POLITECNICO DI TORINO

Department of Environment, Land and Infrastructure Engineering

Master of Science in Petroleum and Mining Engineering

Validation of automatic dispersion curve picking in exploration data

Supervisor: **Student:**

Prof. Laura Valentina Socco   Chalabiyev Elshan

Co-supervisor:

Myrto Papadopoulou

Academic year 2019-2020

Acknowledgements

Contents

List of Figures

List of symbols and acronym

SW - Surface wave

DC - Dispersion Curve

VS - S-wave velocity

1D - 1 dimensional

2D - 2 dimensional

3D - 3 dimensional

f - Frequency

V_{R_i} - Phase velocity of Rayleigh waves

V_{Rstd} - Standard deviation of the phase velocity

A_{norm_i} - Amplitude of the cross-correlation matrix corresponding to the i-th frequency component of the DC

A_u - Uncertainty based on the amplitude of the cross-correlation corresponding to the DC

N_{shot} - Number of shots

$e_{f_{ig}}$ - Uncertainty based on the ignored frequencies

$e_{f_{extra}}$ - Uncertainty based on the extra picked frequencies

$N_{f_{ig}}$ – number of ignored frequency points;

$N_{f_{man}}$ – total number of frequency points picked manually;

$N_{f_{extra}}$ – Number of extra picked frequency points;

$N_{f_{auto}}$ – Total number of frequency points picked automatically.

$N_{f_{auto,effective}}$ – Number of frequency points carrying correct data;

A_{effic} - Effectiveness of DC

S_{ind} - Similarity index

Introduction

A well-established geophysical method for hydrocarbon exploration, which is becoming more and more popular also for other natural resources (e.g. mineral), is seismic reflection. It is based on the analysis of the propagation of seismic body waves, which are included in the seismic data acquired on the field, using proper seismic sources and recording equipment. The reflected waves are highly affected by surface waves (SWs), commonly called the ground roll and characterized by high energy. Ground roll, even though considered as noise in exploration, can be useful to retrieve near-surface velocity models, because SWs are sensitive to the near-surface properties since they propagate only in the shallow layers. The near-surface velocity models valuable in exploration because having high-quality imaging of the near-surface can improve the quality of imaging of deeper exploration targets. For example, the near-surface model can be used to retrieve accurate static corrections (Durrheim and Maccelari 1991; Adam, Milkereit and Mareschal 1998; Eaton, Milkereit and Salisbury 2003; Koivisto *et al.* 2012; Malehmir *et al.* 2017). Static corrections are necessary to correct the travel-times within the near-surface layers, and they are an essential part of the seismic processing workflow. They are used to compensate a near-subsoil layer composed of low seismic velocity material.

One of the positive aspects of using the SW is that they can be extracted straight from the exploration data for further uses thus avoiding the extra cost for additional acquisition of data (Mari 1984; Roy, Stewart, and Dulaijan 2010; Boiero 2013; Douma and Haney 2011; Miao 2016).

To obtain a velocity model from SW, dispersion curves (DC – phase velocity vs. frequency) must be extracted from the seismic records and then inverted to retrieve velocity of the layered subsurface. However, to extract the dispersion curve from the data is necessary to pick the maxima of the energy over a 2D spectrum computed from the data and this step can be , intense and time-consuming and becomes inconvenient if it is performed manually on large-scale

applications and large datasets. It is then a very important task for industry use of SW to implement automatic processing tool to extract DC in a time-effective way and with limited operator cost. .

An autopicking code has been implemented in the framework of the H2020 project Smart Exploration. The performances of this code need to be assessed and the results of the autopicking need to be validated.

The manually picked DCs are considered to be more accurate since they are based on the human observation, and that is why it has been decided to consider the manually picked DCs as the benchmark to evaluate the performance of the autopicking code. However, it is a well-known fact that nowadays, the automatisation of the processes becomes widely required in almost all fields of human activities since it simplifies the workflow and increases the time-efficiency.

In this thesis, we implement a code to efficiently and effectively compare two sets of equivalents (i.e. obtained from the same receiver pairs and sources). One of the DC sets is picked manually by an expert operator, and the other is by a code based on a series of quality controls (QCs) to select the data points. The comparison is aimed at validating the autopicking. The thesis work has focused on the definition of a set of QC parameters to be used to compare the two sets of curves and on the implementation of an automatic comparison tool. The tool can be used to validate the autopicking of a large dataset by comparing a limited subset of random curves which are also picked manually.

This thesis work consists of four chapters:

1. In the first chapter, the theoretical background of the SW methods, with particular emphasis on the propagation characteristics of the SWs and their analysis workflow, are provided. The two-station method and the workflows of the manual and automatic DC picking are provided.

2. The chapter contains a description of the Ludvika mining site, and the dataset used in the current work. Previous research results on the same

site are also presented.

3. The third chapter contains the description of the method implemented in the current thesis, for the comparison of the manually and autopicked DCs.

4. The fourth chapter presents the results of the application of the implemented method to compare the DCs picked manually and automatically from the Ludvika seismic dataset. The results in terms of uncertainties and the time-efficiency are discussed.

Chapter 1. Theoretical background

The chapter introduces the Surface Waves, and their propagation characteristics. Two approaches of inversion are explained and the method for Tomographic approach is described.

1.1 Surface Waves

Surface waves are seismic waves that propagate parallel to the free surface in a layer of limited depth (Richart, Woods, Hall, 1970). Figure 1.1 shows an example of a seismic field record where seismic wave are visible (Drijroningen, 2011), and it is clear that the SWs are characterised as the most energetic event.
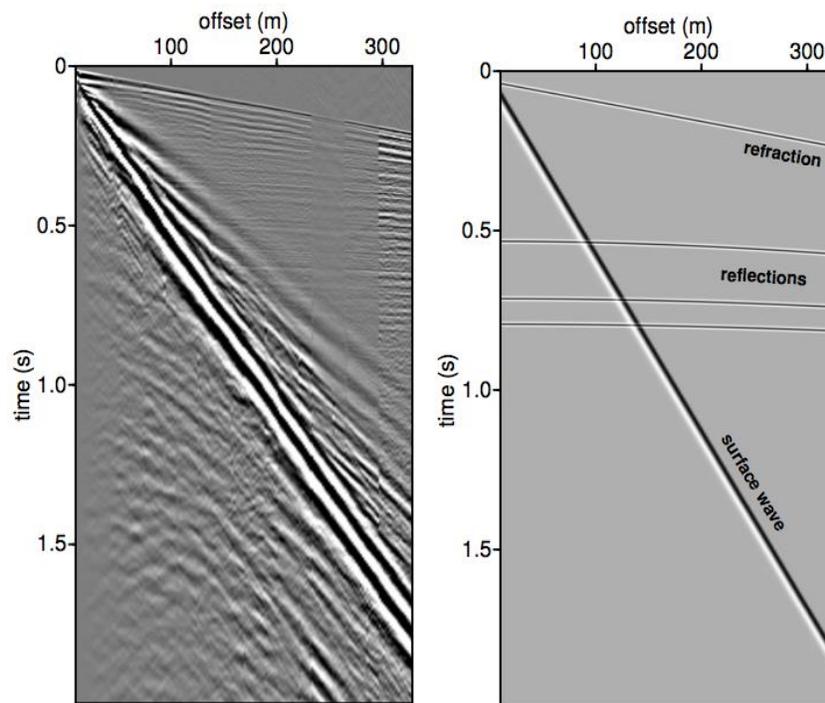
Figure 1. 1. Seismic events obtained from the seismic survey (Drijroningen, 2011).

There exist multiple types of SWs: Love, Stoneley, Rayleigh and etc.

Love wave represents wave which oscillates the particles in horizontal and perpendicular manner to the direction of the wave propagation. This SW type firstly has been described by Augustus Edward Hough Love in 1911.

Stoneley wave (which sometimes is referred to as a Scholte wave) as all the SWs propagates along the interface between two media and is generally used in the sonic logging of the borehole to define the location of fractures and to estimate permeability of the formation.

Rayleigh waves have been defined by Lord Rayleigh in 1885 as the waves that trigger particles to move in an elliptical way within the vertical (to the direction of wave propagation) plane.

1.2 Surface Waves propagation

SW investigation is based on the phenomenon of geometric dispersion, which means that different harmonics of the propagating wave have different wavelengths that define their propagation depths. Furthermore, if the wave propagates in vertically heterogeneous media, the different harmonics (with different wavelengths) will propagate with different phase velocities that depend on the properties of the subsurface (Socco, Foti, and Boiero 2010) (Figure 1.2). So, it becomes clear that the SW propagation velocity is frequency-dependent.

Due to the existence of the geometric dispersion, it is possible to retrieve dispersive characteristics of the layered near-surface subsoil and as a result to retrieve the shear wave parameters. The description of this kind of phenomenon is characterised by DCs, which should be inverted to get the shear wave velocity (Vs) parameters of the subsoil (Bensen et al., 2007; Ritzwoller and Levshin, 1998; Trampert & Woodhouse, 1996).
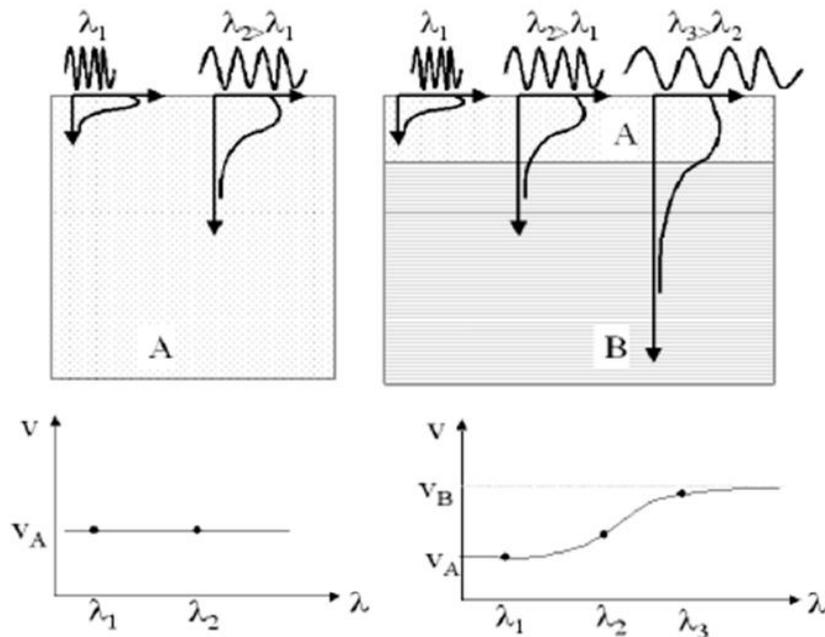
Figure 1. 2.Difference in velocities in various environments. a) Homogeneous environment; b) Heterogeneous environment, where the velocity of wave propagation depends on the wavelength (Strobbia 2003).

1.3 Surface wave analysis

The SW analysis is based on three stages:

- Acquisition
- Processing
- Inversion

There exist several acquisition techniques in the seismic wave surveying and can be based on the use of various seismic survey equipment, including a single receiver, a couple of receivers, or an array of receivers. Depending on the goal, the scale of the survey, the acquisition technique may vary (Socco and Strobbia, 2004).

16

In the processing stage, the gathered data are processed to retrieve DCs, which are given as input to the inversion stage (Figure 1.3).
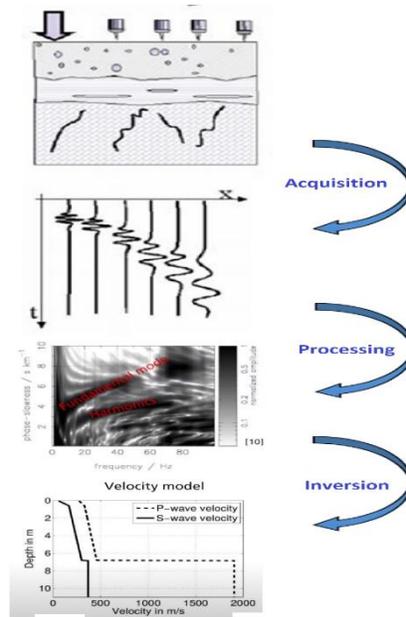
Figure 1. 3.Three main steps of the SW analysis (2003 Strobbia).

There are two kinds of approaches to invert the DCs:

- Non-tomographic;
- Tomographic.

1.4 Non-tomographic approach

The non-tomographic approach is based on the autopicking of the spectral maxima in frequency and slowness ($\omega - p$) or frequency and wavenumber (f-k) domains. One of the methods of processing that uses non-tomographic approach is multichannel SW method. An advantage of the method is the high signal to noise

ratio. On the other hand, it has low potential in detecting lateral variations. Multichannel SW uses the traces of many receives and performs a spectral transform (f-k is an example). The extracted DCs represent the dispersion characteristics of the subsurface portion below the receivers used in the transform. The DC location is considered in this case to be the center of the receiver array used in the spectral transform. They are typically inverted for Vs. The Vs model that we can retrieve is a 1D model, located at the same position as the DC. And this is the reason why they have a low lateral resolution because using only one DC, representing the entire portion below the receivers cannot detect the lateral variability within this subsurface portion. The properties are averaged.

1.5 Tomographic approach

SW tomography is a well-known practice in seismology for estimation of the regional and global velocity model (Zielhuis and Nolet, 1994; Woodhouse et al., 1995; Simons 2002; Shapiro and Ritzwoller, 2002; Yang, 2007; De Ridder and Biondi 2015). Despite the fact the tomography began to grow popular only in the 1950-1960s due to improvements in analytical methods and development in scientific instruments for earthquake measurements (Aki and Richards, 1980; Ben-Menahem, 1995), it has already been used since the 1920s to define the characteristics of the interior of the Earth. In the 1970s, the development of large and flexible networks contributed to a significant study of the layout of the upper mantle (Romanowicz, 2002; Trampert and Woodhouse, 1995). Even though the method was developed for global seismology, it has proven successful in exploration. In 2019 by Da Col performed a 3D SW tomography over a seismic data set from Siilinjärvi (Finland) to expand the knowledge of the dimensions of the mineralised body (Federico Da Col et al., 2019).

The SW tomography can be performed both on active (Socco et al. 2014; Swoboda et al. 2013) and passive (Picozzi et al. 2009) sources. The tomography approach

performed good inversion results for a complicated dataset that has been reprocessed for Smart Exploration from Ludvika mining sites (Papadopoulou et al., 2019).

In tomography approach, the DCs should be obtained for several paths traversing the region of interest. A path coverage as uniform as possible is necessary to achieve a fair resolution of lateral variations. Hence the method that will be used in this thesis for retrieving the DCs for the inversion through the tomographic approach is the two-station method, as introduced by Bloch and Hales,1968 and modified by Yao in 2005.

1.5.1 Two-station method

The method is based on the building the DCs by cross-correlating the signals retrieved from two receivers, and the time lags of the maxima obtained in the matrix from the cross-correlation correspond to the travel times of the propagation modes from one to another receiver. Furthermore, since there are many couples of receivers, the path-average approximation of DCs (Woodhouse, 1974) present a high degree of overlap and high data coverage, which in the end makes SW tomography to be more sensitive to lateral property variation.

