



UNIVERSITÀ DEGLI STUDI DI TRIESTE
DIPARTIMENTO DI MATEMATICA E GEOSCIENZE

XXIX CICLO DEL DOTTORATO DI RICERCA IN
SCIENZE DELLA TERRA E MECCANICA DEI FLUIDI
AUTOMATED REFLECTION PICKING AND INVERSION
APPLIED TO GLACIOLOGICAL GPR SURVEYS

SETTORE SCIENTIFICO-DISCIPLINARE: GEO/11

PH.D. STUDENT
Matteo Dossi

PH.D. PROGRAM COORDINATOR
Prof. Pierpaolo Omari

THESIS SUPERVISOR
Prof. Michele Pipan

THESIS CO-SUPERVISOR
Ass. Prof. Emanuele Forte



ACADEMIC YEAR 2015-2016

Contents

1	Introduction	4
2	Automated picking	12
2.1	State of the art	12
2.1.1	Reflection picking	12
2.1.2	Attribute analysis	13
2.2	Picking algorithm	18
2.2.1	Horizon picking	20
2.2.2	Horizon joining	21
2.3	Automated grouping	23
2.3.1	Phase analysis	25
2.3.2	Energy analysis	31
3	Amplitude inversion	35
3.1	State of the art	35
3.1.1	Glacier monitoring	35
3.1.2	Velocity analysis	40
3.2	Inversion algorithm	43
3.2.1	Amplitude recovery	46
3.2.2	Radiation Pattern	47
3.3	Synthetic data analysis	51
3.4	Field data analysis	53
4	Amplitude sampling	58
4.1	State of the art	59
4.2	Synthetic data analysis	60
4.3	Field data analysis	70
5	Field application	73
5.1	Prevala glacieret	73
6	Conclusion	83

Abstract

The main part of the Ph.D. research consisted in the development and application of an automated picking algorithm, which is designed to quickly and objectively identify the main reflections within a recorded data set, and to extract information such as arrival times, peak amplitudes, and polarities, to be used for quantitative analysis. The algorithm uses attribute analysis to track and mark as a horizon any recorded event characterized by lateral phase continuity, to group together horizons marking different parts of the same reflection, and to select specific reflection phases for further analysis and interpretation. The algorithm is mostly independent from the interpreter, except for a few required input parameters and thresholds, and it differs from other picking methods since it does not rely on any manually picked control point. Moreover, the procedure does not depend on the signal amplitudes, therefore it is able to track any event regardless of its strength or lateral amplitude variations. Similarly to other auto-picking procedures, the algorithm can be sensitive to low signal-to-noise ratio as well as signal interference, therefore appropriate processing may be required before application. Nevertheless, given its independence from the signal amplitudes, the procedure can be applied without any gain function, which may subjectively alter the amplitudes picked for quantitative analysis.

The Ph.D. research also focused on improving an amplitude inversion algorithm designed to recover the subsurface stratigraphy and EM velocity distribution from Ground Penetrating Radar (GPR) data sets. The algorithm iteratively recovers the thickness and EM velocity in each interpreted layer by reconstructing the travel path of each reflected wavelet, using as input the picked amplitudes and travel times of reflections in a GPR profile, among other quantities. In glaciological surveys, the procedure can also be used to recover the internal density and water content of glaciers and snowfields, which are linked to the EM velocity through well-known empirical relations. The improvements to the inversion algorithm included rewriting part of the code in order to greatly reduce the overall computing time, introducing a feature that applies amplitude recovery directly during the inversion itself, and taking into account the changes in radiation pattern for signals propagating in different media.

The effects that the sampling rate selected during data acquisition has on the picking and inversion results were also studied, in order to reduce the peak amplitude error, which affects any further qualitative and quantitative analysis of the recorded data set. Based on the Nyquist-Shannon theorem, the sampling rate is generally set equal to 6 (for GPR) or 8 (for seismics) times the signal central frequency, as a rule of thumb. While the theorem can be used to accurately recover the frequency spectrum of a signal, it is shown that it does not take into account possible signal distortions, which could prevent the accurate sampling of the peak amplitude within a signal phase. Moreover, the peak amplitude error is mostly random in nature, depending on the particular time-position and shape of the recorded signal, besides the selected sampling rate, which prevents any possible prediction or estimation. The research suggested that the sampling rate has to be equal to at least 12 times the central signal frequency, as a rule of thumb, in order to limit the peak amplitude error within 5%. Besides its important theoretical value, such conclusion is useful for many practical problems especially, but not exclusively, for those related to geophysical applications.

In this text the main features and formulas of the developed algorithms are presented, and they are applied to a 3-D glaciological GPR data set, in order to study the internal stratigraphy, density distribution, and water content of an alpine glacieret, while also highlighting both the advantages and limitations of the procedure. The theory, mathematical background, and performance of the algorithms described in this thesis are discussed in greater details in various publications showing several of the results obtained during the Ph.D. research, with applications to both synthetic and field data sets, and to both seismic and GPR surveys in different environments.

Chapter 1

Introduction

The purpose of this thesis is to present an automated picking and inversion procedure, which is designed to accurately and objectively identify the main reflections within Ground Penetrating Radar (GPR) data sets; to characterize them in terms of their arrival times, peak amplitudes, and polarities; and to recover from these and other quantities the internal stratigraphy and EM properties of the subsurface. Also known as georadar, GPR is a non-invasive near-surface geophysical technique based on the propagation of EM signals, which are transmitted into the ground by antennas using frequencies typically in the 10 MHz - 2 GHz range (Jol, 2009). The transmitted signals propagate in the subsurface and are subjected to different processes, such as reflections, refractions, and diffractions, which are caused by discontinuities in the EM properties of the subsurface materials. Some of these signals are then recorded by receiver antennas to form data sets which are used to extract both qualitative and quantitative information regarding the materials and structures through which the recorded signals have propagated. A GPR system can use several transmitting and receiving antennas arranged in different geometries during data acquisition (Fig. 1.1), depending on the particular objectives of the survey, which result in different EM responses from the subsurface and thus require different techniques for signal processing, data imaging, and subsurface characterization (Jol, 2009).

In Common Offset (CO) surveys the GPR system consists of a single transmitter and a single receiver, which are kept at a constant distance (i.e. offset) from each other and are rigidly moved along the survey direction, with measurements usually performed at constant space intervals (Fig. 1.1A). In ground-coupled GPR surveys both the transmitter and the receiver are in contact with, or just above, the ground surface and are moved across the survey area either manually or attached to vehicles. In air-coupled GPR surveys the antennas are instead located at a certain height above the ground and are usually installed on vehicles or aircrafts, thus allowing for high-speed surveys covering large areas at fast acquisition rates. The results are quasi-continuous images of the subsurface, or rather of its response to EM signals, with varying degrees of resolution and penetration, which depend on factors like the signal frequencies, the acquisition parameters, the internal stratigraphy and EM properties of the subsurface, as well as both systematic and external noise sources at the particular time and location of the survey. Common offset (or single-fold) GPR surveys are the most common in terms of acquisition geometry, and they can be used for both qualitative and quantitative analysis. For example, the EM velocity distribution within the profile is usually estimated by studying diffractions through either hyperbolic fitting or migration velocity analysis (Jol, 2009). In this thesis, glaciological CO GPR profiles are used as input of the presented auto-picking and inversion algorithm in order to

study the internal stratigraphy, EM velocity distribution, density distribution, and water content of glaciers from the reflection amplitudes and travel times, among other quantities.

In Common Mid-Point (CMP) surveys the GPR system consists of a transmitter and a receiver which are separated by a varying offset and moved in opposite directions with respect to a constant mid-point (Fig. 1.1B). Similar multi-offset (or multi-fold) geometries are used in Wide-Angle Reflection surveys, also known as Common Shot (Fig. 1.1C) and Common Receiver (Fig. 1.1D) surveys, in which one of the two antennas remains stationary while the other is moved along the survey direction. In order to reduce the acquisition time, it is also possible to use an array of transmitting and receiving antennas, which are geometrically fixed and moved along the survey direction. Another alternative is to acquire several CO profiles along the same path, each with a different offset, and combine them in order to reconstruct CMP surveys at various positions along the path (Forte and Pipan, 2017). Common mid-point GPR surveys are primarily used to recover the vertical EM velocity distribution at specific locations by studying the changes with offset in the amplitude and travel time along the hyperbolic reflections recorded at different depths. Moreover, stacking algorithms can be used to improve the signal-to-noise ratio in CMP data sets, since the signals recorded at different offsets are theoretically reflected by the same points within the subsurface, assuming sub-horizontal reflectors (Fig. 1.1B). Nevertheless, multifold GPR surveys are seldom performed since they require more complex data analysis, due to possible lateral changes in the stratigraphy and topography of the survey areas, and also because data acquisition procedures are usually time consuming and more difficult in case of rough terrains, especially when there are no large movable arrays of fixed antennas and the few available antennas must be rearranged manually (Jol, 2009; Forte and Pipan, 2017).

Transillumination surveys are less common than CO and CMP surveys and they are used to analyze EM signals traveling through the materials between the transmitter and the receiver (i.e. one-way travel time surveys), rather than signals reflected by subsurface discontinuities (i.e. two-way travel time surveys). For example, the transmitter and receiver can be located within two different boreholes in order to study the subsurface materials between them, or they can be placed at the opposite sides of a column or a wall in order to study their structural integrity. The antennas can be arranged in different geometries depending on the particular objectives of the survey. In Zero-Offset Profiling surveys (Fig. 1.1E) the antennas are geometrically fixed, usually placed at the same height, and are rigidly moved along the analyzed boreholes or walls, while in Multi-Offset Gather surveys (Fig. 1.1F) one of the two antennas remains stationary while the other is moved in different positions, thus recording signals from different angles. Changes in the amplitudes and travel times recorded at different positions can be used to highlight features like EM velocity anomalies, internal discontinuities, and attenuation zones, while also taking into account the acquisition geometry, which must be accurately known (Jol, 2009).

In all the different survey geometries, a clear understanding of the signal propagation within the subsurface is of paramount importance for its accurate characterization. Consider a single-frequency, linearly polarized, EM plane wave propagating in the direction given by a vector $\mathbf{r}$, the magnitudes of the electric $E(\mathbf{r}, t)$ and magnetic $B(\mathbf{r}, t)$ fields are derived from Maxwell's equations and depend on the EM properties of the material through which the signal is traveling (Jol, 2009):

$$E(\mathbf{r}, t) = E_0 e^{-\alpha r} e^{i(\beta r - \omega t)} \quad (1.1)$$

$$B(\mathbf{r}, t) = B_0 e^{-\alpha r} e^{i(\beta r - \omega t)} \quad (1.2)$$

with

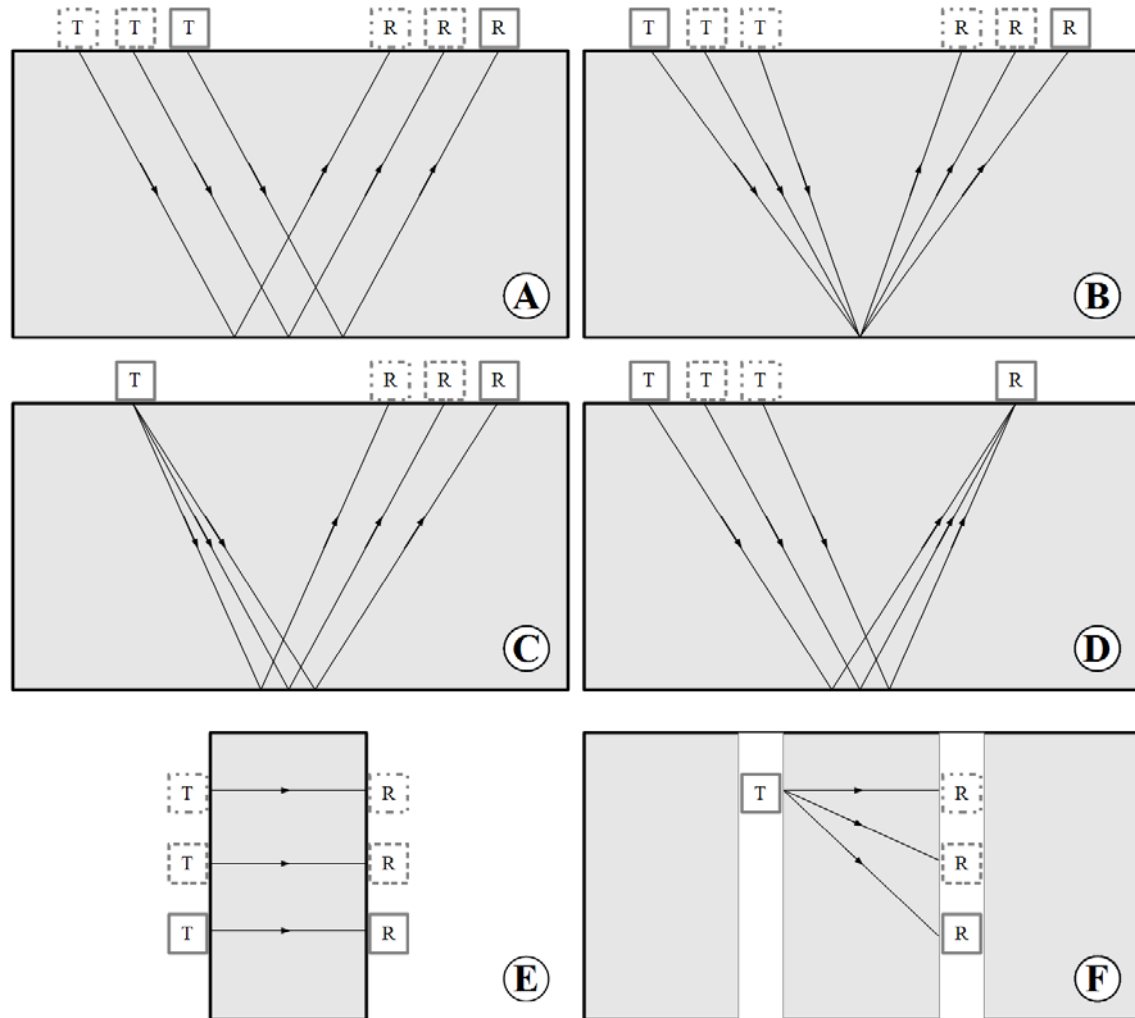


Figure 1.1: Example of possible data acquisition geometries for a GPR system. The figure shows transmitters (T) and receivers (R) arranged in common offset (A), common mid-point (B), common shot (C), common receiver (D), zero-offset profiling (E), and multi-offset gather (F) geometries. The recorded signals are displayed for the different geometries in two-way travel time (A-D) and one-way travel time (E, F) data sets.

$$\alpha = \omega \sqrt{\frac{\mu\varepsilon}{2} \left[\sqrt{1 + P^2} - 1 \right]} \quad (1.3)$$

$$\beta = \omega \sqrt{\frac{\mu\varepsilon}{2} \left[\sqrt{1 + P^2} + 1 \right]} \quad (1.4)$$

$$P = \frac{\sigma}{\omega\varepsilon} \quad (1.5)$$

where ω is the angular frequency of the signal, and σ , μ , and ε are respectively the electric conductivity, the magnetic permeability, and the electric permittivity of the material through which the GPR signal is traveling. The parameters α , β , and P are referred to as the attenuation constant, the phase parameter, and the loss factor, respectively, and they depend on the EM properties of the analyzed materials, which therefore influence the signal propagation, as well as the quality of the recorded GPR data set.

An important parameter is the propagation velocity of EM signals, whose accurate estimation has several applications in signal processing, data imaging, and for both qualitative and quantitative analyses of GPR data sets. The EM velocity v of the GPR signal is linked to the EM properties of the subsurface through the following relation (Jol, 2009)

$$v = \frac{c}{\sqrt{\frac{\mu_r \varepsilon_r}{2} \left[\sqrt{1 + P^2} + 1 \right]}} \quad (1.6)$$

with

$$v = \frac{\omega}{\beta}$$

where c is the EM velocity in vacuum (i.e. about 30 cm/ns), and μ_r and ε_r are respectively the relative magnetic permeability and the relative electric permittivity (a.k.a. the dielectric constant) of the material through which the GPR signal is traveling.

A few examples of the EM responses of different materials are presented in Table 1.1. The table shows the values of the relative permittivity ε_r , the electric conductivity σ , the EM velocity v , and the attenuation factor α , that can be expected from various earth materials when using signal frequencies around 100 MHz. In general, the magnetic response of most geologic materials does not significantly influence the propagation of GPR signals, except in the case of ferromagnetic minerals (Jol, 2009). Therefore, the relative magnetic permeability μ_r is often approximated equal to 1, and thus the quantity is not presented in Table 1.1. Nevertheless, the listed materials cover a large range of values with regards to their relative permittivity and electric conductivity, which results in significant variations in the EM velocity and the signal attenuation. Moreover, the EM response can also greatly vary in several materials for the different frequency components of the propagating GPR signals, with the higher frequencies being more prone to signal attenuation.

