Ground roll extraction using the Karhunen-Loeve transform in the time-frequency domain

Article *in* Geophysics · December 2021

DOI: 10.1190/geo2021-0453.1

1 author:

Alexandr Serdyukov
A.A. Trofimuk Institute of Petroleum Geology and Geophysics SB RAS

Ground roll extraction using the Karhunen-Loeve transform in the time-frequency domain

Aleksander S. Serdyukov[123]

[1] *Trofimuk Institute of Petroleum Geology and Geophysics,*

2 Koptuga ave.,

Novosibirsk 630090, Russia

[2] *N.A. Chinakal Institute of Mining,*

54 Krasny ave.

Novosibirsk 630091, Russia

[3] *Novosibirsk State University,*

1 Pirogova str.

Novosibirsk 630090, Russia

(December 14, 2021)

Running head: **SKL transform**

ABSTRACT

Ground roll suppression is critical for seismic reflection data processing. Many standard methods, i.e., FK filtering, fail when spatially aliased surface wave interference is present in the data. Spatial aliasing is a common problem; receiver spacing is often not dense enough to extract wavenumbers of low-velocity surface waves. It has long been known that the Karhunen-Loeve transform can be used to suppress aliased ground roll. However, the ground roll should be flattened before suppression, which is challenging due to the dispersion of surface wave velocities. I propose to solve this problem via the time-frequency domain.

I apply the S-transform, which was previously shown to perform well in the multichannel analysis of surface waves. A simple complex-valued constant phase shift is a suitable model of surface wave propagation in common-frequency S-transform gathers. Therefore, it is easy to flatten the corresponding S-transform narrow-band frequency surface wave packet and extract it from the data by principal component analysis of the corresponding complex-valued data-covariance matrix. As the result, the proposed S-transform Karhunen-Loeve (SKL) method filters the aliased ground roll without damaging the reflection amplitudes. The advantages of SKL filtering have been confirmed by synthetic- and field-data processing.

INTRODUCTION

Ground roll is the main type of coherent noise in land seismic exploration data. The most common part of the ground roll in vertical component seismic records consists of Rayleigh surface waves. Although high-pass (HP) filtering suppresses the surface waves, it also rejects low-frequency components of reflected signals, which are critical for impedance estimation. In this case, f-k fan filters can effectively remove unaliased ground roll; however, neither f-k filtering nor τ-p filtering can eliminate the spatial aliasing of surface waves. When the amplitude of the ground roll is much stronger than the reflection signals, f-k filters cause severe signal distortion.

Many estimation-based methods were applied in order to enhance ground roll suppression results, including wavelet (Chen et al., 2017) and curvelet (Liu et al., 2018) transforms, lattice filter (Saatcilar and Canitez, 1994), Wiener filtering (Karslı and Bayrak, 2004; Karsli and Bayrak, 2008) and other approaches. Each technique has advantages and disadvantages and is discussed. Development of the ground roll suppression methods that do not harm the reflection energy remains relevant. Here I address the Karhunen-Loeve (KL) transform technique, which is well known and has many applications in seismic data processing and is similar to another popular approach - singular value decomposition filtering (see e.g. (Possidonio and Porsani, 2021)).

Liu (1999) used the KL transform to suppress surface waves, which eliminated spatially aliased noise. The KL transform has been applied to seismic signals for years. Hemon and Mace (1978) used the zero-lag KL transform to enhance horizontal events. Jones and Levy (1987) introduced a lag in the cross-correlation to enhance linear dipping events with the KL transform. Liu (1999) proposed flattening the ground roll using an alignment function, which was selected picked for each 2D common-shot gather. He then modeled ground roll and subtracted it from the original data. However, the ground roll is dispersive, which makes flattening a labor-intensive process. Verma et al. (2016) proposed a more advanced ground roll suppression workflow using a 3D multiwindow KL filter. This method was developed for 3D data processing and comprises a number of nonintuitive steps. It does not make sense to use such ground roll suppression for linear move-out common-shot gathers. Following Liu (1999), I suppress the ground roll from the common shot gathers using the KL transform in the time-frequency domain. The transformation of ground roll data into the time-frequency domain is a proper approach to handle surface wave dispersion.