Surface wave tomography can be done by processing the surface waves present in seismic records acquired for other purposes, typically seismic reflection surveys. The first step consists of selecting couples of receivers which will secure a high DC coverage. The DC coverage that must be granted is:

- In space, so to have a fair distribution of the couples along the investigated line or over the investigated area and many overlaps.

- In terms of wavelengths, to have DCs covering different wavelength ranges, this is ruled by the distance between the receivers of each pair. So, extracting the DCs using couples of receivers with a variety of distance, we achieve high wavelength coverage.

The two-station method is composed of 3 steps:

I.   It is selecting the couples of receivers that are aligned with the shot at a particular azimuth. This step is held by a code that defines the receiver groups that are in line with each shot with an error of azimuthal angle of 1°, which is essential for the accuracy of the retrieved phase velocity. After defining the aligned group of receivers on certain azimuthal angle, the code automatically generates random pairs of receivers with different distances. An example of a random pair selection is demonstrated in Figure 1.4a. Further, the traces of the selected receivers are used for the next steps.

II.   Using multiple Gaussian filters, we calculate the group-velocity envelope for both of aligned receivers. As a result, we pick the high energy events, that will be used as the time-windowing of the trace in order to strengthen the events related to the Rayleigh waves (Dziewonski, Bloch, and Landisman, 1969).

III.   In this step, the time-windowed traces are cross-multiplied frequency by frequency and then transformed from time to phase-velocity domain. On the estimated final cross-multiplication matrix (Figure 1.4b), we manually pick an area where automatically will be searched the maxima.

The 1st step is used only in 3D settings. While in 2D settings that are used in thesis work, all the receivers are aligned with all the shots. In 2D setting configuration we may have several individual shots for the same pair of receivers in line. The matrices of the individual shots can be summed to retrieve the stacked cross-correlation matrix. Summing up the cross-multiplication matrices increases the resolution and S/N ratio.
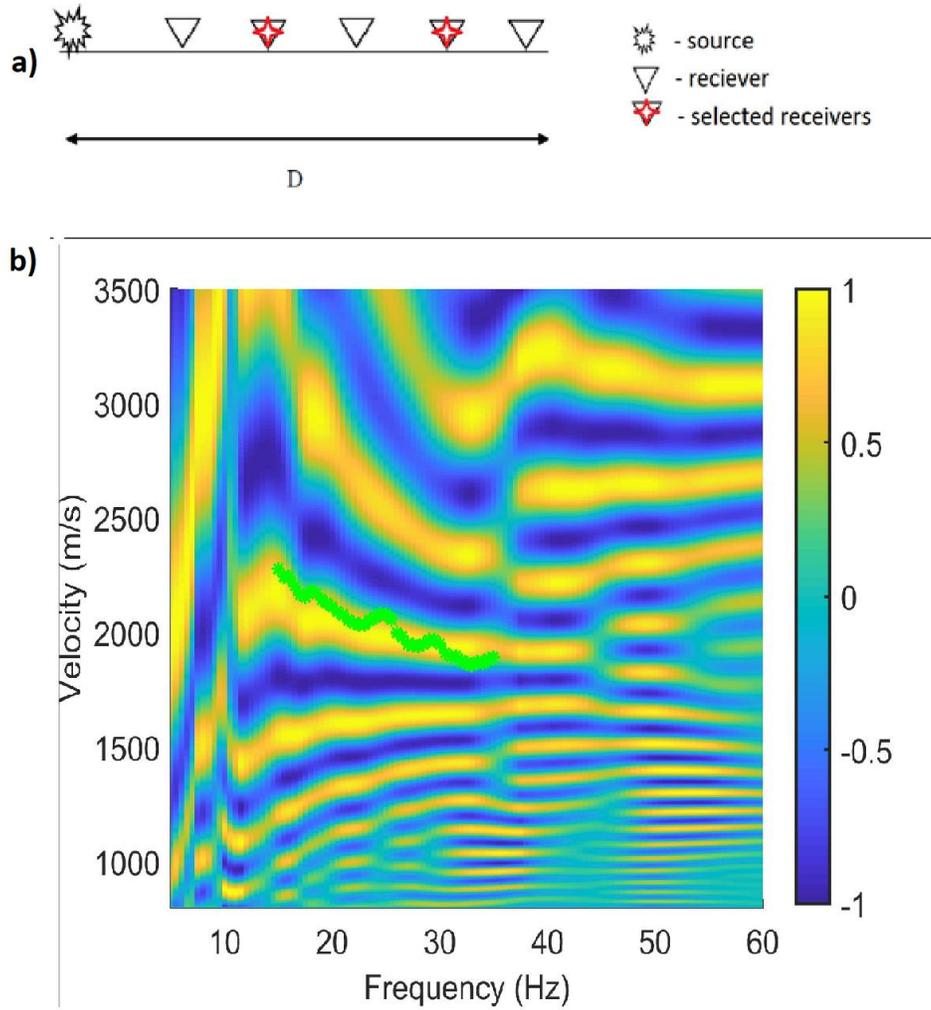
Figure 1. 4. (a) An example of the selection of a pair of receivers, which are indicated with red stars, for the two-station method; (b) a cross-multiplication matrix, represented in phase velocity and frequency domains, the green stars represent reference curve.

Chapter 2. Data

2.1 Dataset

The area belongs to the Bergslagen historical mineral district of south-central Sweden. Blotberget, known for its high-quality iron-oxide mineral deposits (Malehmir et al. 2017).

This Bergslagen district (Figure 2.1) represents banded iron formation, apatite-rich and skarn-type iron-oxide deposits that provide 40% of produced iron ore in this mineral district (Magnusson and Geijer 1944; Stephens et al. 2009). Because of high-quality and abundant iron-oxide deposits of Ludvika mining area, Blötberget is well known and it is one of the reasons for the development of the steel-making industry in Sweden. The site had a history of mining from the 16th century until the late 20th century when the price of steel in Europe dropped. Due to the recent increase in the price of iron in the market, the interest in mineral exploration of this area increased as well. Another positive aspect of the Blötberget deposit site in comparison with other sites in the vicinity is the fact that Ludvika mining area is less affected by urbanisation.

In Blötberget, the mineralisation is reported to reach a known depth of 800-850 meters (Marieset al, 2017), making Ludvika very attractive for further mining operations. The majority of the host rocks are metamorphosed volcano-sedimentary rocks. Dacitic, andesitic, feldspar porphyritic metavolcanic rocks are considered to be rich with the ore (Allen et al. 1996; Ripa and Kubler 2003). The mineralisation of hematite and magnetite comprise 50 % of iron. The ore-bearing unit has a dip of 45° in the south-west direction up to 500m depth, after which it slightly dips till 850 meters.
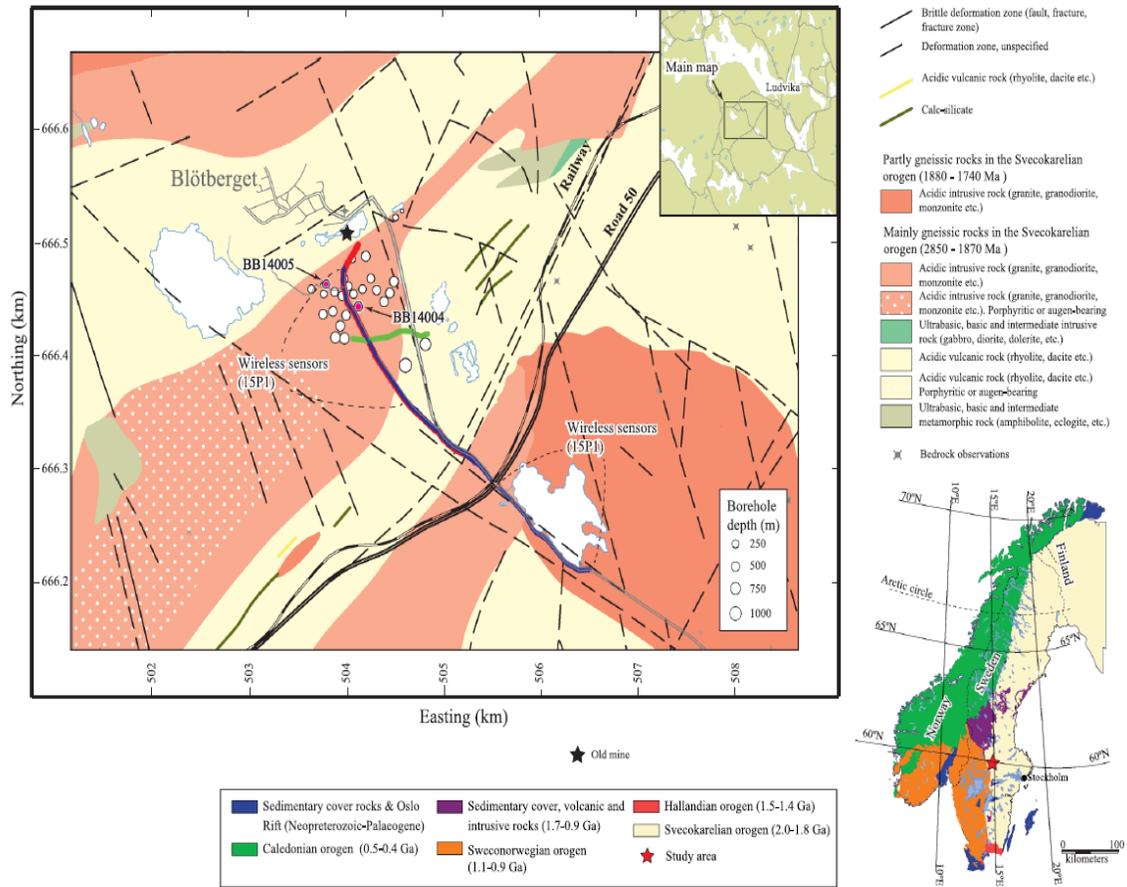
Figure 2. 1. The geological map of the mining area of Blötberget mineral district located in central Sweden showing the main lithological units and seismic profiles (Maries et al. 2017).

2.2 Previous work on the site

In 2015 and 2016 two reflection-seismic campaigns were performed at Ludvika mine, to map the deeper mineralised portions. Campaigns aimed to define a possible extension of the mineralisation for further mining operations. Further conventional processing approaches defined the extension of the mineralised horizons around 300 meters more profound than the already know depth (850 m), that can be considered as a potential resource of additional iron-oxide (Figure 2.2) (Markovic, Maries, Alireza M. et ad. 2019).
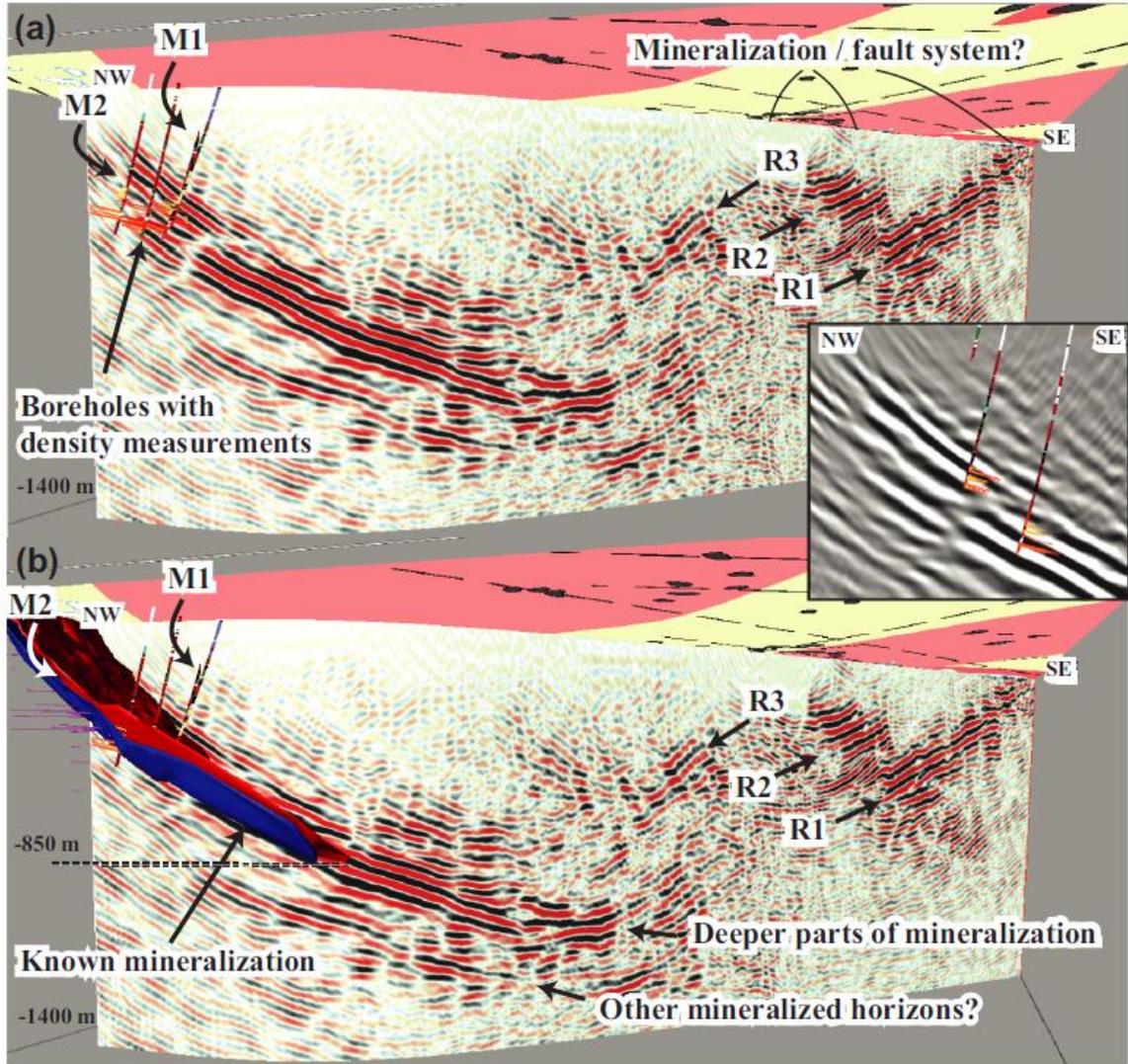
Figure 2. 2.A 3D visualisation of the seismic section with a) boreholes with density measurements on the right-side closer, b) blue and red surfaces that represent 850m of the known portions of the mineralisation and possible extension of it down to 1200 m (Markovic ad et.2019)

Multichannel SW analysis has been performed on the 2016 dataset to estimate the statics (Papadopoulou et al. 2020). Their starting point was based on the detection of the sharp lateral variations to ensure that the extracted DCs are not affected by those variations. They showed that the site presents several sharp lateral variations in the near-surface. However, by applying a workflow, suitable for such

environments, the sensitivity of SW to the near-surface properties can be profitably used to provide high-quality statics.