A material can be considered low-loss with respect to a propagating GPR signal when the loss factor P (Eq. 1.5) has a value lower than 0.1, while it can be considered high-loss when the loss factor is greater than 10 (Seybold, 2005). Among the materials listed in Table 1.1, the ones showing the worst response for GPR surveys are sea water and clays, which present the highest values for signal attenuation and would result in penetration depths of a few meters, if not tens of centimeters in the worst case. Conversely, beside air and distilled water, the materials most favorable toward GPR surveys are dry sand and ice, in which the signal can propagate efficiently with the lowest attenuation, reaching penetration depths in some cases of several hundreds of meters.

EARTH MATERIALS	EM PROPERTIES			
	ε_r	σ [$\frac{\text{mS}}{\text{m}}$]	v [$\frac{\text{cm}}{\text{ns}}$]	α [$\frac{\text{dB}}{\text{m}}$]
Air	1	0	30	0
Distilled Water	80	0.01	3.3	0.002
Fresh Water	80	0.5	3.3	0.1
Sea Water	80	3000	1	1000
Dry Sand	3-5	0.01	15	0.01
Saturated Sand	20-30	0.1-1	6	0.03-0.3
Limestone	4-8	0.5-2	12	0.4-1
Shales	5-15	1-100	9	1-100
Silts	5-30	1-100	7	1-100
Clays	5-40	2-1000	6	1-300
Granite	4-6	0.01-1	13	0.01-1
Dry Salt	5-6	0.01-1	13	0.01-1
Ice	3-4	0.01	16	0.01

Table 1.1: Example of the EM responses expected from different materials for signal frequencies around 100 MHz. The table shows possible values of the relative permittivity ε_r , the electric conductivity σ , the EM velocity v , and the attenuation factor α in different environments. The magnetic properties of most earth materials do not usually influence the propagation of GPR signals in a significant way, except in the case of ferromagnetic minerals, therefore the relative magnetic permeability μ_r is not shown, as it is often approximated equal to 1. Table from Davis and Annan (1989).

A major factor affecting the GPR signal propagation is the presence or absence of either free or bound water within the subsurface, which dominates the EM response in the 10 MHz - 1 GHz frequency range (Jol, 2009). In general, the subsurface is a system characterized by three main phases, namely solid, gas, and liquid, each making up a certain percentage of the total volume and contributing to its average EM properties. Considering the high value of the relative permittivity of fresh water (i.e. 80) with respect to air (i.e. 1) and generic dry geological materials (i.e. 3-8), as shown in Table 1.1, small variations in the pore water content can significantly change the average EM response of the soil. Besides variations in the EM velocity due to the large changes in the average permittivity, the presence of water can significantly increase the signal attenuation within the soil, either through dielectric relaxation, given the high polarizability of water molecules, or through electric conductivity, mostly due to the presence of ions within the water. A clear example is given by comparing the EM responses of dry and saturated sands in Table 1.1, in which the EM velocity decreases from 15 to 6 cm/ns, while the attenuation factor increases from 0.01 to 0.03-0.3 dB/m. Another example is given by clays, which are made up of particular minerals containing variable amounts of water trapped within their structure, and show considerably lower EM velocity and higher attenuation with respect to generic dry materials such as granite and limestone (Table 1.1).

Among the geologic materials, dry snow and ice provide the most favorable conditions for GPR surveys with central frequencies above 1 MHz, due to the low conductivity and the absence of any dielectric and magnetic relaxation processes above this frequency (Jol, 2009), resulting in low attenuation which allows the GPR signal to propagate efficiently through the analyzed air-ice mixture. In fact, ice is a low-loss material with a relaxation frequency in the kHz range and in dry snow the intrinsic attenuation of the signal is negligible, while the dominant parameter determining the dielectric constant is the density (Stiles and Ulaby, 1981). Nevertheless, the presence of liquid water

within glaciers and snowfields can considerably affect their internal EM properties and influence the quantitative analysis of a glaciological GPR data set. In the 10 MHz - 1 GHz frequency range the electric permittivity of air-ice mixtures is approximately independent from the frequency, but it increases with the liquid water content, given the large difference in permittivity between water and ice (Table 1.1), while the signal attenuation tends to increase with both the frequency and the liquid water content (Bradford et al., 2009).

At temperatures around 0°C, depending on the pressure, liquid water can exist within glaciers either along grain boundaries, or in pockets and conduits (Jol, 2009). In the first case, the water content is equal to a few percent in terms of volume and the isolated pockets result in a low conductivity (i.e. below 0.01 S/m), while the dielectric relaxation process of water is centered at a frequency too high (i.e. 9 GHz at 0°C) to significantly affect the GPR signal attenuation. In case pockets or conduits are formed due to a higher liquid water content, the signal penetration can be limited by scattering losses, besides a possible increase in the electric conductivity, which should be taken into account when applying amplitude recovery to the data set. Moreover, the empirical formulas used for the quantitative analysis of the GPR data set, linking the EM properties of the air-ice mixture to its density, should be modified if the presence of liquid water cannot be disregarded. In this thesis, the presence of liquid water is considered negligible within the analyzed alpine glacier, since the GPR survey was performed in the month of May, after the winter accumulation period and before the summer melting period.

The thesis is divided into the following chapters:

1. In chapter 2 the automated picking algorithm is presented. The algorithm uses attribute analysis to track and mark as a horizon any recorded event characterized by lateral phase continuity, to group together horizons marking different parts of the same reflection, and to select specific reflection phases to be used for further analysis and interpretation. In the chapter, the algorithm is applied to a glaciological GPR profile, which is used as an example to highlight the main stages of the procedure. The performance, advantages, and limitations of the picking algorithm are discussed and critically compared with other commonly used methods, which include manual picking, interpolation, auto-picking, surface slicing, and voxel tracking.
2. In chapter 3 the amplitude inversion algorithm is described. The algorithm iteratively recovers the thickness and EM velocity in each subsurface layer by reconstructing the travel path of each reflected wavelet, using as input the reflection amplitudes and travel times picked with the aforementioned auto-picking procedure. In glaciological surveys, the algorithm can also recover the density of air-ice mixtures, which is linked to the EM velocity through well-known empirical relations. The procedure is applied to a synthetic data set in order to assess its performance by comparing the inversion results with the initial model. The inversion algorithm is also applied to the aforementioned glaciological GPR profile in order to recover its stratigraphy, density distribution, and water content. The performance, advantages, and limitations of the inversion algorithm are discussed and critically compared with other commonly used velocity analysis methods, such as hyperbolic diffraction analysis and migration velocity analysis for CO surveys, and velocity-vs-depth analysis for CMP surveys.
3. In chapter 4 the effects that sampling has on the auto-picking and inversion results are analyzed, focusing on sampling-related signal distortions, which affect any further quantitative analysis of the recorded data set. In particular, the dependence of the maximum peak amplitude error on the sampling rate is analyzed for different signals and data sets, resulting in a recommended value of the sampling rate equal to at least 12 times the signal central frequency, which is higher than the commonly adopted standards, in order to limit the peak amplitude error within 5%.

4. In chapter 5 the automated picking and inversion procedure is applied to a 3-D glaciological GPR survey acquired on an alpine glacier, in order to recover its internal stratigraphy, density distribution, total volume, and water content, while also critically analyzing the inversion results from both a geophysical and a glaciological point of view.

In this text the main features and formulas of the developed algorithms are presented, while also highlighting both the advantages and limitations of the proposed auto-picking and inversion procedure. The theory, mathematical background, and performance of the algorithms described in this thesis are discussed in details in the publications listed below, with applications to both synthetic and field data sets, and to both seismic and GPR surveys in different environments. These publications show several of the results obtained during the Ph.D. research, some of which are related to topics not directly analyzed in this thesis.

Publications (2014-2017)

1. Dossi M., Forte E., and Pipan M., “Amplitude variations for propagating GPR signals: analysis of the transmitted wavelet, amplitude decay, and sampling-related distortions”, *Geophysical Journal International*, 2017, submitted
2. Forte E., Dossi M., Pipan M., and Del Ben A., “Automated phase attribute-based picking applied to reflection seismics”, *Geophysics*, vol. 81, no. 2, 2016, pp. V141-V150
3. Dossi M., Forte E., and Pipan M., “Automated reflection picking and polarity assessment through attribute analysis: Theory and application to synthetic and real ground-penetrating radar data”, *Geophysics*, vol. 80, no. 5, 2015, pp. H23-H35
4. Colucci R.R., Forte E., Boccali C., Dossi M., Lanza L., Pipan M., and Guglielmin M., “Evaluation of internal structure, volume and mass of glacial bodies by integrated LiDAR and Ground Penetrating Radar surveys: the case study of Canin Eastern Glacieret (Julian Alps, Italy)”, *Surveys in Geophysics*, vol. 36, no. 2, 2015, pp. 231-252
5. Forte E., Dossi M., Pipan M., and Colucci R.R., “Velocity analysis from common offset GPR data inversion: theory and application to synthetic and real data”, *Geophysical Journal International*, vol. 197, no. 3, 2014, pp. 1471-1483

Conference proceedings (2014-2016)

International conferences - Peer reviewed

1. Dossi M., Forte E., Pipan M., Colucci R.R., and Bortoletto A., “Automated reflection picking and inversion: application to ground and airborne GPR surveys”, in *Proc. IEEE 16th International Conference on Ground Penetrating Radar*, Hong Kong, June 13-16, 2016, pp. 1-6, ISBN 978-1-5090-5181-6
2. Zhao W., Forte E., Dossi M., and Pipan M., “Integrated attribute analysis for improved GPR data interpretation”, in *Proc. IEEE 16th International Conference on Ground Penetrating Radar*, Hong Kong, June 13-16, 2016, pp. 1-5, ISBN 978-1-5090-5181-6

3. Dossi M., Forte E., and Pipan M., “Auto-picking and phase assessment by means of attribute analysis applied to GPR pavement inspection”, in Proc. IEEE 8th International Workshop on Advanced Ground Penetrating Radar, Florence, Italy, July 7-10, 2015, pp. 1-4, ISBN 978-1-4799-6495-6
4. Forte E., Dossi M., Colle Fontana M., and Colucci R.R., “4-D quantitative GPR analyses to study the summer mass balance of a glacier: a case history”, in Proc. IEEE 15th International Conference on Ground Penetrating Radar, Brussels, Belgium, June 30-July 4, 2014, pp. 352-356, ISBN 978-1-4799-6789-6

National conferences - Reviewed by scientific committee

1. Dossi M., Forte E., and Pipan M., “Application of attribute-based automated picking to GPR and seismic surveys”, in Proc. 34th GNGTS Conference, vol. 3, Trieste, Italy, November 17-19, 2015, pp. 141-147, ISBN 978-88-940442-4-9
2. Mocnik A., Dossi M., Forte E., Zambrini R., Zamariolo A., and Pipan M., “Ground Penetrating Radar applications for roads and airport pavements investigations”, in Proc. 34th GNGTS Conference, vol. 3, Trieste, Italy, November 17-19, 2015, pp. 106-112, ISBN 978-88-940442-4-9

Chapter 2

Automated picking

2.1 State of the art

2.1.1 Reflection picking

Common offset GPR surveys produce high-resolution images showing the EM response of the subsurface to the transmitted radar signal. An accurate picking of the reflection amplitudes and travel times is of paramount importance in order to obtain quantitative information from GPR profiles, such as the subsurface stratigraphy and EM impedance contrasts, while the high data density of CO surveys makes the analysis statistically sound. Several different horizon picking methods have been developed in order to improve the interpretation of the recorded data sets, and they include, from least efficient to most efficient (Dorn, 1998), manual picking, interpolation, automated picking, surface slicing, and voxel tracking. These techniques use different assumptions and picking criteria, and the results show different levels of accuracy and dependence on the subjectivity of the interpreter (Dorn, 1998):

1. Manual picking relies entirely on the interpreter’s experience in selecting and manually tracking each reflection while looking for some degree of lateral continuity and similarity within the data set. It is the least effective picking method, both in terms of the objectivity of the results and the time and effort required for the analysis, especially when dealing with large amounts of data.
2. Interpolation is more efficient than manual picking, but it still relies on the initial manual selection of control points (a.k.a. seeds) which are then connected by the algorithm while tracking reflections. Moreover, the procedure assumes that the tracked events are locally smooth, or even linear, and therefore it may result in poor performance in the presence of discontinuities between control points, like for examples faults or interfering events.
3. Automated picking (a.k.a. auto-picking or auto-tracking) tracks recorded events along the analyzed profile, starting from initial control points and then searching for specific features in neighboring traces, which can be considered part of the same event. The picking algorithm described in this thesis belongs to this category, which is later discussed in greater details.
4. Surface slicing involves the visualization of areally limited portions of time slice slabs within 3-D data sets, as well as the interpretation of horizons within these slabs by isolating specific signal attributes (e.g. signal peaks) and selecting elements that fit together. This technique is not fully

automatic and it assumes local continuity, connectivity, and consistency of the signal phase being interpreted. However, it is less sensitive to discontinuities, signal-to-noise ratio, as well as the number of recorded traces in the analyzed 3-D survey.

5. Voxel tracking algorithms follow a particular feature or signal attribute within a 3-D data set, starting from an initial seed volume element (i.e. a voxel), which is manually selected by the interpreter, and then connecting neighboring voxels which satisfy the specified tracking criteria. The procedure can be faster than auto-picking, but it is more sensitive to the signal-to-noise ratio. Similarly to auto-picking, voxel tracking assumes that the analyzed data are locally continuous, consistent, and connected or smooth, while also assuming that a consistent signal phase is being interpreted.

Auto-picking algorithms track events within a recorded profile starting from an initial seed selected by the interpreter, and then looking for specific features in neighboring traces, such as a specific phase, a peak amplitude within a set range, or an arrival time within a certain time interval. If the algorithm finds a feature in a neighboring trace which satisfies the picking criteria, such feature is automatically marked and the process is reiterated in this new trace, otherwise the picking process is interrupted. Dorn (1998) identifies two main classes of auto-picking algorithms:

1. Feature trackers, which search for a specific configuration of samples, defining for example a peak, a trough, or a zero-crossing, within a specified time-window in the neighboring trace. The procedure described in this thesis belongs to this class of auto-picking algorithms.
2. Correlation-based auto-trackers, which use correlations between specific portions of signal from neighboring traces to search for possible matches within a specified time-window. In terms of performance, correlation-based auto-trackers are generally more computationally intensive and robust than feature trackers.

The main purpose of auto-picking methods is to quickly and accurately perform straightforward event correlations along the recorded profiles (Herron, 2000), while allowing the interpreter to manually intervene in more complex areas where the picking procedure could fail. In general, auto-picking algorithms assume that the data are locally continuous, smooth, and consistent, and therefore they can be sensitive to noise and interference. In fact, manual picking and interpolation between control points could be preferable in noisy data sets with respect to fully automated procedures (Herron, 2000).

2.1.2 Attribute analysis

An attribute is a quantitative measure of a signal characteristic of interest, which can be used for both quantitative and qualitative interpretation (Chopra and Marfurt, 2005; 2008). Consider a recorded digital trace made of N amplitude samples A_n defined within the time interval $[0, (N - 1) \Delta t]$, where Δt is the constant sampling interval. In complex trace analysis (Fig. 2.1), the recorded signal is identified as the real part of a complex trace a_n , with the n -th sample given by the following equation

$$a_n = A_n + iA'_n \quad (2.1)$$

with

$$i = \sqrt{-1}$$

$$n = 0, 1, \dots, N - 1$$

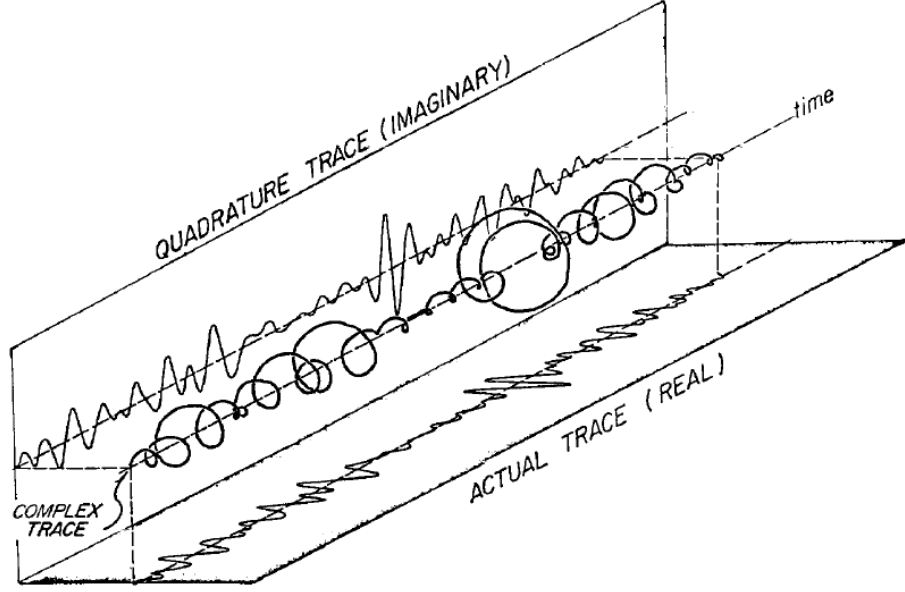


Figure 2.1: Example of a complex trace, defined by a vector with a time-varying modulus and phase. The complex trace can be divided into the actual recorded trace and the imaginary trace, the latter of which is reconstructed by applying a DHT operator to the former, which corresponds to a -90° phase shift. Modified image from Taner et al. (1979).

where A'_n is the imaginary part of the complex trace, also referred to as the quadrature trace, and the subscript n defines a sample recorded at time instant $n\Delta t$.