I use the S-transform (ST) (Stockwell et al., 1996), which has already been applied for surface wave analysis and attenuation in a number of studies (Askari and Siahkoohi, 2008; Askari and Ferguson, 2012; Serdyukov et al., 2019, 2021). Following Serdyukov et al. (2019), I consider a complex-valued ST of a shot gather. The extension of the KL transform to complex seismic signals was introduced by Levy et al. (1983). For a common-frequency ST gather, the surface waves can be considered linear dipping events (assuming a horizontally layered near-surface region). To extract these common-frequency dipping events, I use a slant-KL transform (Jones and Levy, 1987). This transform is based on an adjustment of a data covariance matrix so that it represents time lags, corresponding to some dip angle (slant). I review various slants and select the one that provides the largest value of the first eigenvalue of the adjusted (time-lagged) common-frequency data covariance matrix. This maximizes the energy of the first principal component of the common-frequency data, which is then used to model the extracted surface wave.

METHOD

ST constant phase shift (CPS) model of surface wave propagation

The ST of a signal $h(t)$ is given by:

$$S[h(t)](\tau, f) = \int_{-\infty}^{+\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi f t} dt. \quad (1)$$

The signal $h(t)$ can be reconstructed from its ST (see Schimmel and Gallart (2005) for details):

$$h(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S(\tau, f) e^{+i2\pi f t} d\tau df = \int_{-\infty}^{+\infty} \frac{\sqrt{2\pi}}{|f|} S(t, f) e^{+i2\pi f t} df. \quad (2)$$

Let us observe a single mode of the surface wave. Consider two signals $h_1(t)$ and $h_2(t)$, which are recorded by two receivers of a linear acquisition system (both receivers are located on the same source-receiver line). The ST of the second signal can be expressed in terms of the ST of the first signal (see Askari and Ferguson (2012) for details):

$$S\left[h_2(t)\right](\tau, f) = e^{-i2\pi k(f) l} e^{-\lambda(f) l} S\left[h_1(t)\right]\left(\tau - k'(f) l, f\right), \quad (3)$$

where $k(f)$ is the wavenumber of the surface wave, $\lambda(f)$ is a frequency-dependent attenuation parameter, and l is the distance between the two receivers.

Let us apply the ST to each trace of a shot gather $g(x, t)$ and cut off the frequency axis at each frequency f. Thus, for each common-frequency gather, we obtain a series of 2D complex-valued functions of the time and distance:

$$g^f(x, \tau) = S[g(x, t)](\tau, f, x), \quad (4)$$

which are called *pseudoseismograms*.

According to (3), for time-distance coordinates of a common-frequency pseudoseismogram, the surface-wave packet is expanded along a slanted line with constant phase velocity (i.e., with the same phase shift from trace to trace). To mitigate the influence of the attenuation factor and geometrical spreading factor (surface waves are cylindrical), I consider the normalized pseudoseismograms:

$$\tilde{g}^f(x,\tau) = \frac{g^f(x,\tau)}{\max\limits_{\tau} g^f(x,\tau)}, \tag{5}$$

The model of a signal composed of time-displaced, phase-shifted replications of a source wavelet (the CPS model) was considered by Levy and Oldenburg (1982). Levy et al. (1983) then applied a complex-valued KL transform to accommodate the CPS model. I apply a similar method to extract the surface wave components from the normalized pseudoseismograms (5).

Slant-KL transform

For a number of receivers $x_i, \quad i = 1, \ldots, M$, let us consider a set of normalized pseudoseismogram (5) complex-valued traces: $g_i^f(\tau) = \tilde{g}^f(x_i, \tau)$. Following Jones and Levy (1987), I introduce a time lag Δ, which is expressed in time steps per trace, and produce a set of modified signals:

$$y_i(\tau) = \delta\left(\tau - \Delta(i-1)\right) * g_i^f(\tau), \tag{6}$$

where δ is the Dirac delta function and '$*$' denotes convolution. I consider the complex-valued data $[M \times N]$ matrix $\mathbf{Y}$, with the time-sampled signals y_i from (6) as the rows:

