Two approaches to single-station detection and removal of ground roll from seismic reflection measurements by three component polarization analysis

Pim Lamberts, Department of Earth Sciences, Faculty of Geosciences, Utrecht University
Master of Science Thesis
Supervised by: Sjef Meekes (TNO), Stefan Carpentier (TNO), Hanneke Paulssen (UU)

1. Abstract

The ground roll consists, among others, of Rayleigh waves that are characterized by low velocity, high amplitude, low frequency and retrograde elliptical particle motion in the direction of propagation. The particle motion can be recorded by employing three-component seismic receivers that allow Rayleigh waves and body waves to be differentiated by their nature of polarization.
From covariance analysis and eigendecomposition on a sliding time window polarization attributes are calculated to identify the state and direction of polarization in the data in a single station approach. The polarization attributes are then applied to the data in two polarization filters. The linearity method minimizes all non-linear arrivals in a multiplication method while the ellipticity method separates the elliptically polarized signal before subtracting it from the original data.
Simulations of both polarization filters on synthetic and field data show that an accurate polarization analysis can be made to identify individual arrivals and suppress the ground roll to a great extent.

Contents

2. Introduction

Ground roll or surface waves are source-generated waves, that travel along the surface and near-surface of the Earth. Among others, the ground roll consists of Rayleigh waves which are characterized by low velocity, high amplitude, low frequency, with retrograde elliptical particle motion in the direction of propagation. They can be observed as a fan-shaped zone on seismograms as a result of their dispersive character, an invariable velocity and relative insensitivity to attenuation.
As a result of the high amplitude of the ground roll, its influence on seismic reflection data is significant since it obscures any simultaneous arriving reflected body waves which generally have a relatively low amplitude. This especially holds serious implications for P-waves which are linearly polarized in the direction of propagation and can be observed in a wide band of frequencies and velocities and are considered as the primary indicator of the subsurface.
Therefore, multiple approaches have been employed in the past to filter the ground roll. Although most seismic surveys are currently recorded with three-component (3-C) receivers, the first filters were designed for one-component (1-C) data. Suppression methods for such datasets usually operate in a transformed domain, examples are the frequency-wavenumber (f-k) domain, Tau-P domain and Eigenvector\Principal Component domain. By applying filters a portion of the signal that contains the ground roll can be removed and then be back-transformed to the original domain. Down point of all these transform and spatial domain 1-C methods is that they often generate artefacts due to the large ground roll amplitudes involved and dispersed signals in the data. Invoking two additional spatial dimensions with orthogonal components of particle motion should enhance the separation and thus suppression of ground roll from reflection energy.
De Meersman and Kendall (2005) and Jin and Ronen (2005) applied a covariance analysis to determine the principal directions and polarization in the data and effectively suppress the ground roll in the low frequency band through polarization filtering. Later research by Tiapkina et al. (2012) and Chen et al. (2013) expanded on those ideas.
This research aims to study the possibilities of two 3-C approaches to minimize the ground roll (figure 4), both of which are strongly based on previous research by Maercklin (2007). For each sample polarization attributes are calculated from covariance analysis and eigendecomposition on a sliding time window in a single station approach. Major advantage of this method is that anomalous spikes are smoothed by the time window and the method is independent of spatial sampling.
These polarization attributes are then applied in two polarization filters. The first method is designed to minimize the ground roll by multiplying the original data to the rectilinearity, one of the polarization attributes, and the principal direction of polarization in the linearity method (Maercklin, 2007). The second method isolates the signal above a certain cut-off ellipticity just like the determined threshold ellipticity by Jin and Ronen (2005). The method then removes the elliptical signal by subtracting it from the original data.
The methods will be tested on three datasets. Firstly the methods will be tested on a simple synthetic model for 1. a surface source and 2. a buried source in a densely sampled linear array with 3-C accelerometers. The aim of the model is to provide an idealized polarization analysis before the polarization filters are applied to the third dataset which is field data recorded in a similar array with a drop weight.

3. Description of the test site and its geology

The field data for this research has been shot on a meadow on the outskirts of the Utrecht Science Parc (figure 2), the campus of Utrecht University. Aside from the relatively low amount of expected seismic background noise, the shallow subsurface at this location is of particular interest because of the horizontal stratification and the absence of large faults or other obscuring anomalies (figure 3). Further, the expected large velocity gradient in the top layers should provide ideal conditions for a strong source induced ground roll to develop.

The shallow geology in Utrecht is formed by the Upper North Sea Group (NAM and RGD, 1980) sediments (figure 1). The geological units of which are a sequence of clays and fine-grained to coarse-grained sands. This is in line with the depositional environment which is an alternating sequence of shallow marine, coastal and fluviatile sediments. The sediments at a depth of about 400 meters represent the Breda formation dating back to the early Pliocene while the youngest sediments at the surface are recent Holocene deposits.

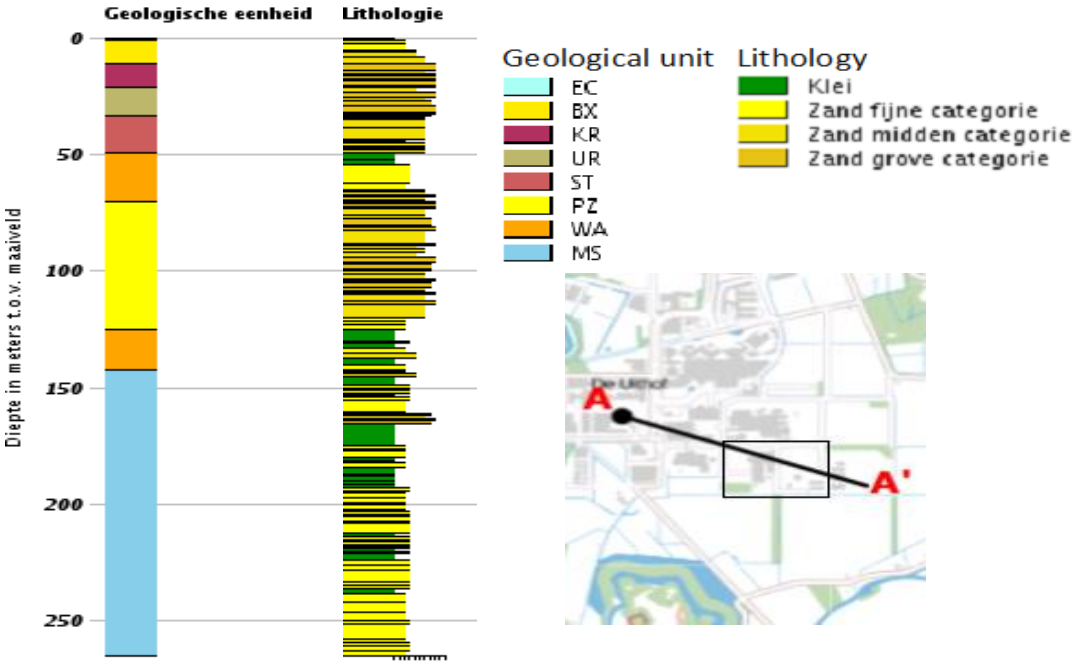

Figure 1. Borehole survey and geological profile of core B32C0489 from location A, provided by Dinoloket (2018). Indicated on the map are the line of the section from figure 3 and the rectangle indicates the aerial photograph from figure 2.

Figure 2. Aerial photograph of the location of the seismic survey. Satellite image provided by Google Maps.

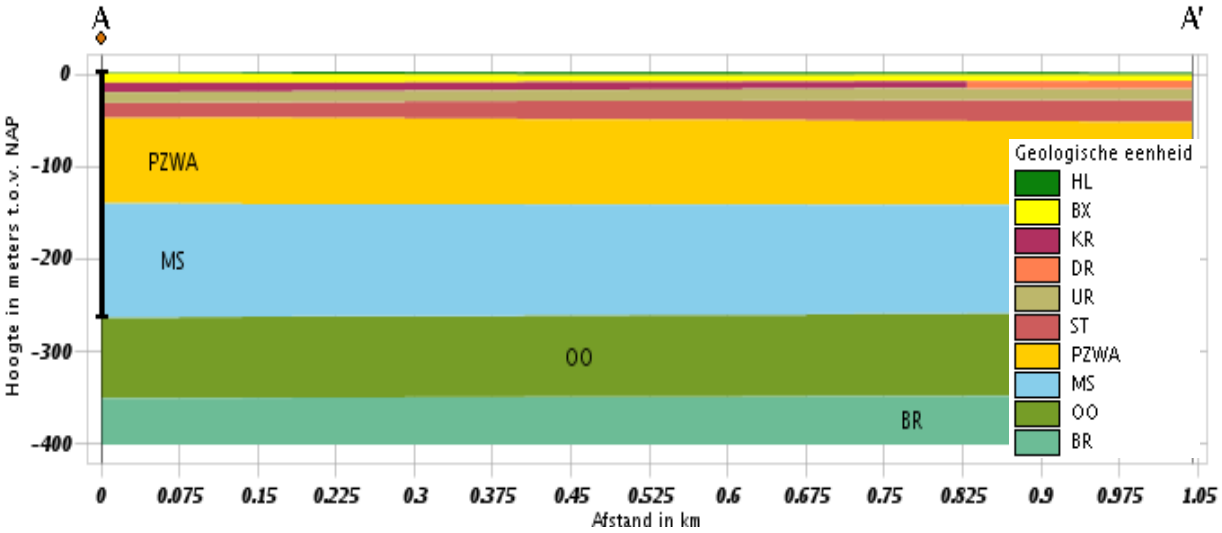

Figure 3. Section from the shallow subsurface beneath the Utrecht Science Parc as indicated in figure 1. The section represents the shallow geology of location of the seismic survey. From Dinoloket (2018).

4. Methods

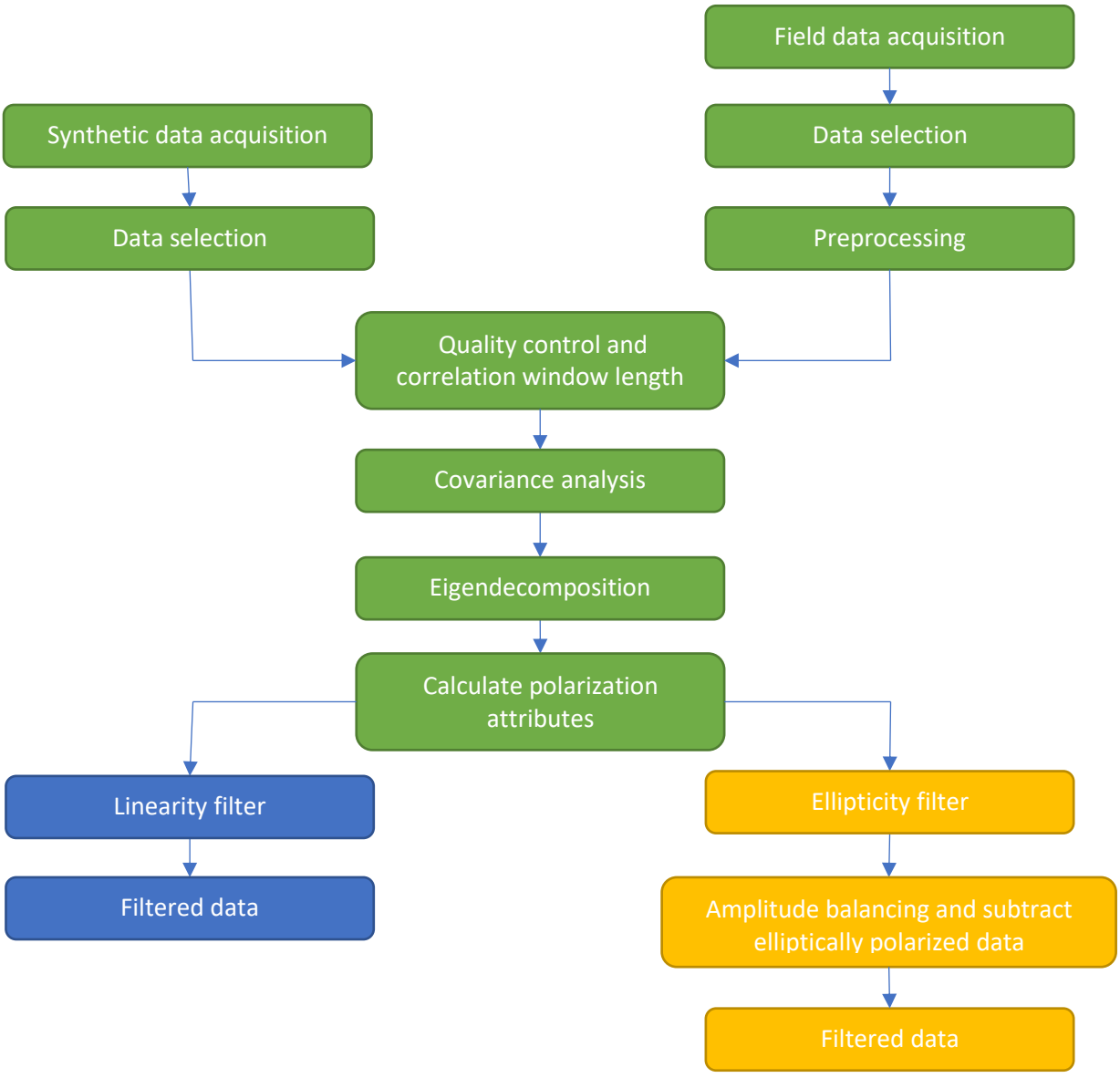

Figure 4. Workflow of the methods described in the following section for the linearity method and ellipticity method.

4.1 Data acquisition

4.1.1 SPECFEM3D Cartesian synthetic data

SPECFEM3D Cartesian (Komatitsch and Tromp (2002a), Komatitsch and Tromp (2002b), Komatitsch et al. (2012)) was employed to model a simplified version of the subsurface in Utrecht to resemble the field data in a dense grid of 200x25x400 meters in an anisotropic, elastic medium (figure 5). The data is, similar to acquisition design in the field, recorded with 62 3-component accelerometers in a linear array with one meter spacing between each station to accurately capture the ground roll and prevent spatial aliasing. The interfaces in the model are perfectly horizontal to resemble the horizontal stratification of the Utrecht subsurface. Furthermore, the amount of interfaces has been kept to a minimum to limit the amount of reflections and wave interference, while the depths of the interfaces are comparable to some of the interfaces observed in figure 3**.** Moreover, to realistically compare the modeled data to the field data the parametrization of the layers was based on typical seismic velocities and densities, from Acoustics of Porous Media (1987) by Bourbié, Coussy, and Zinszner.

Therefore, this simplified model is a suitable approach to not only have an uncluttered case to test the polarization filter but it also provides opportunities to test some variables which could not be altered in the field data, such as comparing a surface source to a buried source to test the influence of burial to the ground roll is an important variable. To keep all other variables identical, both sources are vertical point sources, similar to a drop weight, that apply the same force, although in reality a buried source would be an explosive, omnidirectional type of source.

Depth (m)	Layer	Rho (kg/m^3)	V_p (m/s)	V_s (m/s)
0-2	L1	1500	600	200
2-30	L2	1800	1450	250
30-125	L3	2000	1550	290
125-375	L4	2150	1700	320
375-400	L5	2400	2100	375

Table 1. SPECFEM3D Cartesian density and velocity parameters.

The best results for the synthetic data were obtained employing a basis frequency of 80 Hertz (figure 6). A higher basis frequencies would require an even denser model grid which becomes computationally too expensive. Lower basis frequencies, on the other hand, would result in a narrower spectral range, making the source less similar to the drop weight. Figure 7 shows that the mean power spectra of the synthetic datasets show significant overlap and are characterized by a relatively high ground roll frequency (40-50 Hz).

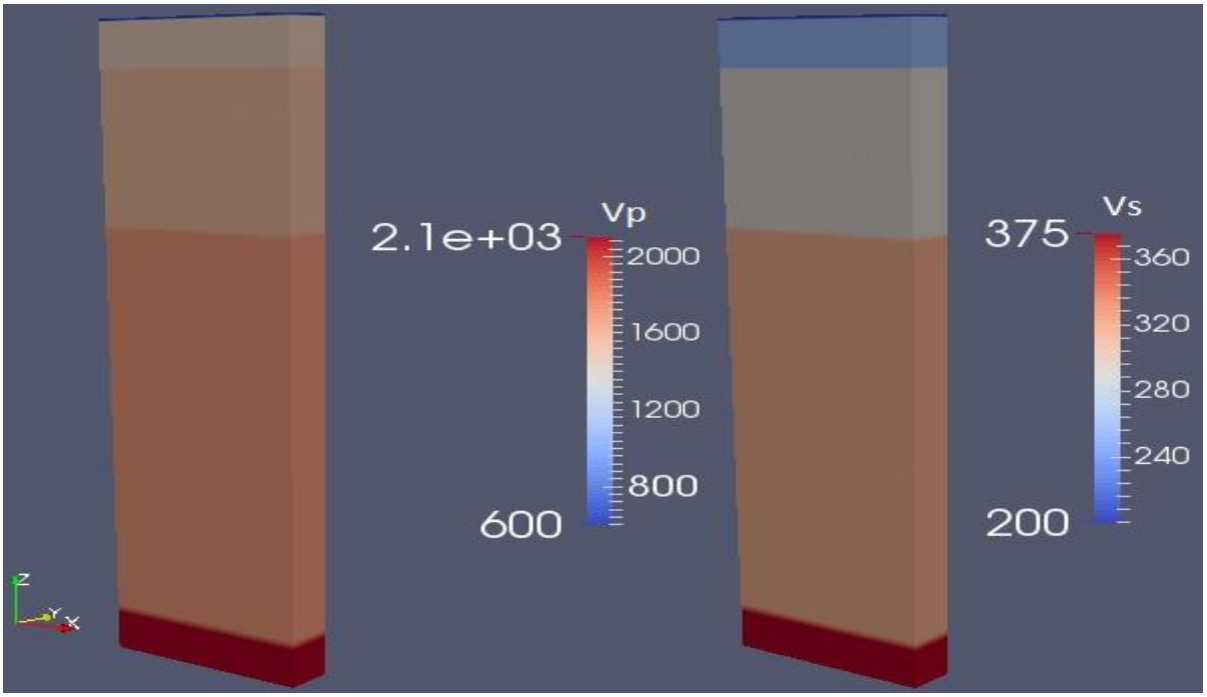

Figure 5. Seismic velocity model applied in the SPECFEM3D Cartesian model. The exact velocity for each layer is listed in table 1.

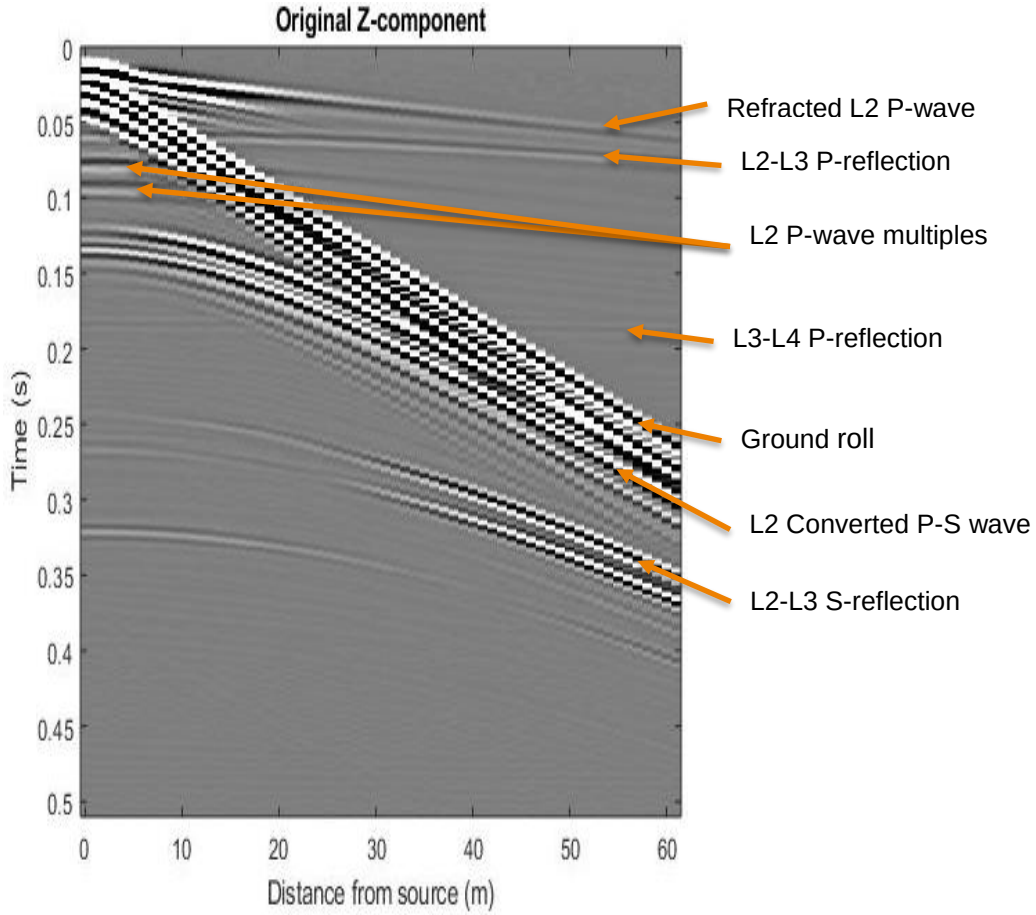

Figure 6. Section from the vertical component of the buried source synthetic dataset with identification of different arrivals based on the model parameters in table 1. For instance, the L2-L3 P-reflection is the reflected P-wave at the 30 m interface.

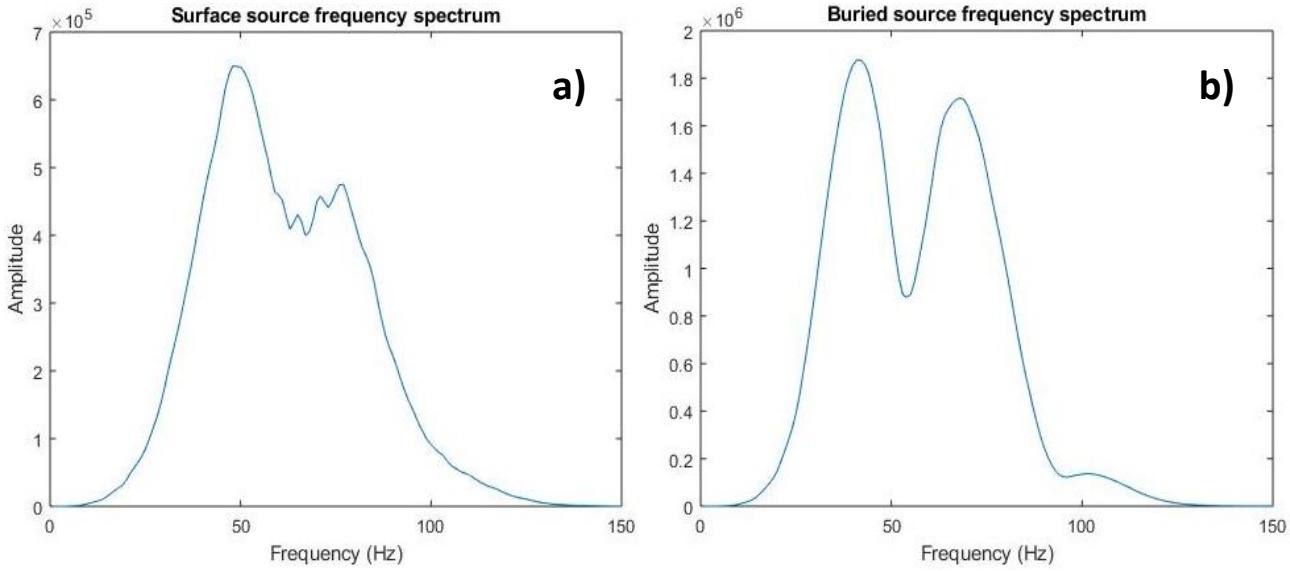

Figure 7. Mean power spectrum of the traces of the synthetic datasets. a) surface source, b) buried source.

4.1.2 Utrecht Field data

The aim of this project is to eliminate the source induced Rayleigh ground roll. Therefore, a weight drop is the ideal source since it is known to induce high amplitude surface waves (figure 8).
The reason for shooting the field data on the outskirts of the Utrecht Science Parc is firstly because of the previously mentioned geological characteristics of the area. Secondly, the amount of seismic

background noise in the environment is expected to be relatively low compared to nearby alternatives. With the nearby heat generator out of operation, the absence of traffic and the minimal influence from the tree line as a result of the mild weather conditions (low wind, no rain) the circumstances were favorable for proper data acquisition. The main setbacks were the nearby cattle farm which caused more background noise than anticipated and the dry weather conditions which reduced coupling of the receivers and source to the subsurface.