The complex trace is defined by a vector (Eq. 2.1) characterized by a time-varying modulus S_n and phase ϕ_n , respectively called the reflection strength (a.k.a. the trace envelope or the instantaneous amplitude) and the instantaneous phase. These two attributes are connected to the recorded and quadrature traces by the following relations

$$A_n = S_n \cos(\phi_n) \quad (2.2)$$

$$A'_n = S_n \sin(\phi_n) \quad (2.3)$$

In the following sections, the main signal attributes used in this text are discussed.

Quadrature trace

The imaginary part A'_n of the complex trace a_n is a convenient mathematical tool, however it does not have any general physical meaning. As it can be noticed from Eqs. 2.2 and 2.3, the quadrature trace is equal to the recorded trace after applying a -90° phase shift to the latter. Assuming that the signal is identically null outside the recorded time interval $[0, (N-1)\Delta t]$, the samples of the quadrature trace are uniquely determined by applying a Discrete Hilbert Transform (DHT) operator to the recorded trace, given by (Taner et al., 1979; Barnes, 1996; 2007):

$$A'_n = \frac{2}{\pi} \sum_{k=-M}^M \frac{\sin^2\left(\frac{\pi k}{2}\right)}{k} A_{n-k} \quad (2.4)$$

with

$$k \neq 0$$

As it can be noticed in Eq. 2.4, the amplitude of the quadrature trace at a certain time instant $n\Delta t$ depends on the amplitudes of the recorded trace defined over the entire $2M\Delta t$ wide time-window centered at time instant $n\Delta t$. Therefore, the quadrature trace depends on the time length of the DHT operator, as well as the recorded trace itself and, consequently, the applied signal processing. In fact, the DHT operator in Eq. 2.4 is generally truncated to a set time-window instead of covering the entire trace (Taner et al., 1979), in order to avoid possible significant distortions in the imaginary trace caused by strong signals A_{n-k} , which can counteract the $1/k$ decay of the DHT operator and become dominant in the calculation of the imaginary trace A'_n for distant unrelated weaker signals.

A clear example is given by the airwave-groundwave interference, which is generally several times stronger than most recorded events in a CO GPR trace, especially if no amplitude recovery as been applied before constructing the imaginary trace, as it is the case for the normalized GPR trace shown in Fig. 2.2A. When calculating the value of the imaginary trace at a certain sampled time instant, the DHT operator performs a weighted summation of the real trace samples located within the selected DHT time-window (Eq. 2.4). The DHT coefficients decrease in modulus with distance from the analyzed sample, which is located at the center of the DHT window, as shown in Fig. 2.2B, which plots the DHT coefficients versus the sample shift (i.e. k in Eq. 2.4). When the operator is covering the entire trace, the distortions are visible in time-intervals containing weaker signals, as shown in Fig. 2.2C, where the imaginary trace is characterized by a negative constant component instead of the expected -90° phase shift with respect to the recorded trace. These effects are removed when using a truncated DHT operator, as shown in Fig. 2.2D, which is equivalent to constructing the imaginary trace from a recorded trace that is identically null outside the DHT window. A valid choice would be a time window $2M\Delta t$ equal to at least the time interval of the transmitted wavelet. In this case, the value of M in Eq. 2.4 is set equal to 25, resulting in a DHT operator made of 51 samples, which correspond to a DHT time window about 20 ns wide. In the extreme case of M set equal to 1, the value of A'_n would be given by a simple subtraction of the two samples around A_n (Eq. 2.4), with effects ranging from imaginary samples constantly close to zero to random behavior, depending on the frequency spectrum of the analyzed signal and the sampling rate. Therefore, the resulting imaginary trace is partially dependent on the time width of the DHT operator, with significant repercussions on all the subsequently derived attributes, which are discussed in the following sections.

Reflection strength

The reflection strength is sensitive to changes in the EM impedance within the recorded GPR profile, and it is calculated as the modulus of the complex trace vector defined in Eq. 2.1, which is given by the following equation

$$S_n = \sqrt{A_n^2 + A_n'^2} \quad (2.5)$$

The reflection strength at a certain time instant depends on the values of both the real and imaginary traces, and therefore it is sensitive to the previously described possible distortions caused

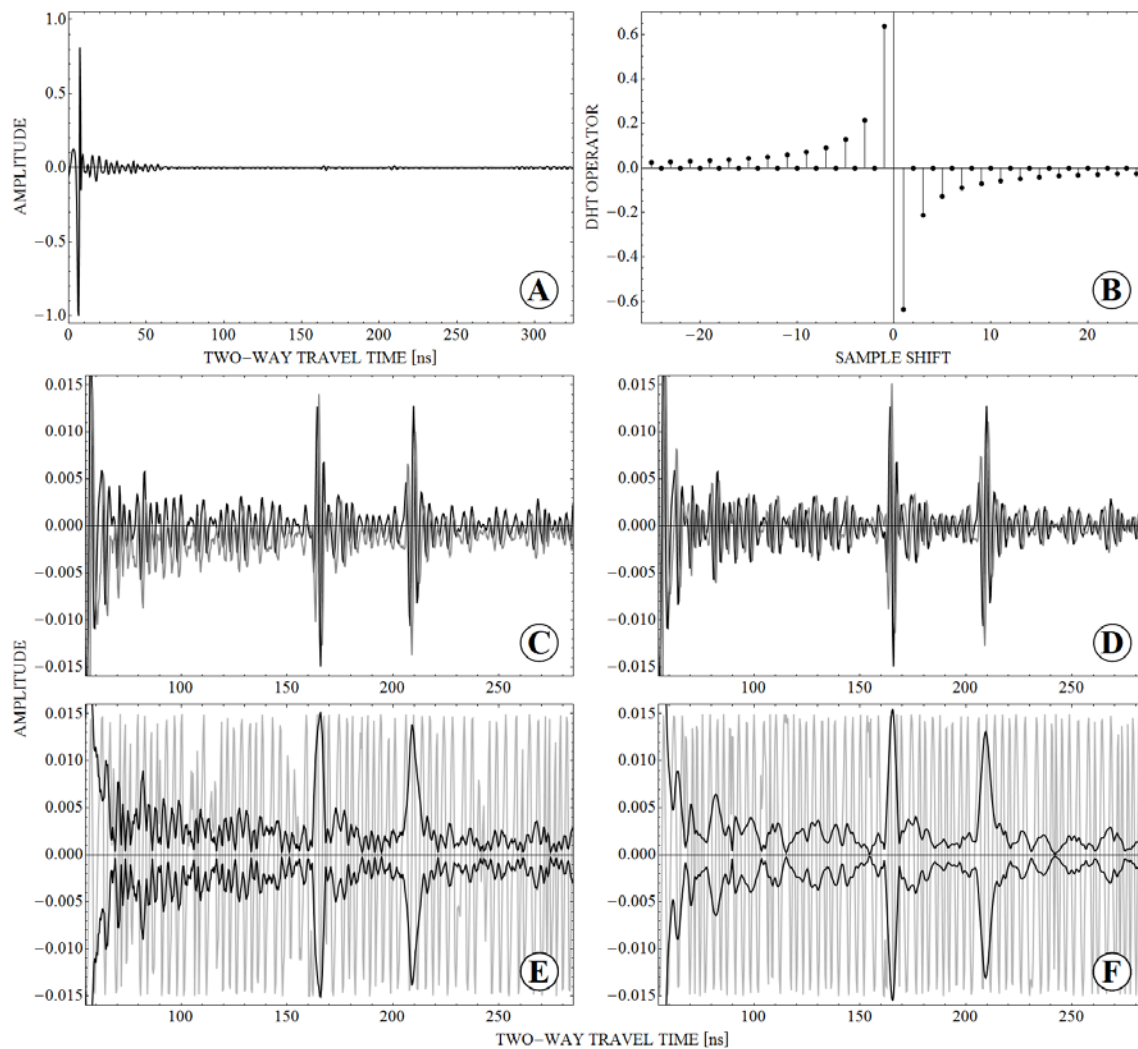


Figure 2.2: Attribute analysis of a recorded GPR trace through the application of DHT operators with different time widths. The normalized GPR trace (A) is dominated by the airwave-groundwave interference, which is several times stronger than any other recorded signal, especially without any amplitude recovery applied. The interference can distort the imaginary trace if the DHT operator (B) covers the entire trace, rather than being truncated. In the former case the attribute analysis results are given in (C, E), while in the latter they are given in (D, F), using a 51 samples (i.e. 19.9 ns) wide DHT operator. The graphs in (C, D) show the constructed imaginary trace (gray) superimposed to the real trace (black) in a specified time interval containing weaker signals. A negative constant component distorting the imaginary trace is visible in (C), caused by distant strong unrelated signals, and it is removed by shortening the DHT operator width (D). The graphs in (E, F) show the trace envelope (black) and the cosine of the instantaneous phase (gray), the latter being normalized to 0.015 for clarity, in the same time interval for the two cases. The distortions of the imaginary trace in (C) cause in (E) more rapid changes in the reflection strength as well as lower cosine phase peaks, while both effects are removed in (F) by truncating the DHT operator.

by the selected DHT window. For example, the constant component distorting the imaginary trace in Fig. 2.2C causes the trace envelope to have high frequency noise in Fig. 2.2E with respect to Fig. 2.2F, which is caused by the aforementioned disconnect between the real and imaginary traces.

The reflection strength can be used to quickly detect the strongest reflections within the recorded profile, by comparing the various local maxima and isolating the main peaks. For example, the trace envelope in Fig. 2.2F clearly shows two main peaks within the analyzed time interval, respectively located at 165 and 210 ns, besides the initial saturated event, which is the tail of an earlier stronger reflection, as well as several other minor peaks. The reflection strength analysis is discussed in greater details in the next section, in which the presented picking algorithm uses the trace envelope to divide the GPR trace into energy packages and automatically identify the various reflections. However, since the reflection strength does not depend on the signal phase, there is no direct correspondence between a reflection strength maximum and a signal peak or trough, or between a reflection strength minimum and a zero-crossing in the analyzed trace. Therefore, the reflection strength alone cannot be used to accurately define the time-space position, peak amplitude, and polarity of a reflection.

Cosine of the instantaneous phase

The instantaneous phase is the phase of the complex trace vector defined in Eq. 2.1, and it is given by the following equation

$$\phi_n = \arctan \left(\frac{A'_n}{A_n} \right) \quad (2.6)$$

The instantaneous phase can improve the interpretation of a recorded profile by removing the signal amplitudes and consequently highlighting the lateral phase continuity of recorded events, such as reflections, diffractions, and even coherent noise, which allows for a better imaging of the subsurface stratigraphy, including discontinuities and interference between different events. The instantaneous phase can also improve automated picking, since reflections are better identified by their lateral phase continuity rather than their amplitude continuity (Barnes, 1996). Moreover, a signal propagating in low or moderate-loss dielectric media can show significant changes in its amplitude spectrum, while the phase spectrum remain practically unaltered for the usual time ranges (Neto and de Medeiros, 2006), which limits the signal dispersion and maintains the time-position of the signal zero-crossings.

The main downside of the instantaneous phase is that the attribute is wrapped between $\pm 180^\circ$ (Eq. 2.6), and therefore it cannot be easily processed for further signal analysis, due to its cyclical discontinuities. This problem can be removed by considering the cosine of the instantaneous phase, also referred to in the following as the cosine phase, which is obtained by rearranging Eq. 2.2:

$$\cos(\phi_n) = \frac{A_n}{S_n} \quad (2.7)$$

In case the reflection strength is null, the cosine phase can be set identically null, in order to avoid indeterminate values in Eq. 2.7. Nevertheless, in real data sets the reflection strength is very rarely null, if ever, since a null value of the reflection strength S_n would require both the real A_n and imaginary A'_n amplitude samples to be null at time instant $n\Delta t$ (Eq. 2.5), and a null value of the imaginary trace A'_n would require the real trace A_n to be null over the entire $2M\Delta t$ wide time-window of the DHT operator centered at time instant $n\Delta t$ (Eq. 2.4), or at least the samples A_n in such time-window would have to cancel each other out in the calculation of A'_n .

The cosine phase profile is conceptually similar to the amplitude profile, behaving as if all peak amplitudes were normalized to 1 in modulus, and therefore it can be further processed like any other recorded data set. In fact, the cosine function removes the $\pm 180^\circ$ discontinuities of the instantaneous phase ϕ_n , turning them into troughs (Eq. 2.7). Moreover, the zero-crossings of the cosine phase coincide with those of the recorded signal, since when the amplitude is null, so is the cosine phase (Eq. 2.7). However, there is not an immediate correspondence between the peaks and troughs of the cosine phase and those of the recorded signal. In fact, the cosine peaks and troughs have a value equal to 1 in modulus only when the reflection strength and the amplitude are equal in modulus (Eq. 2.7), while in any other case they are lower in value. Moreover, the peaks and troughs of the cosine phase coincide with those of the amplitude only at the reflection strength maxima, otherwise they are slightly shifted in the decreasing direction of the reflection strength. This behavior can be observed in Figs. 2.2E-F, which also show the effects of the previously discussed distortions in the imaginary trace. The constant component in the distorted imaginary trace causes in Fig. 2.2E an increase in the reflection strength, which changes or removes several points of tangency with the recorded trace. This causes significant distortions in the cosine phase, with several cosine phase peaks reaching values considerably lower than 1 in modulus, as opposed to the undistorted case shown in Fig. 2.2F.

In the following section, the cosine phase is used to automatically track any recorded event in a GPR profile characterized by lateral phase continuity, regardless of its reflection strength or lateral amplitude variations. The analysis is applied to the glaciological GPR profile shown in Fig. 2.3A, which is hereafter used to highlight the main stages of the proposed picking algorithm, and also to critically analyze its performance, advantages, and limitations. The attribute analysis applied to the single trace in Fig. 2.2D, was extended with the same parameters to the entire profile in Fig. 2.3. The reflection strength profile (Fig. 2.3B) can be used to quickly identify the main reflections within the profile, but it cannot be directly linked to the specific reflection phases, nor it can reliably define the arrival times and polarities of the reflections. The cosine phase profile (Fig. 2.3C) highlights the lateral phase continuity of every recorded event regardless of its strength, therefore it can be used to isolate all laterally coherent events, while disregarding incoherent noise-related signals.

2.2 Picking algorithm

After removing the signal amplitudes, the cosine phase profile highlights the lateral continuity of reflections regardless of their strength or their lateral amplitude variations, with only the signal phases remaining. In this text, the following terms are used:

1. A zero-crossing indicates either a sample in which the signal amplitude is null, or the sample immediately following the actual zero-crossing, when the latter is located between two adjacent samples with opposite polarities.
2. A signal phase (or simply phase) indicates the interval between two consecutive zero-crossings, and therefore in such interval the cosine phase, or equally the signal amplitude, has a constant polarity. A reflection phase instead indicates a set of signal phases picked from adjacent traces and belonging to the same reflection, which have the same position within the reflected wavelet, i.e. they are either the first phase of the reflection, or the second phase, and so on.
3. The arrival time of a signal phase indicates the time instant of the earliest of the two zero-crossings containing such phase, which is equivalent to the first sample within the analyzed phase interval.