$$\mathbf{Y} = \{y_i(\tau_j), \quad i = 1, \ldots, M, \quad j = 1, \ldots, N\} \tag{7}$$

where N is the number of time samples τ_j (in the ST domain) and M is the number of receivers. The slant-KL transform (Jones and Levy, 1987) is expressed in terms of normalized eigenvectors of the $[M \times M]$ covariance matrix $\mathbf{\Gamma}$:

$$\mathbf{\Gamma} = \mathbf{Y}\mathbf{Y}^T = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^T, \tag{8}$$

where 'T' denotes the complex-conjugate transpose, $\mathbf{R}$ is the $[M \times M]$ unitary matrix of column eigenvectors, and the diagonal matrix Λ contains the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$, arranged in decreasing size. The rows ψ_j of the $[M \times N]$ matrix Ψ

$$\mathbf{\Psi} = \mathbf{R}^T\mathbf{Y}. \tag{9}$$

are principal components. These principal components $\psi_j,\ j = 1, \ldots, M$ form an orthogonal basis in the signal space. The KL transform is a projection of the original signals y_i on the reduced signal subspace, which is a linear span of $L < M$ principal components. The corresponding reduced modified signal matrix $\mathbf{Y}_L$ is given as

$$\mathbf{Y}_L = \mathbf{R}_L\left(\mathbf{R}_L^T\mathbf{Y}\right), \tag{10}$$

where $\mathbf{R}_L$ is an $[M \times L]$ matrix composed of the first L column eigenvectors of the covariance matrix (8). Next, one should introduce a reversed time lag Δ and perform back normalization by multiplication to the denominator of equation 5 (which is stored in memory) to obtain the corresponding nonflattened common-frequency pseudoseismogram data.

For the preferred time lag Δ, which is related to a group velocity of the surface wave, it is sufficient to select $L = 1$ to approximate an ideally flattened common-frequency single-mode surface wave package (3). Jones and Levy (1987) showed that the eigenvalues of the covariance matrix are related to the energy content of the corresponding principal components. Assuming the energy dominance of the surface wave, we propose reviewing time lags Δ and selecting the lag that provides the largest first eigenvalue of the covariance matrix (7).

Ground-roll extraction algorithm

The proposed S-transform KL filtering (SKL) ground-roll extraction algorithm includes the following steps:

1. Apply the ST to each trace of the shot gather, and obtain a series of common-frequency pseudoseismograms (4). Select a pseudoseismogram frequency range that contains the main part of the surface wave energy (the range of [0, 20 Hz] is typical for seismic exploration data).

2. For each of the selected common-frequency pseudoseismograms, review time lags corresponding to the specified range of possible values of the group velocity of the surface wave (for the considered frequency). Select the time lag that maximizes the first eigenvalue of the modified (lagged) data covariance matrix.

3. For the selected lag, compute the reduced $L = 1$ modified data matrix (10). Introduce the reverse time lag, perform back normalization, and obtain a slant-KL estimation (model) of the surface wave in the common-frequency ST pseudoseismogram domain.

4. Apply the inverse ST (2) to the obtained pseudoseismogram surface wave model. Reject the obtained ground-roll signal from the data.

The SKL algorithm is illustrated in Figure 1. The raw field data, which contain a strong ground roll, are presented in Figure 1a. The real part of the 10 Hz normalized complex-valued ST pseudoseismogram is shown in Figure 1b. The flattening process is illustrated by grey and black lines in Figure 1b. The optimal lag for the represented common-frequency 10Hz ST gather is 3.46 seconds for the first trace, which corresponds to the surface wave group velocity of 361m/s. The corresponding narrowband surface wave package is focused but mixed with interference (Figure 1b). The flattened pseudoseismogram (using optimal lag, which was found in the second step of the presented SKL algorithm) is presented in Figure 1c. After application of the KL transform to this flattened pseudoseismogram data matrix, I obtain the reduced rank-1 data matrix, which is presented in Figure 1d. The resulting common-frequency ST pseudoseismogram domain model of a surface wave package, presented in Figure 1e, is obtained after adding a reverse lag to reduced data (presented in Figure 1d). Figures. 1d and 1e indicate that only the single largest eigenpair of the complex covariation matrix is sufficient to obtain saturated phase structure. The same operations: looking for the optimal time lag and performing slant KL transform are performed for all pseudoseismograms below 22Hz. After applying the inverse S-transform to the resulting set of the extracted common-frequency ST surface wave packets, the time-domain wide-band ground-roll is obtained. It is presented in Figure 1f.