To accurately capture the ground roll, the data was recorded in the previously mentioned dense linear array of 62 3-component accelerometers with a spacing of 1m that recorded at a sampling rate of 500 µs to optimize resolution and prevent spatial aliasing. The axes of the accelerometers were oriented in a coordinate system containing an inline (X), crossline (Y) and vertical (Z) component with a maximum acquisition deviation of 5°.

The used drop weight is an iron ball-shaped object of 40 kilograms which was dropped from a height of approximately 3 meters above the ground, onto an iron plate of roughly the same weight for better coupling. The GPS times of the shots were recorded with a trigger next to the iron plate to not only track the shot record but also to ensure that the force of impact was strong enough. With this setup the weight drop was placed on multiple locations within the grid. On each location the source was fired five times. The shot records with the source at the end of the array were used for the polarization analysis (figure 13).

Figure 8. Field data acquisition design with the drop weight in front of the linear array.

4.2 Preprocessing of field data

Preprocessing should be kept to a minimum to avoid processing artefacts in later stages of the polarization filter. However, to remove the high frequency ambient noise and the sound wave from the field data (>200Hz), as well as the ultra-low frequency content (<5Hz), a zero-phase bandpass filter is applied to the data (figure 9). Through this configuration of the bandpass filter most of the surface wave energy as well as the body wave energy can be preserved while especially a lot of high frequency noise and the high amplitude ultra-low frequency content are removed to make the frequency content of the field data more analogous to the synthetic data.

Since the main aim of the bandpass filter is to denoise the sections, SPECFEM3D synthetic data does not require any preprocessing because the ambient noise is inconsequential and clipping the spectral range would only result in a loss of useful signal. After the bandpass filter the linear trend is removed from the data by subtracting the mean for each column inside the data matrix. This enables the data analysis in the next section to be focused on the fluctuations in the data balanced to the detrended mean without any bias.

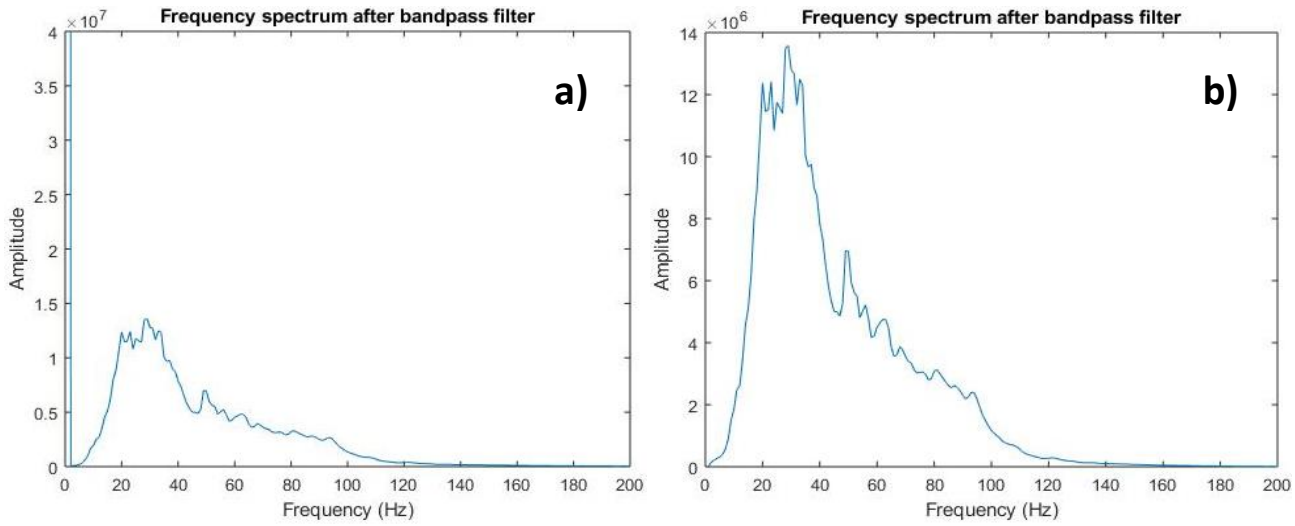

Figure 9. Relative power spectrum of the field data before (a) and after bandpass filtering (b). Low pass frequency is 5 Hz, high pass frequency is 150 Hz. Especially a dominant peak of very low frequency ambient noise (2 Hz) is removed.

4.3 Quality control and correlation window length

Hodograms are used to examine the recorded particle motion and amplitude of seismic waves in three dimensions within a predefined time window. Hodograms, therefore, provide a reliable control on the data quality aside from the seismic sections. P-wave arrivals should record as medium to low amplitude linear hodograms whereas Rayleigh waves display high amplitude elliptical particle motion (figure 10). In time windows where P-waves and Rayleigh waves arrive simultaneously the high amplitude Rayleigh wave dominates the hodogram in a sense that the overprinted P-wave can hardly be detected from the hodogram.

Lastly, an adequate correlation window length for the covariance analysis, discussed in the following section, can be recovered by imaging at least one signal period of the ground roll. A too small correlation window would not capture the entire ground roll which would lead to inaccuracies in defining its state of polarization. A too large correlation window, on the other hand, would cause smoothing of the data, especially of high frequency body waves. Therefore careful determination of the correlation window is essential for adequate polarization analysis. Approximately 1-1.2 times the average ground roll period produces the most accurate results.

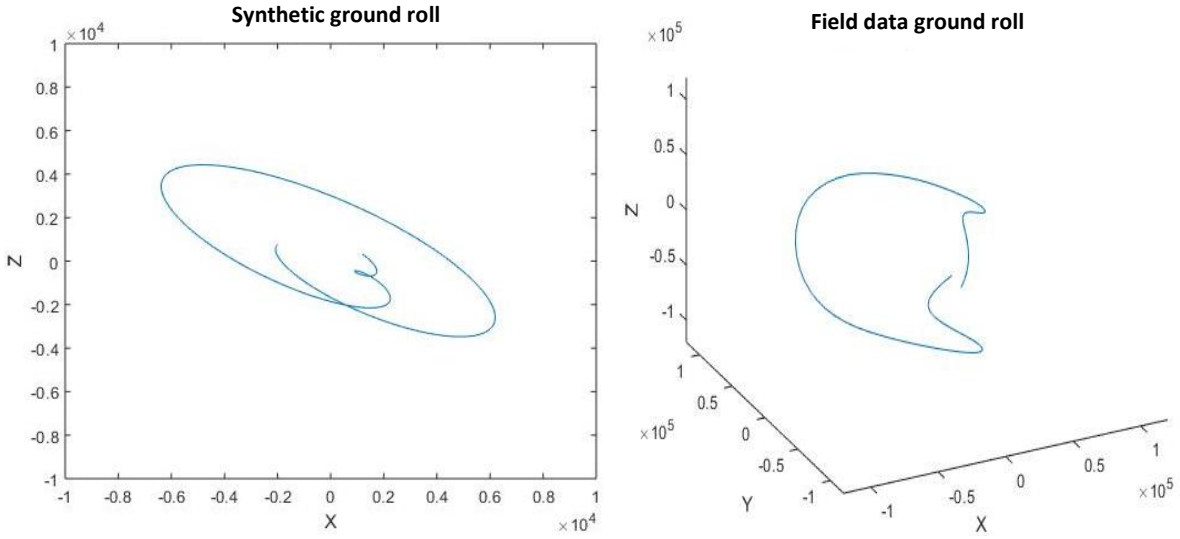

Figure 10. Example of two ground roll hodograms recorded at a source distance of 30m. The left hodogram represents the synthetic ground roll while the right hodogram is the slow ground roll observed in the field data.

4.4 Covariance analysis

In the covariance analysis the data are first sorted as follows:

$$D(z,x,y) = [A_z(t), A_x(t), A_y(t)]$$

With the Z-axis parallel to the vertical component, the X-axis parallel to the inline component and the Y-axis parallel to the crossline component.

Then the covariance matrix is calculated for every recorded time sample, centered in the aforementioned correlation window, for each station individually. However, to keep focus on the properties of each individual time sample the correlation window is shaped like a Hann window. Therefore, the input data for the covariance matrix is a subset of matrix D containing N rows, indicating the amount of time samples inside the Hann window.

On the diagonal position, the covariance matrix consists of the variances of the data components. The covariance between different components in the other positions. The variance on the j'th diagonal position of the covariance matrix represents the squared deviation from the mean of the j'th column vector from the data matrix, within the selected time window. This emphasizes the data with greater deviation from the mean. On the other hand, the covariance analysis determines the correlation between two spatial variables within the selected time window. Hence the element in position (i,j) of covariance matrix M describes the correlation between the column vectors i and j in data matrix D.

$$M(i,j) = \begin{pmatrix} var(X) & cov(X,Y) & cov(X,Z) \\ cov(Y,X) & var(Y) & cov(Y,Z) \\ cov(Z,X) & cov(Z,Y) & var(Z) \end{pmatrix}$$

$$cov(X,Y) = \frac{1}{N}\sum_{k=-L}^{L-1} [X_k(t) - \mu_x]\,[Y_k(t) - \mu_y]$$

$$var(X) = cov(X,X)$$

With L=N/2, indicating the half correlation window length and μ being the mean of the time sequence within the analyzed window.

4.5 Eigendecomposition

The principal directions in the data can be retrieved through an eigendecomposition of the covariance matrix. Hereby the covariance matrix is represented in terms of matrices containing three eigenvalues (λ), as the weights of polarization in a diagonal matrix, and three eigenvectors (v), being the three orthogonal directions of polarization. The eigenvectors and associated eigenvalues are related to the covariance matrix in the equation

$$Mv = \lambda v$$

The retrieved eigenvalues, and corresponding eigenvectors, can be sorted from largest to smallest. The largest eigenvalue (λ_1) is associated with the principal direction of polarization. The values of the two largest eigenvalues are assumed to be almost equal for elliptical polarization whereas linear polarization should result in a large contrast between the first two eigenvalues. Due to the planar nature of the ground roll and linear nature of body waves, the smallest eigenvalue (λ_3) is expected to be negligible, especially for the synthetic model.

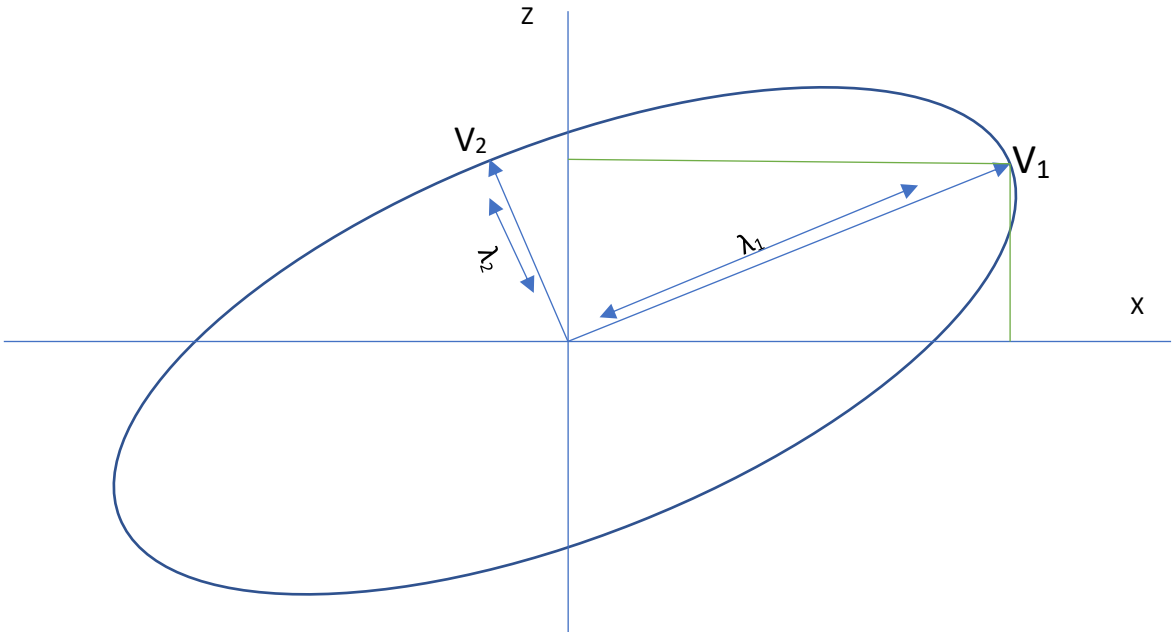

Figure 11: Polarization ellipse to exemplify the rotation from a Cartesian reference frame to an eigenvector reference frame. The direction of the eigenvectors can still be represented in terms of the Cartesian reference frame as indicated by the green lines.

4.6 Polarization attributes

From the eigenvalues and eigenvectors, the polarization attributes are determined such as: the direction of polarization (Dp), rectilinearity (Rl), global polarization parameter (τ) and ellipticity (e_{21}). Except for the directions of polarization, which expresses the direction of the principal eigenvector in terms of the reference coordinate frame (z,x,y), the polarization attributes are all dependent on a certain relation between the eigenvalues.

4.6.1 Direction of polarization

The direction of polarization is captured in the eigenvector (v_1) associated with the strongest eigenvalue (λ_1). The principal eigenvector (v_1) is a unit vector, situated in the reference coordinate frame and can therefore be expressed in terms of j=z, x and y (figure 11):

$$Dp_j(t_i) = \left(\left|v_{i,j}\right|\right)$$

Thus, the principal direction of polarization for each timestep (t_i), centered in the correlation window, always represents the absolute values of the z, x and y component of the principal eigenvector derived in the eigendecomposition. This means that all three components of the principal eigenvector are smaller than one ($Dp_j = [0,1]$).

The direction of polarization does not give any information about the state of polarization because the parameter is not dependent on the eigenvalues. It can merely distinct different wavefields by the dominant direction of polarization.

4.6.2 Rectilinearity

To measure the degree of linear polarization of certain events, the eigenvalues can be applied to calculate the rectilinearity. In the circumstances that the energy is almost completely distributed along a single plane, the rectilinearity can be determined according to the definition of Kanasewich (1981, 1990):

$$Rl = 1 - \left(\frac{\lambda_2}{\lambda_1}\right)^Q \quad 0 < Q \leq 1$$

In case of linear polarization, when $\lambda_1 \gg \lambda_2$, the rectilinearity is approximately 1. Contrarily, in case of elliptical or circular polarization the rectilinearity goes to 0, with $Rl = [0,1]$.

However, if there is significant energy along all three eigenvectors, which may be the case for data acquisition in the field or datasets containing Love and Sh-wave energy, the rectilinearity may be altered according to the definition by Jurkevics (1988):

$$Rl = 1 - \left(\frac{\lambda_2 + \lambda_3}{2\lambda_1}\right)^Q \quad 0 < Q \leq 1$$

Similar to the previous expression, for linear events the rectilinearity is approximately 1 with the principal eigenvalue (λ_1) many times larger than the other eigenvalues, whereas a rectilinearity of 0 in this case implies spherical polarization and the direction of polarization would be undetermined. Nonetheless, for the scope of this research the definition of the rectilinearity from Kanasewich (1981, 1990) is favorable over the second definition.

To emphasize or lessen the distinction between linear and elliptical events the quality factor, Q, can be decreased or increased. The quality factor is a mean to adjust the sensitivity to certain degrees of polarization. For lower values of Q the rectilinearity becomes increasingly more sensitive to linearly polarized events.

4.6.3 Global polarization parameter

Analogous to the rectilinearity, the global polarization parameter can be implemented to identify the state of polarization of arriving phases. The global polarization parameter requires all three eigenvalues and, contrary to the rectilinearity, cannot be influenced by a quality factor. The global polarization parameter is first applied by Samson (1973) in ultra-low frequency magnetic fields and as the name suggests is more commonly applied for large scale seismic events. Therefore, the parameter is less effective as a tool for the small scale shallow survey conducted in this research and will not be applied in the filtering methods. The expression for the global polarization parameters is as follows:

$$\tau^2 = \frac{(1 - \frac{\lambda_2}{\lambda_1})^2 + (1 - \frac{\lambda_3}{\lambda_1})^2 + (\frac{\lambda_2}{\lambda_1} - \frac{\lambda_3}{\lambda_1})^2}{2(1 + \frac{\lambda_2}{\lambda_1} + \frac{\lambda_3}{\lambda_1})^2}$$

A straight line gives τ = 1, a circle τ = 0.5 and a sphere τ = 0 (τ = [0 , 1]).

4.6.4 Ellipticity

Opposite to both rectilinearity and the global polarization parameter, the ellipticity indicates elliptical states of polarization. Therefore, circular ellipticity results in $e_{21} = 1$, whereas for linear polarization $e_{21} = 0$. The expression for ellipticity is basically the exact opposite to the rectilinearity:

$$e_{21} = \left(\frac{\lambda_2}{\lambda_1}\right)^Q \quad Q \leq 1$$

The main aim of ellipticity is not only to distinguish between arrivals like rectilinearity does, but also to enhance the ground roll relative to linear arrivals. This can be amplified by altering the quality factor.

Filtering approaches

To filter the ground roll, the polarization attributes are employed in two opposite approaches. The first approach requires the rectilinearity or global polarization parameter in combination with the directions of polarization to directly remove the ground roll from the data in a multiplication method. The second approach is based on the ellipticity to isolate the ground roll and, after amplitude balancing, subtract the ground roll from the data.

4.7 Linearity method

The linearity method filters the data by multiplication of the data with a weighting function and a directivity function. The weighting function $\widehat{W}$ can be either the rectilinearity or the global polarization parameter:

$$\widehat{W}(t_i) = [F(\lambda_1, \lambda_2, \lambda_3)]^G$$

For this research we employ the rectilinearity, whose effect can be magnified by increasing the exponent G.
The direction of polarization is captured inside the filter operation through the directivity function which applies an exponent H to the direction of polarization.

$$Dp_j(t_i) = (|v_{i,j}|)^H$$

Exponents G and H can be chosen based on filter preferences. The aim is, however, to filter by focusing on the polarization of the ground roll, so the exponent of weighting, G, should be larger than the exponent of directivity, H.
The weighting function and directivity function can both be smoothed by a boxcar time window which is approximately 0.1-0.2 times the correlation window length. Similar to the Hann correlation window, the smoothing window can remove anomalous spikes induced by previous processing steps while being cautious for too much smoothing in the filter operations.
The filtered Z, X and Y components are then determined by the multiplication of the original data with the weighting function and corresponding directivity function:

$$Z_{fl}(t) = Z(t)W(t)^G Dp_Z(t)^H$$
$$X_{fl}(t) = X(t)W(t)^G Dp_x(t)^H$$
$$Y_{fl}(t) = Y(t)W(t)^G Dp_Y(t)^H$$

Because the method is a multiplication of the weighting and directivity function with the data the aim of the method is essentially to minimize the elliptically polarized phases through optimization of the filter operator exponents and the quality factor of rectilinearity. Thus, data can only be removed completely in case of perfect circular or spherical polarization. In such a case $\lambda_1 = \lambda_2$, which means that the rectilinearity and therefore the weighting function would become 0.

4.8 Ellipticity method

The ellipticity method applies the ellipticity to isolate the ground roll and all other elliptical energy through a multiplication of the ellipticity with the data. However, before the multiplication the ellipticity is clipped by a cut-off ellipticity which removes the data less elliptical than the ground roll from the equation through muting.
Because the direction of polarization is disregarded in this method, the ellipticity method focusses solely on the elliptical state of polarization by applying the same weight to each individual component of the data.

$$Z_{ellip}(t) = Z(t)e_{21}(t)$$
$$X_{ellip}(t) = X(t)e_{21}(t)$$
$$Y_{ellip}(t) = Y(t)e_{21}(t)$$

The data is multiplied by a value between the cut-off ellipticity and 1 and therefore requires amplitude balancing before the elliptical data is subtracted from the original data. Therefore, a tapered balancing factor (BF) is employed to more accurately remove the ground roll and enable recovery of overprinted signal while the cut-off between non-elliptical signal and the isolated elliptical signal is reduced. After application of the tapered balancing factor to the elliptical data, the elliptically polarized phases will be subtracted from the original signal:

$$Z_{fl}(t) = Z(t) - Z_{ellip}(t) * BF$$
$$X_{fl}(t) = X(t) - X_{ellip}(t) * BF$$
$$Y_{fl}(t) = Y(t) - Y_{ellip}(t) * BF$$

5. Results

5.1 Data analysis

5.1.1 Synthetic data

As mentioned earlier, the source modeled in the synthetic data from the SPECFEM3D Cartesian model is a vertical force point source at the surface and an identical source at 3 meters depth. From the sections (figure 12) it becomes apparent that the point source in both cases produces a wave-field that is dominated by P- and SV-modes, projected on the Z- and X-component. The energy projected on the Y-axis, on the other hand, is marginal. Only numerical noise is recorded on the Y-component which means that no SH-waves were excited and as a result of that Love waves are absent as well. For this reason the Y-component will not be visualized during further processing steps. Although, however small the influence may be, the processing will continue with all three components.

Looking at the Z- and X-components for both shot records, the resemblance is clear because they employ the same subsurface model. At a first glance the ground roll stands out, as well as the first P- and S-arrivals and some P-wave multiples, indicated in figure 6. However, because of the high amplitude of the ground roll the amplitude of both shot records had to be clipped heavily to make the reflections visible.

Nonetheless, despite the obvious resemblance between both datasets, some distinctions can be observed too. Especially the ground roll and S-reflections appear more distinct and stronger in case of a buried source. Consequently, the dispersive character of the ground roll remains more visible for the surface source model because the ground roll has been influenced less by amplitude clipping.

5.1.2 Preprocessed field data

From the preprocessed dataset several different waveforms can be observed (figure 13). Firstly the ground roll is resolved nicely in the vertical and inline components as the high amplitude waveform. The fan-shaped outline of the ground roll is a result of the dispersive character where especially on the vertical component a fast and slow ground roll can be observed. The ground roll velocity varies from about 100 m/s for the slow ground roll to almost 200 m/s for the fastest ground roll.

Secondly, a high amplitude onset of the guided waves, arriving just before the ground roll can be observed. These guided waves are generated by P- and S-wave multiples in the top layer with high acoustic impedance and is often characterized by the visible shingling caused by reflections of multiples. Lastly, two reflected P-waves can be observed on the vertical component, the first one just before 0.1 second and the second just after 0.1 seconds. These high velocity and relatively low amplitude reflections were resolved by clipping the amplitudes before imaging the sections. Unfortunately, deeper reflections cannot be identified at this point.

This is mainly caused by a low signal-to-noise ratio. Further, a clear bouncing of the drop weight can be observed which creates multiple phases of the ground roll in the data. Unfortunately, the entire shot record bares this signature. Moreover, a strong crossline component is observed as a result of subsurface heterogeneities or an inclination in the metal plate on which the steel ball is dropped. Based on these observations, the distinctions between the field data and synthetic data are evident. Because there are no visible deeper reflections in the data the only recoverable overprinted reflections are the two shallow reflections. In spite of that, the ground roll is dominant throughout the entire section, so both filter methods can be applied to test the ability of filtering the ground roll in low signal-to-noise circumstances.