Papadopoulou et al. (2019) showed, that using SW tomography on the same site, provided a high-resolution VS model (Figure 2.3). Even though the data quality on this site was low, stacking of the cross-correlation matrices of the same pair of receivers led to the extraction of high-quality DCs, with adequate coverage (Figure 2.4). Therefore, the encouraging results of tomography approach in providing the Vs model and additional information about the intersection of the profile by a large-scale structure, probably a fault, has been reviled (black dashed circle in Figure 2.3).
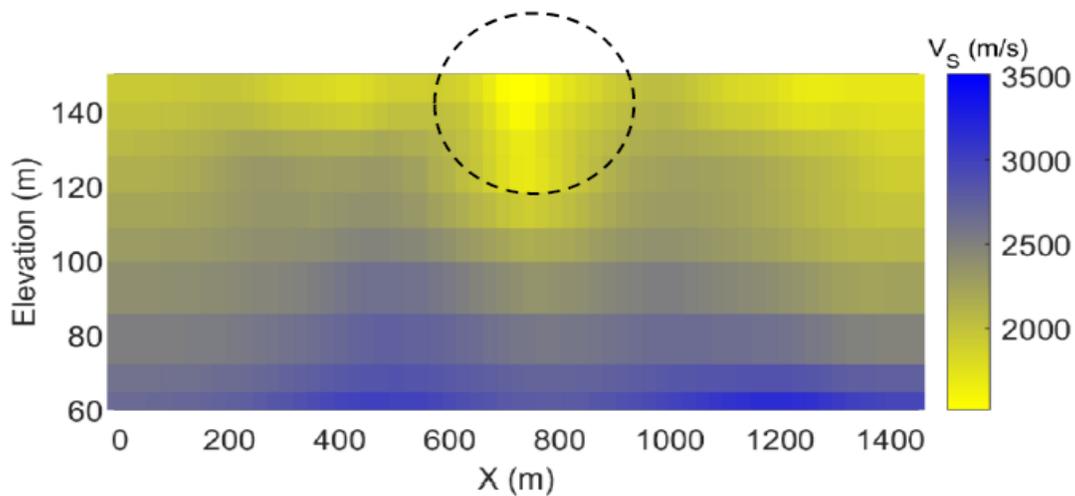
Figure 2. 3. Figure 2.3. Result of the inversion. The strong anomaly caused by a local discontinuity is highlighted with a dashed circle (Papadopoulou et al. 2019).
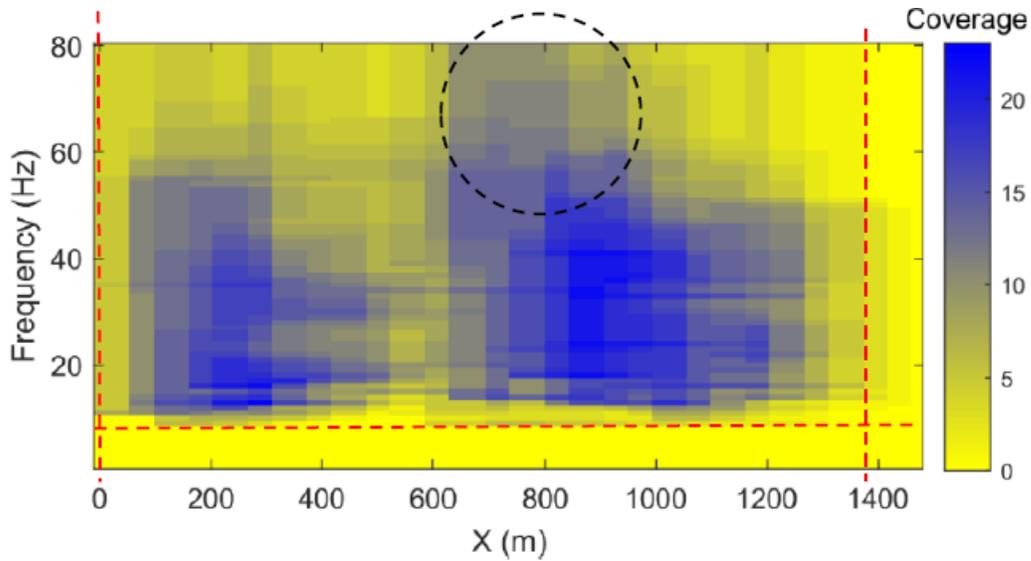
Figure 2. 4. Coverage of the data along the line. Dashed lines indicate the range of the frequencies and positions that provided coverage>0, which also includes the anomaly zone (dashed circle). (Papadopoulou et al. 2019).

2.3 Data Acquisition

The purpose of the 2019 (April 27 – May 18) 3D seismic acquisition that was held by Smart Exploration was to increase the existing knowledge on the lateral and in-depth extent of the mineralisation. It consisted of 1266 10-Hz wireless and cabled receivers, sampling at 1ms, and 1052 shot positions (Figure 2.5). The shots were stroke with 32 kg vibrating source with spacing equal to 10-20 m, while the receiver spacing was also equal 10 – 20 m. In this work, we use a portion of the entire dataset, onward referred to as Line 1 and Line 2 (Figure 2.6a and Figure 7a respectively), that consist of 218 shot positions and 337 receivers; 143 shot positions and 143 receivers correspondingly. Some examples of seismic records from both of the lines with indication of SW are presented in the Figure 2.6b and Figure 2.7b
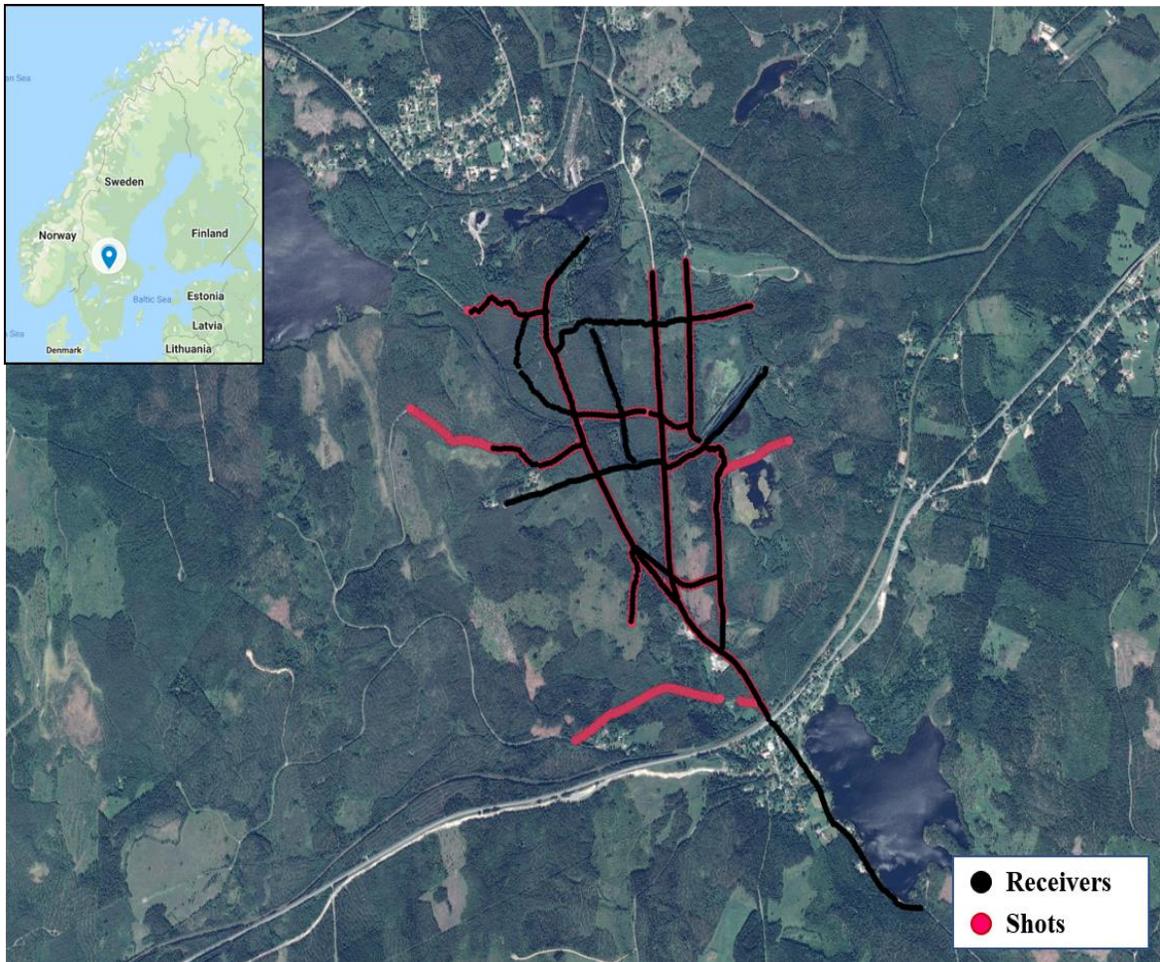
Figure 2. 5. Seismic Acquisition scheme of Ludvika mining area of Blötberget, Sweden. Acquisition date: April 27 – May 18, 2019. Black dots represent receivers, while red dots stand for shots.
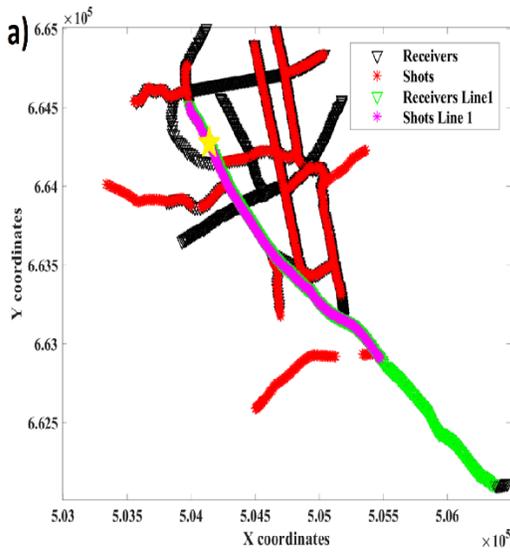
Line 1	
Number of shots	218
Number of receivers	337
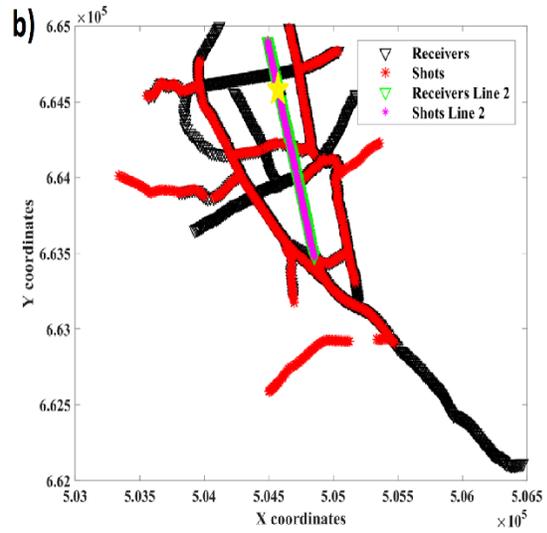
Line 2	
Number of shots	143
Number of receivers	143

Figure 2. 6. Positions of a) line 1 and b) line 2 used in this thesis work. Yellow Stars indicate the position of the source the seismic records of which are shown in the next figure.

Some examples of seismic records from both of the lines with indication of SW are presented in the Figure 2.7.
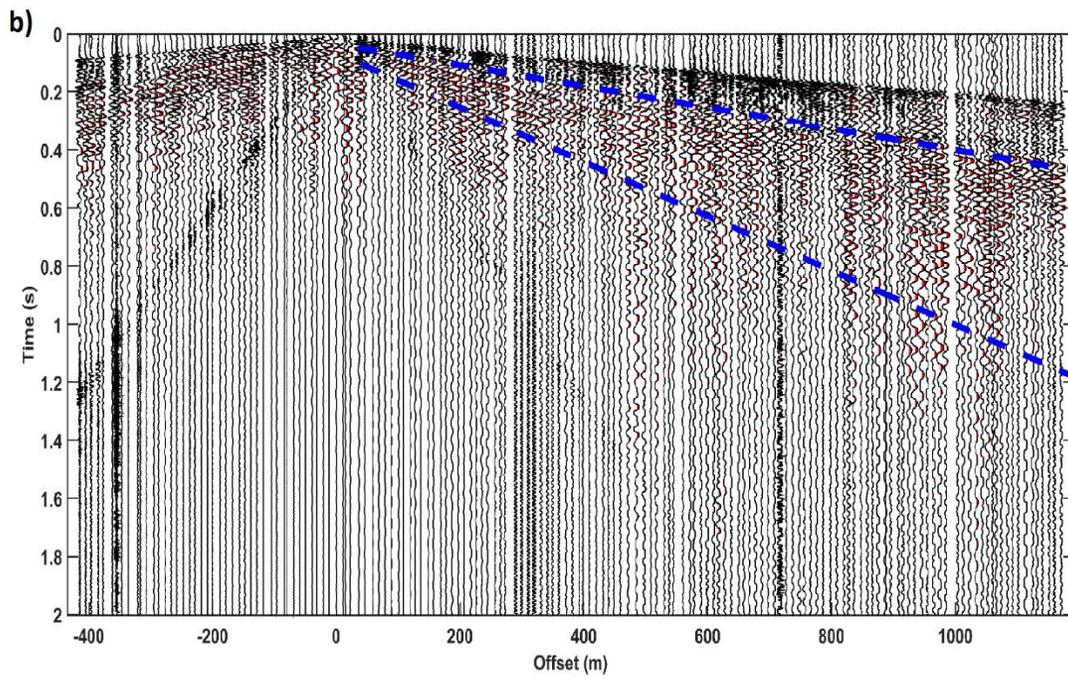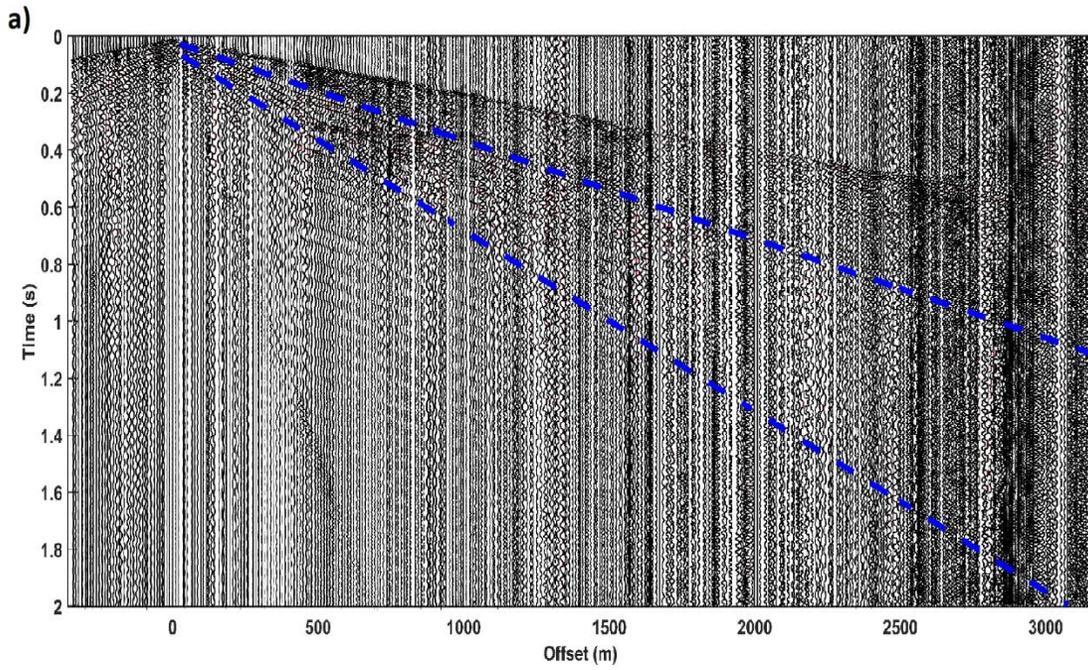
Figure 2. 7. Examples of seismic records obtained from the field. Dashed blue lines highlight SWs.

