, there is a definite positive effect of the further iterations on the data, while in other cases the further iterations did not decrease the quality of the data. In the few cases where the fundamental mode disappeared at higher frequencies, this already happened after a single iteration of SVI.

6 | Conclusions

SVI is a method used to increase the signal-to-noise ratio of the strongest event recorded in seismic data. Most literature discusses application to refraction data, where the refracted wave needs to be isolated to ensure that the right part of the data is boosted. For application to surface-wave data, the standard application should be rethought.

Instead of treating surface-wave arrivals as a single event, each mode should be treated separately. Indeed, when applying SVI to surface-wave data containing a weaker higher mode, it can be observed that only the fundamental mode is boosted and the higher mode mostly disappears from the data. This could be corrected by an iterative application of SVI. To improve the result of iterative SVI for the retrieval of surface waves, we proposed three adaptations to the technique.

The methodological improvement we proposed is that instead of feeding the result of one iteration back into the next, we work with the remainder of the data. This means that the output of SVI is subtracted in a least-squares manner from the input data. The result of this is then used as input for the next iteration. We then saved the results of each iteration separately and summed them together in the end. The other two adaptations we proposed are to use only offsets between a data-dependent minimum and maximum offset before SVI application and to use cross-coherence instead of cross-correlation in the first step of an SVI iteration. We applied the adapted iterative SVI we proposed to three different datasets.

Firstly, we used a synthetic dataset containing Gaussian noise, where strong higher modes were induced. Multiple iterations of SVI showed that in the first iteration the dominant event was boosted, whether the fundamental mode or a higher mode. Later iterations allowed for the higher modes to be picked at a wider frequency band. The Gaussian noise also shifted after the application of SVI to increase with a linear gradient.

Secondly, we used field data with little noise and dominated by surface waves. We showed that the iterative application of SVI increased the signal-to-noise ratio and allowed the fundamental mode to be picked at a wider frequency band.

Thirdly, we used field data that contained strong direct waves other than surface waves and more noise. We showed that iterative SVI cannot help for all locations in this dataset. While some of the spectra lost information about the fundamental mode at higher frequencies. Other shots that were unusable in their original condition due to secondary events were corrected by the iterative SVI, which balances the results. The combination of all iterations provided the largest amount of data.

Thus, we showed that iterative SVI for surface waves can restore a more complete surface-wave wavefield, compared to a single iteration of SVI – it increases the signal-to-noise ratio and generally has a positive, or at least neutral, effect on the quality of the data.

Acknowledgements

We thank Near Surface Observations LLC for providing the Farm dataset, EnergieBeheer Nederland (EBN) for providing the SCAN data and the sponsors of the Delphi Consortium for their support.

Data Availability Statement

The Python code used for processing and the synthetic and SCAN datasets are available upon request, based on permission from the authors.

References

Al-Hagan, O., S. M. Hanafy, and G. T. Schuster. 2014. "Iterative Supervirtual Refraction Interferometry." *Geophysics* 79, no. 3: Q21–Q30.

An, S., and T. Hu. 2016. "Suppression of Seismic Surface Waves Based on Adaptive Weighted Super-Virtual Interferometry." *Science China Earth Sciences* 59, no. 11: 2179–2188.

Bharadwaj, P., G. Schuster, I. Mallinson, and W. Dai. 2012. "Theory of Supervirtual Refraction Interferometry: Theory of SVI." *Geophysical Journal International* 188, no. 1: 263–273.

Boiero, D., and L. V. Socco. 2011. "The Meaning of Surface Wave Dispersion Curves in Weakly Laterally Varying Structures." *Near Surface Geophysics* 9, no. 6: 561–570.

Buchen, P. W., and R. Ben-Hador. 1996. "Free-Mode Surface-Wave Computations." *Geophysical Journal International* 124, no. 3: 869–887.

Gardner, G. H. F., L. W. Gardner, and A. R. Gregory. 1974. "Formation Velocity and Density: The Diagnostic Basics for Stratigraphic Traps." *Geophysics* 39, no. 6: 770–780.

Ghose, R., and J. Goudswaard. 2004. "Integrating S-Wave Seismic-Reflection Data and Cone Penetration Test Data Using a Multiangle Multiscale Approach." *Geophysics* 69, no. 2: 440–459.

Halliday, D., and A. Curtis. 2008. "Seismic Interferometry, Surface Waves and Source Distribution." *Geophysical Journal International* 175, no. 3: 1067–1087.

G Harkrider, D. 1964. "Surface Waves in Multilayered Elastic Media I. Rayleigh and Love Waves From Buried Sources in a Multilayered Elastic Half-Space." *Bulletin of the Seismological Society of America* 54, no. 2: 627–679.

G Harkrider, D. 1970. "Surface Waves in Multilayered Elastic Media. Part II. Higher Mode Spectra and Spectral Ratios From Point Sources in Plane Layered Earth Models." *Bulletin of the Seismological Society of America* 60, no. 6: 1937–1987.

A Haskell, N. 1953. "The Dispersion of Surface Waves on Multilayered Media." *Bulletin of the Seismological Society of America* 43, no. 1: 86–103.

Hassing, S., D. Draganov, and E. Verschuur. 2024. "Surface-Wave Supervirtual Seismic Interferometry: The Ugly, the Bad, and the Good." In *NSG 2024 30th European Meeting of Environmental and Engineering Geophysics*, Volume 2024, Helsinki, 1–5. European Association of Geoscientists & Engineers.