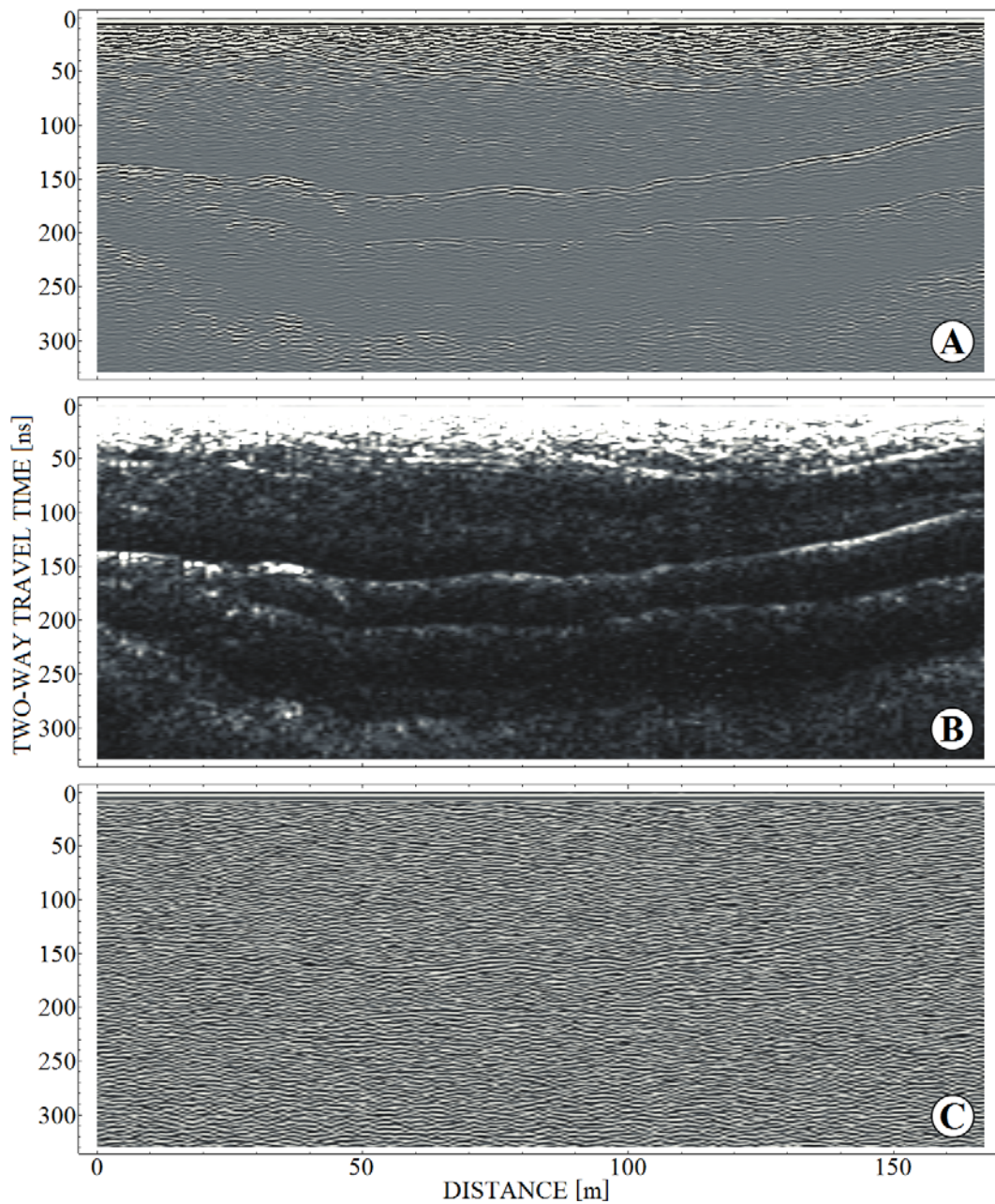


Figure 2.3: Example of attribute analysis applied to a glaciological GPR profile. The figure shows the amplitude (A), the reflection strength (B), and the cosine phase (C) profiles. The reflection strength profile can be used to quickly identify the main recorded reflections, showing two strong reflections at the center of the GPR profile (100-200 ns), as well as a few visible reflections in the shallower region (0-80 ns), which contains several interfering events caused by the greater heterogeneity of recently deposited snow. The cosine phase profiles removes the signal amplitudes and highlights the lateral continuity of every recorded event, regardless of its reflection strength or lateral amplitude variations.

4. A peak cosine phase and a peak amplitude respectively indicate the two samples within a signal phase in which the cosine phase and the signal amplitude are at their maxima, in modulus. As previously discussed, these two samples coincide only when they also coincide with a maximum in the reflection strength, otherwise they are separated by a few samples.

2.2.1 Horizon picking

The auto-picking algorithm presented in this chapter uses the cosine phase to mark as a horizon any laterally coherent event within the recorded profile. After constructing the cosine phase profile, which removes the signal amplitude, any phase can potentially be part of a horizon and therefore the algorithm marks every recorded phase in the profile. Each phase can thus be characterized in terms of its arrival time twt_{ij} , its peak cosine phase C_{ij} , its peak amplitude A_{ij} , as well as its polarity. These attributes are arranged in a matrix form, where the indexes i and j respectively indicate the i -th phase within the j -th trace of the profile, with the total number of phases varying in each trace. From this initial set, those phases considered to be too distorted by either interference or noise to be useful for the picking process can be automatically disregarded by setting thresholds on specific quantities defining their shape. For instance, a phase can be disregarded if its peak cosine C_{ij} has a value considerably lower than 1 in modulus, or if its time interval $[\text{twt}_{ij}, \text{twt}_{i+1j}]$ is either too long or too short with respect to a reference wavelet, which can be separately measured.

After filtering out these distorted phases, the algorithm analyzes the remaining phases singularly, considering each one of them the beginning of a possible horizon, and it connects them across adjacent traces only if they are identified as part of the same horizon. In particular, two phases in adjacent traces C_{ij} and C_{kj+1} are automatically connected if the following conditions are met:

1. The two phases must have the same polarity.

$$\frac{C_{ij}}{|C_{ij}|} = \frac{C_{kj+1}}{|C_{kj+1}|} \quad (2.8)$$

This means that a single phase in a reflection characterized by a change in polarity, as it could be for example in the case of a CMP GPR profile in the transverse magnetic configuration, would be marked with two independent horizons with opposite polarities.

2. The two phases must be the closest possible in terms of arrival times among the possible pairings between the two traces. In particular, they have to be closer than a specified time-window T_P

$$|\text{twt}_{ij} - \text{twt}_{kj+1}| \leq T_P \quad (2.9)$$

The time-window has to be small enough to avoid connecting distant and unrelated phases, while also be large enough to be able to connect phases from the steepest reflections. These requirements can be problematic when analyzing profiles with large trace intervals, which may be characterized by spatial aliasing.

3. The two phases must not be on the opposite sides of an already constructed horizon, in order to avoid crossing it.

Starting from an initial phase C_{ij} , if the algorithm finds a phase C_{kj+1} in the next trace which satisfies the picking conditions, the phase is added to the horizon under construction and the analysis is repeated on this new phase. The procedure is iterated until valid phases can be added to the horizon, and if

no more phases are available the process is interrupted. After constructing all possible horizons, the unconnected phases are deleted from the set of marked phases C_{ij} , and the picking results can be superimposed to the recorded profile for further analysis and interpretation.

In order to avoid having hundreds of horizons just a few traces in length, which are most likely caused by random phase combinations in areas with low signal-to-noise ratio, a minimum horizon length L_H can be specified. Also an optional threshold E_H can be set for the average reflection strength of each horizon, in order to isolate horizons with specific energy values.

There are several factors which can influence the picking results, including:

1. Interference, caused for example by complex subsurface stratigraphy or by the presence of diffraction hyperbolas, which can result in deformed horizons connecting signal phases belonging to different unrelated events.
2. Coherent noise, caused for example by antenna ringing or other external sources, which can result in artificial horizons unrelated to actual subsurface impedance contrasts.
3. Random noise, which can interrupt an horizon under construction by deforming signal phases, and can cause a single long reflection phase to be separated into several consecutive horizons. In areas with low signal-to-noise ratio, the algorithm can create random horizons a few traces wide, which can be removed by setting the aforementioned minimum horizon length.

Given all these factors, appropriate processing may be required before picking, including background removal, band-pass filtering, and migration. However, while migration can be useful in collapsing the hyperbolic diffractions into their apexes and relocating dipping reflections at their correct position, the procedure requires as input an accurate EM velocity distribution for the analyzed GPR profile, which is usually not known in advance and is the very aim of the inversion procedure described in the following chapter. Moreover, the EM velocity distribution within a glaciological GPR profile can show significant vertical and lateral variations, especially in small glaciers and glacierets such as the one shown in Fig. 2.3 (Colucci et al., 2015). Therefore, the application of the migration process using inaccurate velocity distributions, which are often approximated as either constant or slow varying, can result in significant signal distortions, such as the over-migration of diffraction hyperbolas or the creation of artificial long flat horizons in areas with low signal-to-noise ratio.

Nevertheless, since the picking algorithm does not depend on the signal amplitude, it can be applied without the need of amplitude recovery, which allows the original amplitudes to be picked without any subjective distortion caused by the assumptions made by the interpreter with regards to the EM properties of the subsurface, in order to counteract the amplitude decay. Furthermore, the algorithm does not rely on any manually selected control point, which makes the picking results more objective.

An example of automated picking applied to the glaciological GPR profile shown in Fig. 2.3 is given in Fig. 2.4. The minimum horizon length L_H was set equal to 10 traces, therefore the algorithm was able to detect every event, at least 2 m in length, characterized by lateral phase continuity.

2.2.2 Horizon joining

Random noise and interference can cause a long laterally coherent event to be divided into a series of shorter consecutive horizons separated by distorted phases creating gaps generally a few traces wide. The picking algorithm can reduce this effect by automatically connecting consecutive horizons, using other horizons in their vicinity as patches covering the gaps between them. The algorithm analyzes each horizon singularly, treating it as a potential patch and searching for connectable horizons in its

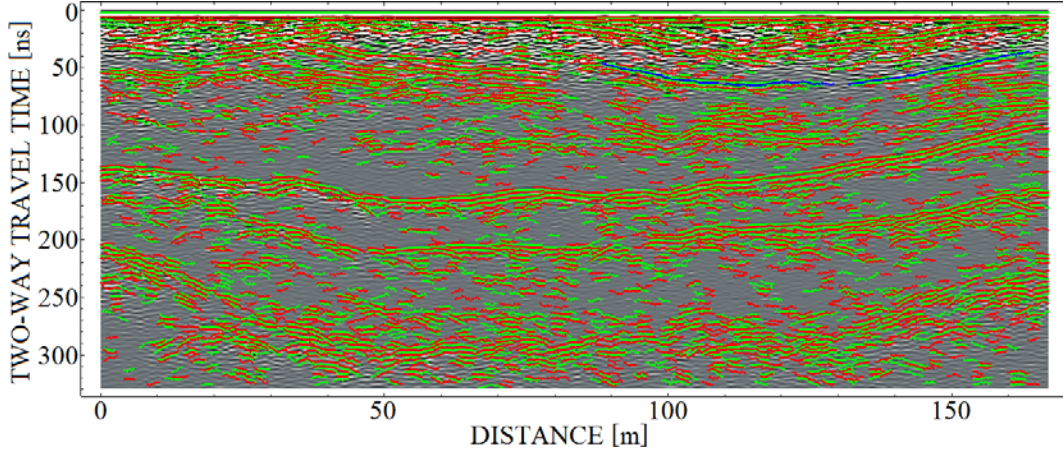


Figure 2.4: Example of automated picking applied to the glaciological GPR profile shown in Fig. 2.3. The minimum horizon length was set equal to 10 traces, which correspond to about 2 m. Positive amplitudes are marked in green, negative amplitudes in red.

vicinity. The algorithm searches both sides of the patching horizon by analyzing the signal phases either above or below the patching horizon in each trace. Consider, for example, two phases C_{ij} and $C_{kj+\Delta j+1}$ within a patching horizon separated by a gap Δj traces wide, and suppose there are two phases immediately above them C_{i-1j} and $C_{k-1j+\Delta j+1}$ respectively identified as the end and the beginning of two potentially connectable horizons. The two consecutive horizons are automatically connected if the following conditions are met:

1. The two horizons must have the same polarity.

$$\frac{C_{i-1j}}{|C_{i-1j}|} = \frac{C_{k-1j+\Delta j+1}}{|C_{k-1j+\Delta j+1}|} \quad (2.10)$$

This means that two horizons marking the same phase of a reflection which changes its polarity would remain separate.

2. The two horizons must be close in terms of arrival time to the patching horizon, at least at their connectable ends. In particular, they have to be closer than a specified time-window T_J , which limits the analysis to a small area centered on the patching horizon.

$$\begin{aligned} |twt_{ij} - twt_{i-1j}| &\leq T_J \\ |twt_{kj+\Delta j+1} - twt_{k-1j+\Delta j+1}| &\leq T_J \end{aligned} \quad (2.11)$$

3. The gap width Δj must not be larger than a specified threshold L_G , in order to avoid connecting unrelated events at the opposite ends of the patching horizon.

$$\Delta j \leq L_G \quad (2.12)$$

4. There must be no other horizon crossing the gap between the two horizons. The algorithm analyzes each phase C_{i-1j} along the gap $[j+1, j+\Delta j]$ directly above the patching horizon, which

belongs to any other horizon, and if an horizon is found to be closer in terms of arrival time to the patching horizons than the potentially connectable ends C_{i-1j} and $C_{k-1j+\Delta j+1}$ of the two analyzed horizons, then this horizon is automatically identified as a crossing horizon, and the two analyzed horizons are therefore not connected.

The same procedure is iterated when searching the other side of the patching horizon, in which case the analyzed phases C_{i+1j} and $C_{k+1j+\Delta j+1}$ would be below the horizon.

The patching process is particularly efficient with horizons marking different phases of the same reflection, which are arranged by construction in a sub-parallel structure. Nevertheless, erroneous connections can occur in more complex structures like, for example, low dipping faults, where unrelated horizons at the opposite sides of the discontinuity can potentially be connected. The algorithm can reduce such errors by preventing a specified number of traces at both ends of the patching horizon from being used in the process. With this constraint, the gap has to be located more toward the center of the patching horizon, so that the two consecutive horizons are more likely to be part of the same coherent structure, rather than to be marking unrelated events within a chaotic transitional zone.

A few examples of possible patching scenarios are given in Fig. 2.5, which shows four parallel sets of consecutive horizons with alternating polarities in five different possible configurations. In the first case, the gap in set 2 can be patched by both the horizons in the sets 1 and 3, while the gap in set 4 can be patched just by the horizon in set 3, assuming that the time-window T_J is set at just about one signal phase interval, so that the sets 1 and 4 would be out of range from each other. In the second case, the horizon in set 2 is able to patch the gaps in the sets 1 and 3, while the gap in set 4 would be out of range using the aforementioned time-window. In the third case, both gaps in the sets 2 and 3 would be patched, respectively using as patching horizons those in the sets 1 and 4. In the fourth case, a dipping fault creates gaps in all four sets, none of which can be patched due to the fact that no horizon is fully covering any gap. In the fifth case, the fault has a lower dipping angle and in this case all gaps would be patched unless the interpreter defines an exclusion zone at both ends of each patching horizon, as previously discussed, in order to preserve the discontinuity.

The patching process is very useful in reducing the total number of constructed horizons, while also increasing their average length, and thus improving the interpretation of the analyzed profile. An example is given in Fig. 2.6, which shows the horizon length distribution before and after the horizon patching process has been applied to the picked horizons shown in Fig. 2.4. In the analyzed GPR profile the algorithm initially picked 3361 independent horizons, then reduced to 2932 (i.e. a 12.8% reduction) by the automated patching process, resulting in a clear shift of the horizon length distribution toward higher values (Fig. 2.6). The maximum gap length L_G was set equal to 5 traces (i.e. about 1 m), while the two outer traces at both ends of each patching horizon were excluded from the process, as previously discussed.

2.3 Automated grouping

The large set of independent horizons automatically constructed by the picking process are not directly useful in defining the subsurface stratigraphy, since they mark every laterally coherent event within the recorded profile. The second part of the algorithm is therefore used to automatically group together horizons marking different phases of the same reflection, while disregarding isolated horizons which are not part of any automatically recognizable structure. There are two possible grouping methods, which use different criteria to detect each reflection and to select specific horizons for interpretation.

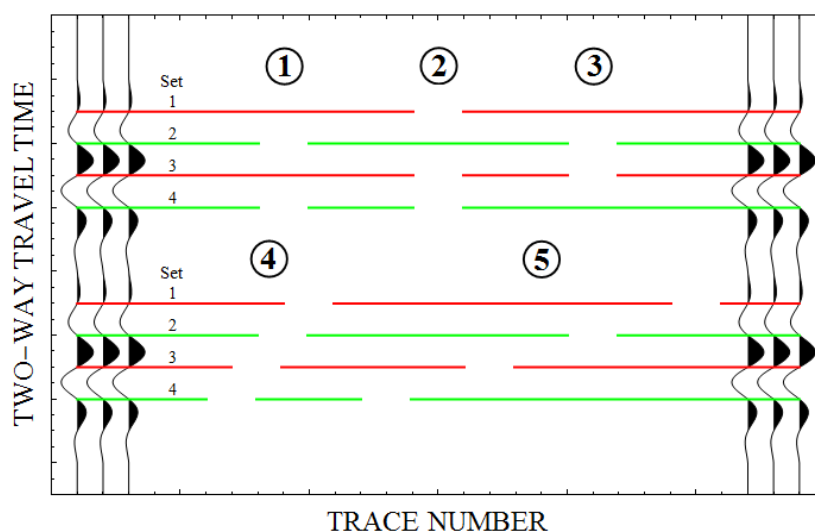


Figure 2.5: Examples of possible horizon patching configurations. The figure shows four parallel sets of consecutive horizons with alternating polarities in five different patching scenarios, which highlight various possible outcomes of the patching process. Modified image from Forte et al. (2016).