Chapter 3. Methods

The chapter deals with DC picking methods. Furthermore, in this chapter, we will outline the advantages and disadvantages of the methods, criteria for picking, and will introduce some QC.

3.1 DC picking

The picking of a dispersion curve is performed on the obtained cross-multiplication matrix. As it can be seen from Figure 1.4, the cross-multiplication matrix reviles several zones with energy maxima. So, it becomes not straightforward the picking of the appropriate maxima on the matrix, therefore, the reference curve comes in. Since we are operating with seismic lines, in which we have quite a lot of aligned receivers, we may apply multichannel analysis. The multichannel analysis processes all the aligned receivers and as an output provides a DC that may give representation about the anticipating DC trend over all the array. Hence, the obtained DC can be used as a reference one.

Not all of the cross-multiplication matrixes are suitable for dispersion curve picking, because of low-quality SWs, related to high levels of recorded noise or energy attenuation. Examples of such low-quality matrices not suitable for processing can be found in Figure 3.1, and as it may be clear that the maxima energy in none of them follows the trend of reference curve.
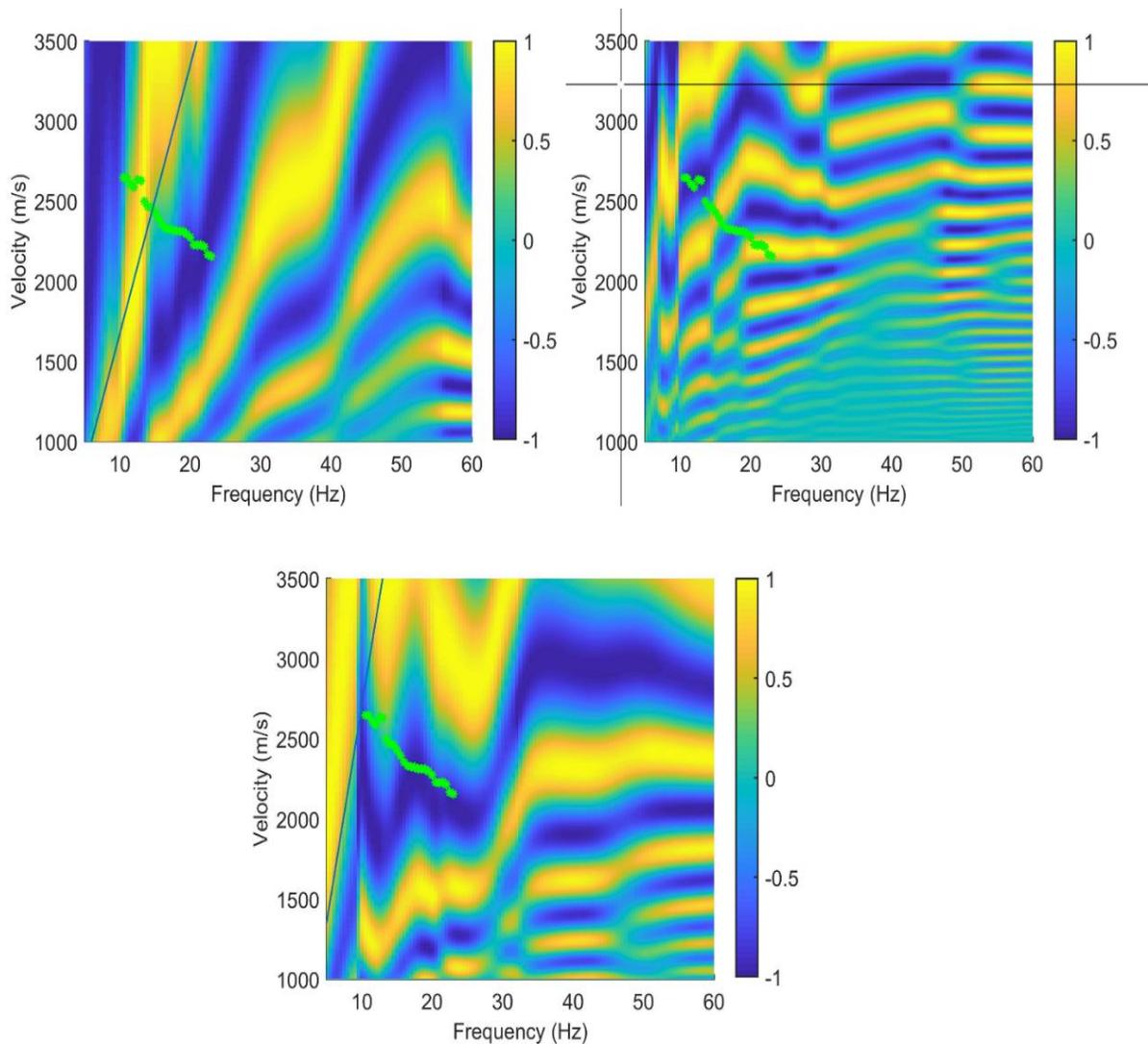
Figure 3.1. *Examples of the rejected from further processing cross-multiplication matrixes.*

There are two possibilities of DC picking:

1. Manual DC picking;

2. Automatic DC picking.

3.1.1 Manual DC picking criteria and advantage

First of all, the operator manually picks on the cross-multiplication matrix the zone closest to the reference curve, that is superposed on the matrix plot, and according to the chosen zone, the program automatically searches for the maxima in that area of the plot. Figure 3.2 demonstrates an example of cross-multiplication matrix on which the manual picking has been performed. From Figure 3.2a, it can be seen that one of the yellow zones of the plot, which stands for high energy zone, nicely coincides with the reference curve. So, the operator clicks on that area. As a result, the automatic maxima searching generates a continuous black line, ignoring the interruption (fading) (a red circle represents that after 45Hz) of the maxima, what is recommended to be cleaned, as it is assumed that the energy maxima after the interruptions. It should be underlined that the energy in the range of 0-15 Hz does not carry any significant information about the DC, therefore, should also be cleaned.

The final picked dispersion curve of the cross-multiplication matrix is represented in Figure 3.2b.
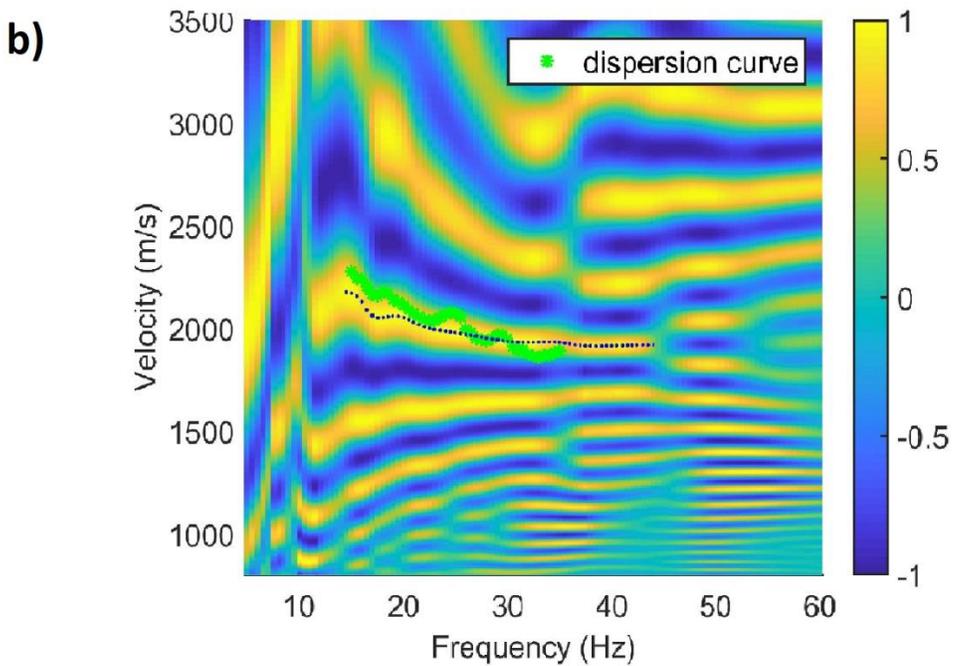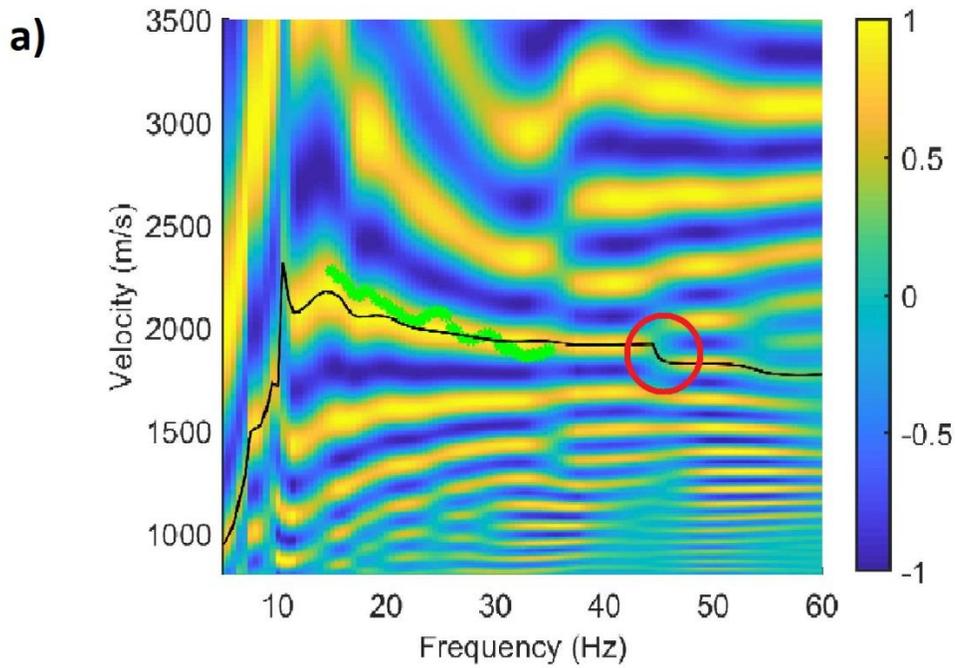
Figure 3.2. *a)An example of DC picking on the cross-multiplication matrix, where the continuous black line represents the automatic maxima searching, and the red circle represents the interruption of the maxima energy; b) blue dots represent the final version of the manually picked DC.*

It should be underlined the advantage that derives from the criteria of the manual picking is the fact that the picking is considered more accurate since the processing of it is performed under human control.

3.1.2 Automatic DC picking criteria

The workflow of the automatic DC picking method is shown in Figure 3.3. For this approach, the cross-correlation matrices of all receiver-pairs and their corresponding shots are automatically computed, based on the two-station method (chapter 2.3) and the matrices of the individual shots are summed to retrieve the stacked cross-correlation matrix. The DC extraction is based on the comparison of the cross-correlation matrix with the reference DC and the path-average DC is picked as the amplitude maxima nearest to the reference curve. A DC is picked on the stacked and the individual matrices, allowing the estimation of the standard deviation of the phase velocity between the stacked and individual DCs (V_{Rstd}), which serves as a measure of the experimental uncertainty of the DC. A second measure of uncertainty (Amplitude uncetainty, A_u), based on the amplitude of the cross-correlation corresponding to the DC is computed as:

$$ A_{u_i} = \frac{1 \big/ A_{norm_i}}{N_{shots}} V_{R_i} \qquad (3.1) $$

where i denotes the frequency components of the stacked DC, A_{norm_i} is the amplitude of the cross-correlation matrix corresponding to the i-th frequency component of the DC, normalized to its maximum amplitude and V_{R_i} its phase velocity. N_{shot} is the number of shots for which the cross-correlation matrices are stacked.

An automatic cleaning of all the DCs is performed, according to which, the parts of the DC falling on wrong (i.e., not corresponding to the fundamental-mode phase velocities) cross-correlation maxima are removed. This is achieved by computing

the gradient of the phase velocities and searching for its local peaks and comparing them with the local peaks of A_u. When these are identical, the frequency band of the DC, following or proceeding the frequency of the common gradient peaks and having an A_u higher than the median A_u of the DC, is deleted.
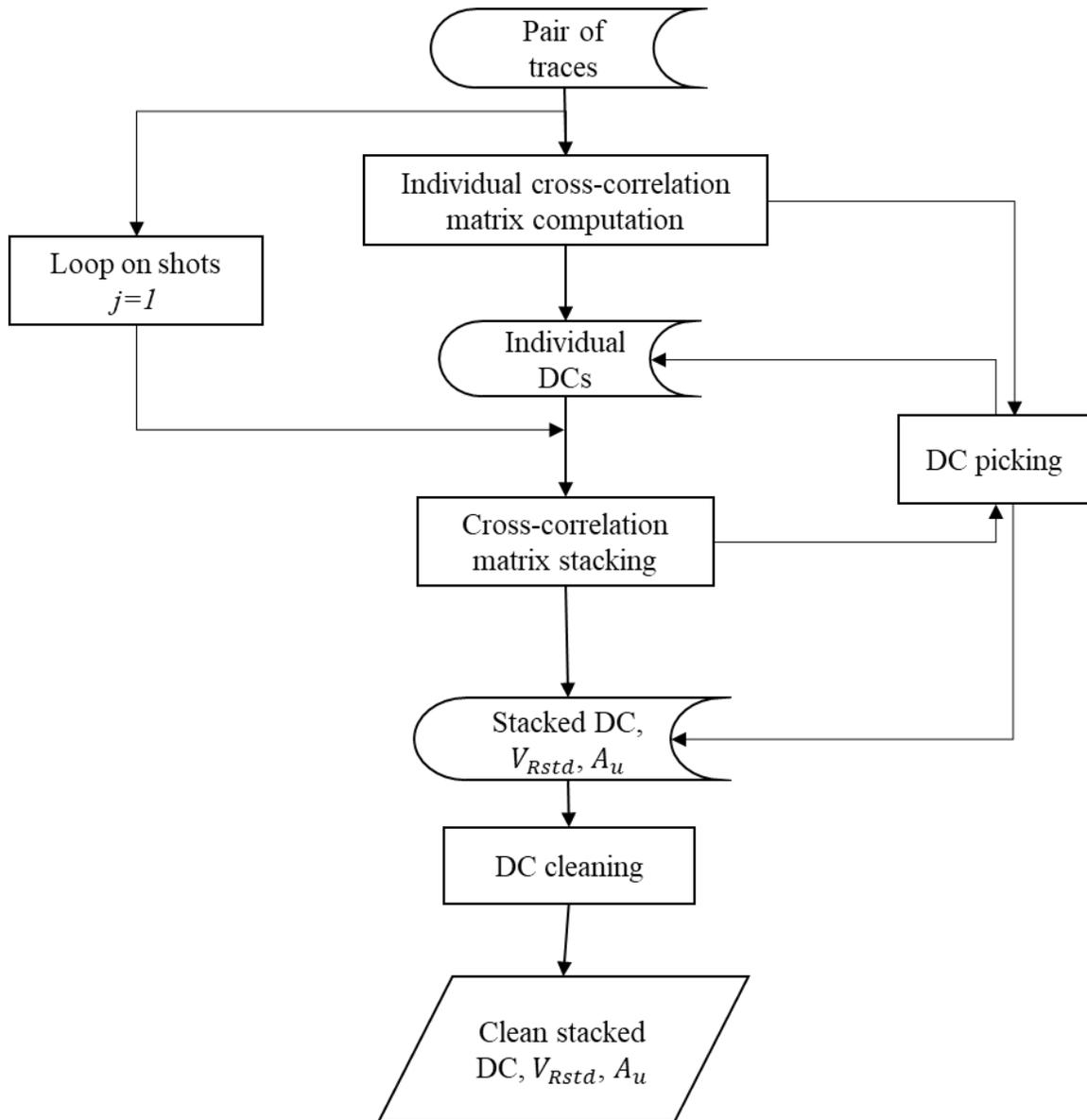
Figure 3.3. Workflow of the automatic DC picking method (Papadopoulou, personal communication)

3.2 Frequency band pitfall

A first possible pitfall of the autopicking method is the wrong frequency band of the DC. We identify two erroneous conditions in the autopicking in comparison with the manual picking:

a) Ignored frequencies;

b) Extra picked frequencies.