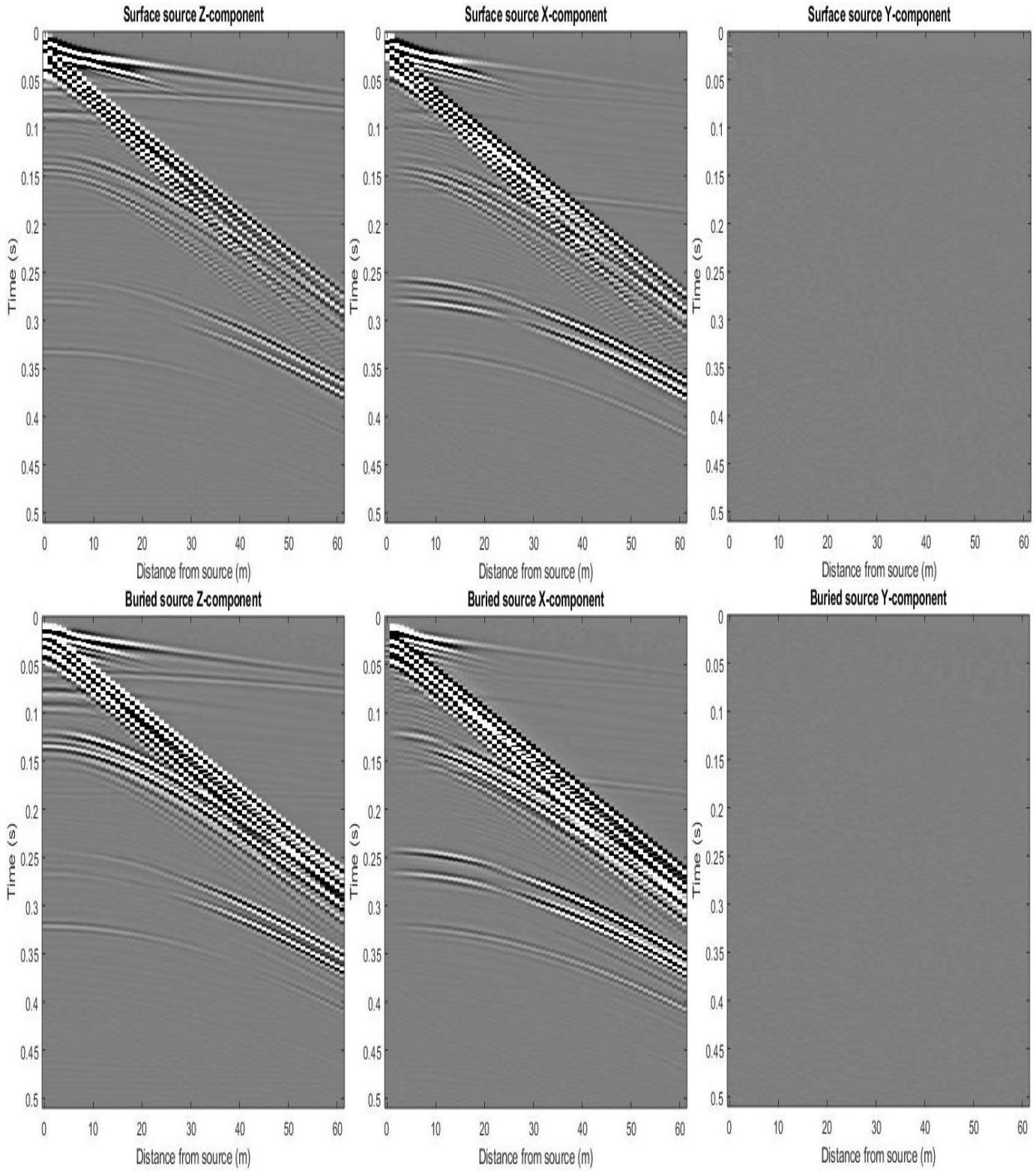

Figure 12: SPECFEM3D Cartesian sections from the individual components for a surface source on the top row and a buried source at 3m depth on the bottom row. The amplitude was clipped equally for both shot records to amplify the reflections.

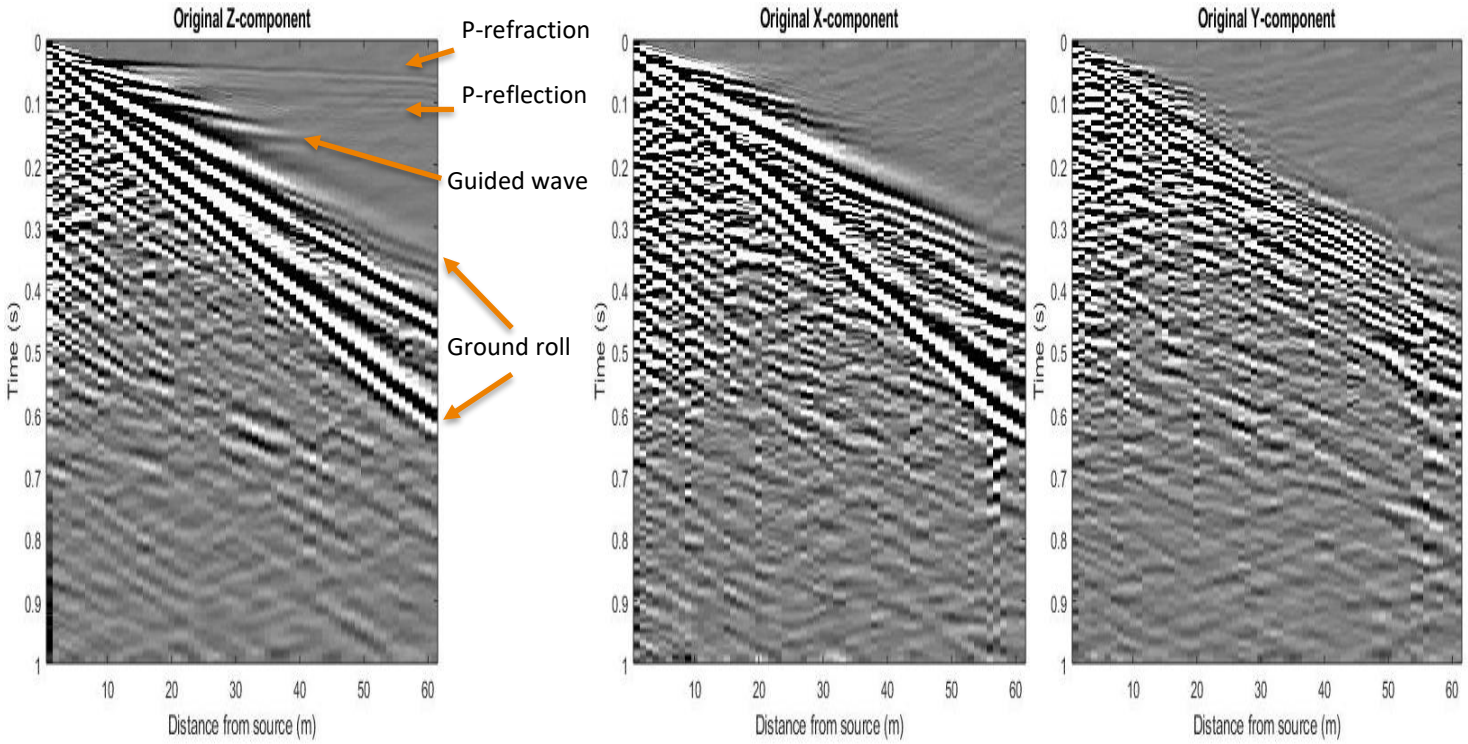

Figure 13. Scaled sections of the three components from the field data after preprocessing with the reflections, guided wave and ground roll fan indicated by the arrows.

5.2 Data control

5.2.1 Synthetic data

By imaging the particle motion of certain arrivals in hodograms (figure 14) the energy distribution, state of polarization, adequate correlation window length and previous observations can be seen. The hodograms in figure 14 represent the particle motions of several arrivals at a source distance of 30 meter with a window length of just over one ground roll wavelength.

Looking at the hodograms from the synthetic data the source depth does not seem to influence the state of polarization or required correlation window length significantly whereas the amplitudes vary greatly. Where the vertical P-arrivals are stronger for a surface source, the S-waves are much stronger for a buried source. Further, the ground roll is even three times stronger for a buried source than a surface source.

Looking at the state of polarization of the arrivals it is clear that the sharp velocity contrast of the top layer causes the body waves to bend sharply to the vertical normal in the top layer when crossing the interface from below. Therefore, the particle motion of P-waves is directed along the vertical component while the particle motion of Sv-waves is along the inline-component.

Since the ground roll is generated through the interference of P- and Sv-waves caused by the sharp velocity gradient at the interface at 2 meters depth, the particle motion of the ground roll is a retrograde planar ellipse, projecting energy along the vertical and inline component.

These observations also verify the previous assumptions made on the basis of visible arrivals on the sections.

5.2.2 Field data

From the previously discussed sections a few observations were made on several phases in the field data. The first of those is the ground roll, which was clearly visible on the inline and vertical component. Judging from the hodogram (figure 15), this assumption is most likely true as the particle motion displays an elliptical, almost circular, rotation on the x- and z-axis, although a clear crossline component is visible too.

The second high amplitude arrival is the guided wave which is linearly polarized in the x-z plane. The observed shingling character in the sections helps to further substantiate the interpretation of the guided wave.
The last discussed arrivals are the expected P-wave reflections observed on the vertical component of the section. The hodogram of these arrivals shows a clear dominant polarization along the vertical axis, so the previously made assumptions on all three arrivals seem to be verified, based on the particle motion and display on the sections.
Looking at the amplitudes in the data, the ground roll is over 20 times stronger than the reflections. This means that the energy contrast between the ground roll and reflections is similar to the energy contrast in the surface source synthetic data.
Though the results from the hodograms look reliable, selection of the correlation time window is not straightforward for the field data. As can be observed from the frequency spectra of the field data and synthetic data (figure 7 and 9), the dominant frequency in the field data is much lower than the frequencies observed in the synthetic data. Therefore, the distinction between the period of low frequency, high amplitude surface waves and high frequency, low amplitude body waves is larger in the field data.
Therefore, the length of one ground roll period is longer for the field ground roll than for the synthetic ground roll, which implies that a longer correlation window length is required for the field data. Consequently, the resulting correlation window length for the field data records one ground roll period, but both reflections can be captured in one time window as well. This may cause smearing in the data in further processing steps. So, a clear trade-off between capturing the polarization of the ground roll and body waves has to be made in case of a broader frequency spectrum.

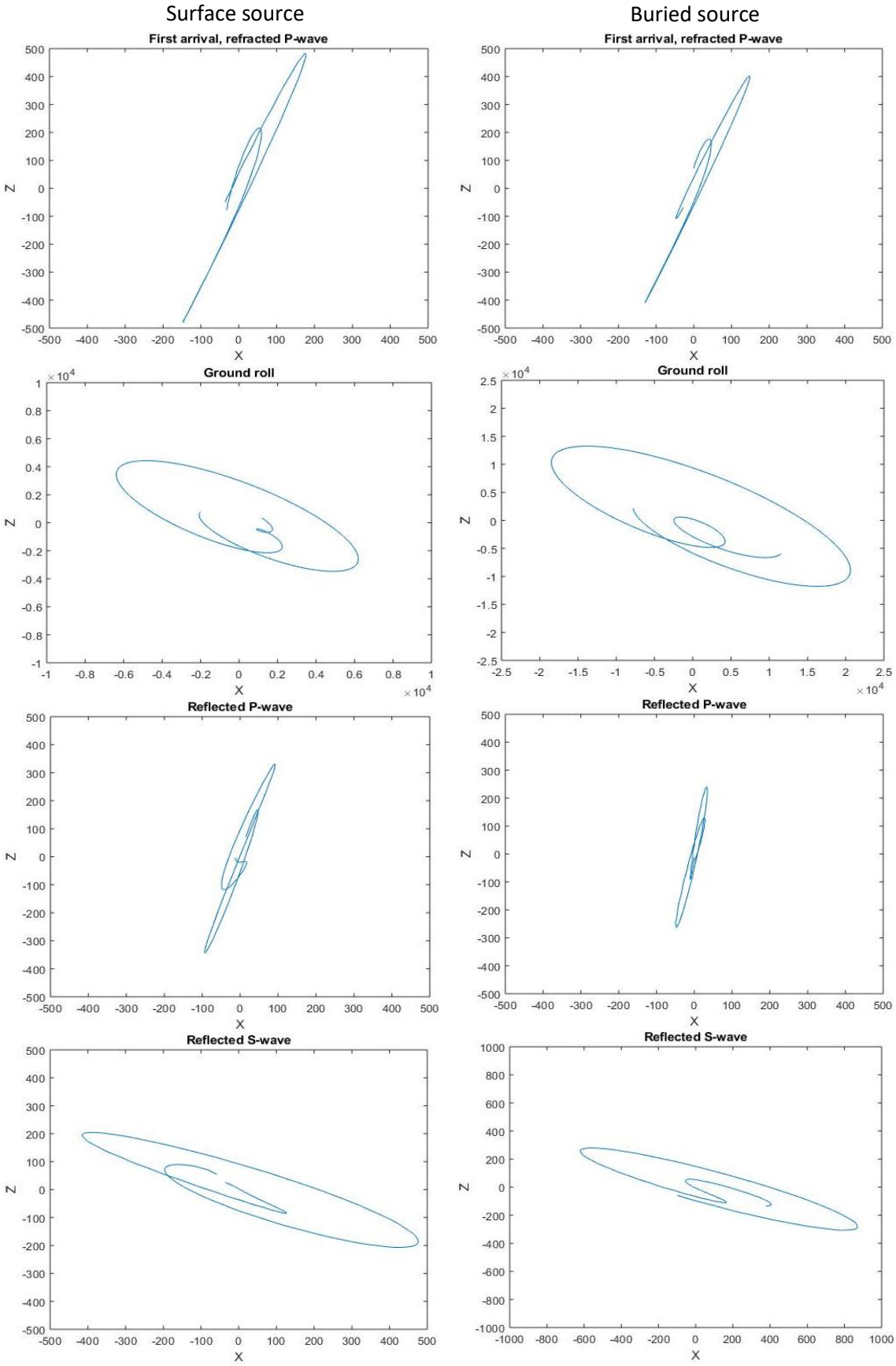

Figure 14. In the left column, surface source hodograms of the first arrival (t=0.027-0.045s), 30m interface P-reflection (t=0.054-0.072s), Ground roll (t=0.126-0.144s) and 30m interface S-reflection (t=0.264-0.282s). In the right column, buried source hodograms of the first arrival (t=0.033-0.051s), 30m reflected P-wave (t=0.057-0.075s), Ground roll (t=0.135-0.153s) and reflected S-wave at the 30m interface (t=0.279-0.297s). The hodograms were all plotted with a window length of just over one ground roll period, corresponding to the correlation window length.

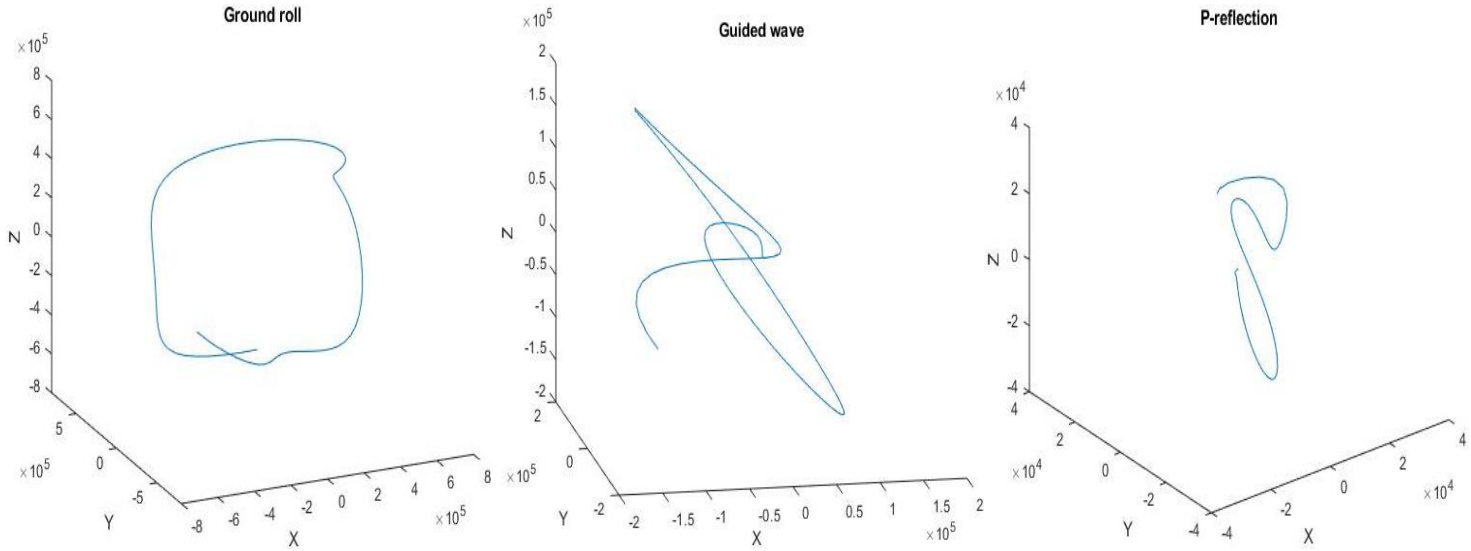

Figure 15. The ground roll (t=0.18-0.22s), guided wave (t=0.065-0.105s) and P-reflection (t=0.02-0.06s) are the field data hodograms for the observed phases in figure 13 at a source distance of 20 meter.

5.3 Polarization attributes

The polarization attributes are calculated from the results of the covariance analysis and eigendecomposition. The resulting polarization attributes, thus, represent an accurate representation of the quality of those steps**.** Additionally, the correlation window length can be examined to verify that the state of polarization in the data is accurately captured while smoothing of high frequency body waves is prevented. The polarization attributes themselves can be proofed by comparing the observed direction and state of polarization in the hodograms to the results from the section below.
To exemplify the polarization attributes the results for the buried source synthetic data are visualized in this section. Further results of the polarization attributes are displayed in appendix A and B. A few of these are also used to clarify the results of the two approaches to the polarization filter.

5.3.1 Direction of polarization

Through the varying principal directions in the data a clear distinction can be made between different phases. Because of the high velocity contrast between the surface layer and the layer underneath, the linearly polarized body waves should be strongly polarized along one component, whereas the principal component of the ground roll most likely has both strong X- and Z- components. Examination of the hodograms previously confirmed those expectations already. This implies that the direction of polarization is a practical tool to separate different types of body waves which enables the user to perform an individual P-wave and S-wave analysis.
As a first example the X- and Z-component of the principal eigenvector for the buried source synthetic data are displayed in figure 16. The ground roll has two relatively strong components with the inline component (x≈0.8) slightly dominant over the vertical component (z≈0.6). The high velocity P-waves, arriving before the ground roll are vertically polarized and the low velocity S-waves, arriving after the ground roll, are clearly polarized along the inline component.
In the bottom part of the sections the signal becomes so weak that numerical noise becomes the dominant signal, causing the direction of the dominant eigenvector to shift rapidly.

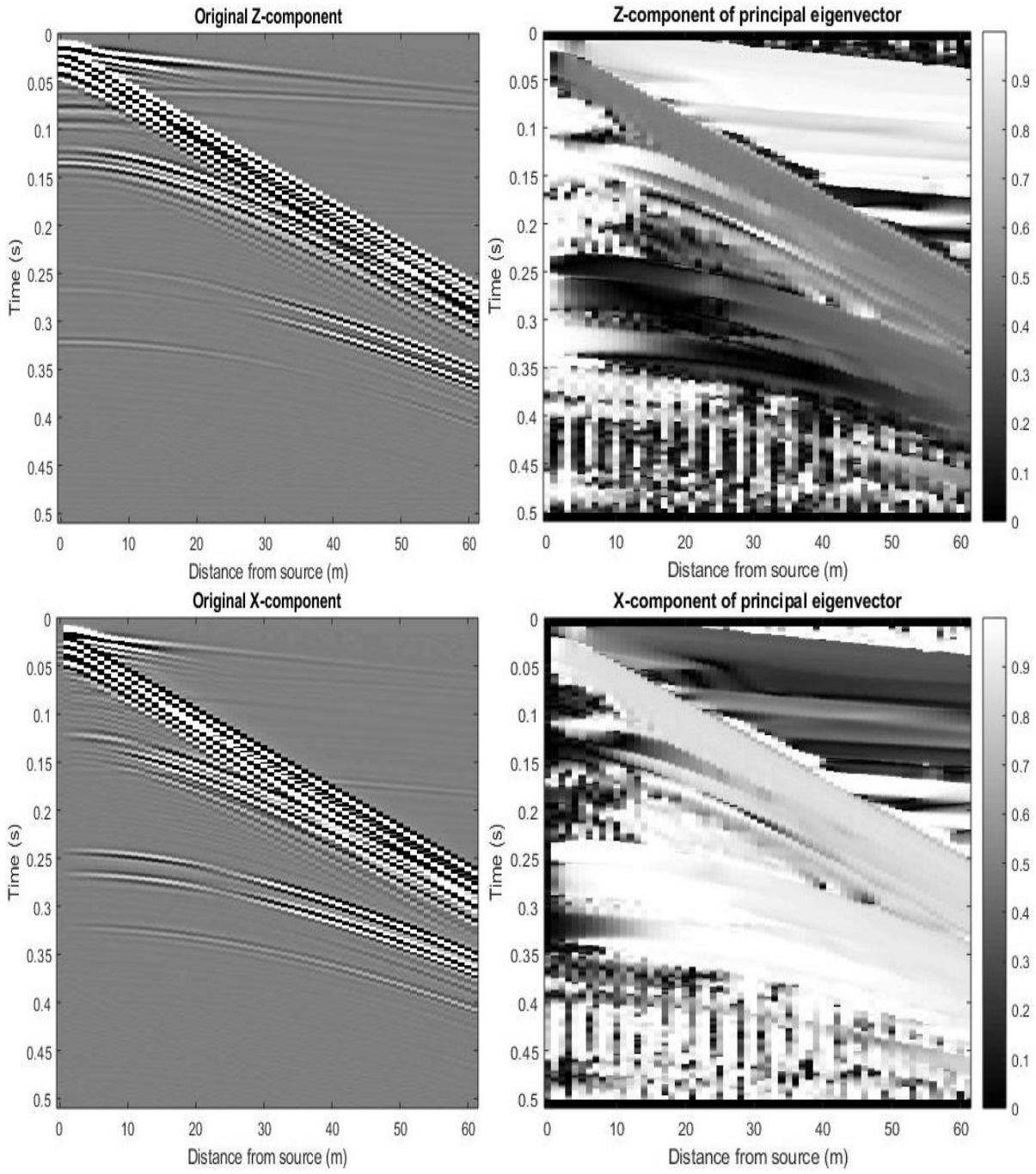

Figure 16. Z- and X-component of the direction of polarization compared to their corresponding original section.

5.3.2 Rectilinearity

Since the rectilinearity is a measure for the linear state of polarization, the results show a clear distinction between linearly polarized body waves and elliptical ground roll. The quality factor (Q) can aid to further amplify this distinction (figure 17). The smaller the quality factor the clearer the distinction between linear and elliptical phases becomes. However, if Q is too small the rectilinearity of non-perfectly linear arrivals start to decay severely which will result in a loss of usable reflections in a rectilinearity based filter. Especially in field data this would cause serious harm to the dataset. Contrarily, if Q is too big the distinction between the ground roll and body waves start to fade which would render rectilinearity filtering useless.