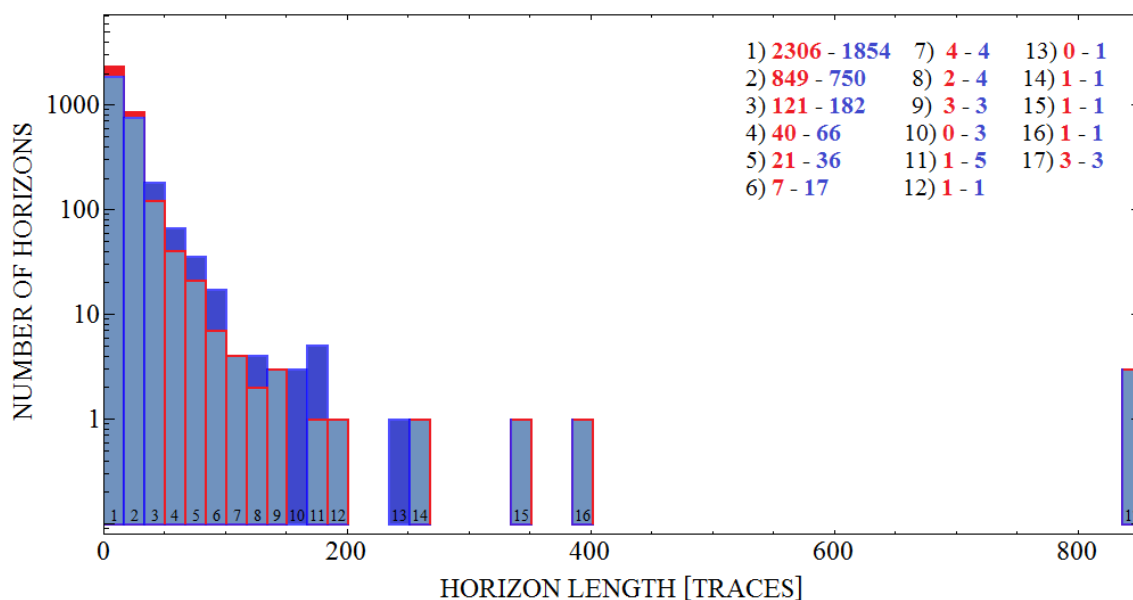


Figure 2.6: Analysis of the horizon length distribution before (red) and after (blue) the patching process for the piking results shown in Fig. 2.4. Each column is about 17 traces wide and their values are given in the top-right corner of the graph, while the gray area indicates overlapping between the two analyzed distributions. The graph shows an expected increase in the average horizon length due to the joining of consecutive horizons, which mark different parts of the same laterally coherent event.

2.3.1 Phase analysis

The phase method analyzes the behavior of the cosine of the instantaneous phase in the vicinity of each horizon, looking for sub-parallel events which can be grouped into the same reflection. These horizons are grouped together by comparing through cross-correlation their reflected wavelets, which can be reconstructed by averaging the cosine phase along each horizon, with the assumption that the reflected signal is laterally stable. In the following, the grouping process is applied to a generic horizon constructed in the trace interval $[J_1, J_2]$, and the analysis is limited to the earlier phases. The same process can be applied to later phases in order to obtain the total number of recognized phases in each reflection, which allows the interpreter to require a minimum number of reflection phases. This constraint can be used to disregard isolated horizons for which too few sub-parallel events have been recognized, and which therefore do not mark any automatically recognizable structure.

After selecting a specific horizon, the algorithm marks earlier phases in each trace up to a specified number N_P , and connects them across traces on the basis of their marking number (i.e. 1 to N_P), forming lines with constant polarities that do not cross one another. There are two obvious choices for the samples (also referred to as reference samples) to be used to mark a signal phase, namely its first zero-crossing and its peak cosine phase, and the choice of reference samples can influence the results of the analysis in particular situations. After all the lines have been constructed, each line is potentially marking a sub-parallel horizon that can be grouped together with the selected horizon, which is itself marked as line 0. In order to verify this, the reflected wavelet is reconstructed by averaging the cosine phase along each line (also referred to as averaging lines), and the resulting shapes are compared by means of cross-correlation. The algorithm creates $N_P + 1$ Averaged Cosine Traces (ACT) using a $2N_T + 1$ samples window centered on each line:

$$ACT_{kn} = \frac{1}{J_2 - J_1} \sum_{j=J_1}^{J_2} \cos(\phi_{(z_{kj}+n)j}) \quad (2.13)$$

with

$$n = -N_T, 1 - N_T, \dots, N_T$$

$$k = 0, 1, \dots, N_P$$

where z_{kj} is the reference sample of the k -th line in the j -th trace, and ϕ_{ij} is the instantaneous phase in the i -th sample of the j -th trace.

The general principle is that if an averaging line is marking a sub-parallel event, the shape of the reflected wavelet should be preserved by constructive averaging, while other external events should cancel each other out. On the other hand, if the line is connecting unrelated phases in an erratic pattern, there should be a destructive averaging in the overall trace, with no recognizable wavelet shape. A necessary condition for the accurate reconstruction of the reflected wavelet is that the signal must be laterally stable along the analyzed reflection, in which case the shape of the ACT resulting from a given reflection phase should not depend on the choice of reference samples for the line marking such reflection phase. It is clear from Eq. 2.13 that if the averaging line is marking either a zero-crossing or a cosine phase peak, the central sample ACT_{k0} of the resulting averaged trace will respectively mark the zero-crossing or the cosine phase peak of the ACT's central phase.

Assuming that the reflected signal is laterally stable, the phase analysis starts with the ACT of the selected horizon, in which the wavelet is considered preserved by construction, since the horizon was automatically tracked using the previously described picking process. In the general case of the

k -th averaging line, once the reflected wavelet is deemed preserved by the algorithm, its shape W_{kn} is extracted from a specified $2N_W\Delta t$ wide time-window centered on the phase preceding the central one in the analyzed k -th ACT:

$$W_{kn} = \text{ACT}_{k(n_P+n)} \quad (2.14)$$

with

$$n = -N_W, 1 - N_W, \dots, N_W$$

where n_P is the reference sample of the phase preceding the central one in the ACT, which corresponds to the central sample W_{k0} of the constructed wavelet, and it marks either a zero-crossing or a cosine phase peak, depending on the initial choice of reference samples.

After extracting the wavelet from the k -th line, the algorithm moves on to the next $(k+1)$ -th line, where the shape of the resulting ACT is compared with that of the wavelet constructed from the k -th line by means of cross-correlation:

$$\Phi_{k+1} = \sum_{n=-N_W}^{N_W} W_{kn} \text{ACT}_{(k+1)n} \quad (2.15)$$

In Eq. 2.15 the wavelet extracted from the ACT of the k -th line should coincide in terms of time-positions with the wavelet supposedly preserved in the ACT constructed from the $(k+1)$ -th line. Specifically, their respective central samples W_{k0} and $\text{ACT}_{(k+1)0}$ coincide in time by construction, and they are both either a zero-crossing or a cosine phase peak, while the remaining phases should coincide if they are part of the same reconstructed signal. The wavelet is considered preserved in the ACT constructed from the analyzed $(k+1)$ -th line if the following conditions are met:

1. The cross-correlation between the wavelets constructed from the $(k+1)$ -th and k -th lines (Eq. 2.15) must reach a minimum threshold $\Phi_{\min}$

$$\Phi_{k+1} \geq \Phi_{\min} \quad (2.16)$$

An example of cross-correlation threshold is a certain percentage of the auto-correlation Ψ_{k+1} of the wavelet extracted from the k -th line, since the objective is to determine whether or not the two cross-correlated signals in Eq. 2.15 are in fact the same:

$$\Psi_{k+1} = \sum_{n=-N_W}^{N_W} W_{kn}^2 \quad (2.17)$$

2. The peak C_P of the ACT's central phase must reach a value $C_{\min}$ close to 1 in modulus

$$|C_P| \geq C_{\min} \quad (2.18)$$

with

$$C_{\min} \approx 1$$

If this condition is violated, the $(k+1)$ -th averaging line is considered to be connecting unrelated phases with greatly different shapes. This requirement avoids particular situations in which the cross-correlation is acceptable (Eq. 2.16), but neither of the two compared shapes in Eq. 2.15 is actually related to the reflected wavelet.

The analysis is iterated in each line until one of the two conditions is violated, and the earliest line satisfying both conditions is automatically identified as the first phase of the analyzed reflection. Under favorable conditions of high signal-to-noise ratio and in absence of interference, the process can be used to automatically recover the polarity of each reflection. Nevertheless, the analysis is sensitive to noise and interference, which can deform the reconstructed wavelets, possibly resulting in horizons from the same reflection identifying different initial phases. Furthermore, close parallel reflections may not be recognized as separate by the algorithm, thus creating a large single group of sub-parallel horizons.

The phase analysis can be applied not only to the phases preceding the selected horizon, but also to the ones following it, in order to estimate the total number of sub-parallel events recognized for each horizon. Through this process, the algorithm can automatically disregard horizons with too few sub-parallel events, which are probably not marking any laterally coherent structure. The phase analysis also allows to select specific reflection phases to be used for further analysis and interpretation.

An example of phase analysis is given in Fig. 2.7, which studies the horizon marked in blue in Fig. 2.4, using zero-crossings as reference samples. The figure shows the cosine phase profile in the vicinity of the selected horizon, which is flattened and centered at relative time 0 for visual clarity, in order to highlight sub-parallel events which can be grouped into the same reflection. The averaging lines marking the phases preceding the selected horizon are superimposed to the cosine phase profile. The resulting ACTs are also shown for the analyzed horizon (graph 1) and the three earlier lines (graphs 2-4), with the cross-correlating wavelets superimposed in gray, each one constructed from the preceding graph. The minimum cross-correlation threshold (Eq. 2.16) in each graph is set at 80% the auto-correlation of the reconstructed wavelet (Eq. 2.17), while the minimum value required for the peak central phase of the ACT (Eq. 2.18) is set equal to 0.8. Both phase analysis conditions are satisfied in graph 2, but not in graph 3, where the two reconstructed signals differ significantly. Therefore the process is automatically interrupted and the selected horizon is identified as the second phase of the analyzed reflection, whose initial phase is marked by the averaging line of graph 2.

The same results are obtained in Fig. 2.8, which is similar to Fig. 2.7, but it uses the cosine phase peaks as reference samples instead of the zero-crossings. In this case, the phase analysis conditions are also satisfied in graph 2, but not in graph 3, therefore the selected horizon is automatically identified as the second phase of the analyzed reflection, while the averaging line of graph 2 is marking the initial phase. In case noise or interference caused distortions in the reconstructed wavelets, thus violating the assumption of a laterally stable signal, the analysis may obtain for the same horizon different results with different reference samples, while different horizons within the same reflection may recognize different initial phases. In case the reflected signal is not laterally stable, the energy analysis discussed in the following section may be preferable for horizon grouping.

The phase analysis was applied to each one of the picked horizons shown in Fig. 2.4, with the same thresholds used in the previous examples and for both choices of reference samples. The procedure was used to automatically identify the main recorded reflections, while disregarding isolated horizons not marking any automatically recognizable structure. As previously discussed, the algorithm selected those horizons in which at least two other sub-parallel events in their vicinity can be grouped into the same reflection, identifying 2201 such horizons (i.e. 75.1% of the total) when using zero-crossings as reference samples, while identifying 2231 such horizons (i.e. 76.1% of the total) when using cosine phase peaks as reference samples. Among these horizons, those automatically identified as part of the first three phases of their respective reflections are shown in Fig. 2.9, for both the zero-crossings (Fig. 2.9A) and the cosine phase peaks (Fig. 2.9B) analyses. We can notice that all the main reflections are accurately identified by the algorithm, while the number of displayed horizons is visibly reduced with respect to Fig. 2.4, thus improving the interpretation.

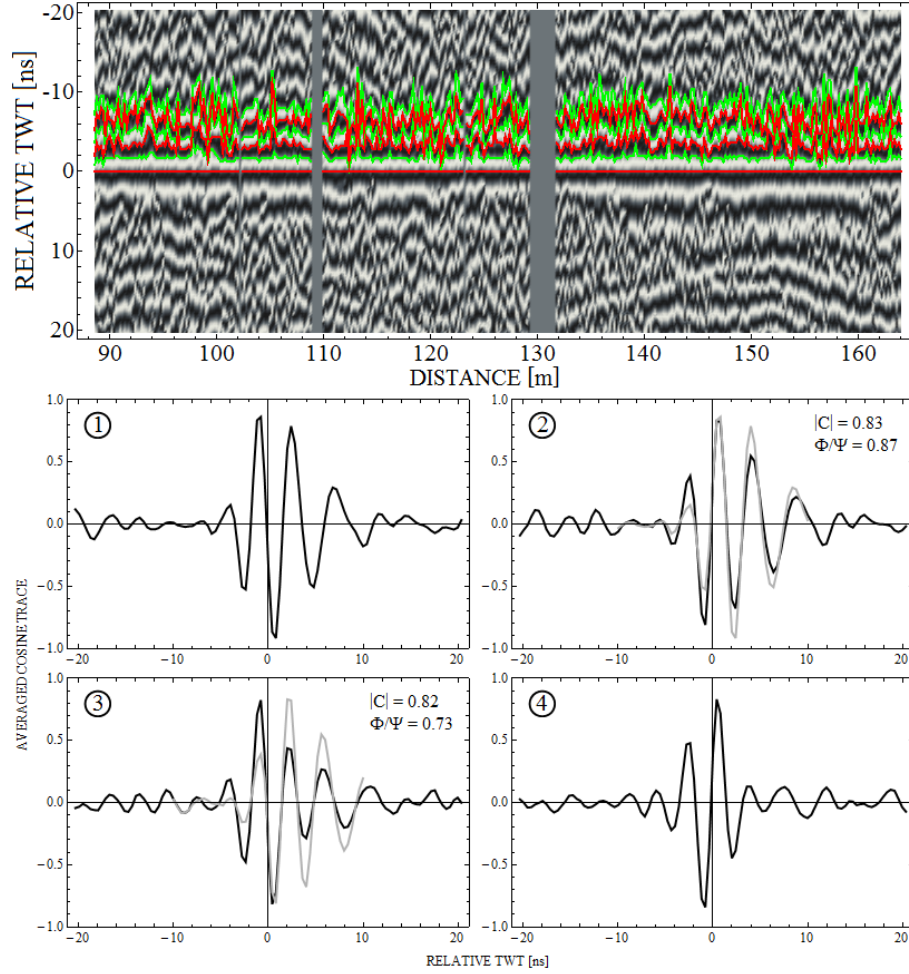


Figure 2.7: Example of phase analysis applied to the horizon marked in blue in Fig. 2.4, using zero-crossings as reference samples. The selected horizon is flattened and centered in the cosine phase profile, in order to better highlight surrounding sub-parallel events. The averaging lines are superimposed to the profile, marking both positive (green) and negative (red) phases. The analysis is limited to the phases preceding the selected horizon, with the graphs showing the averaged cosine traces (black) of the selected horizon (1) and of the three earlier lines (2-4), with superimposed the wavelets (grey) used for cross-correlation, each one constructed from the preceding graph.