When the ground roll is composed of several modes of surface waves, the SKL algorithm should be applied in an iterative manner - steps 1-4 are repeated several (3-5) times. If the shear wave velocity model is composed of several horizontally homogeneous blocks, the iterative application of SKL is also reasonable since the surface-wave package propagates with the different group velocities through each block.

The proposed SKL algorithm may seem to be cumbersome from a computational point of view. Indeed, it requires a time-frequency representation of each trace and multiple searches for eigenvalues. In fact, these issues are not prohibitively critical. The ST can be effectively computed using the fast Fourier transform. The SKL method can probably be accelerated by using a discrete orthonormal ST (Stockwell, 2007). The dimension of the square data covariance matrix (8) equals a number of receivers. For a large number of receivers, it is reasonable to apply SKL filtering in a narrow-offset sliding window manner. There are a number of algorithms that can efficiently find only a few (e.g., only one) largest eigenpairs of large matrices, e.g., the power method. Additionally, the efficiency of the SKL algorithm depends on the proper choice of the bounds of the trial values of lags, i.e., the possible values of the group velocity of the surface wave. Here I do not address these issues related to the implementation of the SKL algorithm. My straightforward and unoptimized MATLAB implementation of the SKL algorithm allows one to carry out 4 iterations of the SKL to 100 traces of 1500 time samples in less than 7 minutes using an Intel Core i7-5960X 3.00 GHz processor.

EXAMPLES

Synthetic data processing

Synthetic normal move-out reflected wave data are generated using a random reflectivity series, which is convolved with a 30 Hz Ricker wavelet. I compute a two-mode Rayleigh surface wave field using the Thomson-Haskell method and add it to the reflected data. The 8 Hz Ricker wavelet is used to generate the ground roll. The resulting synthetic one-shot gather is presented in Figure 2a. The receiver step is 25 m. The f-k spectrum of the synthetic data is shown in Figure 2d. I apply the SKL filtering to the synthetic data and compare it with 20 Hz HP and f-k fan filtering (FK). The HP and FK filter bounds are plotted in Figure 2d as dashed lines. The HP and SKL methods almost completely suppress the ground roll (see Figures. 2b and 2c), while the FK filter result contains some ground roll, which is marked in purple in Figure 2e. The low-frequency components of reflected waves, which are rejected by the HP filter, are shown in Figure 2f. The ground roll, extracted by SKL, is much more pure - almost entirely composed of the surface wave (Figure 2g). I implement SKL filtering 3 times to obtain the presented result. The f-k spectrum of the ground roll extracted by SKL (see Figure 2h) indicates that the two modes of the aliased Rayleigh wave are separated from the reflected waves.

Field data processing

A land seismic shot record is presented in Figure 3a. The data was gathered in Khanty-Mansiysk Autonomous Okrug, Western Siberia, Russia. The explosive seismic source was used. The receiver step is 25 m. The time images in Figures. 3a, 3b, 3d, and 3e are obtained after the automatic gain control (AGC). The ground-roll suppression is applied separately for the right- and left-hand sides of the shot gather (both parts then come back together).

Similar to the synthetic data, I compare the SKL, HP and FK filtering results. The HP and FK filter bounds are plotted in Figure 3d as dashed lines. The SKL, HP and FK ground-roll suppression results are presented in Figures. 3b, 3d and 3e, respectively. The result of FK filtering is strongly distorted. The ground roll, which is suppressed by the proposed SKL algorithm, is presented in Figure 3c. This result indicates that SKL suppresses only surface waves and does not affect reflected waves. At first glance, the result of applying HP filter(Figure 3e), looks more pure than the SKL result (Figure 3b). However, Figure 3f, which represents the ground roll extracted by HP filter, indicates that HP filter removes low-frequency components of reflected waves. This leads to defocusing of seismic images of some reflective interfaces (compare the sections of seismograms highlighted in purple in

Figure 3).