Nonetheless, for all values of the quality factor previously identified arrivals can be distinguished based on their state of polarization. The previously identified P- and S-arrivals are clearly distinct as linearly polarized, even though they appear to have been smoothed slightly as a result of the correlation window. The ground roll is clear as well, although for larger quality factors the distinction between the ground roll and linearly polarized arrivals becomes relatively small. The elongated

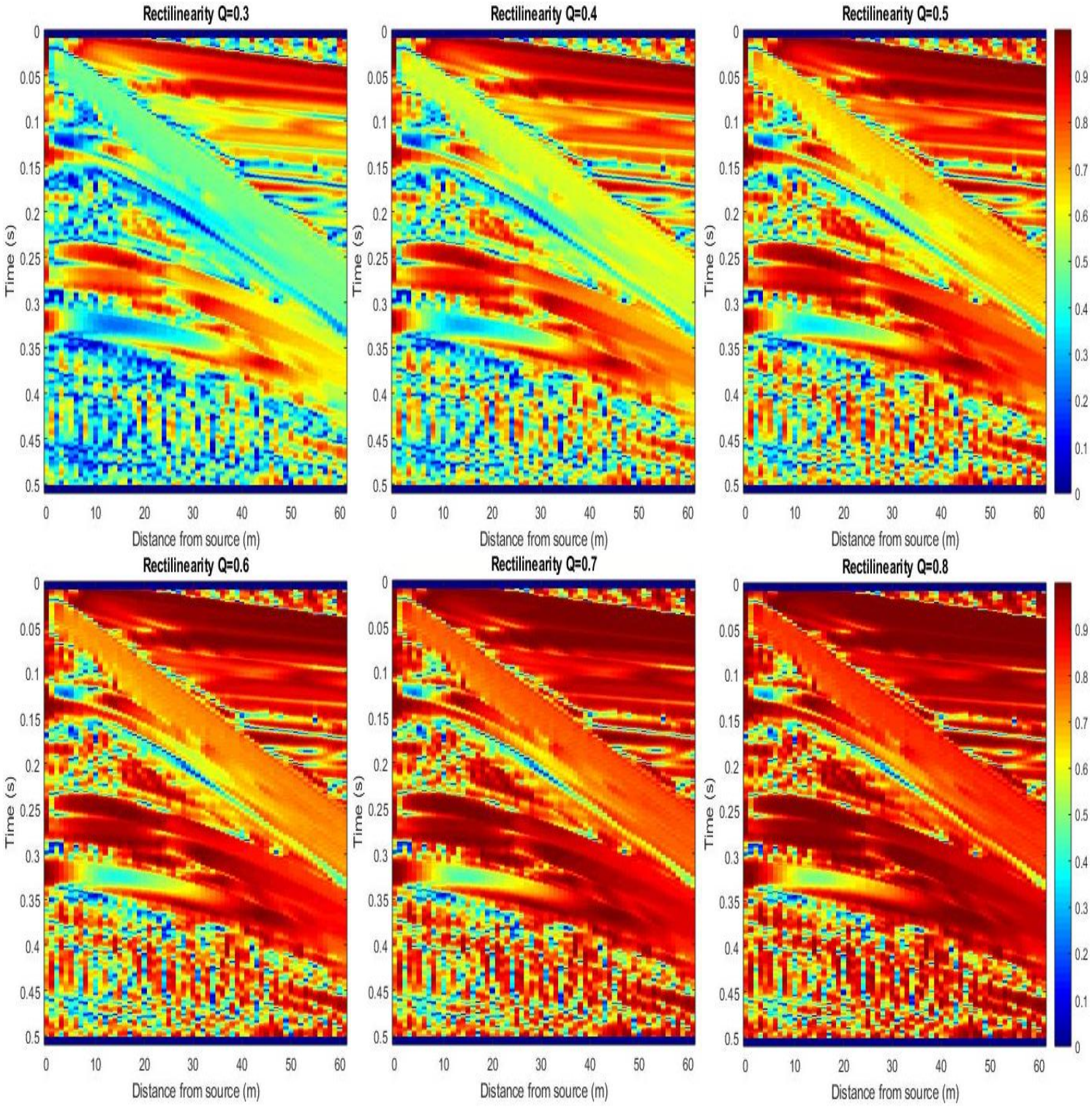

Figure 17. Representation of the influence of the quality factor (Q) on rectilinearity.

elliptical polarization observed in the ground roll hodogram (figure 14) already suggested that the rectilinearity of the ground roll would be relatively high for high quality factors.
Not only the surface waves show non-rectilinear behavior, but also two as S-wave identified arrivals temporarily demonstrate a very low rectilinearity. The change from rectilinear to elliptical state of polarization is especially clear in the arrival just after 0.3 seconds for a source distance of 5 to 30 meters. The arrival from about 0.12 seconds shows similar behavior for a source distance of 2 to 10 meters. The state of polarization for both arrivals clearly changes from rectilinear to elliptical and then back to rectilinear. Looking at the direction of the principal eigenvector from the previous section (figure 16), the change of polarization is visible as well and appears to have been caused by a shift from vertical to inline polarization as distance from the source increases. This appears to be a P- to S-wave conversion as a result of the angle of incidence at the 2 meter interface from below. However, the elliptical particle motion accompanying the shift in principal direction cannot be explained.
Another interesting observation that can be made is an almost circular recording visible as a dark blue line following the ground roll and converted S-wave. The model or direction of polarization does not suggest an elliptically polarized arrival is supposed to be recorded at that specific part of the section. Therefore, the elliptical particle motion may be an artefact of wave interference between the ground roll and the converted S-wave. Especially interesting about this observation is that the recorded data will definitely be removed in both polarization filters while it may actually be just an artefact of wave interference.

5.3.3 Global polarization parameter

From a first look at the result from the global polarization parameter (figure 18) the section resembles the higher order rectilinearity sections to a great extent.
Though, the global polarization parameter does resolve the identified arrivals adequately, in this case the rectilinearity is favorable because of the ability to alter the quality factor and the possibility to only invoke a third component if necessary.

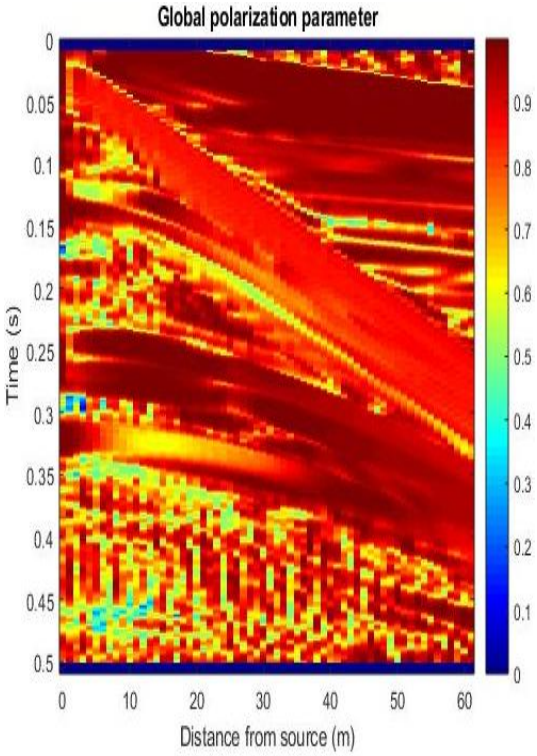

Figure 18. Global polarization parameter

5.3.4 Ellipticity

Since the ellipticity is the factorized division of the two dominant eigenvalues, the ellipticity is basically an opposite operation of the rectilinearity. Consequently, similar statements about the quality factor are applicable for both methods, so the quality factor can again be altered to better distinguish between elliptical and linear arrivals. Where the linear arrivals have a rectilinearity close to 1, they have an ellipticity close to 0. Similarly for the elliptical to circular arrivals the state of rectilinearity decreases depending as the ratio between the principal eigenvalues increases whereas the ellipticity simultaneously increases.
As a result of this opposite approach, in this case a small quality factor amplifies the elliptical nature of certain arrivals such as the ground roll while a high quality factor impedes polarization analysis (figure 19).
Because the ellipticity is used in a polarization filter with a completely different approach, the quality factor does not necessarily have to maximize the ground roll relative to the linear arrivals. The filter takes all arrivals above a certain value or state of ellipticity, so the quality factor is the key component to determine which portion of the data is going to be filtered.

The visible waveforms discussed in the results of the rectilinearity section are similarly visible in the results below. The same counts for the unexpected highly elliptical arrivals which will also be attenuated by the filter.

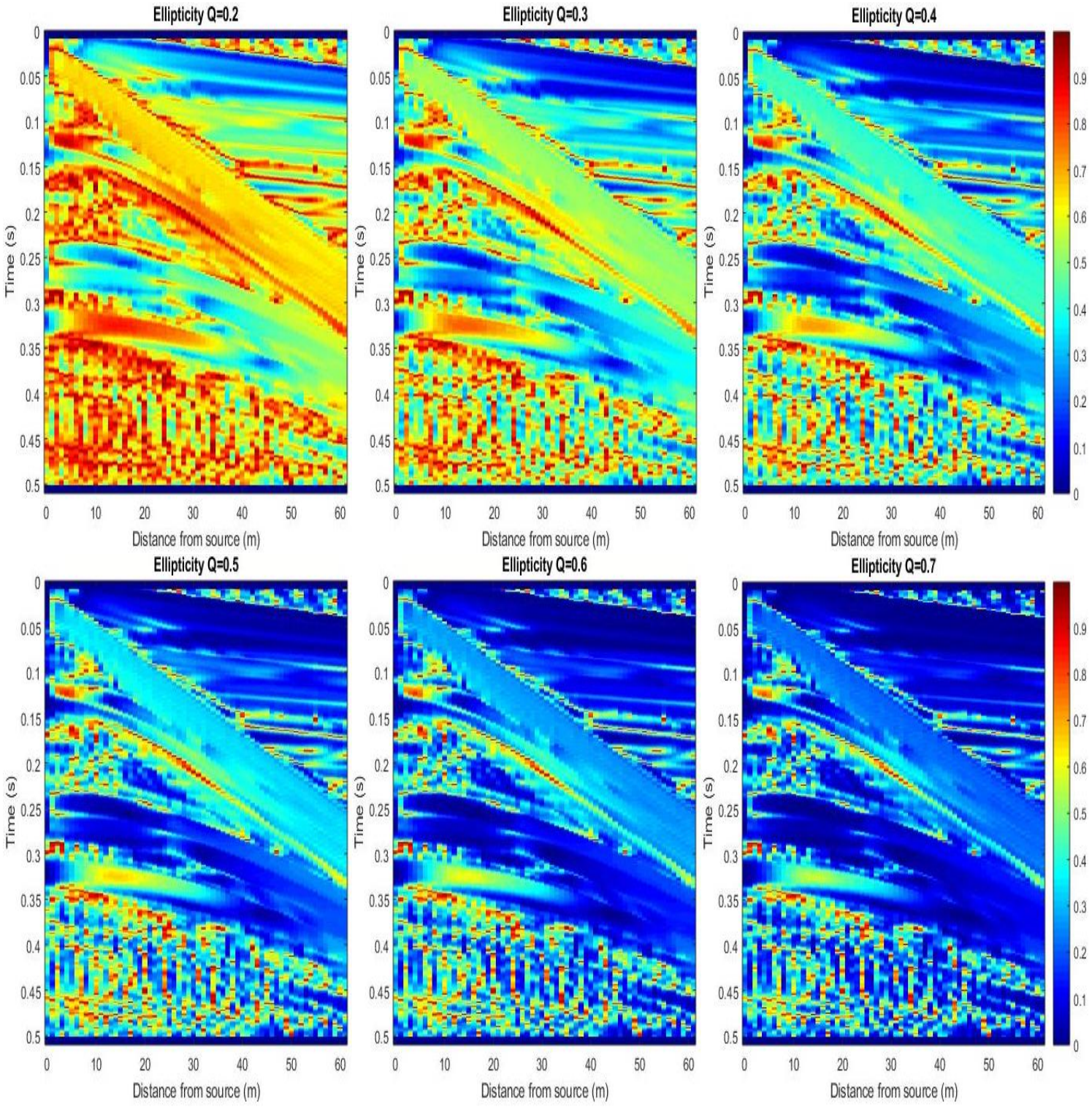

Figure 19. Representation of the influence of the quality factor (Q) on ellipticity.

5.4 Linearity method

As mentioned in the methods section, the linearity method consists of applying a weighting function and a directivity function which are multiplied with the original data.

The rectilinearity is employed with a quality factor Q=0.4. For this quality factor the linear to semi-linear arrivals should be preserved well while the ground roll (Rl= 0.5-0.6) and other elliptical arrivals are distinctive to a degree where they can largely be removed by the exponent of weighting.

Expected results are thus that filtering strongly by rectilinearity removes the parts of the data that contain non-linear arrivals while those parts that only contain linear arrivals are only mildly affected.

The component of the principal eigenvector is multiplied with the corresponding component of the original data. The direction of polarization focusses the filter on arrivals with the principal direction of polarization along other components. If the principal component of an arrival is parallel to one of the spatial axes the arrival will be unharmed by the direction of polarization along that axis.

So the linearity method invoking both the rectilinearity and directivity leaves especially linearly polarized reflections parallel to a specific component unaffected. For instance the linearly polarized P-waves arrive with a vertical angle of incidence and would thusly be unaffected by the filter on the Z-component while the S-waves, ground roll and all other energy on this component is, to a large extent, removed. Conversely, the deviation from the vertical axis from these P-arrivals should be removed as a result of the weak x- and y-component in the principal eigenvector.

From these statements it is clear that if a body wave is not polarized along a spatial axis, the direction of polarization becomes inapplicable and the filter should completely rely on the weighting function instead.

Figure 20 is a visual representation of the filtering process. Variations in the result of this sample to sample multiplication can be made by varying the exponents of the rectilinearity (G) and direction of polarization (H) or in other words, the weighting function and directivity.

The choice of the exponent values are essential for the linearity method to function properly and specifically remove data based on personal preference. For example larger exponents lead to stronger filtering. Likewise, the exponents can be altered relative to each other, based on whether the state or direction of polarization is considered more important. To illustrate the implications of the exponents to the filtering, multiple simulations were run for each dataset.

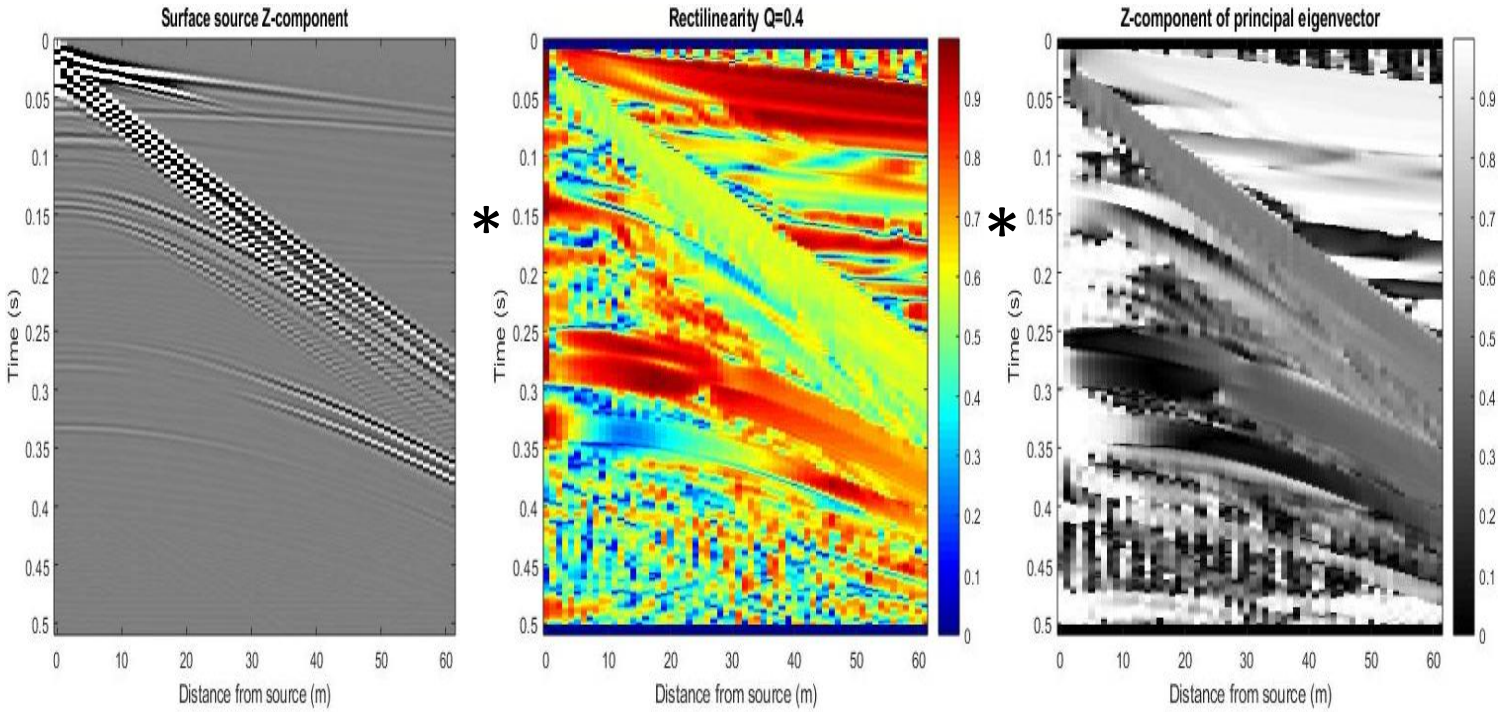

Figure 20. Visual representation of the multiplications in the linearity method. The Rectilinearity and the component of the principal eigenvector represent the weighting function and directivity function.

5.4.1 Synthetic data

Surface source

The results for the linearity method applied to the surface source synthetic dataset are within the line of expectation (figures 21 and 22). As was observed from the hodograms, the source induced ground roll is in both absolute and relative energy less dominant over the section in case of a surface source relative to a buried source. This should aid the effectiveness of the filter and create better results for lower exponents of the filter operations. Nonetheless, the ground roll is still severely dominating the desired reflections. From the results in figures 21 and 22 can be observed what the impact is of different configurations of the exponents of weighting and directivity. For instance, to observe the effect of filtering based on the weighting function compared to directivity filtering. From these first two sections (figures 21a and b and 22a and b) can be seen that the results for the vertical component appear very much alike while the results on the inline component look very different. Clearly the linear arrivals from the body waves are well preserved in the rectilinearity filter while on the components of the principal eigenvalue the P-waves are only preserved on the vertical component and removed on the inline component and vice versa for the S-waves. Directing the attention to the ground roll, most surface wave energy can already be removed by invoking an exponent of G=5 to the weighting function. The surface waves are clearly too dominant along the inline component to properly remove the ground roll on the x-component by directivity filtering whereas the values of the vertical component of the principal eigenvalue and rectilinearity are approximately the same (figure 20). Therefore, the filter has about the same effect on the ground roll for both filter operations on the vertical section.

Combination of the filter operations (figures 21c & 22c) then results in an almost complete removal of the vertical component of the ground roll. The directivity removes the S-wave energy while the linear P-arrivals are preserved by both weighting and directivity. For the body waves on the inline component the opposite effect is caused by the directivity. The combination of the filter operations does, however, not cause the inline ground roll to be removed more accurately since the inline directivity function hardly affects the ground roll.

Therefore, mere removal of surface waves and other elliptically polarized observations can be accomplished by employing a filter based on only the rectilinearity (figures 21d & 22d). For almost complete removal of the surface waves on both components as well as a clear separation of the body waves a strong weighting filter in combination with a mild directivity filter will accurately separate different wave modes (figures 21e & 22e).

These observations can be validated by visualizing the remaining amplitudes and polarization state of the initially imaged arrivals. By comparison of the hodograms before and after filtering a precise analysis can be made on the impact of the linearity method on the data. From previous observations it is straightforward that only the linearly polarized phases which line up perfectly to one component are completely unaffected by the linearity filter when both filter operators are implemented. Therefore, it is obvious that even the linearly polarized P-waves lose a substantial amount of their energy. In spite of that, the elliptically polarized arrivals are attenuated more severely and depending on the configuration of the filter operators the results vary significantly. From figures 23-25 it becomes clear that focusing completely on the state of polarization to equal filtering of both weighting and directivity leads to substantially different results, but they were all able to remove the ground roll to a great extent. From the filtered sections it was already evident that a linearity filter based on the direction of polarization leaves a large portion of the ground roll almost unaffected because of its dominant inline energy. The main result from the directivity filter is that it has completely linearized the body waves along a preferred axis and that only a clear inline component of the ground roll remains.

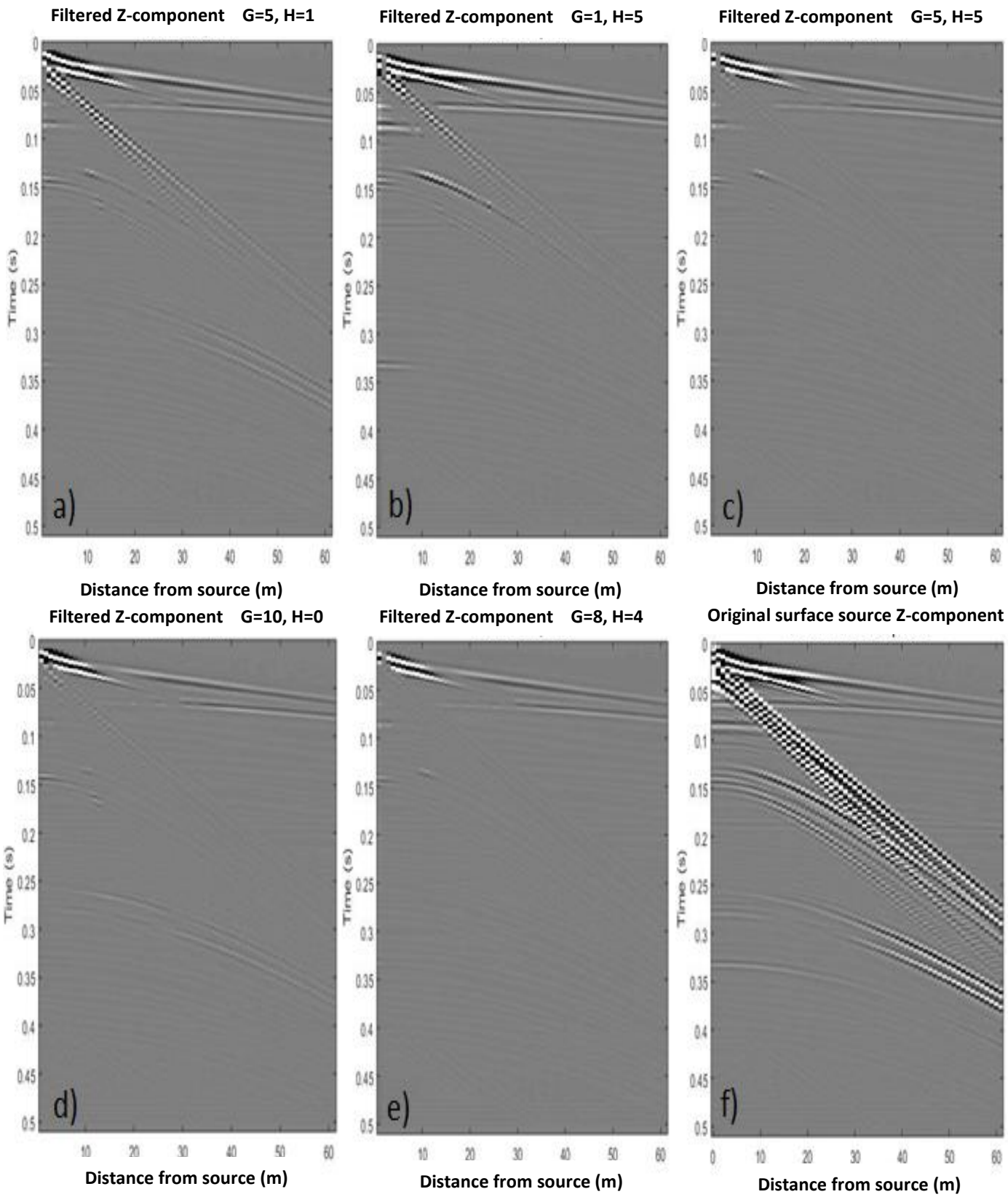

Figure 21. Results of the linearity method for the vertical component from the surface source synthetic data. Quality factor Q=0.4 for all filtered results. a) rectilinearity filtering G=5, H=1. b) Directivity filtering G=1, H=5. c) Linearity filtering with equal filter parameters G=5, H=5. d) Strong rectilinearity filtering G=10, H=0. e) Linearity filter with dominant rectilinearity G=8, H=4. f) Original data. All filtered results were similarly imaged with clipped amplitudes to remove the dominant amplitudes at near offset. The amplitudes of the original data were clipped to a twice as high amplitude.