In the case of ignored frequencies, we may lose information that can be potentially useful. In contrast, the case of extra frequencies means that the automatic code picked DC with some frequency values that have been considered as not acceptable and have been cleaned by the operator in manual picking. In Figure 3.6a, we see an example of the dispersion curve and we compare it with its corresponding manually picked DC. The ignored frequencies are circled in red on the manually picked DC (Figure 3.4 a), while on the DC (Figure 3.4 b) the green circle highlights the extra picked frequencies.
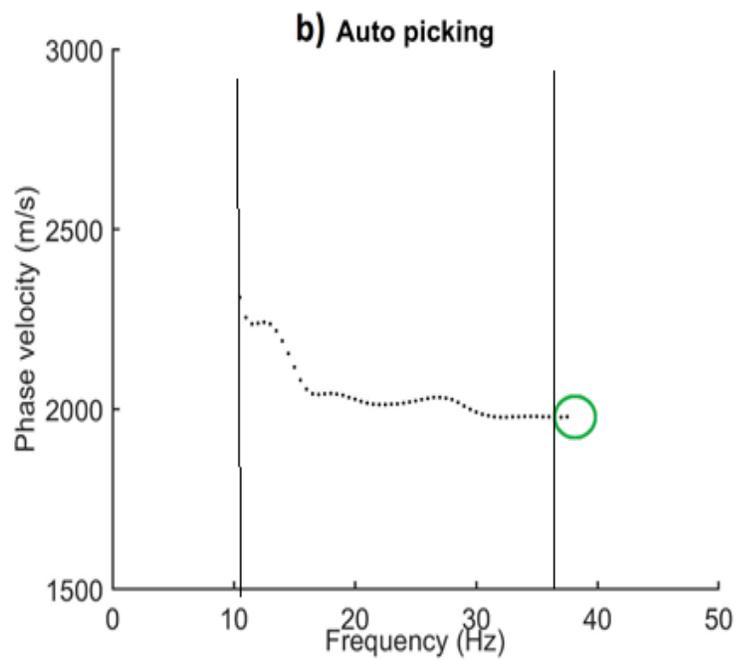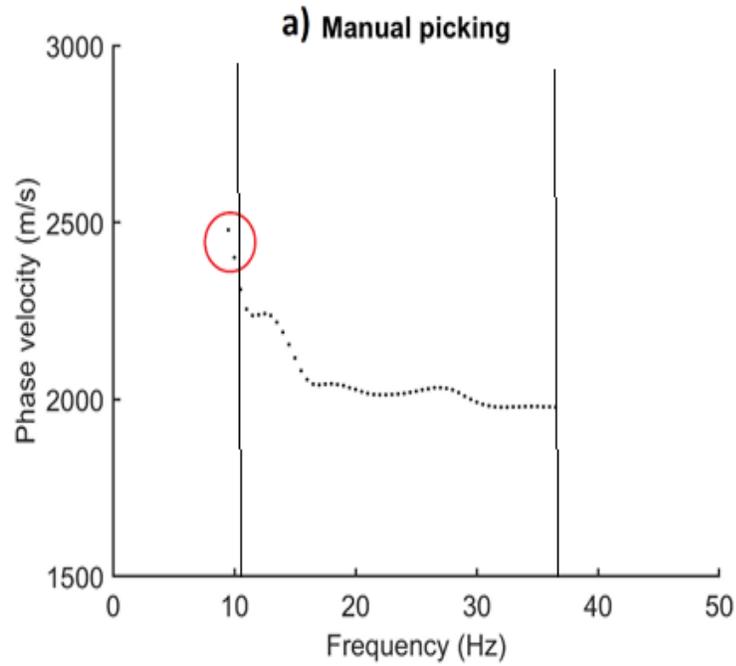
Figure 3.4. *An example of the a) manually and b) DCs corresponding to the same shot and receiver pair. The red circle indicates the ignored frequencies; the green circle indicates the extra picked frequencies.*

3.2.1 Quality control based on the frequency band

The code that has been implemented for the Comparison is provided in Annex 1. To measure the discrepancies in the frequency band of the automatic and manually picked DCs, we perform the following calculations:

a) Uncertainty based on the ignored frequencies ($e_{f_{ig}}$):

$$e_{f_{ig}} = \frac{N_{f_{ig}}}{N_{f_{man}}} * 100\% \qquad (3.2)$$

b) Uncertainty based on the extra picked frequencies ($e_{f_{extra}}$):

$$e_{f_{extra}} = \frac{N_{f_{extra}}}{N_{f_{auto}}} * 100\% \qquad (3.3)$$

Where: $N_{f_{ig}}$ — number of ignored frequency points;

$N_{f_{man}}$ — total number of frequency points picked manually;

$N_{f_{extra}}$ — number of extra picked frequency points;

$N_{f_{auto}}$ — total number of frequency points picked automatically.

3.3 Estimation of the effectiveness of DC

The comparison implies the estimation of the portion of each autopicked DC in terms of common frequency components with manually picked DC. The part of manual and autopicked DCs with intersecting frequency components is represented in the Figure 3.6 between two vertical lines.

$$N_{f_{auto,effective}} = N_{f_{auto}} - N_{f_{extra}} \qquad (3.4)$$

$$A_{effec} = \frac{N_{f_{auto,effective}}}{N_{f_{auto}}} * 100\%$$ (3.5)

$N_{f_{auto,effective}}$ – number of common frequency points between manual and autopicking.

A_{effic} - effectiveness of DC.

3.4 Estimation of the similarity index

The similarity index (S_{ind}) computes the similarity of the effective part $(N_{f_{auto,effective}})$ of the autopicked DC to the total frequency band of the manually picked DC. The maximum value of which is defined to be equal to 1 and the formula implemented to calculate it is the following:

$$S_{ind} = \frac{N_{f_{auto,effective}}}{N_{f_{man}}}$$ (3.6)

For example, the S_{ind} of DCs shown in the Figure 3.6 is equal to the ration of the 25 frequency intersecting components($N_{f_{auto,effective}}$) to the 27 manually picked frequency components. S_{ind} = 0.9 which is fairly high index.

Chapter 4. Results

This chapter deals with the application of the Comparison Test that has been introduced in the previous chapter on real data from the Ludvika site. First of all, we are defining the existing uncertainties in frequency band. Furthermore, based on those uncertainties the comparison test is performed on the DCs. The application of the automatic DC picking on two seismic lines and result of the inversion of autopicked DC are also provided in this chapter.

Before application of the comparison test by me there have done a manual and automatic picking of two sets of data from Line 1 and Line 2.

4.1 Quality control of the dataset

The comparison test between the outcomes of manual and automatic DC picking methods, as described in Chapter 3 was applied on 100 and 200 randomly chosen receiver pairs for Line 1 and Line 2, respectively. For both lines, the separation between the receivers chosen for the test was equal to 10m-500m. With the given range of receivers' separation, it is possible to have 17556 and 6615 receiver pairs on Line 1 and Line 2 respectively.

From all the processed seismic data of the 100 receiver pairs of Line 1 both of methods rejected 13 cross-correlation matrixes while from other 87 cross-correlation matrixes the DCs have been saved. The result of manual picking is presented in Figure 4.1.

The DCs retrieved from the autopicking of the tested 100 receiver pairs are presented in Figure 4.2.
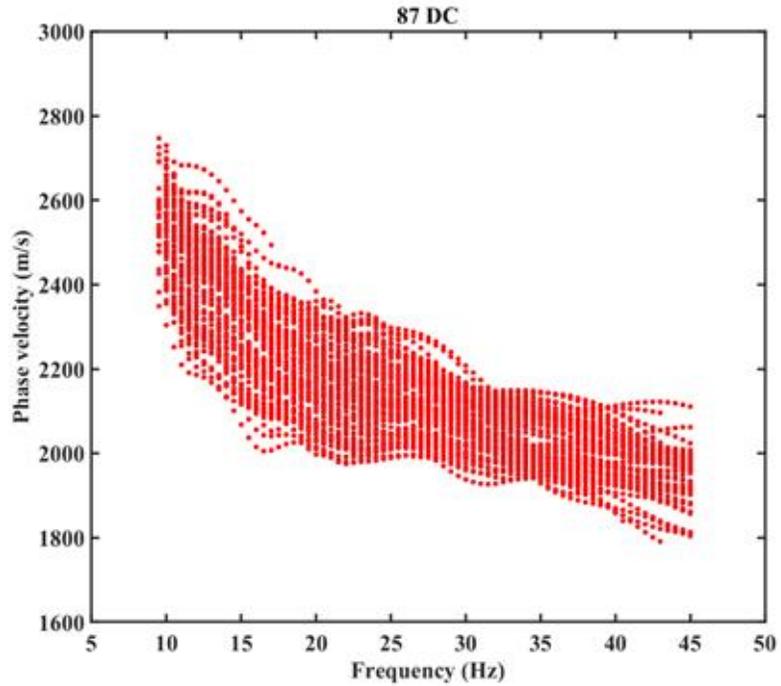
Figure 4.1. *Manually picked DCs from 100 receiver pairs' seismic data relevant to Line 1.*
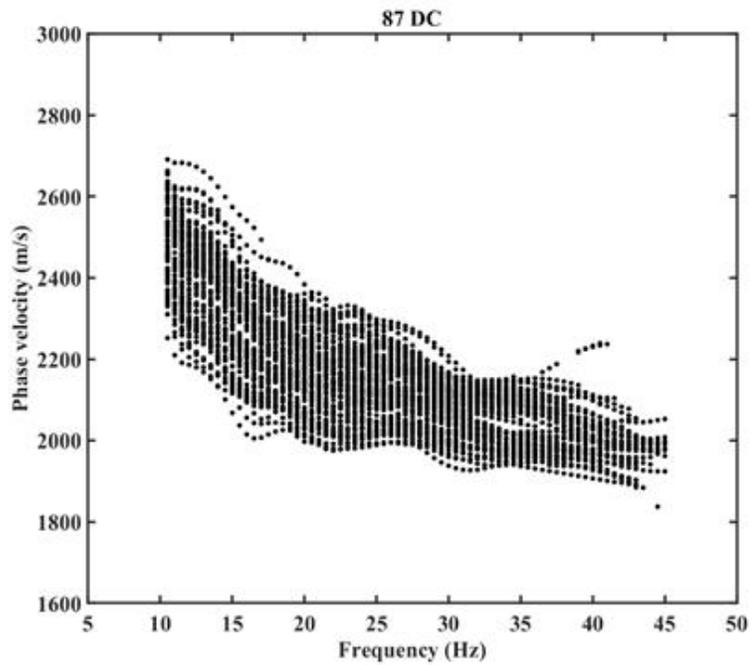
Figure 4.2. *DCs relevant to those 87 receiver pairs from which the manually picked DCs have been saved. Line 1.*
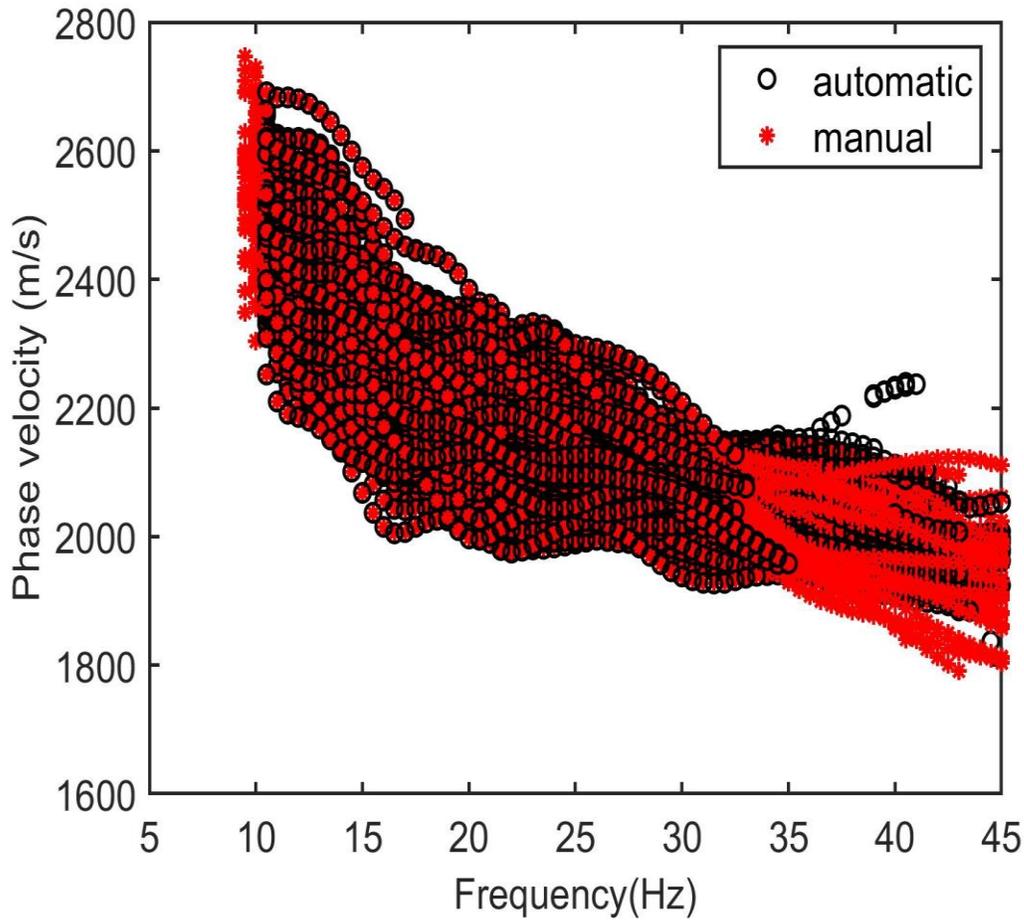
Figure 4.3. *Superposition of 87 DCs retrieved by manual and autopickings. Where the red colour characterises manually picked DCs, while black ones.*

The superposition of the curves (Figure 4.3), as preliminary comparison reviles a good matching between two methods. In the following sections, there will be presented the implementation of the comparison test of these DCs.