The f-k spectrum of the right-hand side of the data is shown in Figure 4a. I plot only the right half of the FK plane for the considered split-spread dataset for simplicity. One can observe the aliased ground roll, interfering with the reflected wave. Actually, the FK filter could be enhanced by more accurate selection of the area to be filtered using polygons instead of fan area (see Figure 4a). However, the selection of filtered area is a human-labor process. The SKL method has the potential to provide stronger filtering as it is capable of separating the waves whose f-k images are intersecting. The f-k image of the ground roll, which is removed from the right-hand side of the seismic gather by SKL, is presented in Figure 4b. One can also compare the results of applying HP and SKL filters in terms of average Fourier amplitude spectrum, which is presented in Figure 4c. In particular, the average Fourier spectrum after applying SKL filter (red dotted line in Figure 4c), indicates that SKL passes a large amount of low-frequency reflection wave components.

CONCLUSION

The slant Karhunen-Loeve (KL) transform is known to be a suitable method for suppression of the spatially aliased ground roll. However, the surface waves should be flattened before the KL transform, which is laborious and problematic due to their dispersion. I propose a new ground roll extraction method, which solves this problem. The S-transform (ST) time-frequency representation of common-shot data is considered. The ST is a proper method to handle the surface wave dispersion. I flatten the ground roll in the time-frequency domain and apply slant KL transform to common-frequency ST complex-valued pseudo-seismograms. Such kind of a complex-valued seismic trace processing is well suited for extracting constantly phase-shifted ST common-frequency surface wave packets.

The presented results of aliased ground roll suppression demonstrate that the proposed S-transform slant KL method (SKL) can be effective in practice and provide better results than high-pass and FK fan filtering approaches.

One of the main concerns about the proposed method may be its computational complexity. My experiments show that the efficiency of SKL is sufficient for reflection data processing.

ACKNOWLEDGEMENTS

The research was supported by the Russian Science Foundation (project no. 20-77-10023). I would like to thank the editor-in-chief, Dr. John Etgen, the associate and assistant editors of this paper, Ayse Kaslilar and Joakim Blanch, and three anonymous reviewers for their comments and corrections that improved the manuscript.

DATA AND MATERIALS AVAILABILITY

Data available on request from the authors

REFERENCES

Askari, R., and R. J. Ferguson, 2012, Dispersion and the dissipative characteristics of surface waves in the generalized S-transform domain: Geophysics, **77**, V11–V20.

Askari, R., and H. R. Siahkoohi, 2008, Ground roll attenuation using the s and x-f-k transforms: Geophysical Prospecting, **56**, 105–114.

Chen, X., W. Chen, X. Wang, and W. Wang, 2017, Sparsity-optimized separation of body waves and ground-roll by constructing dictionaries using tunable q-factor wavelet transforms with different q-factors: Geophysical Journal International, **211**, 621–636.

Hemon, C., and D. Mace, 1978, Use of the karhunen-loeve transformation in seismic data-processing: Geophysical Prospecting, **26**, 600–626.

Jones, I., and S. Levy, 1987, Signal-to-noise ratio enhancement in multichannel seismic data via the karhunen-loève transform: Geophysical prospecting, **35**, 12–32.

Karslı, H., and Y. Bayrak, 2004, Using the wiener–levinson algorithm to suppress ground-roll: Journal of Applied Geophysics, **55**, 187–197.

———, 2008, Ground-roll attenuation based on wiener filtering and benefits of time-frequency imaging: The Leading Edge, **27**, 206–209.

Levy, S., and D. Oldenburg, 1982, The deconvolution of phase-shifted wavelets: Geophysics, **47**, 1285–1294.

Levy, S., T. Ulrych, I. Jones, and D. Oldenburg, 1983, Applications of complex common signal analysis in exploration seismology, *in* SEG Technical Program Expanded Abstracts 1983: Society of Exploration Geophysicists, 325–328.