Khosro, Anjom, F., T. J. Browaeys, and L. V. Socco. 2021. "Multimodal Surface-Wave Tomography to Obtain S- and P-Wave Velocities Applied to the Recordings of Unmanned Aerial Vehicle Deployed Sensors." *Geophysics* 86, no. 4: R399–R412.

Kimman, W. P., and J. Trampert. 2010. "Approximations in Seismic Interferometry and Their Effects on Surface Waves: Approximations in Seismic Interferometry." *Geophysical Journal International* 182, no. 1: 461–476.

Luo, Y., J. Xia, J. Liu, Q. Liu, and S. Xu. 2007. "Joint Inversion of High-Frequency Surface Waves With Fundamental and Higher Modes." *Journal of Applied Geophysics* 62, no. 4: 375–384.

Mallinson, I., P. Bharadwaj, G. Schuster, and H. Jakubowicz. 2011. "Enhanced Refractor Imaging by Supervirtual Interferometry." *The Leading Edge* 30, no. 5: 546–550.

Nazarian, S., and K. H. Stokoe II. 1985. "In Situ Determination of Elastic Moduli of Pavement Systems by Spectral-Analysis-of-Surface-Waves Method: Practical Aspects." Technical Report 0-368-1F, Center for Transportation Research, University of Texas at Austin, Austin, TX.

Park, C. B., R. D. Miller, and J. Xia. 1998. "Imaging Dispersion Curves of Surface Waves on Multi-Channel Record." In *SEG Technical Program Expanded Abstracts 1998*, 1377–1380. Society of Exploration Geophysicists.

Park, C. B., R. D. Miller, and J. Xia. 1999. "Multichannel Analysis of Surface Waves." *Geophysics* 64, no. 3: 800–808.

Place, J., D. Draganov, A. Malehmir, C. Juhlin, and C. Wijns. 2019. "Crosscoherence-Based Interferometry for the Retrieval of First Arrivals and Subsequent Tomographic Imaging of Differential Weathering." *Geophysics* 84, no. 4: Q37–Q48.

Rehling, J., G. Zaalberg-Metselaar, R. Peenstra, J. Van Den Akker, and M. Ter Borgh. 2023. "The Dutch SCAN Geothermal Exploration Programme — The first three years of 2D seismic acquisition and (re)processing." *First Break* 41, no. 1: 33–43.

Robinson, E., and S. Treitel. 1980. "Maximum Entropy and the Relationship of the Partial Autocorrelation to the Reflection Coefficients of a Layered System." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28, no. 2: 224–235.

Slob, E., D. Draganov, and K. Wapenaar. 2007. "Interferometric Electromagnetic Green's Functions Representations Using Propagation Invariants." *Geophysical Journal International* 169, no. 1: 60–80.

Socco, L., and C. Strobbia. 2004. "Surface-Wave Method for Near-Surface Characterization: A Tutorial." *Near Surface Geophysics* 2, no. 4: 165–185.

Socco, L. V., S. Foti, and D. Boiero. 2010. "Surface-Wave Analysis for Building Near-Surface Velocity Models — Established Approaches and New Perspectives." *Geophysics* 75, no. 5: 75A83–75A102.

T Thomson, W. 1950. "Transmission of Elastic Waves Through a Stratified Solid Medium." *Journal of Applied Physics* 21, no. 2: 89–93.

Thorbecke, J. W., and D. Draganov. 2011. "Finite-Difference Modeling Experiments for Seismic Interferometry." *Geophysics* 76, no. 6: H1–H18.

TNO - Geologische Dienst Nederland. n.a. "Subsurface data | DINOloket." https://www.dinoloket.nl/en/subsurface-data.

Verschuur, D. J., and A. J. Berkhout. 1997. "Estimation of Multiple Scattering by Iterative Inversion; Part II, Practical Aspects and Examples." *Geophysics* 62, no. 5: 1596–1611.

Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel. 2010. "Tutorial on Seismic Interferometry: Part 1 — Basic Principles and Applications." *Geophysics* 75, no. 5: 75A195–75A209.

Wapenaar, K., and J. Fokkema. 2006. "Green's Function Representations for Seismic Interferometry." *Geophysics* 71, no. 4: SI33–SI46.

Xia, J., R. D. Miller, C. B. Park, and G. Tian. 2003. "Inversion of High Frequency Surface Waves With Fundamental and Higher Modes." *Journal of Applied Geophysics* 52, no. 1: 45–57.

Xu, Z., C. Juhlin, L. Han, and F. Zhang. 2017. "Reconstruction of Surface Waves Using Super Virtual Interferometry: An Example From the Forsmark Site, Sweden." *Geophysical Journal International* 211, no. 1: 120–140.

Zywicki, D. J., and G. J. Rix. 1999. "Frequency-Wavenumber Analysis of Passive Surface Waves." In *12th EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems*, Atlanta, United States, 75–84. European Association of Geoscientists & Engineers.

Zywicki, D. J., and G. J. Rix. 2005. "Mitigation of Near-Field Effects for Seismic Surface Wave Velocity Estimation with Cylindrical Beamformers." *Journal of Geotechnical and Geoenvironmental Engineering* 131, no. 8: 970–977.