For Line 2 both of the picking methods rejected 135 cross-correlation matrices and picked 65 DCs related to identical receiver pairs. Figure 4.4 shows the superposition of the retrieved DCs.
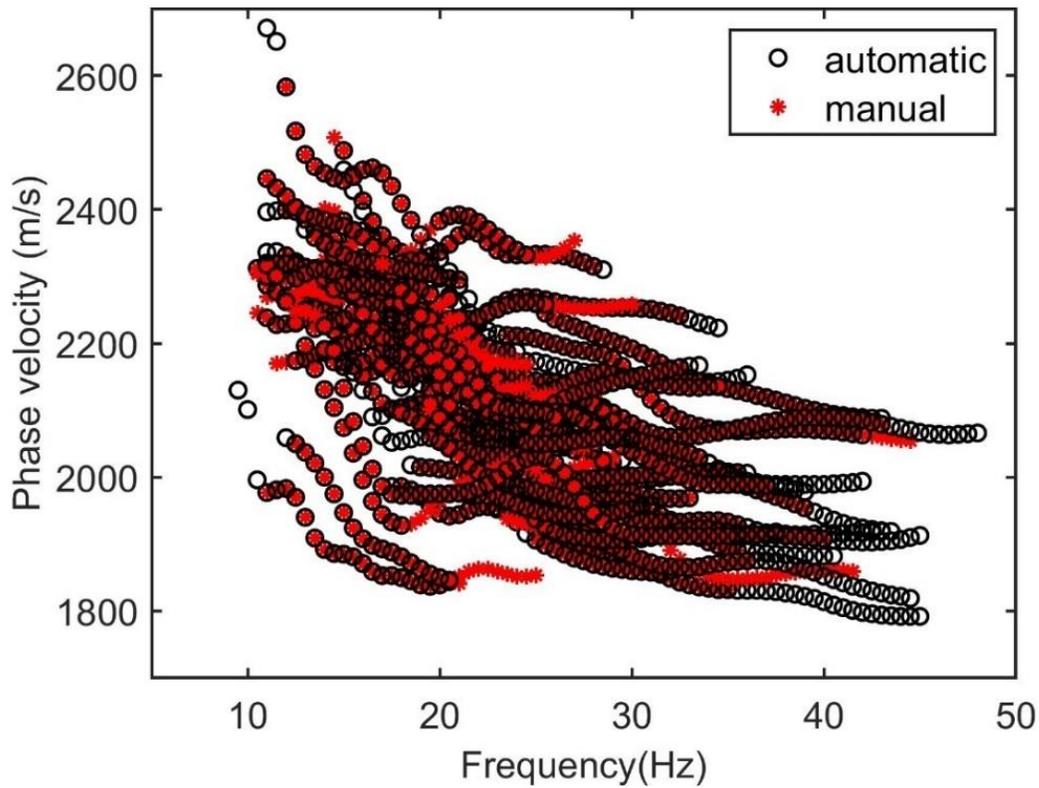
Figure 4. 4. Superposition of 65 DCs retrieved by manual and automatic DC pickings from the seismic data of 200 receiver pairs allocated on Line 2.

4.1.1 Frequency domain uncertainties

The implemented MATLAB code (Annex 1), according to Section 3.5, defines the existing uncertainties in the frequency domain of autopicking in terms of ignored and extra picked frequencies. The percentage of those uncertainties per each DC is presented in Figure 4.5.
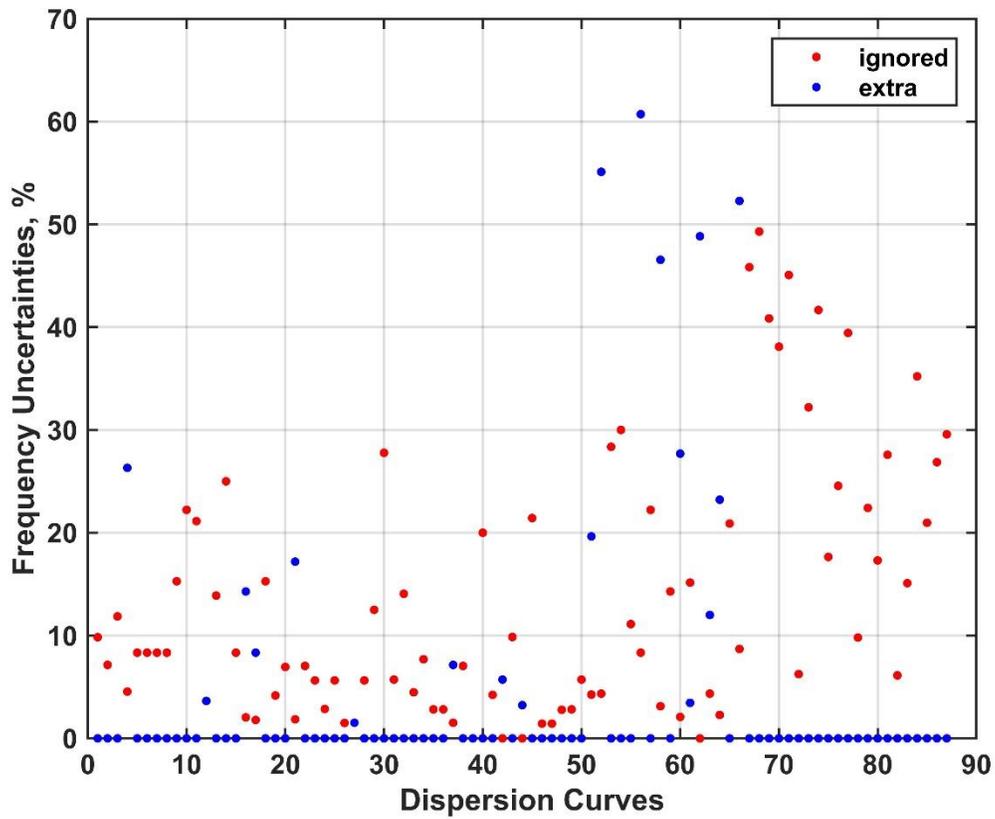
Figure 4. 5. Distribution of the frequency uncertainties on the 87 DCs. Red and blue dots represent the percentage of ignored and extra picked frequencies, respectively.

The result of the performed frequency band uncertainties' test on the dataset of Line 2 is given in Figure 4.6.
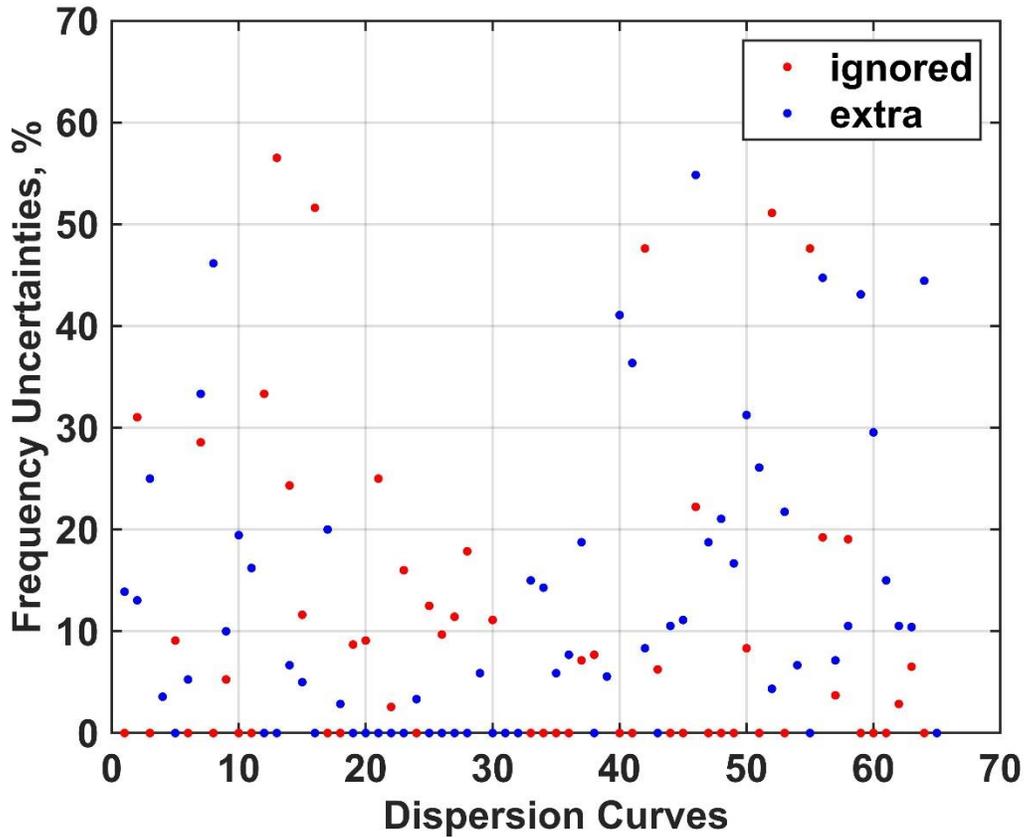
Figure 4. 6. Extra and ignored frequencies by automatic DCs of Line 2.

It is evident that uncertainties mostly do not exceed 30%. The reason that these uncertainties exist can be related to the autopicking algorithm developing by Papadopoulou. In autopicking algorithm, the frequency components of DCs that have high uncertainties in terms of amplitude and phase velocity get ignored to be picked by algorithm.

4.1.2 Effectiveness of the DC

The effectiveness of the DCs of Line 1 and Line 2 is computed using equation 3.4 and is plotted as histograms in Figure 4.7 a) and b) respectively.

From the plot, we see that 75 out 87 (86%) of DCs of Line 1 show more than 90% of effectiveness. Which means that 86% of the autopicked DCs have more than 90% of common frequency components. Nevertheless, if we compare with the results of Line 2 (Figure 4.7 b) we may see that the common components are negligibly less, it may be related to the lower quality of the gathered data.
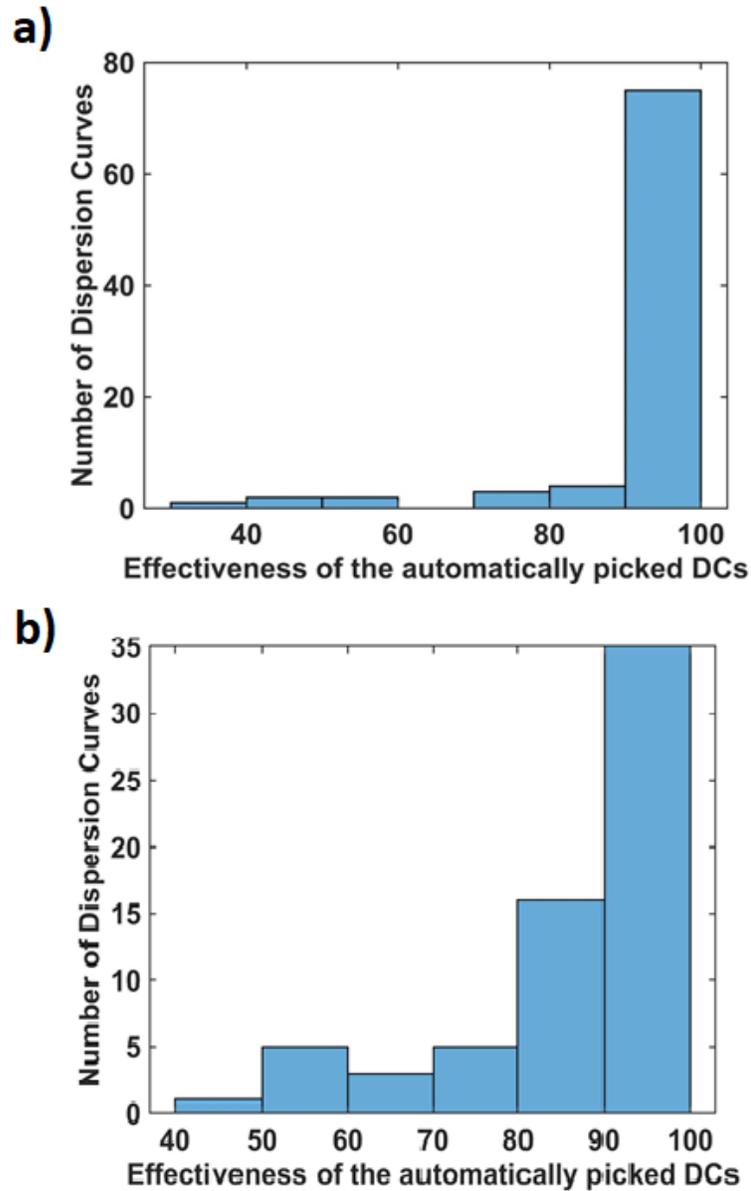
Figure 4. 7. Effectiveness of the DCs for a) Line 1 and b) Line 2.