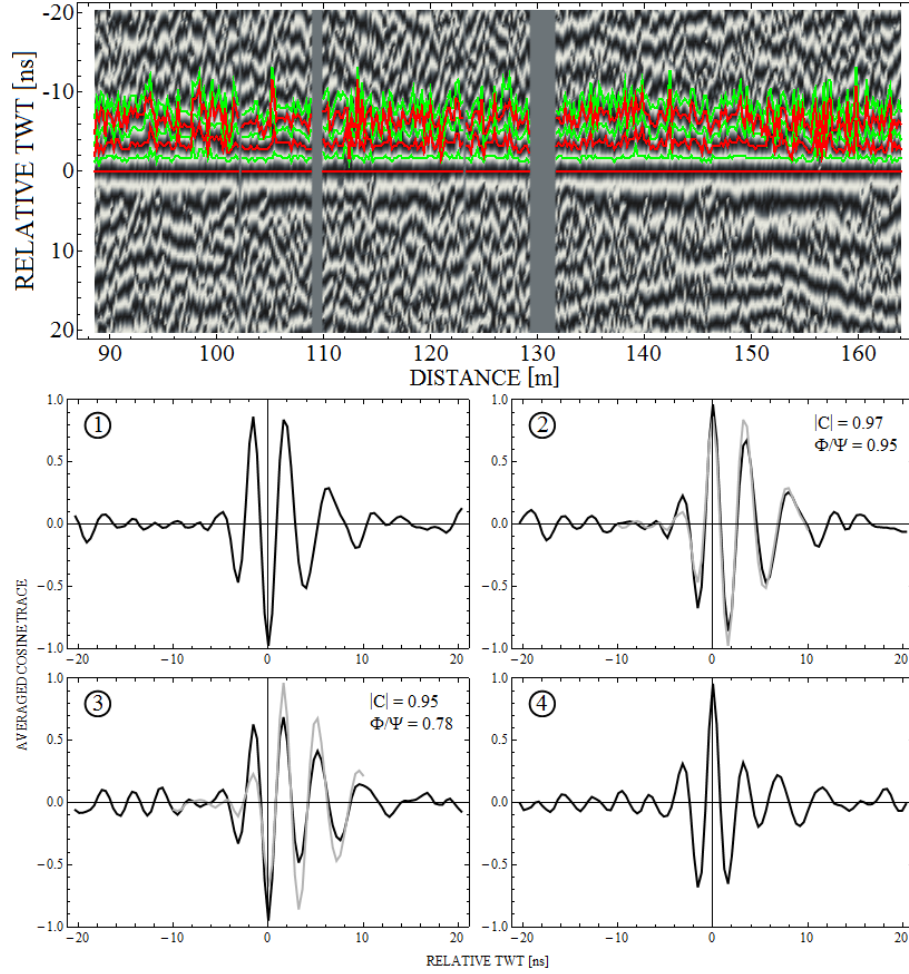


Figure 2.8: Example of phase analysis applied to the horizon marked in blue in Fig. 2.4, using cosine phase peaks as reference samples. The selected horizon is flattened and centered in the cosine phase profile, in order to better highlight surrounding sub-parallel events. The averaging lines are superimposed to the profile, marking both positive (green) and negative (red) phases. The analysis is limited to the phases preceding the selected horizon, with the graphs showing the averaged cosine traces (black) of the selected horizon (1) and of the three earlier lines (2-4), with superimposed the wavelets (grey) used for cross-correlation, each one constructed from the preceding graph.

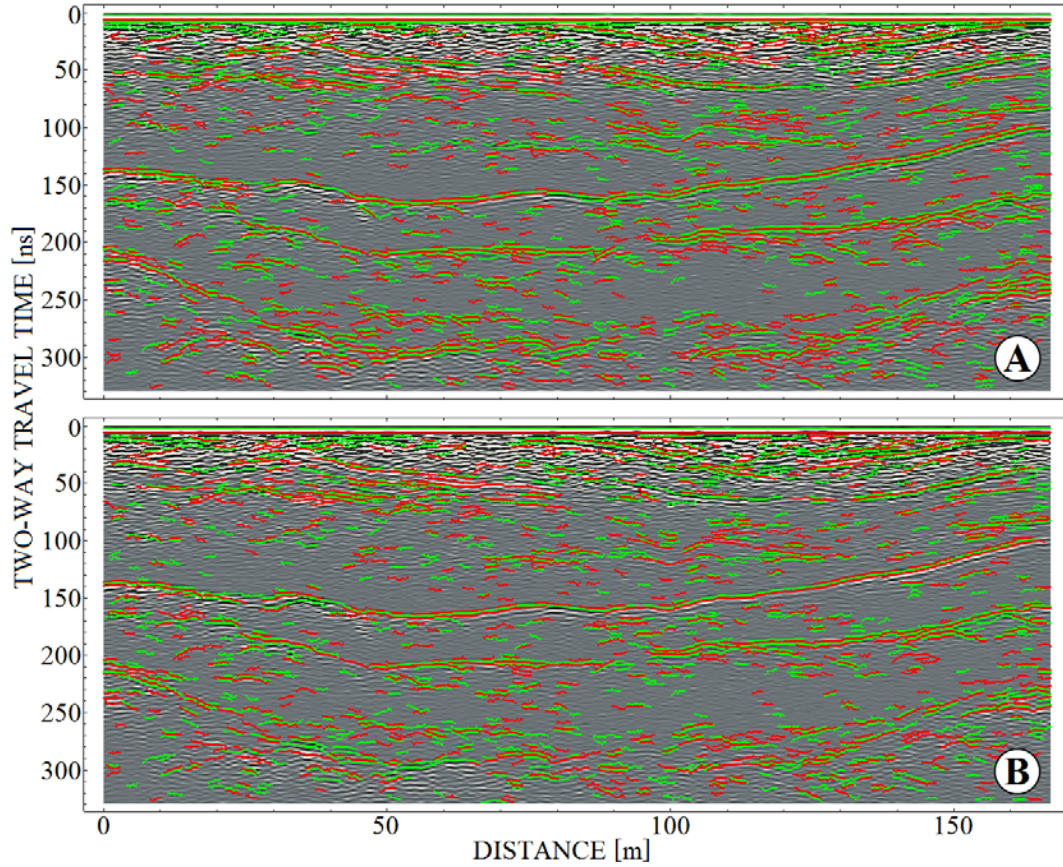


Figure 2.9: Example of automated phase analysis applied to the picked horizons shown in Fig. 2.4, using both zero-crossings (A) and cosine phase peaks (B) as reference samples. The figure shows those horizons automatically identified as marking one of the first three phases of reflections having at least three recognizable phases, which means that each horizon must have at least two sub-parallel events in its vicinity that can be grouped into the same reflection. Positive amplitudes are marked in green, negative amplitudes in red.

2.3.2 Energy analysis

The basic assumption of the phase analysis procedure is that the reflected signal is laterally stable, so that its shape can be accurately reconstructed by averaging the cosine phase along a picked horizon and be comparable with other wavelets reconstructed from horizons marking different phases of the same reflection. In case the accurate reconstruction of a wavelet is prevented by lateral changes in the reflected signal caused for example by noise, interference, or signal processing, the horizon grouping can be performed by using the energy analysis as an alternative. In this method, the recorded reflections are automatically identified by dividing the recorded profile into energy packages, which are defined by analyzing the various peaks of the reflection strength. Subsequently, horizons marking different phases of the same reflection are automatically grouped together by searching for sub-parallel horizons within the same energy package. Specific horizons from the different reflections can then be automatically selected by applying either temporal or energy thresholds.

The algorithm initially identifies each energy peak by marking all the reflection strength maxima and minima in each trace. A main energy peak may contain several local maxima and minima caused by small variations of the reflection strength, which must be filtered out. Consider two consecutive maxima of the reflection strength S_k and S_{k+1} , such that $S_k < S_{k+1}$, which are separated by a minimum s_k within a given trace. The algorithm regards the two maxima to be marking separate peaks if the minimum between them has a value lower than a specified percentage of the weaker of the two maxima

$$s_k < \chi_S S_k \quad (2.19)$$

with

$$0 \leq \chi_S \leq 1$$

If the condition in Eq. 2.19 is violated, the two maxima are automatically identified as part of the same peak, therefore the weaker of the two S_k is disregarded as a local maximum, together with the local minimum s_k between them. The algorithm iteratively analyzes each minimum in the recorded trace, from the strongest to the weakest in terms of reflection strength, removing in each cycle the local maxima and minima until only the actual maxima remain, separated by minima which define the boundaries of each energy package. An example of division of a single GPR trace into energy packages is given in Fig. 2.10, which shows the marked reflection strength maxima and minima before and after the application of the automated filter. In the analysis, the value of χ_S in Eq. 2.19 is set equal to 0.5, which means that a minimum has to have a reflection strength lower than 50% of the weaker of the two adjacent maxima, in order to be recognized as separating two independent energy peaks.

The procedure illustrated in Fig. 2.10 was iteratively applied with the same parameters to each trace of the profile shown in Fig. 2.3B, resulting in the reflection strength maxima and minima shown in Fig. 2.11A. We can notice that the energy packages are well defined along the main reflections, which are characterized by higher lateral continuity, while areas without clearly defined structures or with low signal-to-noise ratio are characterized by a more random distribution of the reflection strength maxima and minima. We can also notice that the minima marking the boundaries between energy packages in Fig. 2.11A do not necessarily form continuous lines which can be automatically picked, due to the high variability of the reflection strength creating several peaks with a wide range of energy values. Therefore, horizons within the same reflection are grouped together through a statistical analysis, which compares their arrival times with the time-space positions of the reflection strength minima, with the algorithm searching for sub-parallel horizons within the same energy package. The algorithm singularly analyzes each horizon, from the strongest to the weakest in terms of reflection strength, comparing it with other horizons located in its vicinity, and adding it to an existing horizon

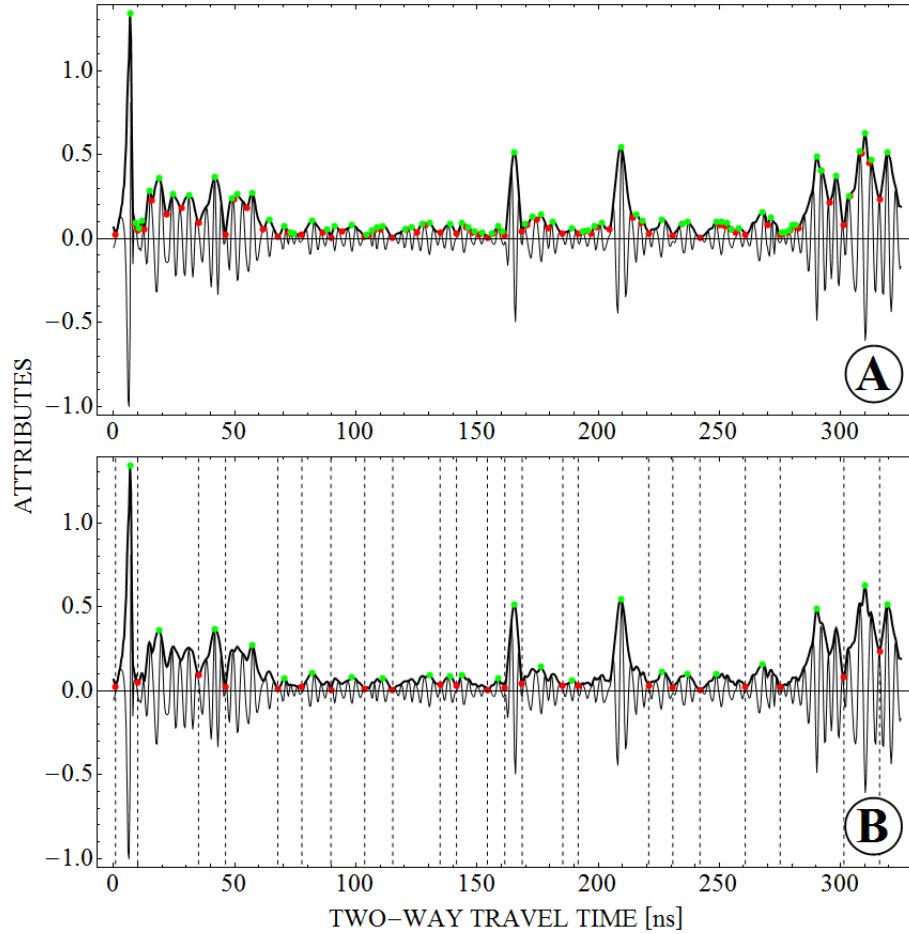


Figure 2.10: Example of energy analysis applied to the recorded GPR trace shown in Fig. 2.2, which is divided into energy packages. The figure shows the normalized recorded signal (thin black line), its reflection strength (thick black line), and the reflection strength maxima (green dots) and minima (red dots) before (A) and after (B) the automated filtering. The threshold used for the automatic recognition of a minimum was set at 50% of the weaker of the two adjacent maxima (Eq. 2.19). A linear gain function was applied to the trace, in order to counteract the dominant airwave-groundwave interference and thus better visualize the recorded reflections. Also, the vertical dashed lines in (B) are used to highlight the boundaries of the automatically identified energy packages.

group if the compared horizons are compatible. If neither of two compatible horizons belongs to an existing group, a new group is created to which other horizons can be added. Consider two horizons respectively L_1 and L_2 traces in length, such that $L_1 < L_2$, that share a certain number Δj of traces. These two horizons are automatically grouped together if the following conditions are met:

1. The two horizons must have opposite polarities, as expected from a reflected wavelet made of several phases with alternating polarities.
2. The number of shared traces must be higher than a specified percentage of the length of the shorter of the two horizons

$$\Delta j > \chi_L L_1 \quad (2.20)$$

with

$$0 \leq \chi_L \leq 1$$

This condition prevents the grouping of horizons which share just a few traces at their ends.

3. The number of traces N_t in which the difference in arrival times between the two horizons is less than a specified time interval, must be higher than a specified percentage of Δj

$$N_t > \chi_t \Delta j \quad (2.21)$$

with

$$0 \leq \chi_t \leq 1$$

4. The number of traces N_m in which a reflection strength minimum is between the cosine phase peaks of the two horizons, must be lower than a specified percentage of Δj

$$N_m < \chi_m \Delta j \quad (2.22)$$

with

$$0 \leq \chi_m \leq 1$$

If there is a certain number of shared traces in which either one of the two analyzed horizons contains gaps caused by the previously described horizon patching process, this number has to be subtracted from Δj in Eqs. 2.21 and 2.22.

5. If one of the two horizons already belongs to an existing group, the other one has to be weaker in terms of their average reflection strength in order to be added to the group. This constraint improves the procedure by avoiding the grouping of horizons belonging to close but separate reflections, which would have too few reflection strength minima between them to be automatically separated using just the previous conditions.

After all compatible horizons have been grouped together, specific horizons from each energy package can be selected for interpretation. Unlike the phase analysis, which can detect the first phase of a reflection under favorable conditions, the energy method cannot be used to determine the polarity of a reflection. Specifically, we can notice in Fig. 2.10B that the first signal phase within a given energy package does not necessarily correspond to the first phase of the main reflected wavelet within such energy package, especially when multiple close reflections are recorded. Nevertheless, the horizons close to the peak in each energy package can be selected by setting either temporal thresholds, isolating those horizons within a specified time-window centered on the reflection strength maxima in each trace, or

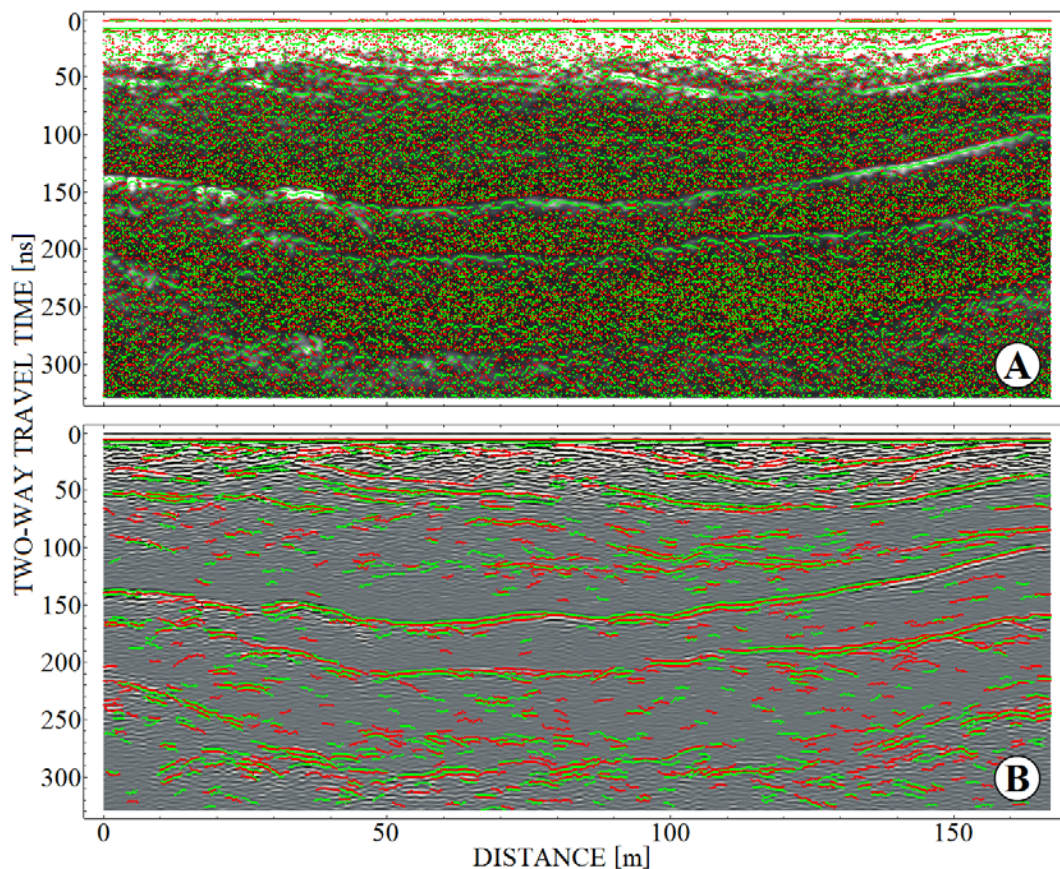


Figure 2.11: Example of automated grouping applied to the picking results shown in Fig. 2.4. The figure shows (A) the maxima and minima of the reflection strength used to define the boundaries of each energy package, and (B) the resulting grouped horizons having a reflection strength equal to at least 50% the corresponding energy peaks in each trace.

energy thresholds, isolating those horizon characterized by a reflection strength higher than a certain percentage of the corresponding energy peak in each trace. This selection is statistical in nature, since the reflection strength maxima also do not necessarily form continuous lines that can be automatically picked (Fig. 2.11A), nevertheless it can be used to automatically identify the main reflections and to reduce the total number of displayed horizons, thus improving the interpretation.