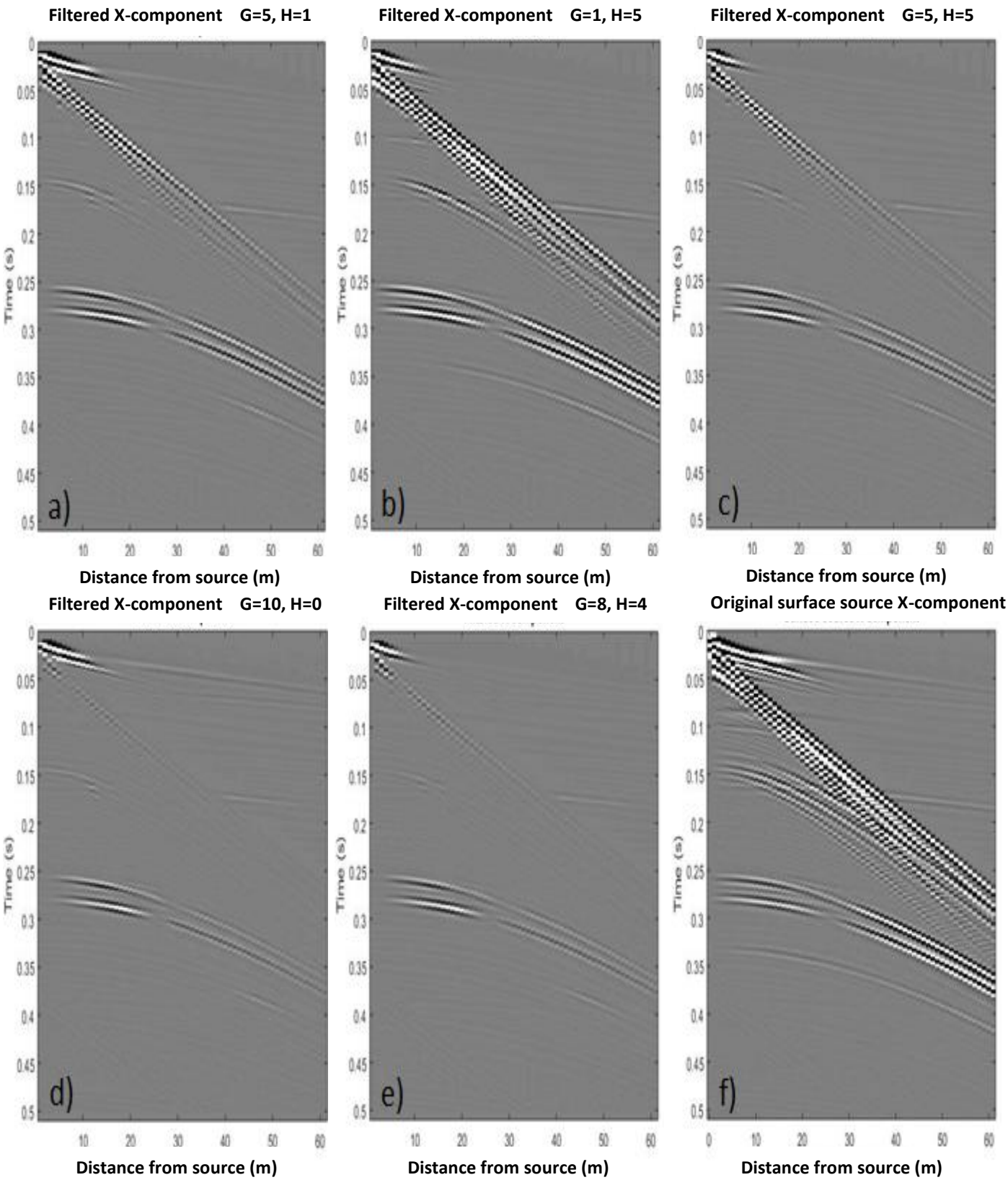

Figure 22. Results of the linearity method for the inline component from the surface source synthetic data. Quality factor Q=0.4 for all filtered results. a) rectilinearity filtering G=5, H=1. b) Directivity filtering G=1, H=5. c) Linearity filtering with equal filter parameters G=5, H=5. d) Strong rectilinearity filtering G=10, H=0. e) Linearity filter with dominant rectilinearity G=8, H=4. f) Original data. All filtered results were similarly imaged with clipped amplitudes to remove the dominant amplitudes at near offset. The amplitudes of the original data were clipped to a twice as high amplitude.

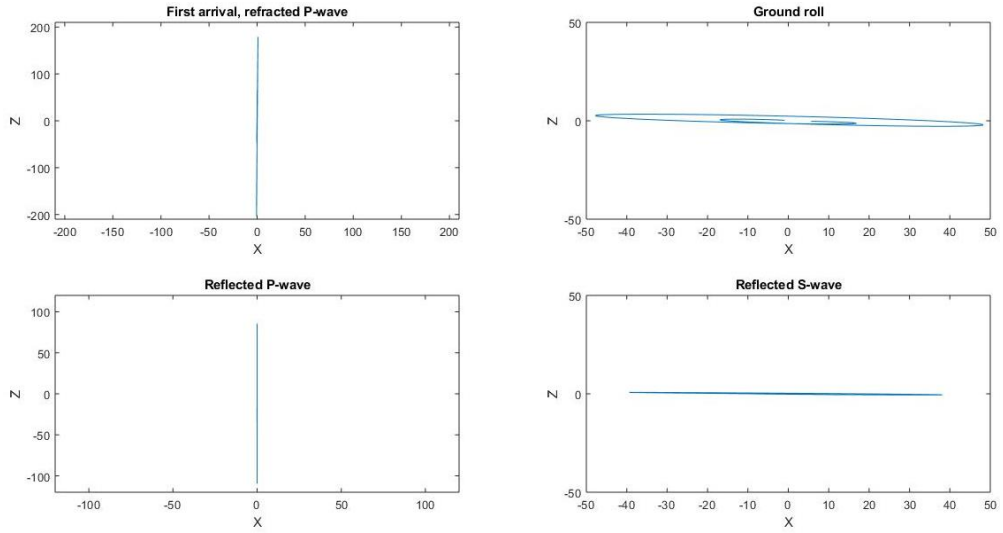

Figure 23. Hodograms of the same arrivals imaged in figure 14, after applying the linearity method with a dominant weighting function (G=8) and a mild directivity filter (H=4) for the surface source synthetic data.

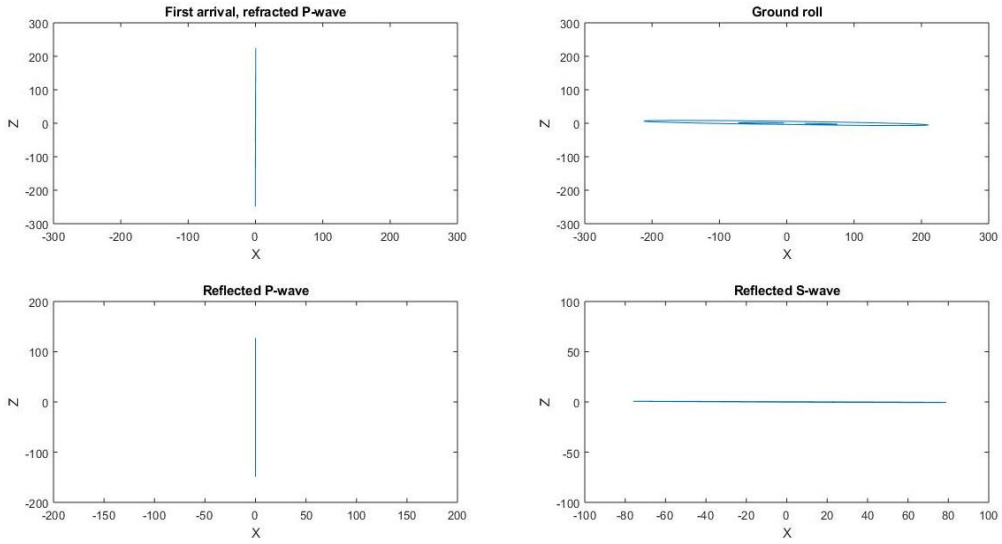

Figure 24. Hodograms of the same arrivals imaged in figure 14, after applying the linearity method with equal weighting and directivity exponents (G=H=5) for the surface source synthetic data.

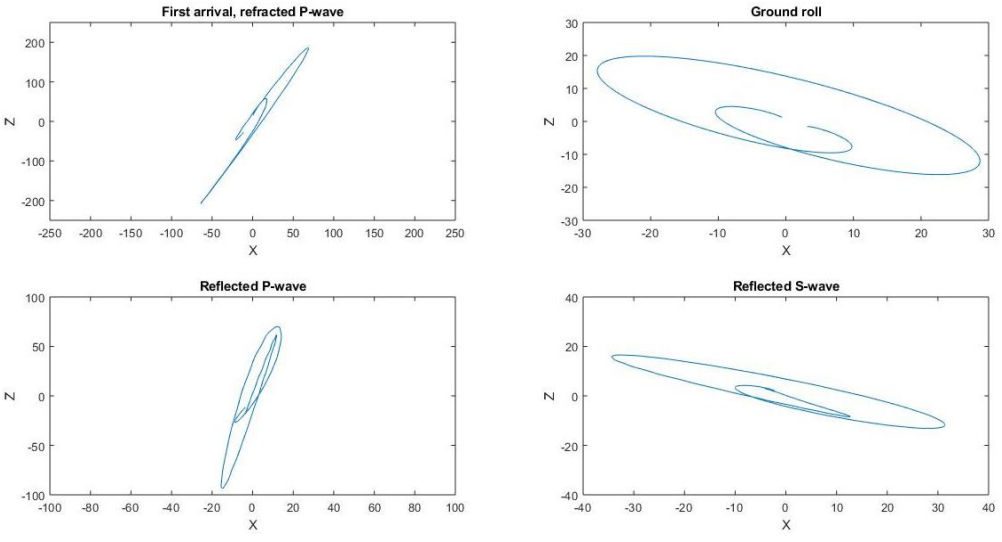

Figure 25. Hodograms of the same arrivals imaged in figure 14, after applying the linearity method based on only the weighting function (G=10, H=0) for the surface source synthetic data.

On the other hand, the state of polarization of the individual arrivals is preserved by the weighting function (figure 25), making it essentially a polarity based amplitude attenuation while doing no further harm to the data. This can be seen by implementing a strong weighting function to the data. By this method all three components are multiplied by the same value of rectilinearity for each time and station. This results in a clear amplitude cut-off, especially for the ground roll and S-waves which are more elliptical.
Then for the linearity filter with equal operators (figure 24) a distinct ground roll components along the x-axis remains, as expected, and due to the directivity the arrivals are all linearly polarized towards their strongest component of the principal eigenvalue. Therefore, an alteration of the filter components towards more dominant weighting (figure 23), can remove both the state of polarization and a large portion of the ground roll.

Buried source
It was already observed that compared to the surface source, the S-wave and surface wave energy are two to three times stronger and the P-wave energy is a slightly weaker. Consequently, the linearity filter will be less effective, because the state of polarization is approximately the same for both sources, thus the filter operations are approximately the same for both sources as well.
For unaltered exponents of weighting and directivity the distinction between the effectiveness of the linearity method to both synthetic datasets is obvious. For equal parameters the ground roll signature over the sections is much more visible because of the amplitude contrast. Because the linearity method applies similar multiplications to both datasets, the initial amplitude contrast remains unchanged after filtering. Therefore, the filter operators would have to be more powerful for the buried source to reduce the ground roll equally. However, if the exponents were to be increased even further, more P-wave energy would be removed as well, leading to an overall decimation of energy.
Nonetheless, a lot can be observed from the sections (figures 26 & 27) and hodograms (figures 28-30) of the buried source data as well. Especially that the linearity method does not remove enough surface wave energy for mild filter operators and not even for the case of equal filter operators (figure 28) where particularly a lot of inline component ground roll persists. For stronger weighting functions the results already become significantly better as the first P-arrivals become the dominant energy (figures 29 & 30).
For the buried source the weighting function seems even more effective in minimizing the ground roll compared to the directivity function then it was for the surface source. For even stronger weighting operators the ground roll could be removed even further while the P-waves would remain relatively strong. Therefore, the relative energy distribution would mainly be affected and what remains would be low amplitude, linearly polarized arrivals.

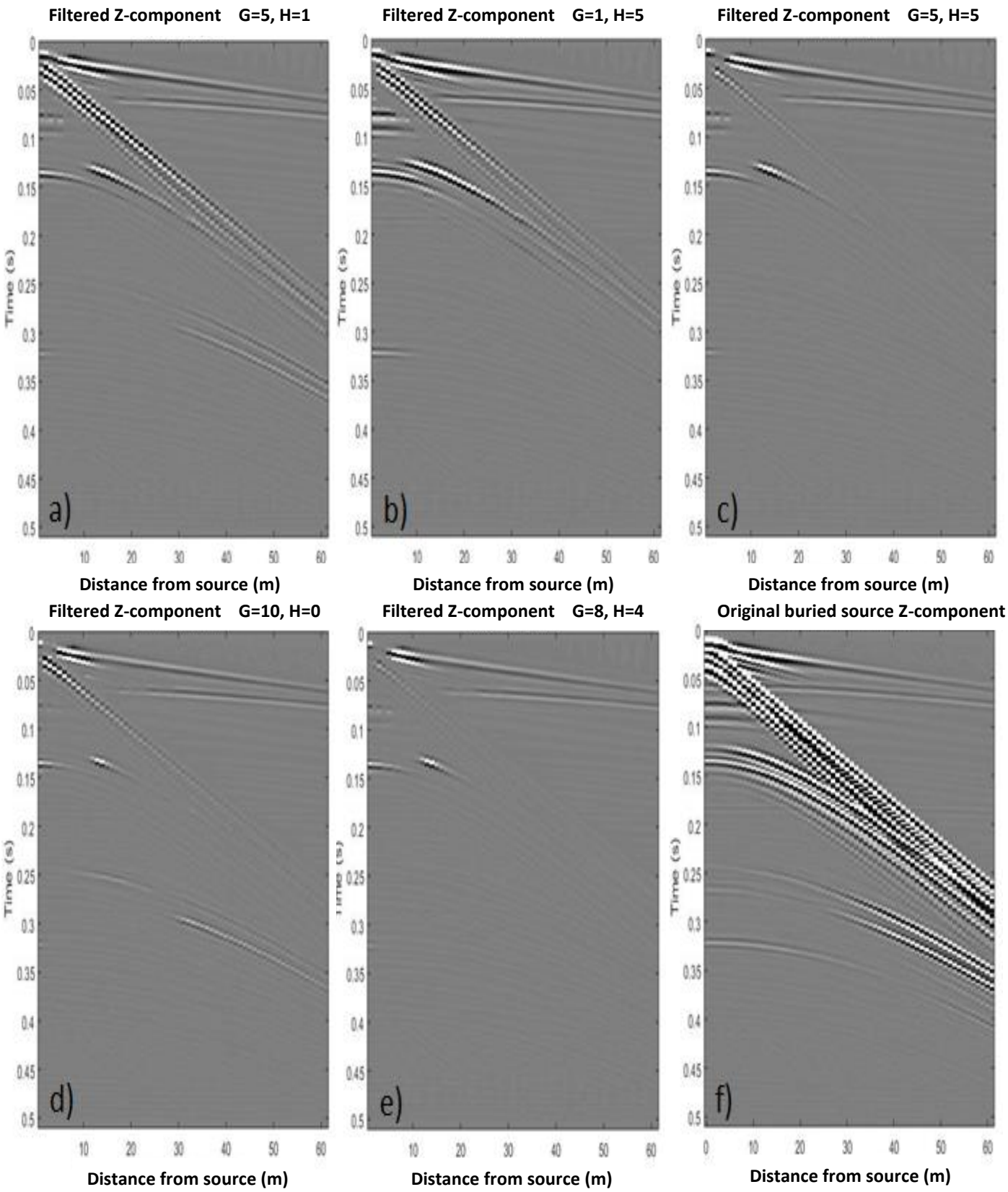

Figure 26. Results of the linearity method for the vertical component from the buried source synthetic data. Quality factor Q=0.4 for all filtered results. a) rectilinearity filtering G=5, H=1. b) Directivity filtering G=1, H=5. c) Linearity filtering with equal filter parameters G=5, H=5. d) Strong rectilinearity filtering G=10, H=0. e) Linearity filter with dominant rectilinearity G=8, H=4. f) Original data. All filtered results were similarly imaged with clipped amplitudes to remove the dominant amplitudes at near offset. The amplitudes of the original data were clipped to a twice as high amplitude.

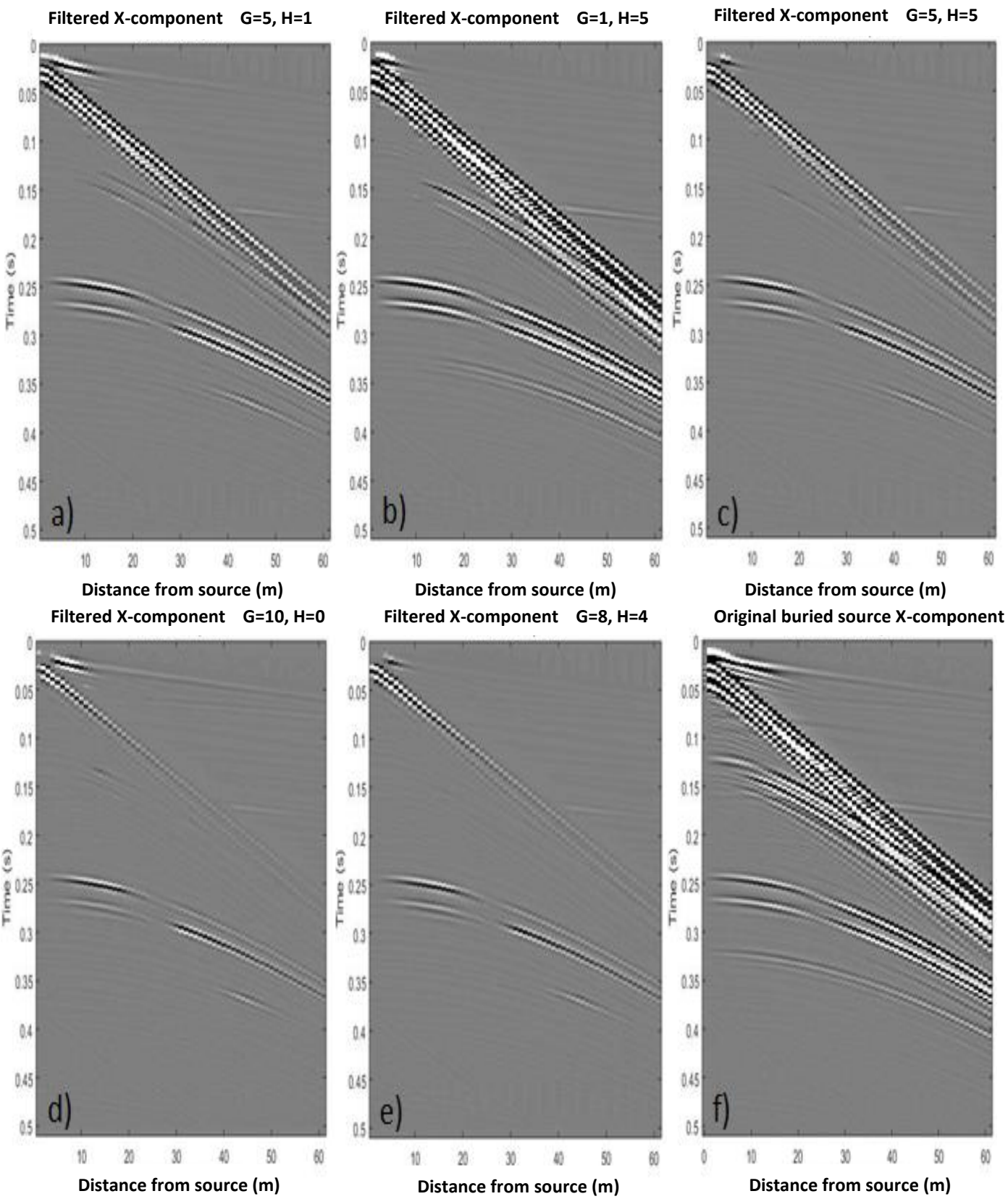

Figure 27. Results of the linearity method for the inline component from the buried source synthetic data. Quality factor Q=0.4 for all filtered results. a) rectilinearity filtering G=5, H=1. b) Directivity filtering G=1, H=5. c) Linearity filtering with equal filter parameters G=5, H=5. d) Strong rectilinearity filtering G=10, H=0. e) Linearity filter with dominant rectilinearity G=8, H=4. f) Original data. All filtered results were similarly imaged with clipped amplitudes to remove the dominant amplitudes at near offset. The amplitudes of the original data were clipped to a twice as high amplitude.

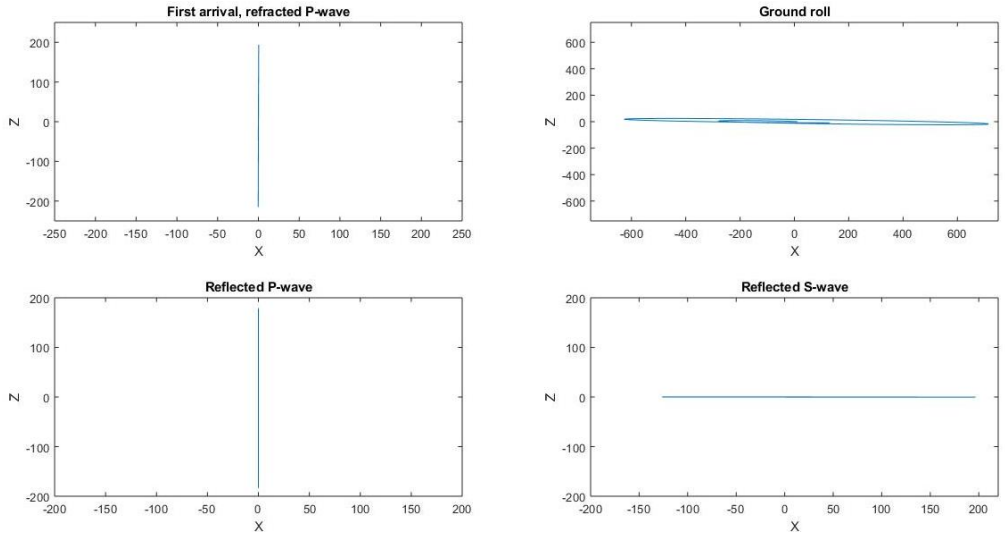

Figure 28. Hodograms of the same arrivals imaged in figure 14, after applying the linearity method with equal weighting and directivity exponents (G=H=5) for the buried source synthetic data.

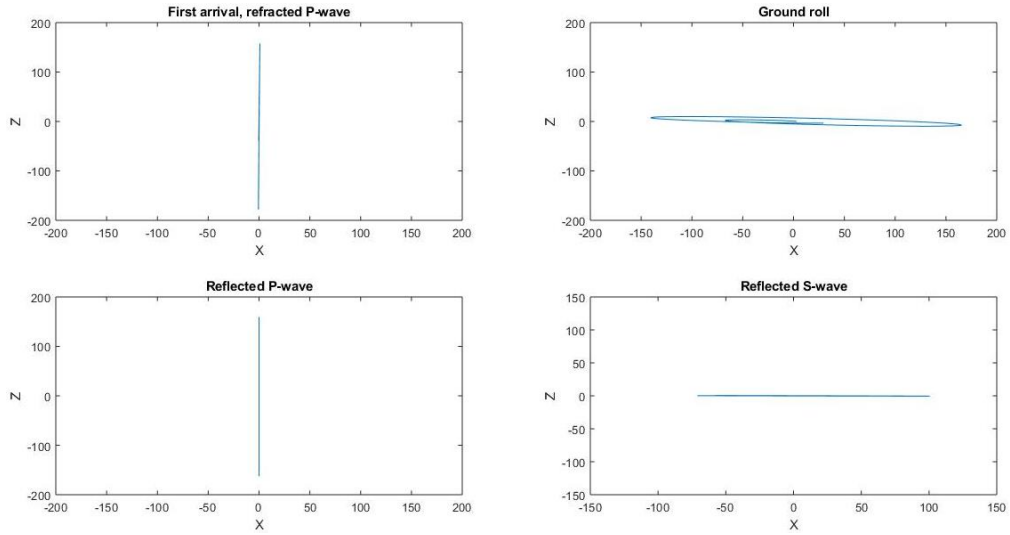

Figure 29. Hodograms of the same arrivals imaged in figure 14, after applying the linearity method with a dominant weighting function (G=8) and a mild directivity filter (H=4) for the buried source synthetic data.

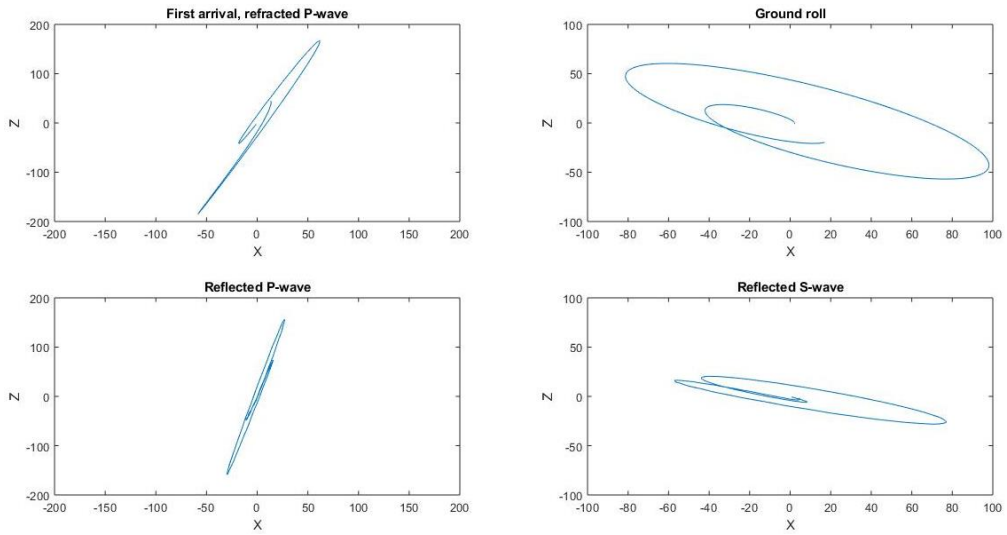

Figure 30 Hodograms of the same arrivals imaged in figure 14, after applying the linearity method based on only the weighting function (G=10, H=0) for the buried source synthetic data.