Liu, X., 1999, Ground roll supression using the karhunen-loeve transform: Geophysics, **64**, 564–566.

Liu, Z., Y. Chen, and J. Ma, 2018, Ground roll attenuation by synchrosqueezed curvelet transform: Journal of Applied Geophysics, **151**, 246–262.

Possidonio, M., and M. J. Porsani, 2021, A combined method using singular spectrum analysis and instantaneous frequency for the ground-roll filtering: Geophysical Journal International, **226**, 446–455.

Saatcilar, R., and N. Canitez, 1994, The lattice filter in ground-roll suppression: Geophysics, **59**, 623–631.

Schimmel, M., and J. Gallart, 2005, The inverse s-transform in filters with time-frequency localization: IEEE Transactions on signal processing, **53**, 4417–4422.

Serdyukov, A., A. Yablokov, A. Duchkov, A. Azarov, and V. Baranov, 2019, Slant f-k transform of multichannel seismic surface wave data: Geophysics, **84**, A19–A24.

Serdyukov, A. S., A. V. Azarov, A. V. Yablokov, T. V. Shilova, and V. D. Baranov, 2021, Reconstruction of seismic signals using s-transform ridges: Geophysical Prospecting.

Stockwell, R. G., 2007, A basis for efficient representation of the s-transform: Digital Signal Processing, **17**, 371–393.

Stockwell, R. G., L. Mansinha, and R. Lowe, 1996, Localization of the complex spectrum: the *S* transform: IEEE transactions on signal processing, **44**, 998–1001.

Verma, S., S. Guo, T. Ha, and K. J. Marfurt, 2016, Highly aliased ground-roll suppression using a 3d multiwindow karhunen-loeve filter: Application to a legacy mississippi lime survey: Geophysics, **81**, V79–V88.

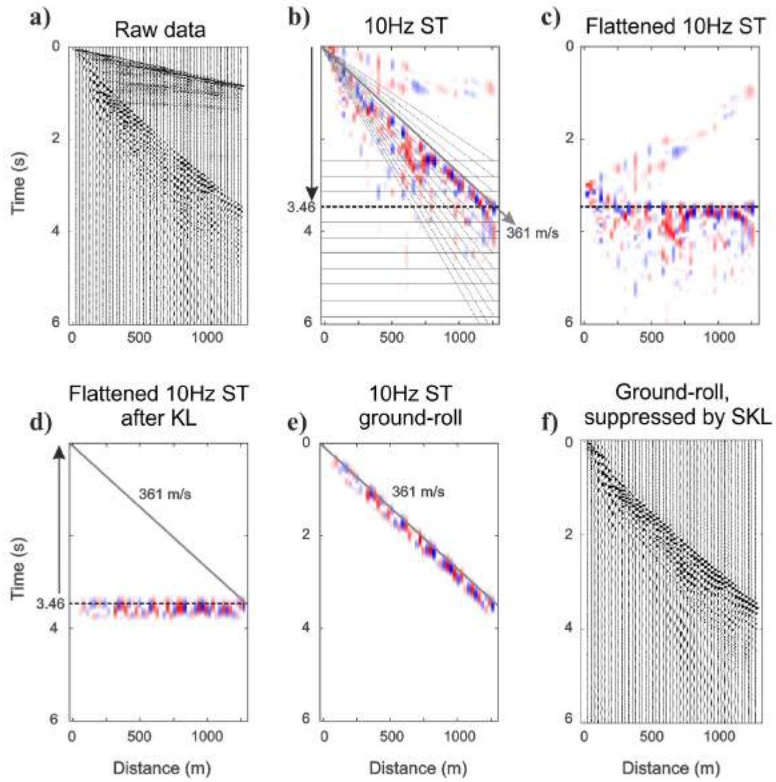

Figure 1: Figure 1: Karhunen-Loeve (KL) filtering in the S-transform common-frequency domain: a) raw field data; b) real part of the 10 Hz common-frequency gather; the optimal time lag (865 time samples for the first trace) is shown by dotted black line c) real part of the flattened 10 Hz common-frequency gather; d) real part of the rank-1 approximation of the optimal flattened common-frequency gather; e) extracted 10 Hz ground roll S-transform component; f) the ground roll suppressed by one-time application of the SKL method