An example of horizon grouping results obtained using the energy packages defined in Fig. 2.11A is given in Fig. 2.11B, which shows those horizons that have a reflection strength equal to at least 50% the corresponding energy peaks in each trace. The algorithm identified 1136 such horizons (i.e. 38.7% of the total), and they were selected among horizons which were automatically grouped together using values of the parameters χ_L , χ_t , and χ_m in Eqs. 2.20, 2.21, and 2.22, respectively equal to 0.7, 0.7, and 0.3. We can notice that the main reflections are accurately identified, while other horizons in areas with no automatically recognizable structures are mostly disregarded.

Chapter 3

Amplitude inversion

3.1 State of the art

3.1.1 Glacier monitoring

The study of the internal stratigraphy, density distribution, total volume, and water content of glaciers and snowfields, as well as their variations with time, can be useful to better understand glacial structures and processes. Glacier monitoring offers several applications in many fields, which include (Richardson and Reynolds, 2000; Zemp et al., 2009):

1. Assessment and mitigation of the risk associated with glaciers and glacier-related hazard events. These events include flooding, caused for example by a sudden discharge of water under pressure from a glacier, or by an outburst from a proglacial lake, either ice or moraine dammed; avalanches, caused by the downward slide of large masses of snow, ice, and debris; and glacier surges, which are rapid increases in the rate of glacier flow. Flooding and avalanches can directly impact people, livestock, and infrastructures within minutes or hours, while glacier surges can cause significant loss of land and last for several months or even years.
2. Resources management, in particular for hydropower generation and fresh water reservoirs. Glacier fluctuations, that is the temporal variations in the solid water content, can cause water supply shortages during low flow conditions, as well as excessive flow during large water discharges. Glacier fluctuations are linked to changes in the climate and can have time scales of several years or even decades.
3. Climate monitoring, in particular the glacier mass balance offers a direct and undelayed response to local atmospheric conditions.

Annual and seasonal mass balance measurements are commonly acquired using direct glaciological methods, which consist of nets of snow probes, ablation stakes, and snow pits distributed over the entire surface of the glacier (Godio, 2009; Zemp et al., 2009). Snow probe surveys insert stakes into the subsurface in order to estimate the thickness of the superficial snow layer, while ablation stakes are used to measure the thickness variations of glaciers and snowfields over a certain period of time. Snow pits (a.k.a. trenches) are excavated usually a few meters deep, depending on the overall compaction and thickness of the shallowest snowpack, and they are used to directly probe the free water content,

the metamorphism of the snow crystals, and the density of the glacier at different depths. Static penetration tests can also be performed in order to study the compaction and cohesion of the air-ice mixture (Godio, 2009). All these direct measurements are then interpolated in order to create a model of the entire glacier. Nevertheless, this manual approach has several disadvantages, specifically

1. It requires a significant amount of time and manpower, especially with regards to the placement of snow probes and ablation stakes, as well as the excavation of snow pits;
2. It is not always applicable, particularly in the case of glaciers situated in remote or unreachable locations, which are often characterized by rough topography;
3. It is normally limited to the shallowest part of the glacier, where the less dense snow can be more easily probed, as opposed to the deeper and more compact firn and ice;
4. It is unsuitable to study large areas, where the limited number of snow probes, ablation stakes, and snow pits usually results in low data densities, which prevent the accurate recovery of possible significant lateral thickness and density variations.

Zemp et al. (2009) list the following sources of both systematic and random error in the estimation of the mass balance of a glacier:

1. The accuracy of stake readings, snow probing, and density measurements of the air-ice mixture;
2. The spatial distribution and density of the stake and pit network;
3. The method used for the interpolation between the measurement points and the consequent extrapolation of values within unprobed areas, such as crevasses or debris-covered zones;
4. Special problems related to internal accumulation and calving;
5. Any changes in the glacier area that are usually not subject to annual measurement.

More accurate thickness variation measurements can be acquired through geodetic methods, which include photogrammetry, satellite image analysis, GPS measurements, and LiDAR surveys. As opposed to in-situ point measurements acquired over a changing glacier surface with the previously described direct methods, geodetic measurements repeated over time can record topographic changes over the entire surface of a glacier based on one or more fixed external points of reference (e.g. the surrounding bedrock), resulting in a more accurate estimation of the glacier's volumetric changes with respect to direct methods. Nevertheless, the mass balance cannot be accurately recovered without integrating density data with the reconstructed volume difference, while making assumptions on the internal density distribution of the glacier can lead to large inaccuracies in the resulting model. In fact, the internal density distribution is generally assumed constant or slow-varying, with a few constraints obtained by interpolating values sampled in localized snow pits, and which are therefore limited to the shallowest part of the glacier. However, the density distribution can show significant vertical and lateral variations, especially in small glaciers and glacierets (Colucci et al., 2015), and it is the result of two main density forming processes (Harper and Bradford, 2003; Godio, 2009):

1. The primary densification process develops at the surface during deposition, and it is influenced by atmospheric conditions like precipitation rate, air temperature, size and type of snow crystals, and wind action reworking the deposited crystals;

2. The secondary densification process develops after deposition, due to internal thermomechanical phenomena, which include the metamorphism, melting, and refreezing of the air-ice mixture over the entire glacial volume, with changing values of temperature and pressure.

In the last few decades, glaciological GPR surveys have increasingly been used to accurately estimate the internal stratigraphy, density distribution, total volume, and water content of glaciers and snowfields (Annan et al., 1994; Previati et al., 2011). The acquisition of GPR data sets is less time-consuming and requires less effort than the positioning and monitoring of ablation stakes or the excavation of snow pits, while also providing greater data densities. In particular, the large number of recorded traces, covering the entire surface of a glacier, makes their quantitative analysis statistically sound. These advantages are even more apparent when considering airborne GPR surveys, which can cover large areas regardless of the surface topography.

Glaciers and snowfields are particularly favorable with respect to GPR surveys, since non-magnetic and low-conductive media like air-ice mixtures allow the signal to propagate efficiently throughout their entire volume, with the remaining main limitation being amplitude decay caused by wavefront expansion, assuming that the internal liquid water content can be considered negligible. In these conditions Eq. 1.6 is significantly simplified, with the relative permeability μ_r being equal to 1 and the loss factor P becoming negligible, and the EM velocity v can therefore be directly linked to the relative permittivity ε_r of the analyzed medium through the following relation (Annan et al., 1994)

$$v \approx \frac{c}{\sqrt{\varepsilon_r}} \quad (3.1)$$

After reconstructing the internal EM velocity distribution of a glacier, Eq. 3.1 can be used to recover the relative permittivity distribution, which in turn can be connected to the density distribution through well-known empirical relations. In general, a glacier is a system characterized by three main phases, namely ice, air, and liquid water, besides possible internal debris, each making up a certain percentage of the total volume and contributing to its average internal properties. In the general case, the relative permittivity ε_r and the density ρ of a glacier are calculated as volumetric weighted averages of the relative permittivity and the density of its three main components, according to the Complex Refractive Index Model (CRIM), given by (Birchak, 1974)

$$\varepsilon_r = \left[\sqrt{\varepsilon_{r,\text{ice}}} (1 - \Phi) + \sqrt{\varepsilon_{r,w}} \Theta_W + \sqrt{\varepsilon_{r,\text{air}}} (\Phi - \Theta_W) \right]^2 \quad (3.2)$$

$$\rho = \rho_{\text{ice}} (1 - \Phi) + \rho_w \Theta_W \quad (3.3)$$

with

$$\varepsilon_{r,w} = 88 \quad \varepsilon_{r,\text{ice}} = 3.18 \quad \varepsilon_{r,\text{air}} = 1$$

$$\rho_w = 1000 \text{ kg/m}^3 \quad \rho_{\text{ice}} = 917 \text{ kg/m}^3 \quad \rho_{\text{air}} \approx 0 \text{ kg/m}^3$$

where Φ is the porosity of the air-ice mixture and Θ_W is the liquid water content, both expressed as percentages of the total volume. These quantities have to be known before applying Eqs. 3.2 and 3.3, which constitutes an important uncertainty factor.

In case the liquid water content and the porosity can be disregarded, the system can be approximated as made up of a single solid phase, and the relative permittivity can be directly linked to the density using Robin's equation (Robin, 1975; Kovacs, 1995)

AIR-ICE MIXTURE	$\rho \left[\frac{\text{kg}}{\text{m}^3} \right]$	ROBIN		LOOYENGA	
		ε_r	$v \left[\frac{\text{cm}}{\text{ns}} \right]$	ε_r	$v \left[\frac{\text{cm}}{\text{ns}} \right]$
Fresh snow	100	1.2	27.7	1.2	27.8
	200	1.4	25.7	1.3	25.9
	300	1.6	23.9	1.5	24.2
	400	1.7	22.4	1.8	22.7
Firn	500	2.0	21.1	2.0	21.3
	600	2.3	19.9	2.2	20.1
	700	2.5	18.9	2.5	18.9
Ice	800	2.8	17.9	2.8	17.9
	900	3.1	17.0	3.1	17.0

Table 3.1: Analysis of the EM velocities estimated for different air-ice mixtures, from fresh snow to pure ice, using Robin (Eq. 3.4) and Looyenga's (Eq. 3.5) empirical relations.

$$\varepsilon_r = (1 + 0.845\rho)^2 \quad (3.4)$$

where the density is expressed in g/cm^3 .

In case of low internal temperature and pressure, which allow the liquid water content to be disregarded, the glacier can be approximated with a two-phase system made of air-ice mixtures, and the relative permittivity is linked to the density using Looyenga's formula (Looyenga, 1965)

$$\varepsilon_r = \left[\frac{\rho}{\rho_{\text{ice}}} (\sqrt{\varepsilon_{r,\text{ice}}} - 1) + 1 \right]^3 \quad (3.5)$$

Considering a N-layered system, after calculating the densities ρ_n and thicknesses Δz_n of each layer, the corresponding water equivalent WE can be estimated, that is the thickness per unit area of the water layer resulting from the total melting of the air-ice mixture, given by (Lundberg et al., 2006)

$$\text{WE} = \sum_{n=1}^N \frac{\rho_n}{\rho_w} \Delta z_n \quad (3.6)$$

Table 3.1 shows the values of the relative permittivity and the EM velocity resulting from Eqs. 3.4 and 3.5, for density values ranging from fresh snow to pure ice. It can be noticed that a significant increase in density of the air-ice mixture, from 100 to 900 kg/m^3 , corresponds to a relatively smaller decrease in EM velocity, from 27.8 to 17.0 cm/ns . Therefore, an incorrect estimation of the EM velocity distribution within a glacier has a greater effect on the reconstructed density distribution, rather than the internal stratigraphy. Specifically, a change in the EM velocities corresponds to a greater change in the water equivalent than in the total glacier thickness, with fixed travel times.

1-D Modelling

In order to better highlight the effects that a changing EM velocity has on the estimated thickness and water equivalent of a glacier, consider a 1-D glaciological model consisting of a single homogeneous glacial layer over a semi-infinite base layer, and a monostatic GPR system, i.e. a system with a null transmitter-receiver offset, recording a single basal reflection with a certain two-way travel time. The

TWT [ns]	EM VELOCITY [$\frac{\text{cm}}{\text{ns}}$]													
	17.0	19.0	21.0	23.0	25.0	27.0	Δv	17.0	19.0	21.0	23.0	25.0	27.0	
	THICKNESS [m]							WATER EQUIVALENT [m]						
100	8.5	9.5	10.5	11.5	12.5	13.5	0.1 0.2 0.5 1.0	7.63	6.59	5.49	4.34	3.15	1.91	
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$		$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	
	0.05	0.05	0.05	0.05	0.05	0.05		0.14	0.12	0.11	0.10	0.09	0.08	
	0.10	0.10	0.10	0.10	0.10	0.10		0.28	0.25	0.22	0.19	0.17	0.15	
	0.25	0.25	0.25	0.25	0.25	0.25		0.70	0.61	0.54	0.48	0.43	0.38	
	0.50	0.50	0.50	0.50	0.50	0.50		1.40	1.23	1.09	0.96	0.86	0.77	
200	17.0	19.0	21.0	23.0	25.0	27.0	0.1 0.2 0.5 1.0	15.25	13.18	10.99	8.69	6.30	3.83	
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$		$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	
	0.1	0.1	0.1	0.1	0.1	0.1		0.28	0.25	0.22	0.19	0.17	0.15	
	0.2	0.2	0.2	0.2	0.2	0.2		0.56	0.49	0.43	0.39	0.34	0.31	
	0.5	0.5	0.5	0.5	0.5	0.5		1.40	1.23	1.09	0.96	0.86	0.77	
	1.0	1.0	1.0	1.0	1.0	1.0		2.79	2.46	2.17	1.93	1.72	1.54	
300	25.5	28.5	31.5	34.5	37.5	40.5	0.1 0.2 0.5 1.0	22.88	19.77	16.48	13.03	9.45	5.74	
	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$		$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	$\pm$	
	0.15	0.15	0.15	0.15	0.15	0.15		0.42	0.37	0.33	0.29	0.26	0.23	
	0.30	0.30	0.30	0.30	0.30	0.30		0.84	0.74	0.65	0.58	0.52	0.46	
	0.75	0.75	0.75	0.75	0.75	0.75		2.10	1.84	1.63	1.45	1.29	1.51	
	1.50	1.50	1.50	1.50	1.50	1.50		4.19	3.68	3.26	2.89	2.58	2.30	

Table 3.2: Analysis of the effects of the uncertainty of the EM velocity on the estimated thickness and water equivalent of a single homogeneous glacial layer. The table shows three examples with different arrival times of the basal reflection, which are then combined with six different EM velocities in the glaciological range (Table 3.1) to calculate the total thickness and water equivalent of the analyzed layer. The analysis highlights how much an increase in the EM velocity uncertainty (Δv) affects the uncertainty of the inverted quantities. The densities of the air-ice mixture were estimated using Looyenga's formula (Eqs. 3.1 and 3.5), while the uncertainties were calculated by applying to each equation the propagation law for the maximum error.

uncertainty of the EM velocity has significant consequences on the estimation of the total thickness and water content of the analyzed layer, given the variability of the density of the air-ice mixture estimated from such EM velocity (Table 3.1).

The estimated thickness and water equivalent, as well as the corresponding uncertainties, are shown in Table 3.2 for three different arrival times of the basal reflection in the synthetic GPR data set, six different EM velocities assumed within the analyzed layer (i.e. six different densities, Table 3.1), and four different velocity uncertainties set for each case. The densities of the air-ice mixture were estimated using Looyenga's formula (Eqs. 3.1 and 3.5), while the uncertainties of the inversion results were calculated by applying to each equation the propagation law for the maximum error.

As expected, the uncertainty of the water equivalent is larger than that of the thickness, since it depends on the calculation of the electric permittivity (Eq. 3.1), the density (Eq. 3.5), as well as the layer thickness itself (Eq. 3.6). More importantly, a larger increase in the uncertainty of the water equivalent can be noticed with the increasing uncertainty of the EM velocity, since small EM velocity changes corresponds to relatively larger density variations (Table 3.1). The linearity between the estimated thickness and the EM velocity for fixed travel times, in the analyzed normal incidence

model, also causes the thickness uncertainty to be independent from the EM velocity, while it is linearly dependent on the uncertainty of the EM velocity. On the other hand, Table 3.2 shows large variations in the uncertainty of the water equivalent with different EM velocities, given the corresponding large changes in the estimated densities of the air-ice mixture within the layer (Table 3.1).

3.1.2 Velocity analysis

The accurate estimation of the subsurface velocity distribution is very important for several applications, including signal processing, data imaging, and both qualitative and quantitative analysis of the recorded data set (Jol, 2009). Signal processing and data imaging procedures include time-depth conversion, which recovers the internal stratigraphy of the subsurface; amplitude recovery, which is essential for any amplitude-related quantitative analysis; and migration, which is used to correct the spatial position of dipping and curved reflectors, as well as to collapse hyperbolic diffractions at their true scatterer location. Examples of quantitative analysis include the aforementioned density distribution estimation from glaciological surveys, given the previously discussed relation between the EM velocity and the density of air-ice mixtures (Table 3.1). Another example is the analysis of soil moisture content, given the stark difference in the EM response of water with respect to other earth materials within the soil, particularly due to its high permittivity and subsequent low EM velocity (Table 1.1). This contrast can cause small variations in pore water content to greatly change the EM properties of the soil.