4.3 Similarity Index

By applying equation 3.5, we computed the similarity between the autopicking and the manual picking DCs. The histogram of the similarity index is presented in Figure 4.8. It shows that 50 out of 87 DCs (57.5%) have an index of similarity more than 0.9, and only 9 DCs show similarity indexes less than 0.7. The average similarity of all autopicked DCs to the manually picked DCs is equal to 0.87, which can be considered as a reasonably high index.
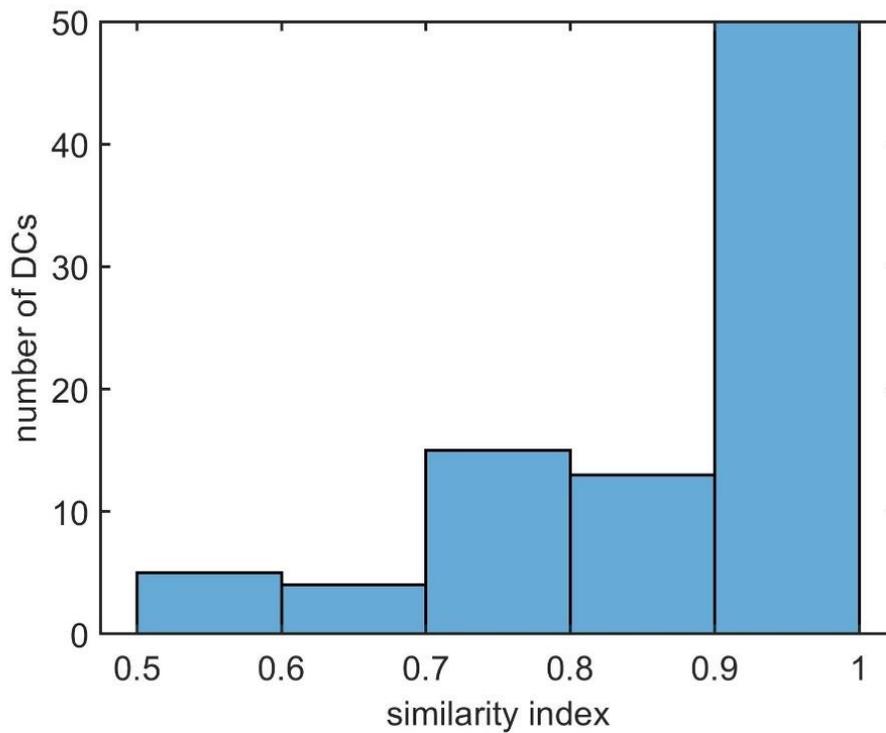
Figure 4. 8. . Similarity index of DCs.
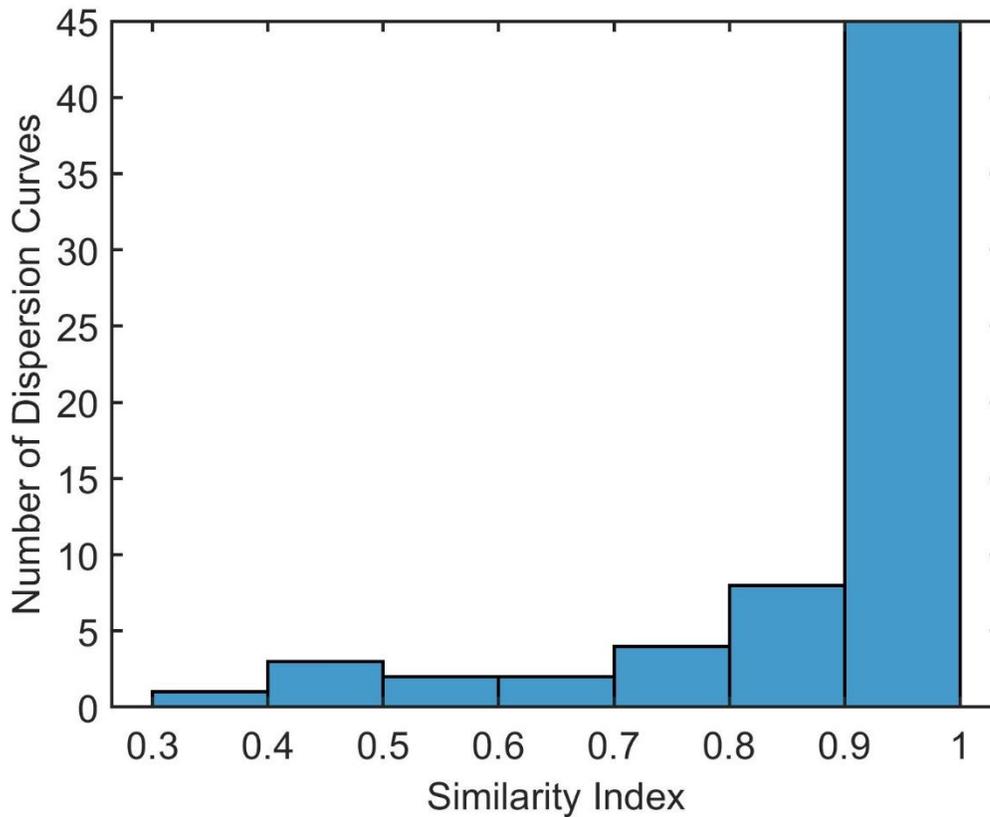
Figure 4. 9. . Similarity indices of the manual and automatic DCs from Line 2.

Around 82% of the DCs demonstrate similarity index of more than 0.8 (Figure 4.9), while the average value of the similarity index is 0.89.

Samples of the automatically and manually picked DCs with the lowest similarity indexes of 0.55 and 0.53 are given in Figure 4.10 a) and b) respectively. From it becomes evident that the reason for the low similarity index is a rejection of some part of the frequency components by the autopicking. Rejection of the frequency components by autopicking is the pitfall of the autopicking algorithm.
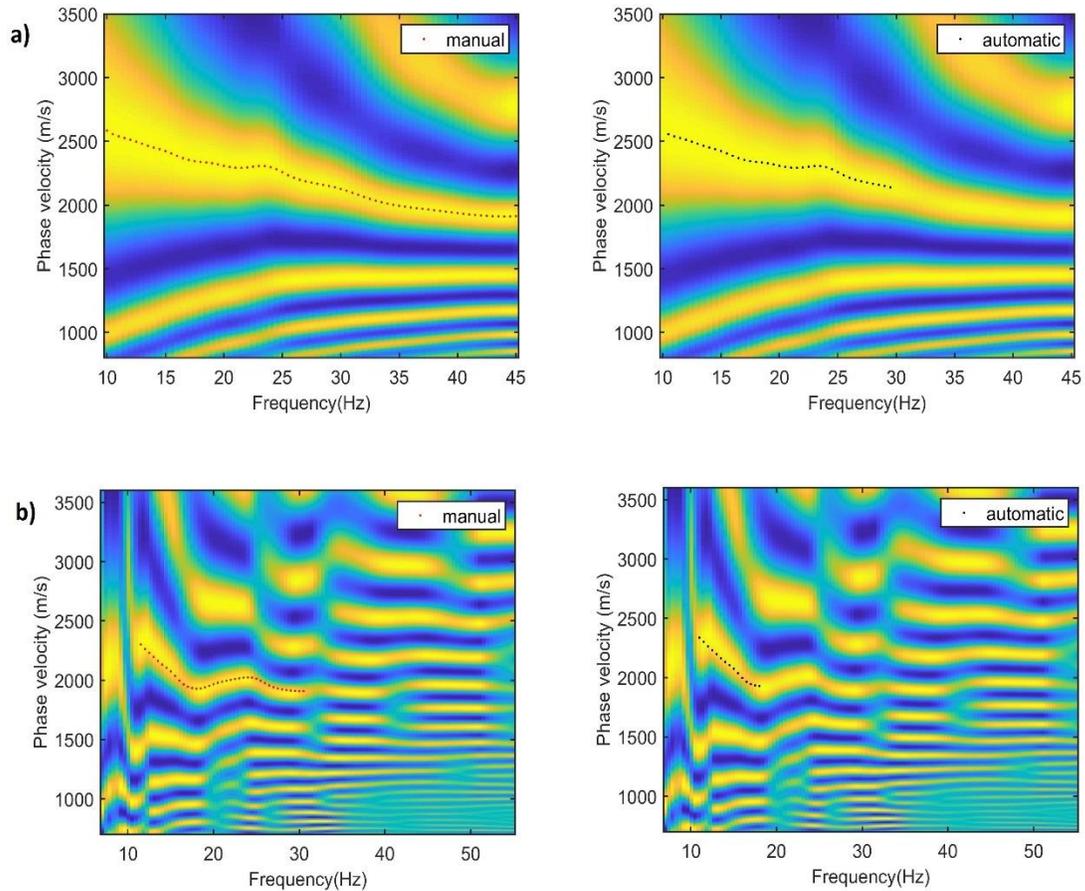
Figure 4. 10. Manually and DCs with a)0.55 and b)0.53 similarity indexes.

4.4 Time efficiency

An additional comparison which is carried out is the determination of the time efficiency of the autopicking by comparing the time required for the automatic dispersion picking and the time spent for the manual processing of the DCs.

The comparison, which shows a considerable advantage of the automatic DC picking over the manual one, is the comparison of the time that is spent for both of the methods. The timing results for the manual and automatic DC pickings for both Lines are the following:

	Time for Manual picking (s)	Time for Autopicking (s)
Line 1	2135	769
Line 2	8448	1142

Table 4.1. *Time spent on manual picking and autopicking of DCs for Line 1 and Line 2.*

For Line 1 the autopicking is almost three times faster than the manual one. Moreover, it should also be noticed that the timing results are related to the time that has been spent only just for 100 receiver pairs' dataset that has been chosen for testing.

Since the total dataset consists of 17556 receiver pairs, its manual processing, through the extrapolation would have been 104 hrs. If we consider 8 hours for each working day of the operator, we conclude that with the manual processing of Line 1 would last approximately 13-14 days. In comparison, the autopicking will deal with it just 1-2 days.

The time spent for manual picking on Line 2 is equal to 8448 seconds, while the autopicking coped with the task in 1142 seconds. Since on Line 2, there has been processed two times more DCs it took several times more time for manual processing; therefore, it can be an excellent example illustrating the high efficiency of the autopicking.

Another advantage of the autopicking is the fact that the operator is not required, it is all done by the machine without manpower.

.

4.5 Automatic DC picking of all the Line 1 and Line 2

The performed comparison test showed reasonably accurate results of the

automatic DC picking. Hence, it can be considered to be suitable for application on the whole dataset both of the Line1 and Line 2, and the results of the DCs are presented in the Figure 4.11 and Figure 4.12 respectively (Papadopoulou, personal communication):

Line 1: Out of 56616 possible pairs, 2485 DCs;

Line 2: Out of 12720 possible pairs, 1214 DCs.

It is evident that the difference is to big between picked and possible pairs(each of which corresponds to one pair). This difference may be related with poor seismic data. Nevertheless, the picked curves are enough to get good inversion results.
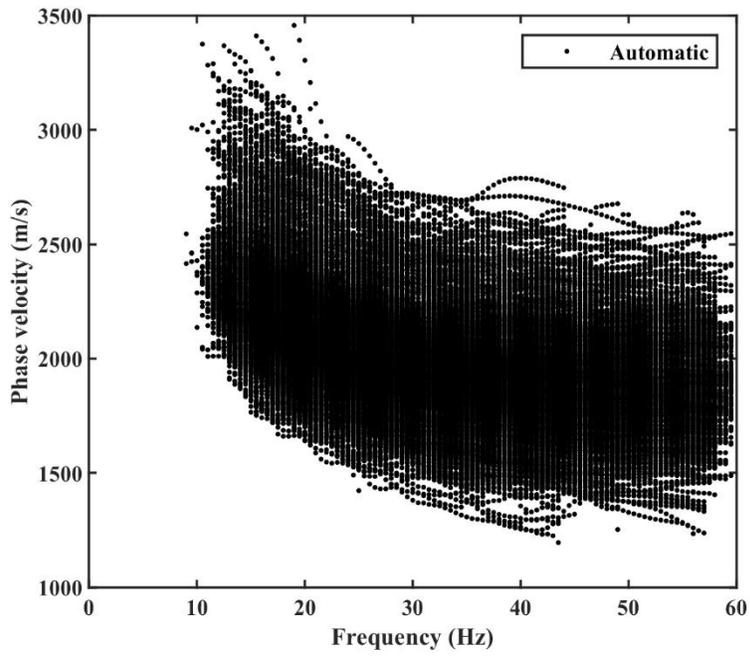
Figure 4. 11. Performance of the automatic DC picking on Line 1.
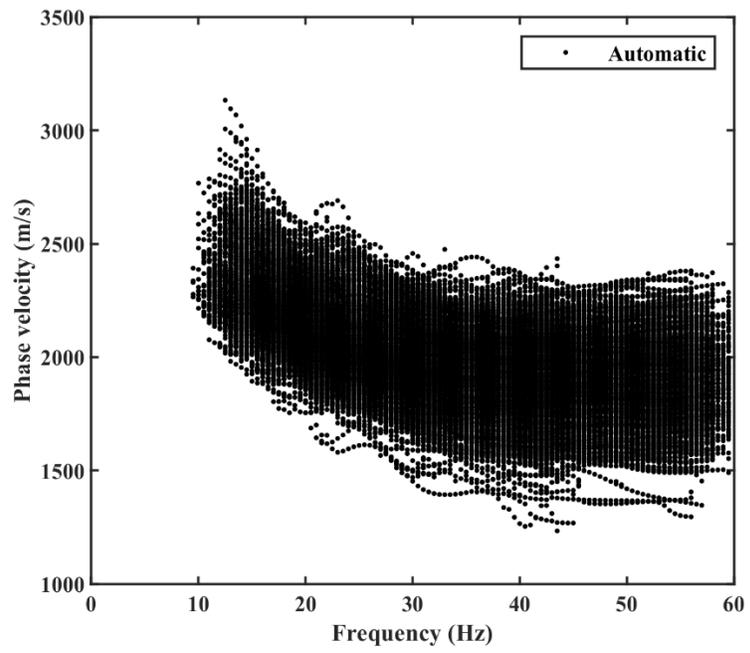
Figure 4. 12. DCs from Line 2

4.6 Inversion

The DCs from both of the lines were inverted with the SW tomography algorithm. The results of the inversion of Lines 1 and 2 are shown in Figure 4.16 a and b, respectively.

The inversion results for Line 1 (Figure 4.13.a) shows good agreement with the results achieved by Papadopoulou et al. 2019 in Figure 2.3 and confirms substantial lateral variations of the velocity along the line.
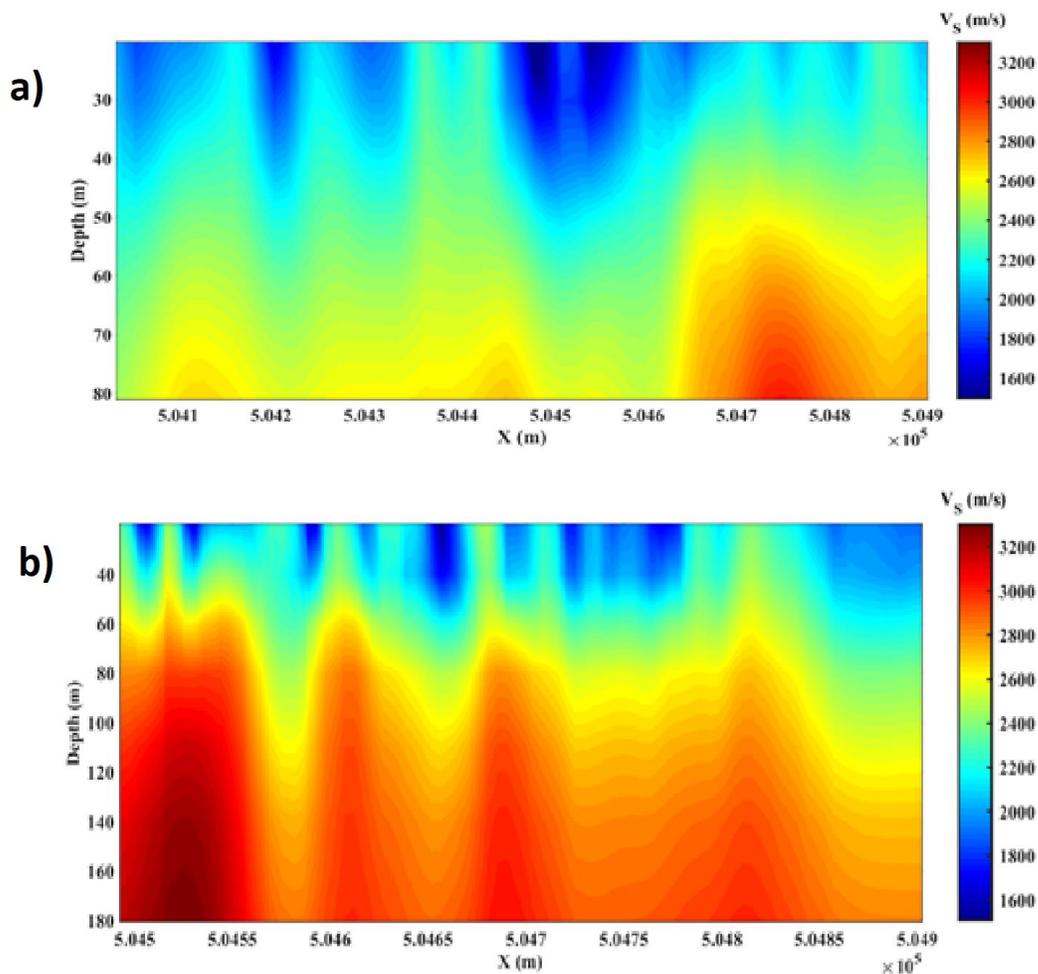
Figure 4. 13. Shear wave velocity profiles retrieved from the inversion of DCs of a) Line 1 and b) Line2 (Papadopoulou, personal communication).