5.4.2 Field data

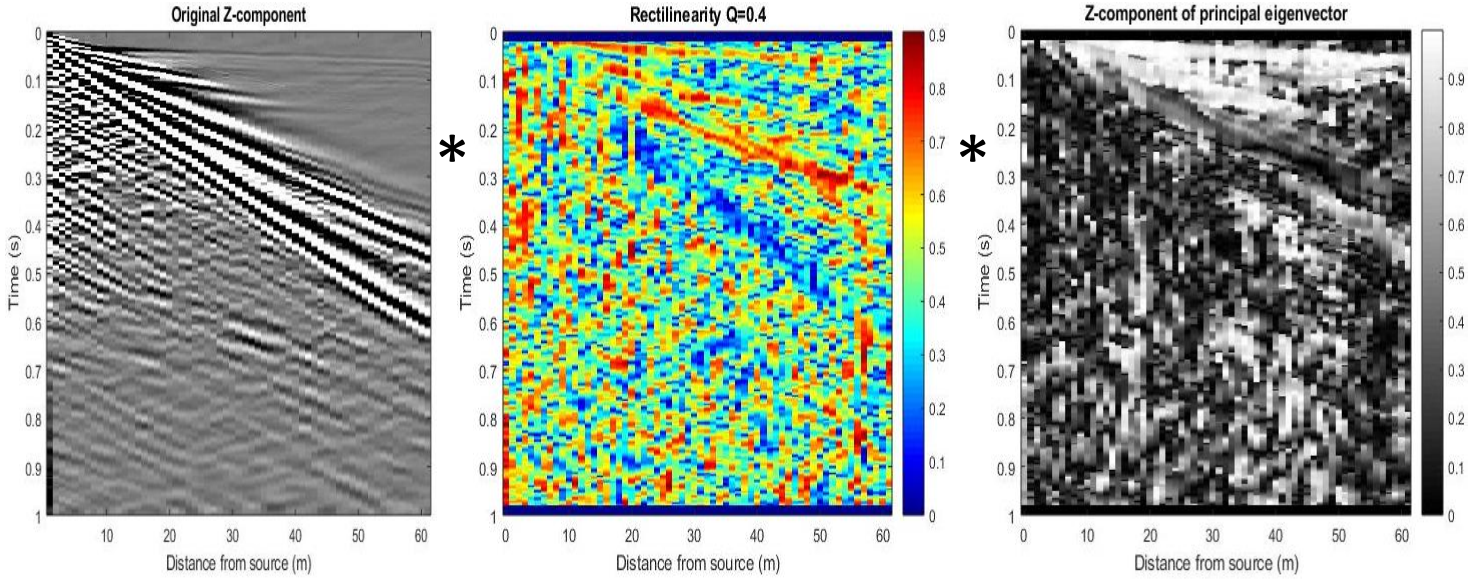

Figure 31. Sample to sample filter process for the vertical component of the field data.

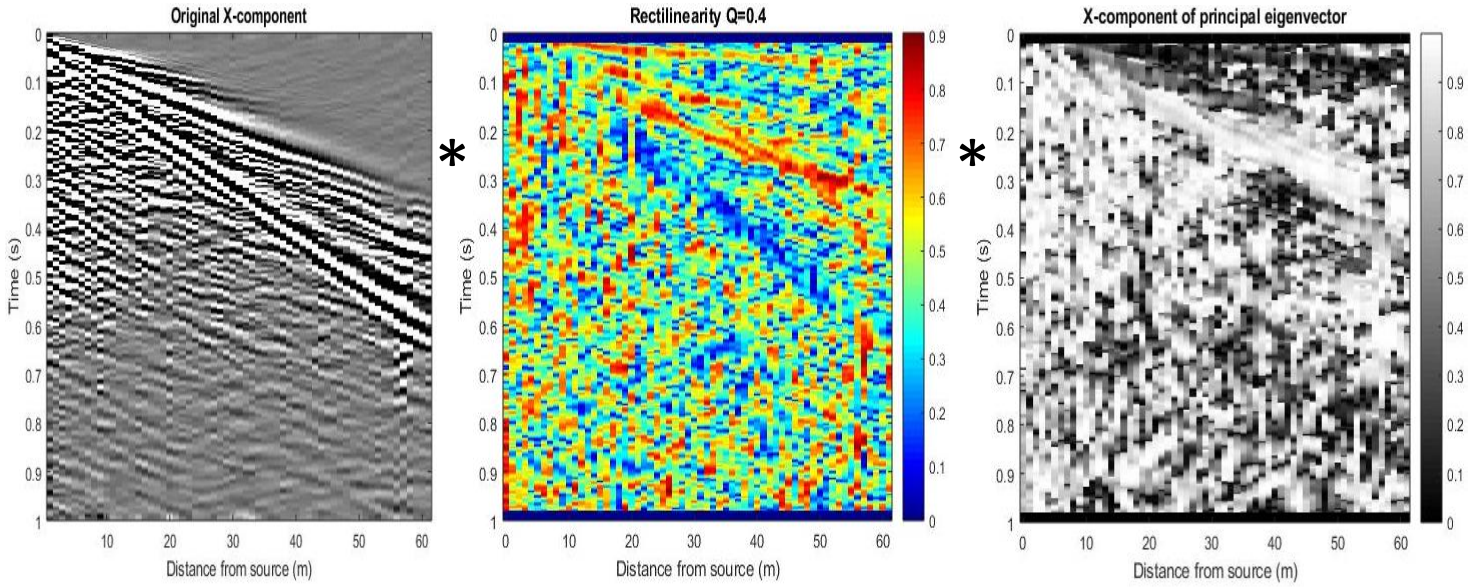

Figure 32. Sample to sample filter process for the inline component of the field data.

As can be seen from the direction of the principal eigenvector, for most samples the principal direction of polarization is along the X-component, although the P-arrivals are clearly visible on the Z-component (figures 31 & 32). The ground roll cannot be identified clearly from the direction of the principal eigenvectors. Therefore, the exponent of directivity should be selected carefully because the directivity will most likely be relatively ineffective for the inline component and destructive on the vertical component, except for the two P-arrivals. Furthermore, the guided waves are clearly polarized along the X-component.

From the rectilinearity can be observed that the ground roll is very distinct. Judging from the hodograms of the data control section the ground roll has a high degree of ellipticity and the rectilinearity confirms this apprehensive knowledge for most of the section. Especially close to the source location the ground roll becomes harder to distinguish from other arrivals.

This has serious implications for the exponent of weighting, because the ground roll throughout most of the section can be minimized by small exponents, but to remove the full ground roll the exponent

would have to be stronger which would result in a loss of data. To exemplify this, the sections have again been filtered for multiple configurations of the filter operators (figures 33 & 34).
From those sections it is clear that, looking at the exponents of the filter operators, minimization of the ground roll from the field data requires less heavy filtering than the synthetic ground roll. Despite the higher initial amplitude of the ground roll in the field data, the state of ellipticity makes it relatively easier to minimize the surface waves.
Further, the rectilinearity has a relatively strong influence on the two P-waves as well. From the rectilinearity plot can be observed that most of the refracted P-wave is visible while the reflection is resolved poorly. Therefore, filtered sections with a strong weighting function lose relatively more P-wave energy than directivity dominated sections.
Moreover, the guided waves are very persistent along the inline component because of their high amplitude relatively linear behavior along the inline component. The vertical component of directivity can quite accurately remove the guided waves.

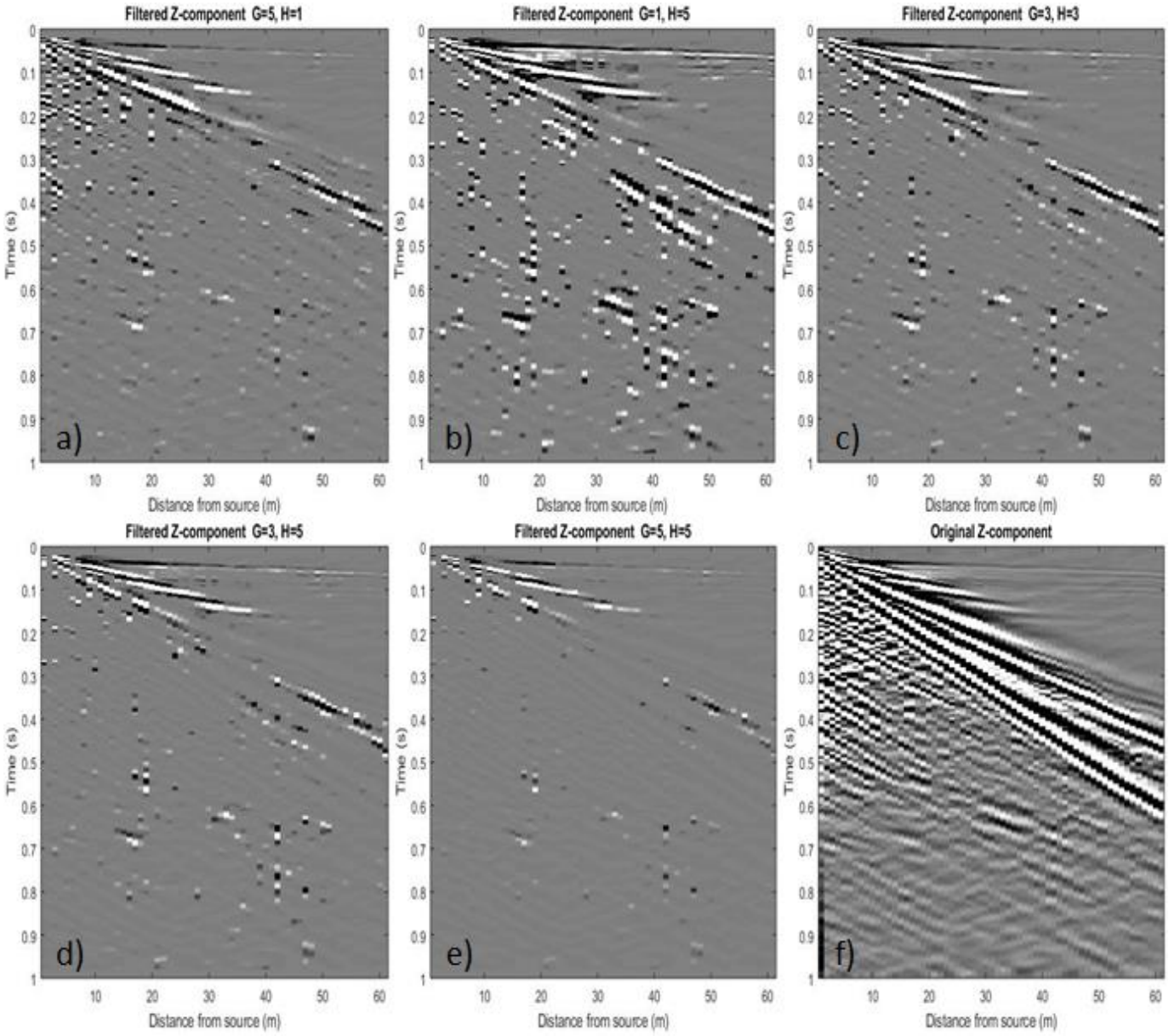

Figure 33. Results of the linearity method for the vertical component of the field data. Quality factor Q=0.4 for all filtered results. a) rectilinearity filtering G=5, H=1. b) Directivity filtering G=1, H=5. c) Linearity filtering with equal filter operators G=3, H=3. d) Dominant directivity filter G=3, H=5. e) Strong linearity filter with equal filter operators G=5, H=5. f) Original data. All filtered results were similarly imaged with clipped amplitudes to remove the dominant amplitudes at near offset. The amplitudes of the original data were clipped to a ten times higher amplitude.

Judging from all sections the strongest linearity filter (G=5, H=5) removes too much data to an extent where only certain parts of the refraction and guided wave are distinguishable and as opposed to the synthetic data the vertical reflections in the field data are best preserved by directivity filtering, especially in combination with a small weighting function to attenuate the ground roll.
The inline component, however, can be denoised more accurately with a stronger weighting function since the corresponding directivity component has much less effect and despite the absence of clear reflections on the inline component, the ground roll can be minimized robustly.
Furthermore, similar to the synthetic data, the energy over the entire section is reduced substantially for any configuration of the filter operators in the linearity method and the overprints of the reflections could not be recovered. Nevertheless the linearity method is an accurate approach to distinguish between different arrivals for the field data as well.

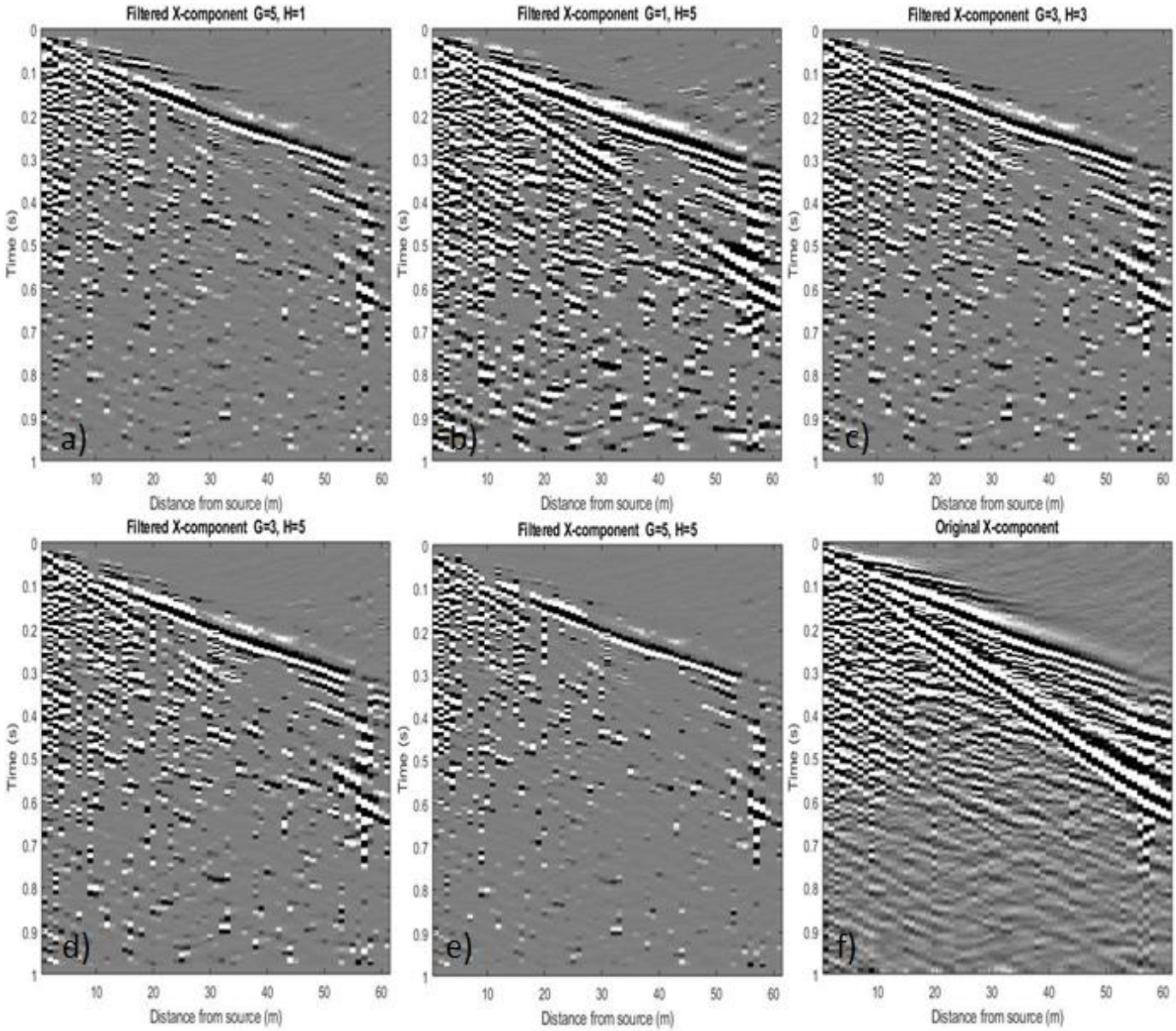

Figure 34. Results of the linearity method for the inline component of the field data. Quality factor Q=0.4 for all filtered results. a) rectilinearity filtering G=5, H=1. b) Directivity filtering G=1, H=5. c) Linearity filtering with equal filter operators G=3, H=3 d) Dominant directivity filter G=3, H=5. e) Strong linearity filter with equal filter operators G=5, H=5. f) Original data. All filtered results were similarly imaged with clipped amplitudes to remove the dominant amplitudes at near offset. The amplitudes of the original data were clipped to a ten times higher amplitude.

5.5 Ellipticity method

The ellipticity method isolates the elliptical data by multiplying the original data with the polarization attribute ellipticity (e_{21}). First the quality factor (Q) is chosen to make an accurate distinction between linear and elliptically polarized signal. Then a cut-off ellipticity is chosen to isolate the elliptical arrivals and the linear arrivals are muted. To prevent sharp cut-offs caused by muting, the edges of the muted data are smoothed.

5.5.1 Synthetic data

Surface source

Since the aim is to remove the ground roll and other nonlinear arrivals, the cut-off ellipticity is chosen by empirical determination of the ground roll ellipticity. The quality factor of ellipticity can aid to better separate the linear signal from the ground roll. Then from this plot the ellipticity of the ground roll is determined as the cut-off ellipticity. Since the ellipticity of the ground roll for the surface source is 0.40, all less elliptical arrivals are muted with smoothing along the cut-off ellipticity to an ellipticity of 0.33 (figure 35). From this separation, on the ellipticity plot can be observed that the complete ground roll is separated as well as the inexplicable high elliptical phase in the P-S conversions and a substantial amount of the random noise. The linear arrivals, on the other hand are clearly zero.

Multiplication of the ellipticity with the original signal and amplitude balancing then results in a section with merely elliptical arrivals (middle sections of figure 36). These are subtracted from the original data.

However, mere subtractions would completely remove the ground roll and overprinted reflections all together, so to be able to recover any overprinted signal, the effect of the reflection on the ground roll has to be determined. In this case, the combined energy of the vertical P-waves with the ground roll actually causes the ground roll to be recorded as more elliptical. This is caused by the stronger polarization in the X-component of the ground roll. However, since the amplitude contrast is so large, the P-wave does not alter the ellipticity of the ground roll enough to be able to accurately parameterize the ellipticity filter to be able to clearly recover the P-waves. Only the first P-wave reflection could be recovered minimally by leaving a small portion of energy of arrivals with a slightly higher ellipticity than the average ground roll. Additional effect of this method is that more ground roll signal remains throughout the entire section.

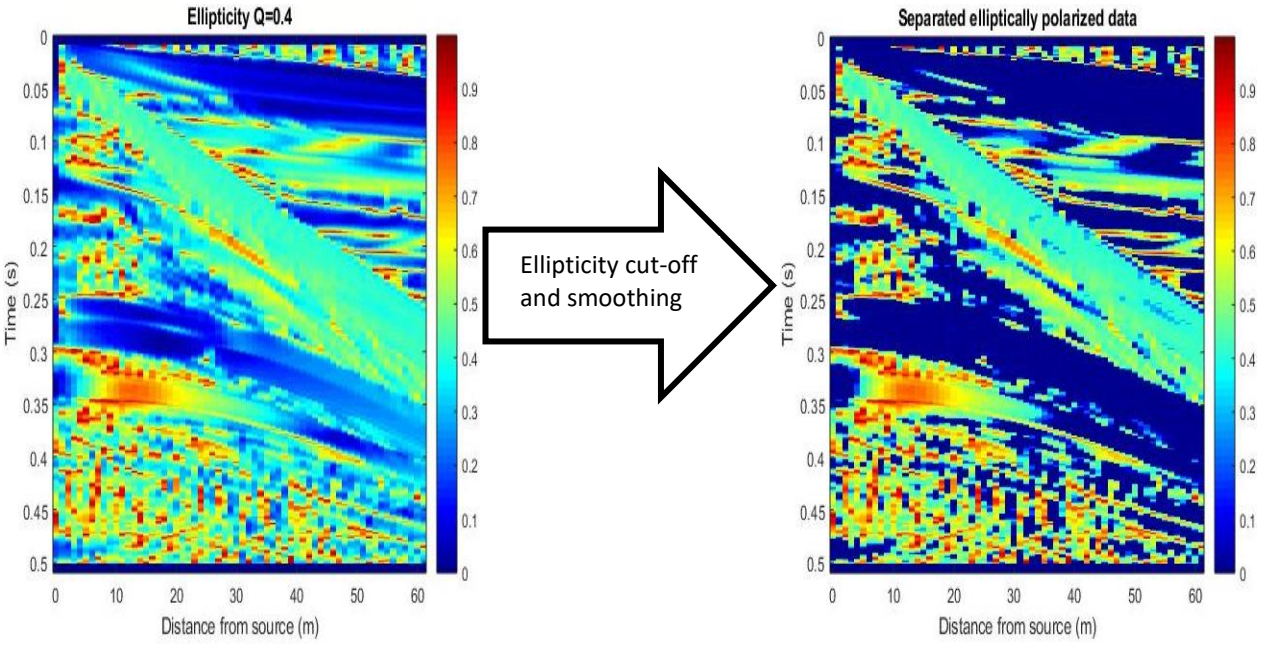

Figure 35. Ellipticity cut-off and smoothing before multiplication of the ellipticity with the original dataset. Cut-off ellipticity for synthetic data is 0.4.

Based on these previous assumptions, the S-waves, or inline polarized waves, would have a linearizing effect on the hodogram of the ground roll. However, clear S-wave reflections arrive after the ground roll and are therefore unaffected by the filter. The only overprinted arrival with clear linear polarization along the inline component is the high velocity horizontal reflection at about 0.18s. Though the arrival is especially weak compared to the ground roll, some parts have clearly linearized the data enough to pass the filter.

Nevertheless, the recovered energy is basically a multiplication of the original signal with a factor, so although it appears as if some parts of the linear arrival have been recovered, the state of polarization remains in fact unchanged. Therefore, it is arguable to call these recovered bits of data overprinted body waves.

The non-overprinted body waves, on the other hand, are mostly unharmed by the filter as a result of the ellipticity clipping. The removed body wave energy is mainly caused by interference of body waves or the edges have been smoothed due to the correlation window length.

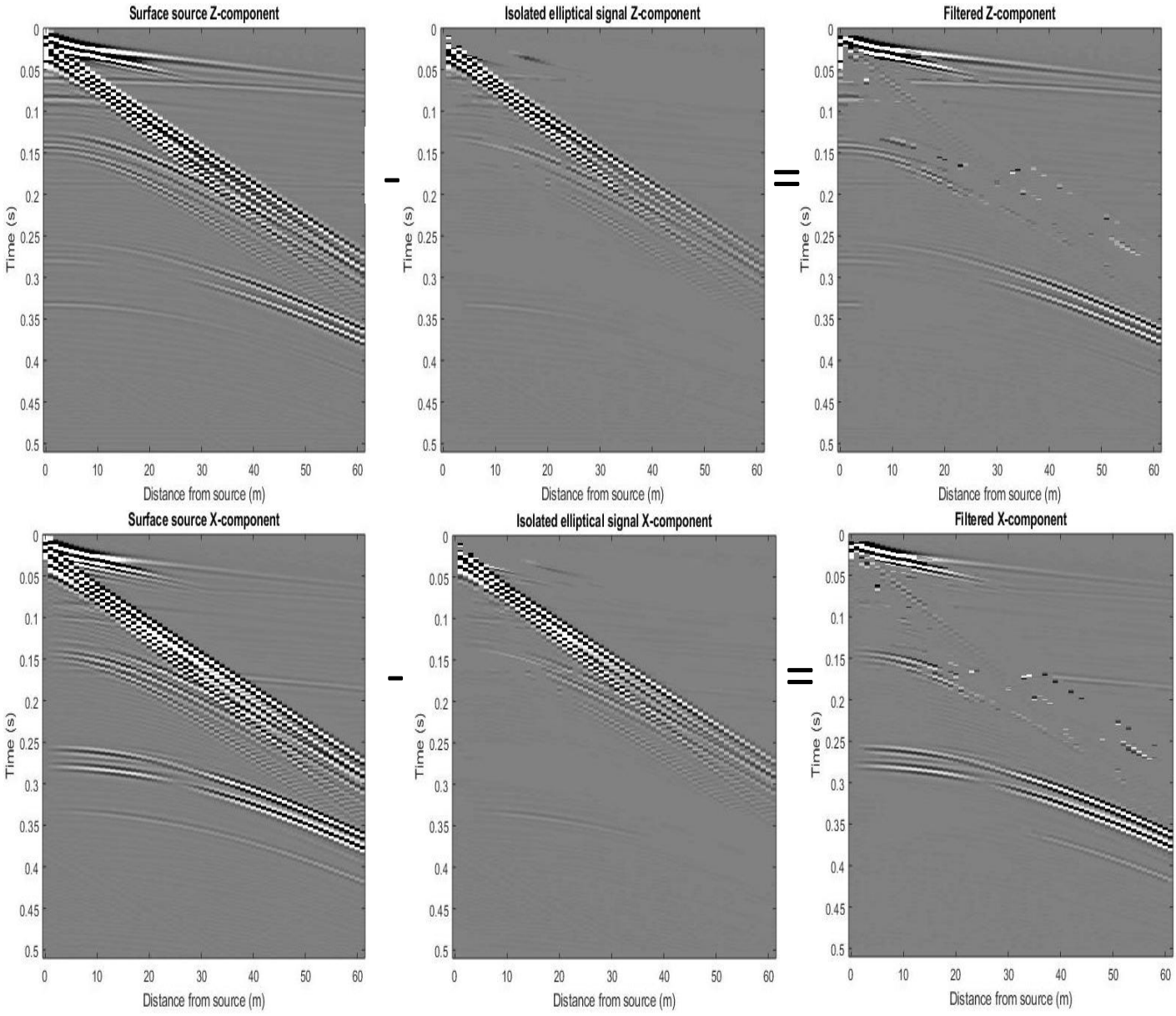

Figure 36. Ellipticity method applied to the vertical (top) and inline (bottom) component of the surface source synthetic dataset. The section containing the elliptically polarized signal is subtracted from the original section.