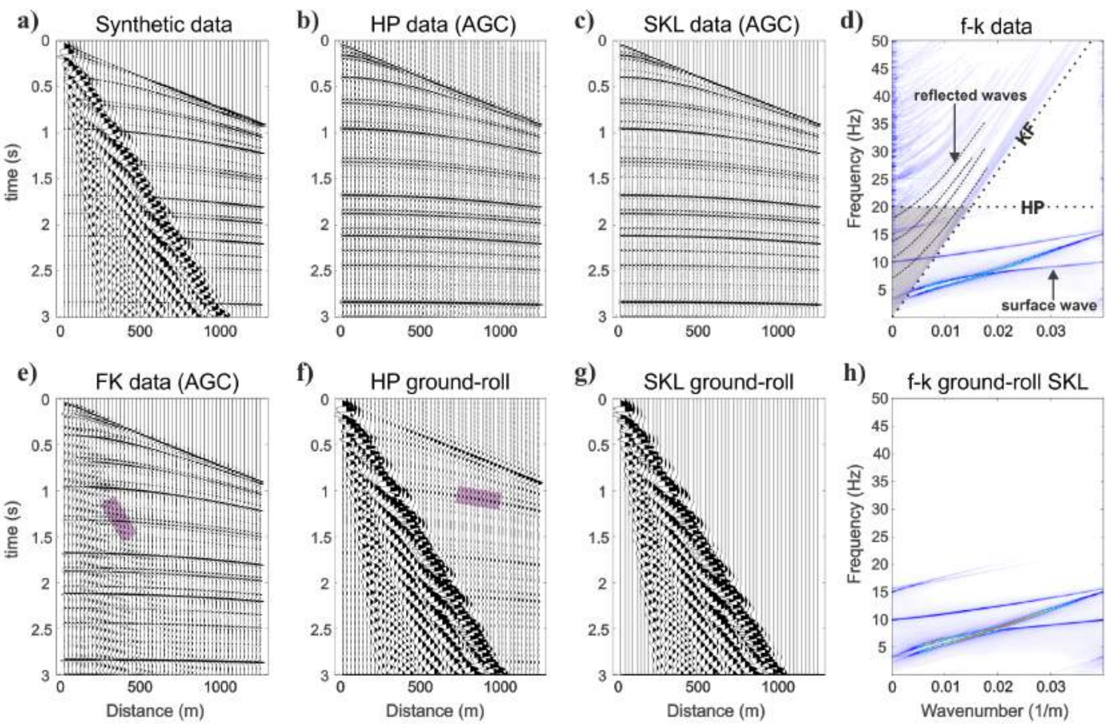

Figure 2: Figure 2: Synthetic data ground-roll attenuation results: a) synthetic common-shot gather; b) result of the high-pass (HP) (20 Hz) filter after automatic gain control (AGC); c) result of the proposed S-transform KL filter (SKL) after AGC (the SKL is performed below 20Hz); d) f-k spectrum of the synthetic data, where the f-k fan filter (FK) and HP filer bounds are plotted by dotted lines; the aliasing area is highlighted by grey e) result of applying the FK filter; f) ground roll, suppressed by HP g) ground roll, extracted by SKL; h) f-k spectrum of the ground roll, suppressed by SKL