Conclusion

It is a well-established fact that the SW analysis has essential value in the near-surface subsoil investigation. And in this thesis work as a part of the SW analysis, the workflow of the processing of 2D seismic tomography analysis has been analysed. The analyses aimed at performing QC on the automatisation of DC processing.

The 2D seismic data acquired from 2 seismic lines located in Ludvika mining area, Sweden, which is known for its high-quality iron-oxide mineral deposits were processed.

The test consisted in comparison of the automatic DC picking with the manual DC piking. The first step that has been undertaken towards the testing was manually picking of the DCs from a seismic data of a certain number of receiver pairs that have been randomly picked from the seismic lines. Later on, to the same set of seismic data of the same receiver pairs, there has been applied the automatic DC picking. Both picking methods have saved the equal amount of DCs from identical receiver pairs, 87 DCs for the Line 1 and 65 DCs for Line 2. This fact can be considered as a good starting point for a comparison test, since for the further steps to be able to compare the quality of the picked DCs they should be derived from the same receiver pairs.

The comparison was based on two types of uncertainties in the frequency bandwidth: ignored and extra picked frequencies by the autopicking. There have been implemented two parameters to identify the degree of influence of those uncertainties to the quality: 1) estimation of the effectiveness of the autopicked DC and 2) the similarity index. As a result of the application of those QCs, it is possible to conclude that the autopicking performs good quality in picking the DCs, since the calculated parameters demonstrated fairly high quality of the autopicking.

Our test also demonstrated the time-efficiency of the autopicking which consists in the speed of processing. In contrast, manual processing is a long-term procedure. Moreover, the autopicking requires only computational time and not operator time;

hence it is by far cheaper and potentially subject to optimisation than manual picking.

References

Malehmir, Alireza, Raymond Durrheim, Gilles Bellefleur, Milovan Urosevic, Christopher

Hayashi, K., and H. Suzuki, 2004, CMP cross-correlation analysis of multichannel surface-wave data: Exploration Geophysics, 35, 7–13.

Socco, L.V., D. Boiero, S. Foti, and R.Wisén, 2009, Laterally constrained inversion of ground roll from seismic reflection records: Geophysics, 74, no.6, G35–G45.

Grandjean, G., and A. Bitri, 2006, 2M-SASW: Multifold multichannel seismic inversion of local dispersion of Rayleigh waves in laterally heterogeneous subsurfaces: application to the Super-Sauze earthflow, France: Near Surface Geophysics, 4, 367–375.

Socco, Laura Valentina, Sebastiano Foti, and Daniele Boiero. 2010. "Surface-Wave Analysis for Building near-Surface Velocity Models - Established Approaches and New Perspectives." Geophysics 75 (5). https://doi.org/10.1190/1.3479491.

Mari, J. L., 1984. "Estimation of Static Corrections for Shear-Wave Profiling Us-Ing the Dispersion Properties OfLovewaves." Geophysics 49: 1169–79.

Roy, Soumya, Robert R Stewart, and K Al Dulaijan. 2010. "S-Wave Velocity and Statics from Ground-Roll Inversion." The Leading Edge 29 (10): 1250–57. https://doi.org/10.1190/1.3496915.

Boiero, D, P Marsden, V Esaulov, A Zarkhidze, and P Vermeer. 2013. "Building a Near-Surface Velocity Model in the South Ghadames Basin: Surface-Wave Inversion to Solve Complex Statics SW Inversion to Solve Complex Statics." SEG Houston 2013 Annual Meeting, 1811–15.

Douma, Huub, and Matthew Haney. 2011. "Surface-Wave Inversion for near-Surface Shear-Wave Velocity Estimation at Coronation Field." SEG San Antonio 2011 Annual Meeting, 1411–15.

Miao, X., D. Zheng, L. Zi, Zhou Z., and Meng G. 2016. "Robust Multimodal Surface-

Wave Inversion for Shallow Velocity and Shear Statics." SEG Technical Program, Dallas, USA, Expanded Abstracts, 4956–60.

Zielhuis, Alet, and Guust Nolet. 1994. "Deep Seismic Expression of an Ancient Plate Boundary in Europe." Journal of Chemical Information and Modeling 53 (9): 1689–99. https://doi.org/10.1017/CBO9781107415324.004.

Trampert, Jeannot, and John H. Woodhouse. 1995. "Global Phase Velocity Maps of Love and Rayleigh Waves between 40 and 150 Seconds." Geophysical Journal International 122 (2): 675–90. https://doi.org/10.1111/j.1365-246X.1995.tb07019.x.

Simons, Frederik J., Rob D. van der Hilst, Jean-Paul Montagner, and Alet Zielhuis. 2002. "Multimode Rayleigh Wave Inversion for Heterogeneity and Azimuthal Anisotropy of the Australian Upper Mantle." Geophysical Journal International 151 (3): 738–54. https://doi.org/10.1046/j.1365-246X.2002.01787.x.

Shapiro, N. M., and M. H. Ritzwoller. 2002. "Monte-Carlo Inversion for a Global Shear-Velocity Model of the Crust and Upper Mantle." Geophysical Journal International 151 (1): 88–105. https://doi.org/10.1046/j.1365-246X.2002.01742.x.

Ridder, S. A.L. De, and B. L. Biondi. 2015. "Near-Surface Scholte Wave Velocities at Ekofisk from Short Noise Recordings by Seismic Noise Gradiometry." Geophysical Research Letters 42 (17): 7031–38. https://doi.org/10.1002/2015GL065027.

Yang, Yingjie, Michael H. Ritzwoller, Anatoli L. Levshin, and Nikolai M. Shapiro. 2007. "Ambient Noise Rayleigh Wave Tomography across Europe." Geophysical Journal International 168 (1): 259–74. https://doi.org/10.1111/j.1365-246X.2006.03203.x.

Aki, K, and PG Richards. 1980. Quantitative Seismology. Vol. II.

Ben-Menahem, A., 1995. "A Concise History of Mainstream Seismology: Origins, Legacy, and Perspectives." Bulletin - Seismological Society of America 85 (4): 1202–25. https://doi.org/10.1016/0148-9062(96)81706-2.

Romanowicz, Barbara. 2002. "Inversion of Surface Waves: A Review."

International Geophysics 81 (PART A): 149–73. https://doi.org/10.1016/S0074-6142(02)80214-5.

Trampert, Jeannot, and John H. Woodhouse. 1995. "Global Phase Velocity Maps of Love and Rayleigh Waves between 40 and 150 Seconds." Geophysical Journal International 122 (2): 675–90. https://doi.org/10.1111/j.1365-246X.1995.tb07019.x.

Da Col Federico, Myrto Papadopoulou, Emilia Koivisto, Łukasz Sito, Mikko Savolainen, and Laura Valentina Socco. 2019. "Application of Surface-wave Tomography to Mineral Exploration: A Case Study from Siilinjärvi, Finland." Geophysical Prospecting 68 (1): 254–69. https://doi.org/10.1111/1365-2478.12903.

Socco, Laura Valentina, Daniele Boiero, Paolo Bergamo, and Flora Garofalo. 2014. "Surface Wave Tomography to Retrieve near Surface Velocity Models," 2013–18.

Swoboda, U, S Uchtmann, K Limbrock, R Elsen, D Orlowsky, K Telenga, and F Koschare, A and Wollnik. 2013. "Seismic and Geoelectric Investigation of Historical Mining Structures-A Case Study." Near Surface Geoscience 2013-19th EAGE European Meeting of Environmental and Engineering Geophysics.

Picozzi, M., S. Parolai, D. Bindi, and A. Strollo. 2009. "Characterization of Shallow Geology by High-Frequency Seismic Noise Tomography." Geophysical Journal International 176 (1): 164–74. https://doi.org/10.1111/j.1365-246X.2008.03966.x.

Bloch, S., and A.L Hales. 1968. "New Techniques for the Determination of Surface Wave Phase Velocities." Bulletin of the Seismological Society of America 58 (3): 1021–34.

Yao, Huajian, Guoming Xu, Liangbao Zhu, and Xiang Xiao. 2005. "Mantle Structure from Inter-Station Rayleigh Wave Dispersion and Its Tectonic Implication in Western China and Neighboring Regions" 148: 39–54. https://doi.org/10.1016/j.pepi.2004.08.006.

Knopoff, L., S. Mueller, and W.L Pilant. 1966. "Structure of the Crust and Upper

Mantle in the Alps from the Phase Velocity of Rayleigh Waves." Bulletin of the Seismological Society of America 56 (5): 1009–44.

Socco, L.V., and C. Strobbia. 2004. "Surface-Wave Method for near-Surface Characterization: A Tutorial." Near Surface Geophysics 2 (4): 165–85. https://doi.org/10.3997/1873-0604.2004015.

Richart, F. E, R. D Woods, and J. R Hall. 1970. Vibrations of Soils and Foundations. Vol. 414.

Steeples, Don W., 2005. "Shallow Seismic Methods." In Hydrogeophysics, 215–51. Dordrecht: Springer, Netherlands. https://doi.org/10.1007/1-4020-3102-5_8.

Trampert, J., and J. H. Woodhouse. 1996. "High-Resolution Global Phase Velocity Distributions." Geophysical Research Letters 23 (1): 21–24. https://doi.org/10.1029/95GL03391.

Yang, Yingjie, Michael H. Ritzwoller, Anatoli L. Levshin, and Nikolai M. Shapiro. 2007. "Ambient Noise Rayleigh Wave Tomography across Europe." Geophysical Journal International 168 (1): 259–74. https://doi.org/10.1111/j.1365-246X.2006.03203.x.

Bloch, S., and A.L Hales. 1968. "New Techniques for the Determination of Surface Wave Phase Velocities." Bulletin of the Seismological Society of America 58 (3): 1021–34.

Landesman, M., A. Dziewonski, and Y. Satô. 1969. "Recent Improvements in the Analysis of Surface Wave Observations." Geophysical Journal International 17 (4): 369–403. https://doi.org/10.1111/j.1365-246X.1969.tb00246.x.

Malehmir.2017. Developing cost-effective seismic mineral exploration methods using a land streamer and a drop hammer.

Geijer P. and Magnusson N.H. 1944. De Mellansvenska J¨arnmalmernas Geologi. Geological Survey of Sweden.

Stephens M.B., Ripa M., Lundstr¨om I., Persson L., Bergman T., Ahl M. et al. 2009. Synthesis of the Bedrock Geology in the Bergslagen Region,

Fennoscandian Shield, South-Central Sweden. Sveriges geologiska unders ¨ okning (SGU).

Allen R.L., Lundstrom I., Ripa M. and Christofferson H. 1996. Facies analysis of a 1.9 Ga, continental margin, back-arc, felsic caldera province with diverse Zn-Pb-Ag-(Cu-Au) sulfide and Fe oxide deposits, Bergslagen Region, Sweden. Economic Geology 91, 979– 1008.

Ripa M. and K¨ ubler L. 2003. Apatite-Bearing Iron Ores in the Bergslagen Region of South-Central Sweden. Sveriges Geologiska Unders ¨ okning.

Magdalena Markovic, Georgiana Maries, Alireza Malehmir, Julius von Ketelhodt, Emma Bom, Monica Schon and Paul Marsden. 2019. "Deep reflection seismic imaging of iron-oxide deposits in the Ludvika mining area of central Sweden."

Annex 1: Code for Comparison Test

```matlab
clear all
clc


files=dir('*.mat');

for i=1:length(files)
    cd 'C:\Users\Desktop\Folder_Name'
    fname=files(i).name;
    freqtot=[];
    freqtot1=[];
    freqtot2=[];
    stack_dc=[];
    stack_dc1=[];
    stack_dc2=[];

    load(fname)
    if isempty(freqtot1)
        freqtot1=NaN;
    end
     if isempty(freqtot2)
        freqtot2=NaN;
     end
     if isempty(stack_dc1)
        stack_dc1=NaN;
    end
    if isempty(stack_dc2)
        stack_dc2=NaN;
    end

    % Intersects freqtot1 & freqtot2,
    % and ia & ib are the indexes of intersected values in freqtot1 and freqtot2
    % correspondingly. If ia=1 or ib=1 It does not mean that they have
    % intersection, it has done because MatLab needs integers in further
    % calculations.

  [C,ia,ib] = intersect(freqtot1,freqtot2,'rows');

if isempty(ia)
    ia=1;
end
if isempty(ib)
    ib=1;
end

%gives from each freqtot the values that are not found the intersection
dif1=setdiff(freqtot1,C);
```

```matlab
dif2=setdiff(freqtot2,C);

%percentage calculation according to the freqtot-s
a_ignored=length(dif1)/length(freqtot1)*100;
a_ignored_all(:,i)=a_ignored;
a_extra=length(dif2)/length(freqtot2)*100;
a_extra_all(:,i)=a_extra;

%extracting velocity values corresponding to ia and ib indexing
stack_dc1_new=stack_dc1(ia,1);
stack_dc2_new=stack_dc2(ib,1);

%stack_dc_fraction shows the reliability of the autopicking in point
%of phase velocity

stack_dc_comparison=nonzeros(stack_dc1_new==stack_dc2_new);
stack_dc_jumping=nonzeros(stack_dc1_new~=stack_dc2_new);

if isempty(stack_dc_jumping)
    stack_dc_jumping=0;
end
stack_dc_jumping_all(:,i)=stack_dc_jumping;
stack_dc_fraction= length(stack_dc_comparison)/length(stack_dc2)*100;
stack_dc_fraction_all(:,i)=stack_dc_fraction;
similarity_index= length(stack_dc_comparison)/length(stack_dc1);
similarity_index_all(:,i)=similarity_index;
dc_number(i)=i;

fprintf('file name: %s\nignored_freq: %.2f%%\nextra_freq:
%.2f%%\nefficience_of_autopicking: %.2f%%\nsimilarity_index:%.2f\n\n',...
    fname, a_ignored, a_extra,stack_dc_fraction,similarity_index)
save(fname,'stack_dc1','stack_dc2','dif1','dif2','freqtot1','freqtot2','stackdisp','ia','ib','a_ignored','a_extra',...

'stack_dc1_new','stack_dc2_new','stack_dc_comparison','stack_dc_fraction','similarity_index')

end

average_similarity=sum(similarity_index_all)/length(similarity_index_all)

figure
for i=1:length(files)
    fname=files(i).name;
    load(fname)
    axis xy
    plot(freqtot1,stack_dc1,'*r')
    plot(freqtot2,stack_dc2,'ok')
    legend('automatic','manual')
    xlabel('Frequency(Hz)')
    ylabel('Phase velocity (m/s)')
    hold on
end
```

```matlab
figure
plot(a_ignored_all,'r.')
hold on
plot(a_extra_all,'b.')
legend('ignored','extra')
xlabel('DCs')
ylabel('Frequency Uncertainties, %')

figure
histogram(stack_dc_fraction_all)
xlabel('Efficiency of the  DCs')
ylabel('Number of DCs')

figure
histogram(similarity_index_all)
xlabel('Similarity Index')
ylabel('Number of DCs')
```