Buried source

To isolate the elliptical signal for the buried source, the same parameters were used as for the surface source because the ground roll in both models is equally elliptical. So, for the same quality factor (Q=0.4), the cut-off ellipticity in this case is again 0.40 with the same smoothing of the edges of the ground roll to an ellipticity of 0.33 (figure 37). From previous observations was concluded that the body wave energy in both models is quite similar and the ground roll for the buried source is about three times stronger.

Thus, it is evident that the overprinted reflections are even harder to recover than for the previous case since the influence of the overprinted body wave on the state of polarization of the ground roll has become even smaller.

However, because the strong P-S converted wave in this model is much more intertwined with the ground roll, the influence of a body wave of on the ground roll is visible and clearly some small pieces of the converted wave were recovered (figure 38). Nevertheless, the state of polarization of the converted wave is already near the cut-off ellipticity, so it would be unrealistic to expect for this arrival to alter the ground roll enough to pass the cut-off ellipticity completely.

The other, unconverted, body waves are much more linear and are clearly unaffected by the ellipticity while the ground roll is seemingly unaffected by the body waves. Hence, the overprinted body waves cannot be recovered for the current energy contrast.

Nonetheless, because of the constant state of polarization of the ground roll throughout the entire section, the cut-off ellipticity to remove the band of surface waves could again be determined accurately.

Based on both synthetic datasets the results are very similar. Aside from the amount of energy, the ground roll obtains the same properties for both models, so the filter basically has the same effect on both datasets. The body wave reflections are disregarded by the filter, so they do not lose any energy. The P-S waves lose a significant amount of energy, just like the ground roll which is filtered almost completely in both sections. Unfortunately most of the overprinted reflection energy could not be recovered in both sections.

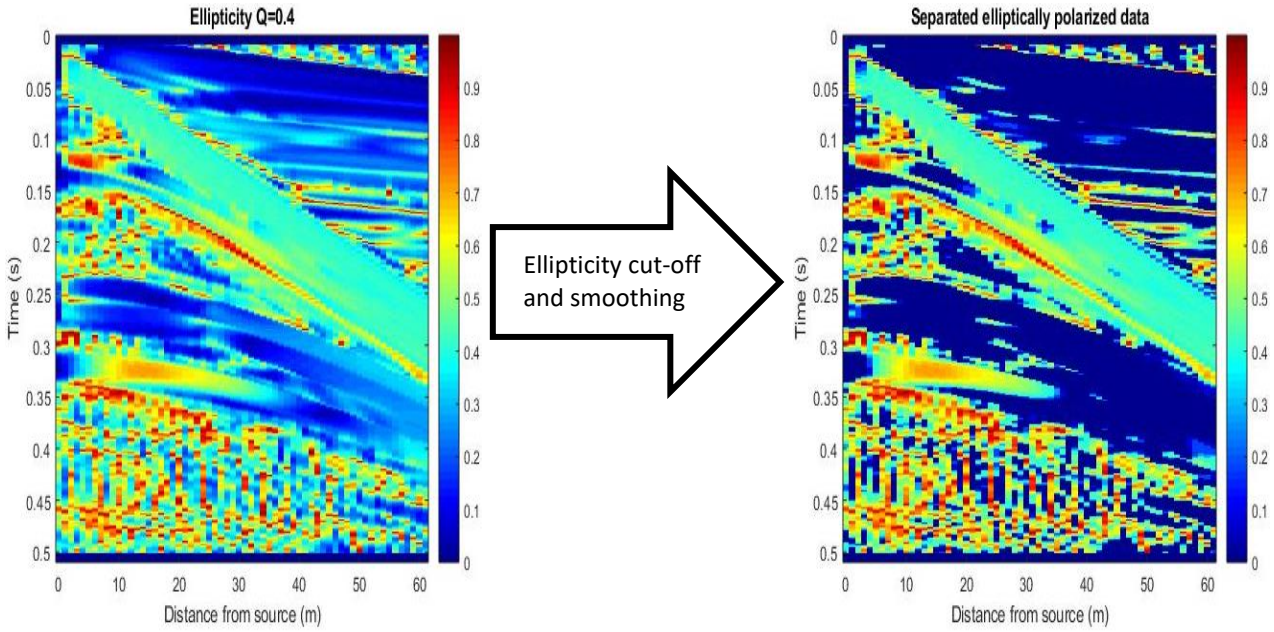

Figure 37. Ellipticity cut-off and smoothing before multiplication of the ellipticity with the original dataset. Cut-off ellipticity for synthetic data is 0.4.

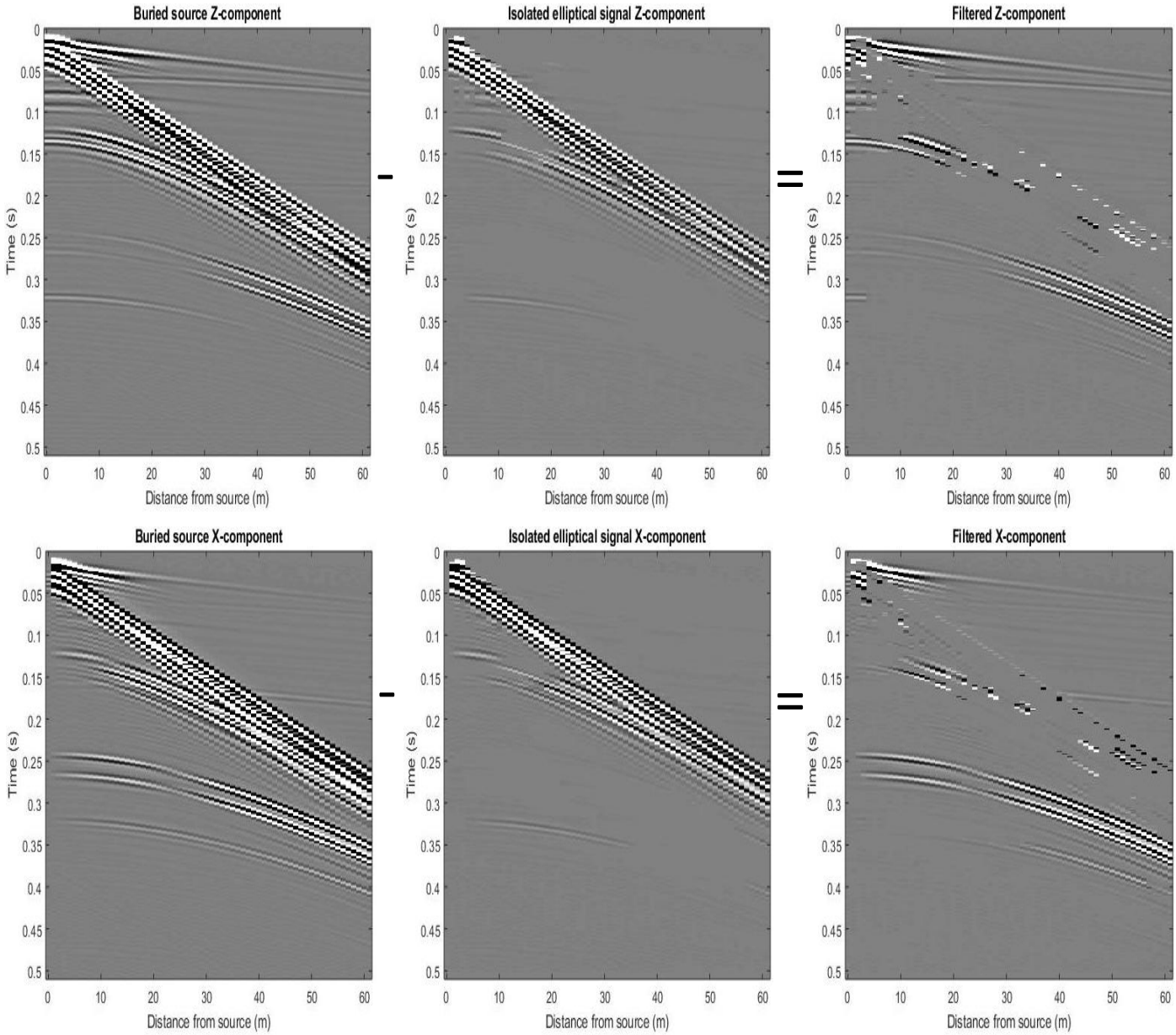

Figure 38. Results of the ellipticity method for the vertical and inline component of the buried source. The amplitude balanced elliptical signal is subtracted from the original signal.

5.5.2 Field data

From previous analysis of the field data and its polarization attributes could already be determined that the ground roll is dominant, but rather inconsistent throughout the section. Specifically in the middle part of the section the ground is highly elliptical while in the latter and initial part of the section the ground roll is recorded as more linear signal and is much less distinct.

This has serious implications on the cut-off ellipticity, because to accurately capture the entire ground roll would mean that a lot of other signal is filtered as well. However, to conserve this data some major sections of the ground roll might not be affected by the filter.

To better separate the ground roll, initially the quality factor of ellipticity is of importance to amplify the separation of different phases. Therefore, a quality factor of Q=0.5 is more suitable in this case than a quality factor of 0.4 and because of the higher state of ellipticity of the field data ground roll, the ellipticity cut-off is again 0.40 with similar smoothing as the synthetic datasets.

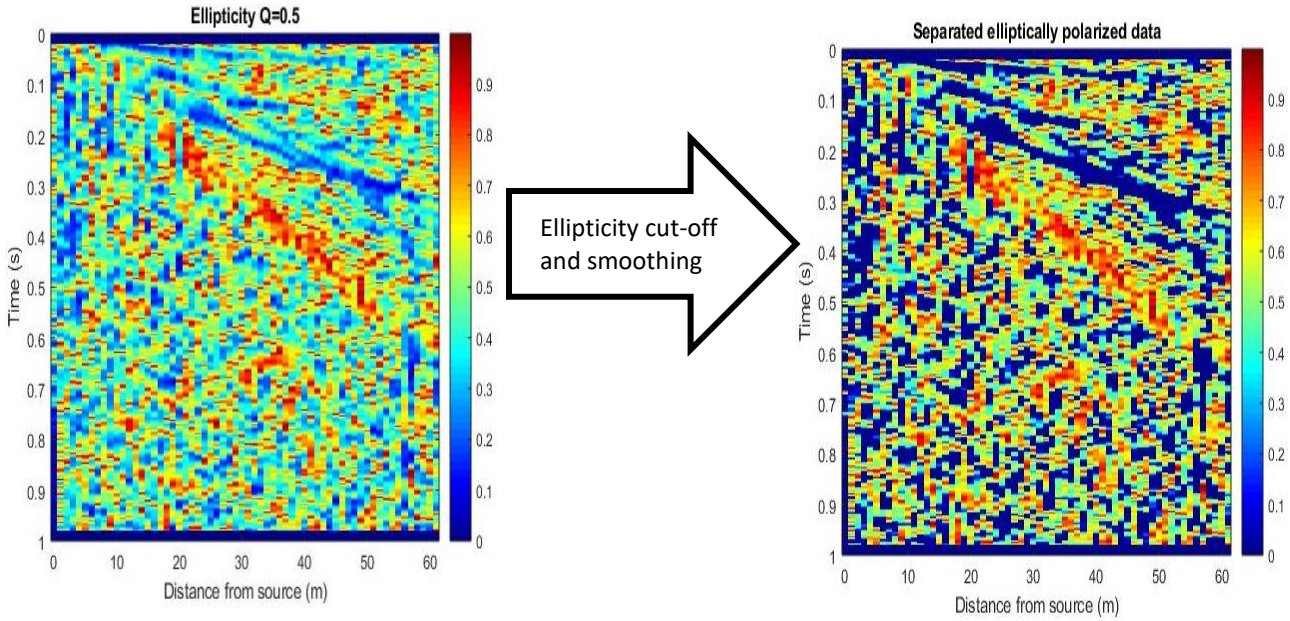

Figure 39. Ellipticity cut-off and smoothing of the field data before multiplication of the ellipticity with the original dataset. Cut-off ellipticity for the field data is also 0.4, but the quality factor is 0.5.

The ellipticity plot after cut off and smoothing (figure 39) shows that most of the ground roll is incorporated into the filter whereas most guided wave energy is regarded as linearly polarized signal. The onset of the vertical P-wave reflections is clearly disregarded by the filter as well, but as the distance to the source increases, the P-waves apparently lose too much energy relative to the ambient background noise to be distinctly recognized as linear arrivals. Nonetheless, the refraction, can accurately be identified as linearly polarized up to a source distance of 50 meters and will largely be unaffected.

Multiplication of the ellipticity with the original data then results in a section with most of the ground roll cone well defined throughout the majority of the section (middle sections of figure 40). Further, the P-waves at greater distance from the source are vaguely visible as well on the Z-component. Lastly, a lot of random noise and artefacts as a result of the bouncing of the drop weight are isolated as well on all three components.

Thereafter, subtraction of the isolated signal from the original data generates the filtered section for all three components of the field data from which a significant portion of the original data has been removed by the ellipticity filter. The ground roll in particular is removed quite accurately, as well as a significant portion of ambient noise and the vertical P-wave reflection further away from the source. As a result of this vigorous filter mostly guided wave energy and a the majority of the P-refraction remains.

The overprinted portion of the P-waves could not be recovered either, for the same reason as the overprinted body waves in the synthetic datasets; the body wave energy relative to the surface wave energy is too small to considerably alter the state of polarization of the ground roll.

In particular for the field data, due to the low signal-to-noise ratio, the results of the ellipticity method represent a trade-off between accurately removing the ground roll and preserving reflections and other semi-linear phases.

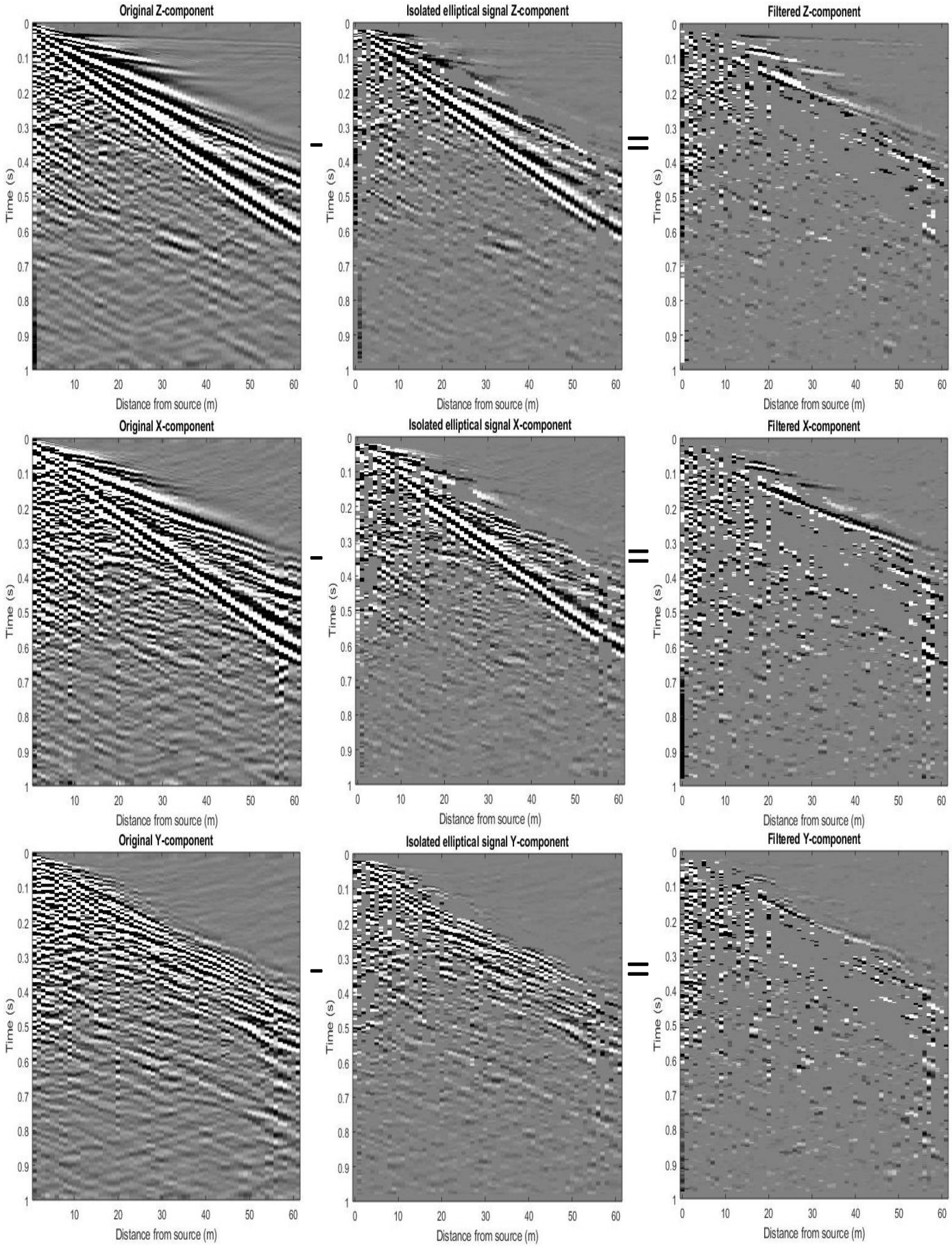

Figure 40. Results of the ellipticity method on the three orthogonal components of the field data.

6. Discussion

Since the aim of the project is to remove the ground roll by employing a polarization filter, the initial aim for the data acquisition was to obtain datasets with a very dominant ground roll and to test our polarization filters on comparable synthetic data and field data.
Therefore, the subsurface SPECFEM3D Cartesian model was designed to represent a simplified model of the actual subsurface model from the survey with the objective to analyze and optimize the effectiveness of the polarization filters to synthetic data before applying it to actual field data. However, due to poor field data quality a comparison to the synthetic data would be ungrounded aside from the dominant ground roll. The poor data quality is most likely caused by poor receiver coupling due to the dryness of the soil and the bouncing effect of the drop weight. Moreover, the amount of ambient noise was much higher than anticipated. Despite the mentioned favorable conditions in the data acquisition section, the nearby horse stables are probably the main cause for the amount of noise in the data.
Even so, in spite of the lack of similarities between the field data and synthetic data, the influence of the source location is a second important reason to employ the synthetic model. Since the model itself incorporates a high velocity and density gradient at the 2 meter interface to amplify the ground roll, a buried source at a depth of 3 meters, just below the interface, is an accurate approach to test whether a buried source can invoke a similar ground roll as a surface source. However, the induced ground roll from the buried source is about three times stronger than the surface source ground roll. Possible explanations for this significant energy contrast could be that a substantial amount of energy is lost for the source at the free surface or that a source just below the low velocity layer has an amplifying effect on the ground roll when crossing the interface, sparking stronger P- and S-wave interference then the surface source does. Nevertheless, both of these hypotheses cannot be proven, so the actual reason for the energy contrast is unknown.
What is clear is that the modeled vertical sources produce reflected wavefields that are dominated by P- and Sv-arrivals where, aside from the one clear Sv-reflection, most of the Sv-energy in these wavefields is observed in P-Sv conversions. These conversions clearly coincide with P-arrivals at interfaces at a deviant angle from the normal. This phenomenon is accurately captured by the polarization attributes where a clear shift in dominant direction of polarization can be observed in the same period when the arrival is recorded from linear to elliptical and back to linear again.
The strong velocity contrast of the top layers does not only constitute favorable conditions for the ground roll, it has consequences for the body waves as well. Body waves transmitted across an interface into a medium with lower wave velocities bend towards the vertical axis, or the normal. This implies that upon transmission from below the body waves bend strongly toward the normal and are recorded as almost vertical arrivals, meaning that P-waves are polarized along the vertical component and Sv-waves along the horizontal inline component, as can be observed from the hodograms. However, the velocity contrast for the S-waves is in both relative and absolute sense a lot smaller than the P-wave velocity contrast. Therefore, the S-waves bend less strongly to the normal when crossing the interface from below and are thus less strongly polarized along a recording axis which has serious implications on the principal direction of polarization. As a result of that, the directivity component in the linearity filter would remove more S-wave energy than P-wave energy. Thus, a prerequisite to accurately distinguish body waves by their principal direction of polarization is a strong velocity contrast between the top strata.
A second condition to accurately capture the body waves along with the surface waves is the correlation window which is dependent on the broadness of the frequency spectrum to determine whether the high frequency arrivals are not smoothed and the window embodies at least one low frequency period. In this case, the field data has a broader frequency spectrum than the synthetic data. As a result of that, the difference between one period of a Rayleigh wave and P-wave is substantially larger in the field data which causes implications on the correlation window length. For the linearity or ellipticity method to work the correlation window needs to be at least one low

frequency ground roll period. Consequently, the high frequency body waves are recorded less accurately as was observed in the field data rectilinearity and ellipticity plots. So, there is always a trade-off between accurately capturing the ground roll while not doing too much harm to the reflections. The Hann window helps to somewhat minimize this effect by correlating the data to a half cosine to amplify the data point centered in the correlation window while preventing smoothing.
Because the frequency spectrum of the synthetic data is much narrower, the correlation window is much more capable to capture the individual body wave arrivals as well as the ground roll. This also explains the results seen in the polarization analysis from the rectilinearity and ellipticity**.** As long as the data quality is high enough and the frequency spectrum allows for individual arrivals to be resolved, a very accurate polarization analysis can be made.
Especially the polarization analysis on the synthetic data indicates that the rectilinearity and ellipticity are very capable methods to distinguish between arrivals. The linear and elliptical arrivals are resolved accurately for the synthetic datasets while the field data does not match the criteria for accurate polarization analysis. This is caused by the inconsistency of the ground roll and weakness of the reflection energy throughout the section in addition to the broad frequency spectrum. Although the relative amplitude of the reflection signal compared to the ambient noise appears to be a more dominant factor in properly capturing the linearity of the reflections than the frequency spectrum. Nevertheless, the principal direction of the field data reflections is along the vertical component which means that the reflections are obvious, but because of the strong relative noise component the reflections are recorded as nonlinear arrivals.

On these grounds, it is not surprising that the results of the linearity method and ellipticity method vary considerably between the synthetic and field data. Further, the results of both methods vary considerably too.
The linearity method has the ability to clearly distinguish between different arrivals by their state or primary direction of polarization. Though, because the method of Maercklin (2007) utilizes multiplications, the linearity method can only minimize the ground roll energy while simultaneously a significant amount of body wave energy is removed as well. Instead, similar to the ellipticity method, a cut-off ellipticity could have been applied to leave the linear arrivals unharmed, which could significantly increase the effectiveness of the rectilinearity component. However, the linearity method also employs the direction of polarization which cannot simply be clipped. Previous observations already showed that the principal component of the synthetic ground roll is strongly polarized along the X-component and also has a relatively strong Z-component, so the linear arrivals would have to be polarized strongly along a specific component. Further, in the field data the ground roll is even more elliptical and the two principal eigenvectors are almost just as strong and hypothetically might have been almost parallel to the reference frame. Therefore, a cut-off on the direction of ellipticity would be ineffective in many cases especially if body waves are recorded at a less normal angle of incidence.
Nonetheless, the results from the linearity method already showed that the directivity is a less accurate method to minimize the ground roll than the rectilinearity and accordingly should be the smaller contributor to the linearity method in most cases. However, if the purpose of the filter is to also separate the body waves, the directivity is an ideal tool to contribute to the rectilinearity in the filter. Prime requirement for the separation to work is for the body waves to be strongly polarized along a specific axis as a result of a strong velocity contrast between the top two layers. Directivity filtering then removes the Sv-waves from the vertical component and the P-waves from the inline component.
Moreover, as mentioned before, the principal direction of polarization is the only polarization attribute that could accurately distinguish the field data reflections despite poor data quality. Therefore, the directivity can be a very useful tool, but should always be parameterized depending on the data quality and principal direction in the data.