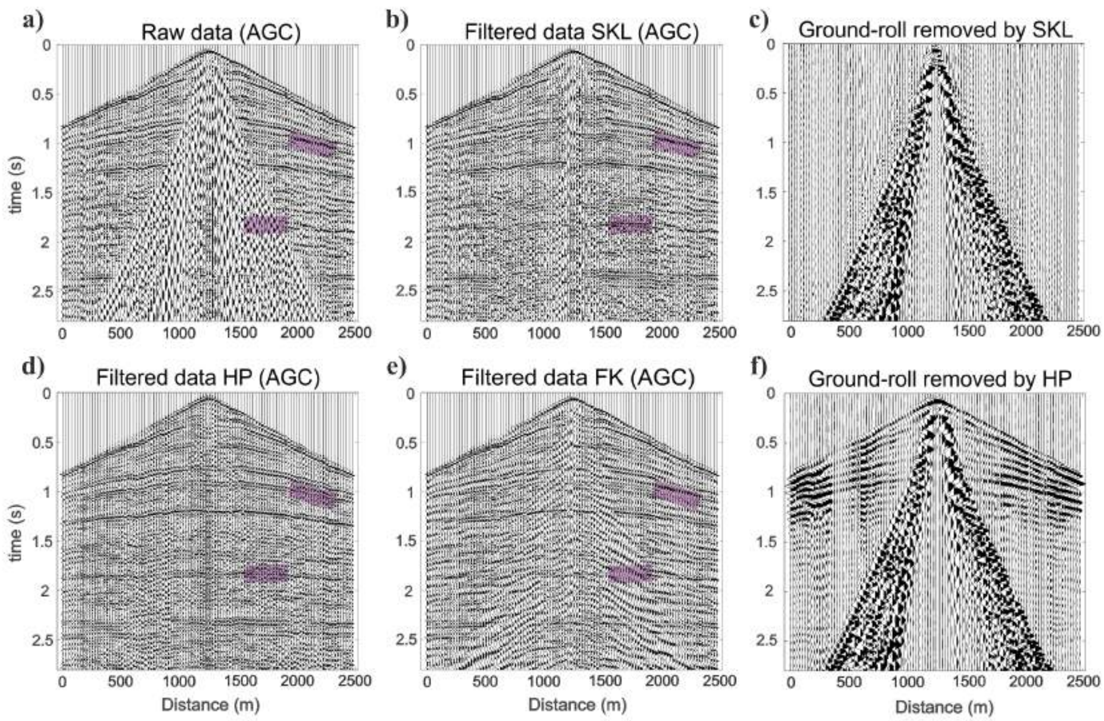

Figure 3: Figure 3: Field data ground-roll attenuation results: a) raw data after AGC; b) result of the proposed SKL after AGC; c) ground roll, suppressed by SKL (the SKL is performed below 22Hz); c) result of applying the HP (22 Hz cutoff) filter after AGC; d) the result of applying FK filter; e) ground roll, suppressed by HP filter.

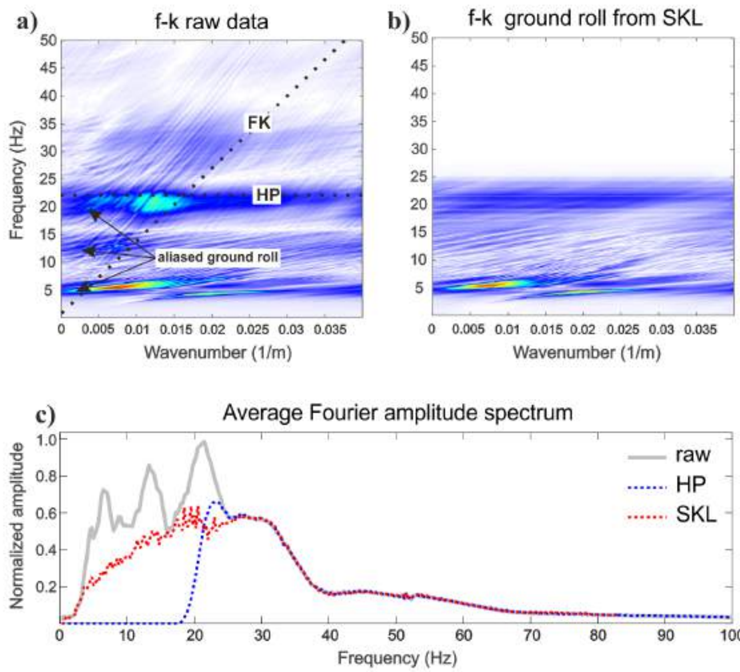

Figure 4: Figure 4: Field data spectrum before and after processing a) f-k spectrum of the field data b) f-k spectrum of the ground roll, suppressed by SKL; c) average Fourier amplitude spectrum of the raw (grey line) data and after applying high-pass filter (blue line) and SKL filter (red line)