On the other hand, if the sole purpose of the filter is to capture the state of polarization in the dataset, the directivity can be neglected, because the direction of polarization alters the state of polarization by linearizing the data to a specific component.
The results of the linearity method based on a combination of both rectilinearity and directivity then show dominant P-waves on the vertical component and Sv-waves on the inline component while all other wavefields are minimized. Since the individual body waves are already dominant on one specific component section, the results for the linearity filter with dominant rectilinearity are somewhat similar, except that the small amount of energy of the linear wavefields projected on other components remains.
Moreover, the overprinted reflections could not be recovered for whichever configuration of the filter components. This essentially causes the method to have severe limitations, because similar to other methods such as f-k or τ-p filtering, important reflection data is lost. The overprinted body waves could not be recovered in this method because the initial energy of the overprinted phases is too small relative to the ground roll energy. As a result of that the state of ellipticity was left unharmed and the filter operations removed the overprinting ground roll just as strongly as the regular ground roll.
The ellipticity method suffers from the same data limitations as the linearity method. Because of the cut-off ellipticity the linear data is unaffected by this filter and the elliptical data isolated. However, in further processing the recovery of overprinted body waves is again performed by polarization analysis to identifying the body wave signature inside the ground roll.
Amplitude balancing based on the elliptical state of the arrival proved that the overprints could only minimally be recovered. This is because of the amplitude deficit between body waves and surface waves as well as variable ellipticity of the ground roll. Therefore, the influence of the body waves on the ground roll is not high enough to clearly alter the ellipticity out of the ground roll range.
If the relative amplitude between body waves and surface waves were smaller, the ellipticity method may very well be able to capture the signature of linear arrivals inside the ellipticity of the ground roll by accurate tapering.
However, the primary advantage of the ellipticity over the linearity method is that the individual linear arrivals are preserved perfectly if the data quality is high enough. The results for the synthetic data are satisfactory aside from the unrecovered overprints. An additional advantage of the polarization filter over methods like f-k or τ-p filtering is that the polarization filter works locally in the x-t domain without any global transformation effects and artefacts that f-k and τ-p filtering suffer from. On the other hand, for the low quality field data the linearity method is more effective in preserving reflections because the filter applies two parameters to distinguish between arrivals.
To further improve the polarization filter, invoking three extra components to the data may help to further distinguish the ground roll from the body waves in a 6-component dataset. Previous research by Edme et al. (2013), Barak et al. (2014) and Sollberger et al. (2017), showed that recording of three extra rotational component with special rotaphones can aid to further distinguish between the ground roll and other arrivals. Especially the elliptical rotation of the ground roll around the crossline component could be of interest to further separate different wavefields and may be able to better recover the overprinted reflections.

7. Conclusion

From the results of the polarization analysis based on the polarization attributes can be concluded that different wavefields can accurately be distinguished by their state and direction of polarization as long as data quality is high enough. Especially for the synthetic datasets the polarization analysis is very effective and could resolve almost all individual arrivals precisely based on either the rectilinearity, ellipticity or the principal direction of polarization. However, the results also showed that a low quality dataset, like the applied field record, and a broad frequency spectrum pose serious limitations to the polarization analysis and filters.

Nonetheless, the results from the polarization filters prove that the ground roll can be removed to a large extent purely by polarization filtering. Especially the ellipticity method can accurately remove the ground roll without affecting the reflections. Aside from that the linearity method can even provide individual P- and S-wave analyses if the body waves are polarized along a specific spatial axis, although the downside to the method is that all wavefields lose a significant amount of energy.
Taking all into consideration, the provided polarization filters are useful tools to separate individual wavefields and remove the ground roll, but unfortunately accurate recovery of overprinted reflections by mere polarization analysis on 3-component data cannot be achieved by the represented methods.

8. References:

Barak, O., Herkenhoff, F., Dash, R., Jaiswal, P., Giles, J., de Ridder, S., Brune, R., and Ronen, S. (2014). Six-component seismic land data acquired with geophones and rotation sensors: Wave-mode selectivity by application of multicomponent polarization filtering. The Leading Edge, 33, pp. 1224-1232, doi: http://doi.org/10.1190/tle33111224.1.

Bourbie, T., Coussy, O., and Zinszner, B. (1987). Acoustics of porous media. Gulf Publishing Company, United States.

Chen, H. F., Li, X. Y., Qian, Z. P., & Zhao, G. L. (2013). Robust adaptive polarization analysis method for eliminating ground roll in 3C land seismics. Applied geophysics, 10(3), pp. 295-304.

Cliet, C. and Dubesset, M. (1988). Polarization analysis in three-component seismics. Geophysical Transactions, 34, pp. 101–119.

De Meersman, K., Kendall, R. (2005). A complex SVD-polarization filter for ground roll attenuation on multi-component data. EAGE 67th Conference & Exhibition.

Dinoloket (2018). Data and geology of the Dutch subsurface, Core B32C0489, from https://www.dinoloket.nl/ondergrondmodellen-betaversie

Edme, P., Daly, M., Muyzert, E., and Kragh, E. (2013). Side scattered noise attenuation using rotation data: 75th Annual International Conference and Exhibition, EAGE, Extended Abstracts.

Jin, S., Ronen, S. (2005). Ground roll detection and attenuation by 3C polarization analysis. EAGE 67th Conference & Exhibition.

Jurkevics, A. (1988). Polarization analysis of three-component array data. Bulletin of the seismological society of America, 78(5), pp. 1725–1743.

Kanasewich, E.R. (1981). Time sequence analysis in geophysics. The University of Alberta Press.

Kanasewich,E.R.(1990). Seismic noise attenuation, volume 7 of Handbook of Geophysical Exploration. Pergamon Press, Oxford.

Komatitsch, D., Tromp, J. (2002a), Spectral-element simulations of global seismic wave propagation-I. Validation, Geophysical Journal International, 149 (2) , pp. 390-412,
doi: http://doi.org/10.1046/j.1365-246X.2002.01653.x

Komatitsch, D., Tromp, J. (2002b), Spectral-element simulations of global seismic wave propagation–II. Three-dimensional models, oceans, rotation and self-gravitation, Geophysical Journal International, 150 (1) , pp. 303-318

Komatitsch, D., Vilotte, J.-P., Tromp, J., Ampuero, J.-P., Bai, K., Basini, P., Blitz, C., Bozdag, E., Casarotti, E., Charles, J., Chen, M., Galvez, P., Goddeke, D., Hjorleifsdottir, V., Labarta, J., Le Goff, N., Le Loher, P., Lefebvre, M., Liu, Q., Luo, Y., Maggi, A., Magnoni, F., Martin, R., Matzen, R., McRitchie, D., Meschede, M., Messmer, P., Michea, D., Nadh Somala, S., Nissen-Meyer, T., Peter, D., Rietmann, M., de Andrade, E.S., Savage, B., Schuberth, B., Sieminski, A., Strand, L., Tape, C., Xie, Z., & Zhu, H. (2012). SPECFEM3D Cartesian v2.0.2 [software], Computational Infrastructure for Geodynamics, doi: http://doi.org/NoDOI, url: https://geodynamics.org/cig/software/specfem3d/

Maercklin, N. (2007) -Supolar and supofilt: Su programs for polarization analysis and filtering of three-component data.

NAM and RGD (Nederlandse Aardolie Maatschappij and Rijks Geologische Dienst), (1980). Stratigraphic nomenclature of the Netherlands. Verh. Kon. Ned. Geol. Mijnbouwk. Gen., 32, 77 pp.

Samson, J. C. (1973). Descriptions of the polarization states of vector processes: applications to ULF magnetic fields. Geophys. J. R. Astr. Soc., 34(4), pp. 403–419.

Sollberger, D., Schmelzbach, C., Van Renterghem, C., Robertsson, J., & Greenhalgh, S. (2017). Automated, six-component, single-station ground-roll identification and suppression by combined processing of translational and rotational ground motion. *SEG Technical Program Expanded Abstracts 2017*, pp. 5064-5068.

Tiapkina, O., Landrø, M., Tyapkin, Y., & Link, B. (2012). Single-station SVD-based polarization filtering of ground roll: Perfection and investigation of limitations and pitfalls. Geophysics, 77(2), pp. V41-V59.

Appendix A: Polarization attributes synthetic data

Effect of the quality factors on surface source rectilinearity and ellipticity.

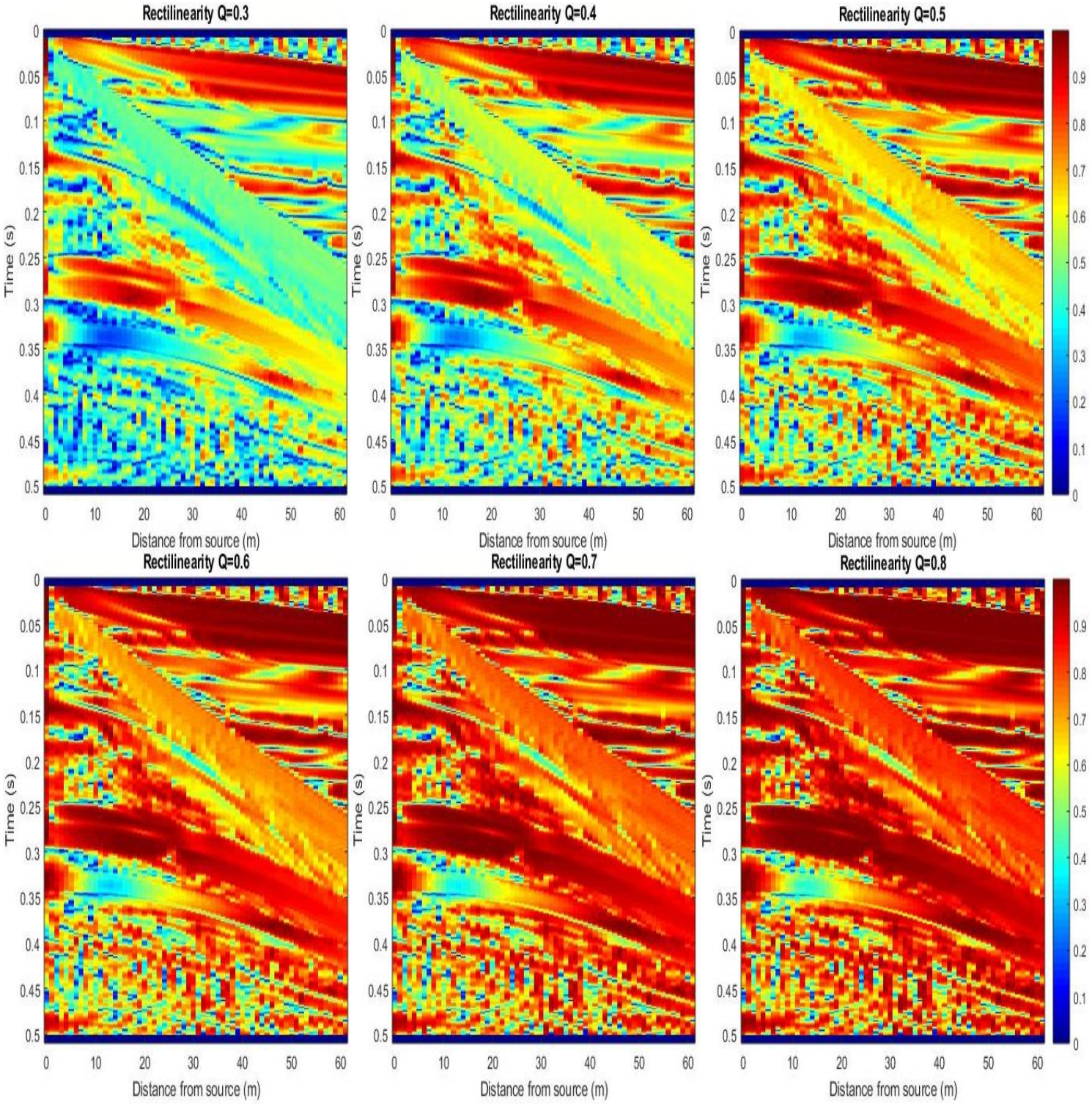

Figure 41. The effect of the quality factor on the rectilinearity for the surface source synthetic dataset.

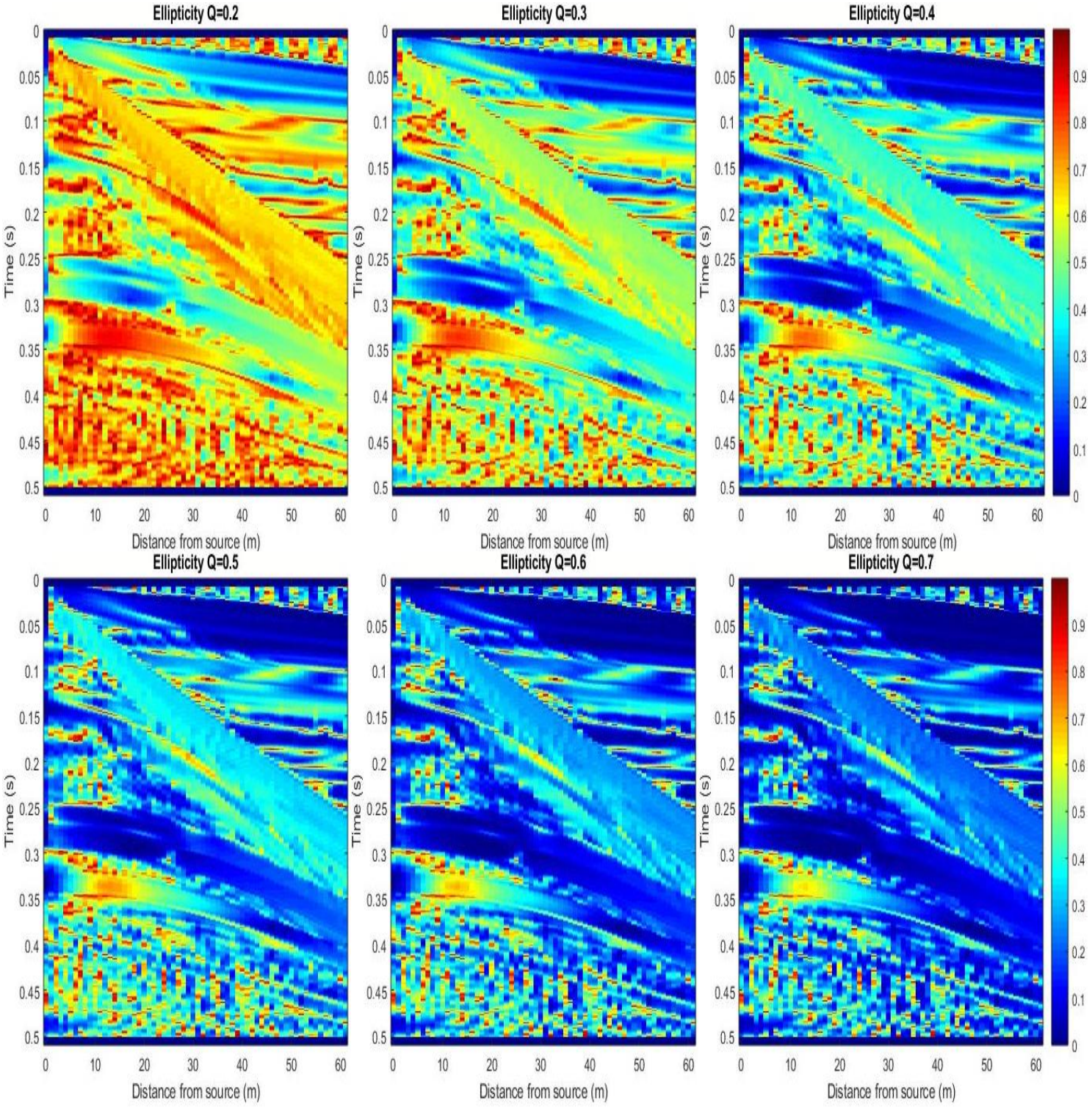

Figure 42. The effect of the quality factor on the ellipticity for the surface source synthetic dataset.

Appendix B: Polarization attributes field data

Effect of the quality factors on field data rectilinearity and ellipticity

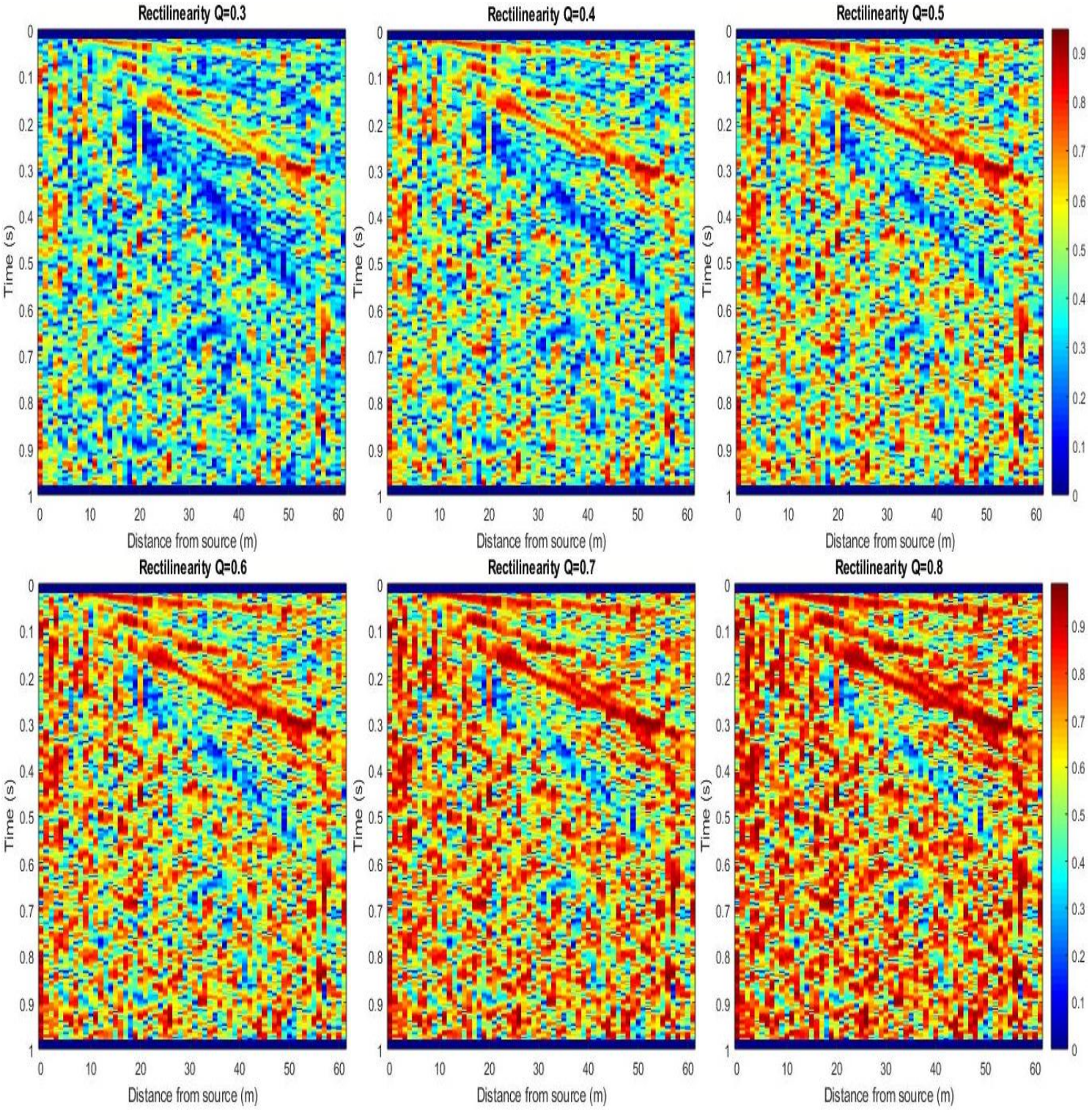

Figure 43. The effect of the quality factor on the rectilinearity for the field data.

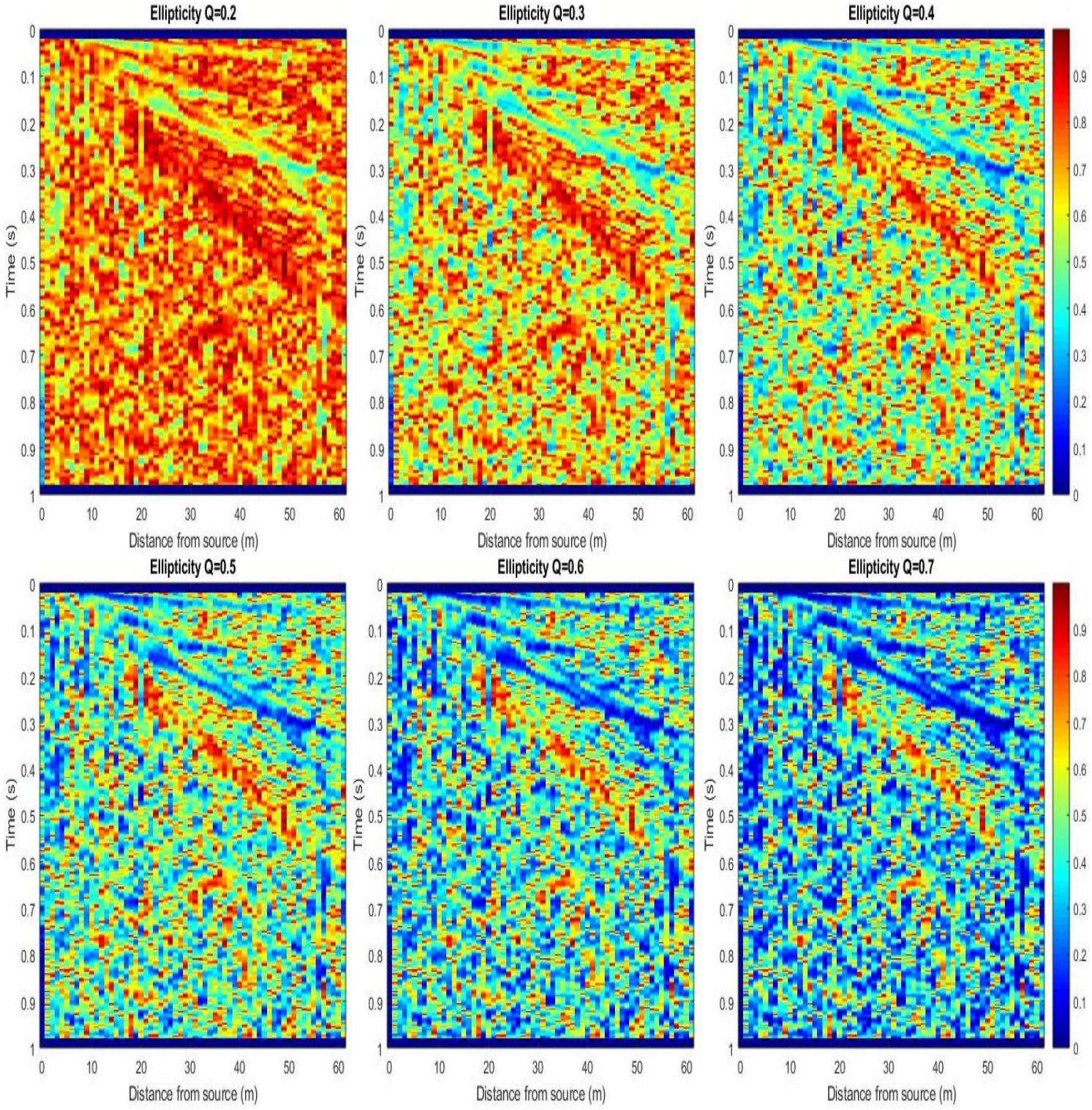

Figure 44. The effect of the quality factor on the ellipticity for the field data.