There are several methods commonly used to extract the velocity distribution from a GPR data set, which differ in terms of their acquisition geometry, data density, and accuracy of the results.

Common Mid-Point surveys

The estimation of the subsurface EM velocity distribution from CMP GPR profiles is commonly based on the Normal Move-Out (NMO) correction analysis, which searches for a velocity distribution that can be used to correctly flatten all the recorded hyperbolic reflections. Specifically, given the thicknesses h_i and the EM velocities v_i of the first n layers in a 1-D system, the arrival time twt_n of the signal reflected by the n -th interface is approximated by the following equation (Tillard and Dubois, 1995)

$$\text{twt}_n = \sqrt{\frac{x^2}{\bar{v}_n^2} + t_n^2} \quad (3.7)$$

with

$$\bar{v}_n^2 = \frac{\sum_{i=1}^n v_i h_i}{\sum_{i=1}^n \frac{h_i}{v_i}}$$

$$t_n = 2 \sum_{i=1}^n \frac{h_i}{v_i}$$

where t_n is the two-way travel time in the normal incidence case (i.e. the zero-offset case), and $\bar{v}_n$ is the root mean square velocity of the first n layers.

The average EM velocities between the surface and the various planar reflectors can be recovered by fitting Eq. 3.7 to the series of reflection hyperbolas in a CMP data set. These average values can then be transformed into a 1-D velocity model consisting of a vertical step function, which divides the

subsurface into various layers with constant EM velocities. The velocity analysis can be done either manually or through processes such as the semblance analysis (Jol, 2009), and the accuracy of the constructed velocity model can be verified through a NMO analysis.

Nevertheless, the accuracy of the velocity model recovered through the analysis of CMP data sets can be sensitive to the following factors:

1. Limited values of the offset. The approximation of a 1-D system is acceptable for small values of the transmitter-receiver offset, while on a larger scale the non-homogeneity of the layers and the lateral changes in the stratigraphy can significantly distort the shape of the reflections with respect to the ideal hyperbolic case (Eq. 3.7). On the other hand, velocity estimations are more accurate at large offsets, in which the reflection curvature is better defined.
2. Presence of dipping reflectors. In case of a dipping reflector, fitting Eq. 3.7 to the reflection hyperbola would lead to an overestimation of the average propagation velocity, unless appropriate corrections are applied (Tillard and Dubois, 1995). In this case, a geometrical term containing the reflector dip, which is usually not known in advance, should be added to Eq. 3.7.

In terms of accuracy, Jol (2009) states that the CMP analysis tends to produce approximate velocity values, with errors and variance equal to at least 10%, thus justifying the use of either constant or slow-varying EM velocity distributions for interpretation purposes. Moreover, Tillard and Dubois (1995) state that repetitive field measurements could be useful in reducing calculation uncertainties, however they also state that the procedure is insufficient to ensure errors of less than 10%, when calculating EM velocities using the NMO correction analysis (Forte et al., 2013). In glaciological surveys, considering the large changes in the calculated densities of air-ice mixtures with respect to relatively smaller EM velocity variations (Table 3.1), such velocity uncertainties can lead to significant errors when estimating the density distribution and water content of glaciers (Table 3.2).

Common Offset surveys

The estimation of the EM velocity distribution from CO GPR profiles is commonly based on the analysis of hyperbolic diffractions caused by the presence of scatterers within the subsurface. The simplest method is hyperbola matching, in which an hyperbolic function is directly fitted to the recorded travel times of the scattering hyperbolas, whose opening angles depend on the EM velocity distribution above each scatterer. Specifically, given the horizontal position x and the one-way travel time t measured at a given point of the fitting hyperbola, the average EM velocity between the surface and the analyzed diffractor is given by the following equation (Moore et al., 1999)

$$\bar{v} = \frac{x - x_0}{\sqrt{t - t_0}} \quad (3.8)$$

where x_0 and t_0 are respectively the horizontal position and the one-way travel time measured at the top of the hyperbola. The transmitter-receiver offset is assumed null in Eq. 3.8, nevertheless the equation can be considered a good approximation for large depth-to-offset ratios (Moore et al., 1999).

The presence of point scatterers at different depths can be used to construct a model of the EM velocity distribution within the recorded profile, by assigning different velocities to the various layers between diffractors and then adapting these velocities to the stratigraphy as defined by the recorded reflections. Nevertheless, the accuracy of the velocity model recovered through the hyperbola matching method can be sensitive to the following factors:

1. Distortions of the diffraction hyperbolas. If a scattering object cannot be approximated as point-like due to elongation in a particular direction, the simple scattering assumed in Eq. 3.8 is no longer valid and distortions caused by a combination of diffractions and reflections would produce an erroneous velocity estimate. Also, the EM velocity calculated with Eq. 3.8 is an average of the vertical velocity changes above a diffractor under the assumption of a 1-D system, while any lateral variability cannot be resolved and will appear as distortions in the hyperbolic shape, depending on the depth-to-offset ratio (Moore et al., 1999). Moreover, unless 3-D data sets are available, the hyperbola matching method assumes that the scatterers lie directly beneath the survey line, therefore any diffraction hyperbola generated by offline scatterers would result in erroneous velocity estimates (Murray et al., 2007).
2. Horizontal width of the hyperbolas. Depending on the total energy scattered, the depth of the diffractor, and the spatial trace interval used during the survey, the number of recorded traces in which the diffraction hyperbola is clearly defined may be limited. In this case, the function used for hyperbolic matching would be fitted to only a few traces around the diffractor, resulting in greater uncertainties of the calculated EM velocities (Eq. 3.8), due to the large number of acceptable hyperbolas with different opening angles matching the data.
3. Distribution of the point scatterers within the subsurface. The accuracy of the constructed velocity model depends on the actual presence and regular distribution of point scatterers all over the subsurface producing undistorted and spatially wide hyperbolas. Moreover, the resulting EM velocity distribution must be adapted to the stratigraphy derived from the recorded reflections, which mark clear impedance and EM velocity contrasts. While hyperbola matching can be the only available velocity analysis technique for some urban environments, where no clearly defined reflection may be present, in the case of glaciological GPR surveys the method will not necessarily provide accurate EM velocity estimates over the entire glacier volume, in which scattering debris are not uniformly present (Bradford and Harper, 2005).

Similarly to the aforementioned CMP analysis, Jol (2009) states that the approximate velocity values estimated using hyperbolic matching have an associated error equal to at least 10%, thus justifying the use of either constant or slow-varying EM velocity distributions for interpretation purposes. As previously discussed, such velocity uncertainty can lead to significant errors when calculating the density distribution and water content of glaciers (Table 3.1).

Another possible method that can be used to recover the EM velocity distribution from CO GPR surveys involves migration analysis. In signal processing, migration uses as input the velocity distribution in order to place the recorded energy at its correct point of origin within the profile. Specifically, the procedure collapses the hyperbolic diffractions and moves dipping reflections at their correct spatial position, in order to accurately reconstruct the subsurface stratigraphy. Migration velocity analysis is a technique that recovers the subsurface EM velocity by searching for the velocity distribution that most accurately collapses diffractions and maximizes coherence along complexly dipping reflections within the recorded profile (Bradford and Harper, 2005). As opposed to hyperbolic matching, the migration analysis does not depend on the assumption of a 1-D system, therefore the estimated EM velocity distribution can directly take into account the subsurface stratigraphy as defined by the recorded reflections. Nevertheless, similarly to hyperbolic matching, the accuracy of the constructed velocity model depends on the actual presence and regular distribution of undistorted and clearly defined diffractions within the recorded profile, while the point-like scatterers need to be located directly beneath the survey line in the case of 2-D GPR data sets.

3.2 Inversion algorithm

The presented inversion algorithm iteratively recovers the thickness and EM velocity in each layer in a GPR trace, by reconstructing the travel path of each reflected wavelet. The procedure is independently applied to each trace in a CO GPR data set, therefore the trace index, which was used to describe the auto-picking procedure in chapter 2, is dropped in this section from each symbol, while the remaining index indicates either the reflection, the layer, or the interface number, depending on the specific quantity considered. The inversion algorithm makes the following assumptions with regards to the analyzed subsurface and the propagation of the reflected wavelets:

1. The GPR signal propagates as a plane wave, in particular the inversion algorithm is based on the principles of geometrical ray theory.
2. The reflectors are approximated as plane parallel (i.e. 1-D system, Fig. 3.1) in the vicinity of each GPR trace position. This assumption is necessary, since the subsurface stratigraphy is not usually known in advance, and it can be applied in most subsurface conditions, especially for reflections with high depth-to-offset ratios. Specifically, in most cases the incident angle can be considered small, and at small incident angles the reflection coefficients do not show significant changes, independently of the specific EM impedance contrasts, therefore this approximation does not significantly affect the inversion results (Forte et al., 2014a).
3. Each layer is considered isotropic, homogeneous, lossless, and non-dispersive. These conditions do not exactly represent average earth materials, but they are acceptable approximations of subsurface environments which are favorable to EM wave propagation, and therefore to GPR surveys, such as glaciers and snowfields.
4. The reflection amplitudes depend only on the subsurface EM impedance contrasts, while the amplitude decay caused by processes like intrinsic attenuation, signal scattering, and wavefront expansion is either corrected or disregarded. In particular, wavefront expansion can be corrected either before the inversion, by applying to the data set a gain function based on the available information regarding the subsurface, or during the inversion process itself, by using the iteratively inverted EM velocities to calculate the distance traveled by the reflected wavelets.

The inversion algorithm requires the following input data:

1. The transmitter-receiver offset, which is known.
2. The EM velocity in the shallowest layer, which can be estimated through either CMP analysis or, in the case of glaciological GPR surveys, direct density measurements (Table 3.1).
3. The peak amplitude of the incident wavelet at the first interface, also referred to in the following as the reference amplitude, which can be measured by separately recording the wavelet initially transmitted into the ground.
4. The peak amplitudes and arrival times of the reflected wavelets, which can be automatically picked using the previously described procedure.

The inversion algorithm can theoretically be applied to both ground-based and airborne GPR surveys, with the difference that in the latter case the shallowest layer is given by the air space between the antennas and the ground surface. The general principle of the inversion process is that, given the

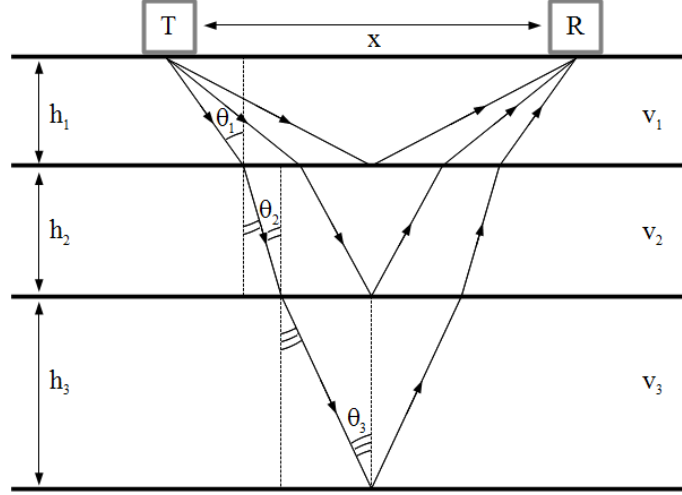


Figure 3.1: Not-to-scale model showing the assumptions made by the inversion algorithm with regards to the subsurface and the propagation of the GPR signal. The symbols indicate the transmitter T, the receiver R, the offset x , the incident angles θ_i for the third reflection, and the thicknesses h_i and EM velocities v_i of each layer. Modified image from Forte et al. (2014a).

thicknesses of the first $n-1$ layers and the EM velocities in the first n layers, the n -th cycle recovers the thickness of the n -th layers and the EM velocity in the $(n+1)$ -th layer, by reconstructing the travel path of the n -th reflection.

In the first cycle the algorithm analyzes the signal reflected by the first interface, and the calculation is simplified by the fact that the EM velocity in the first layer has to be known. In this case, the thickness h_1 of the first layer, the incident angle θ_1 , and the reflection coefficient R_1 at the first interface are given by the following equations:

$$h_1 = \frac{1}{2} \sqrt{(v_1 \text{twt}_1)^2 - x^2} \quad (3.9)$$

$$\theta_1 = \arctan \left(\frac{x}{2h_1} \right) \quad (3.10)$$

$$R_1 = \frac{A_1}{A_{i1}} \quad (3.11)$$

where twt_1 is the two-way travel time of the first reflection, and A_{i1} and A_1 are the incident and reflected amplitudes at the first interface, respectively. The EM velocity v_2 in the second layer is then calculated using the Snell equation:

$$v_2 = \frac{\sin(\theta_2)}{\sin(\theta_1)} v_1 \quad (3.12)$$

with

$$\theta_2 = \arctan \left[\frac{1 + R_1}{1 - R_1} \tan(\theta_1) \right] \quad (3.13)$$

In the n -th cycle, the algorithm analyzes the signal reflected by the n -th interface, whose arrival time $\text{tw}t_n$ is known, while the thicknesses h_i of the first $n-1$ layers and the EM velocities v_i in the first n layers are also known. The thickness h_n of the n -th layer is thus calculated as the only positive solution of the following third-degree equation, which is obtained by rearranging the two-way travel time equation for the n -th reflection in a 1-D system (Eq. 3.7):

$$ah_n^3 + bh_n^2 + ch_n + d = 0 \quad (3.14)$$

with

$$a = \frac{4}{v_n} \quad (3.15)$$

$$b = \frac{4}{v_n^2} \sum_{i=1}^{n-1} v_i h_i + 8 \sum_{i=1}^{n-1} \frac{h_i}{v_i} \quad (3.16)$$

$$c = \frac{x^2}{v_n} + \frac{8}{v_n} \left(\sum_{i=1}^{n-1} v_i h_i \right) \left(\sum_{i=1}^{n-1} \frac{h_i}{v_i} \right) + 4v_n \left(\sum_{i=1}^{n-1} \frac{h_i}{v_i} \right)^2 - v_n \text{tw}t_n^2 \quad (3.17)$$

$$d = x^2 \sum_{i=1}^{n-1} \frac{h_i}{v_i} + 4 \left(\sum_{i=1}^{n-1} v_i h_i \right) \left(\sum_{i=1}^{n-1} \frac{h_i}{v_i} \right)^2 - \text{tw}t_n^2 \sum_{i=1}^{n-1} v_i h_i \quad (3.18)$$

Given the assumptions of plane parallel reflectors and homogeneous layers, the travel path of the n -th reflection within the subsurface is geometrically fixed and symmetric with respect to the mid-axis between the transmitter and the receiver (Fig. 3.1). In particular, the sum of the horizontal projections of the travel path of the GPR signal in each layer is equal to the transmitter-receiver offset, while the incident angle at the k -th interface and the transmitted angle at the $(k-1)$ -th interface are equal in a 1-D system (Fig. 3.1), and they are small for high depth-to-offset ratios. After applying these geometrical constraints to the Snell equation, the incident angles θ_k at all the interfaces for the n -th reflection are given by the following equation (Forte et al., 2014a):

$$\theta_k = \arctan \left(\frac{xv_k}{2 \sum_{i=1}^n v_i h_i} \right) \quad (3.19)$$

with

$$k = 1, 2, \dots, n$$

Notice that Eq. 3.19 is a generalization of Eq. 3.10, used in the first cycle.

The reflection R_k and transmission T_k coefficients for the first $n-1$ interfaces in the n -th reflection are then calculated using the Fresnel equations in the TE antenna configuration:

$$R_k = \frac{\sin(\theta_{k+1} - \theta_k)}{\sin(\theta_{k+1} + \theta_k)} \quad (3.20)$$

$$T_k = 1 + R_k \quad (3.21)$$

with

$$k = 1, 2, \dots, n-1$$

The reflection coefficient R_n for the n -th interface is then calculated from the incident A_{i_n} and reflected A_{r_n} amplitudes at that interface, which are recovered from the reference amplitude A_{i_1} and the amplitude A_n of the n -th recorded reflection using the symmetry of the travel-path (Fig. 3.1):

$$R_n = \frac{A_{r_n}}{A_{i_n}} \quad (3.22)$$

with

$$A_{i_n} = A_{i_1} \prod_{i=1}^{n-1} T_i \quad (3.23)$$

$$A_{r_n} = \frac{A_n}{\prod_{i=1}^{n-1} (2 - T_i)} \quad (3.24)$$

After calculating the reflection coefficient for the n -th interface, the EM velocity v_{n+1} in the $(n+1)$ -th layer is calculated using the Snell equation:

$$v_{n+1} = \frac{\sin(\theta_{n+1})}{\sin(\theta_n)} v_n \quad (3.25)$$

with

$$\theta_{n+1} = \arctan \left[\frac{1 + R_n}{1 - R_n} \tan(\theta_n) \right] \quad (3.26)$$

Notice that Eqs. 3.25 and 3.26 are generalizations of Eqs. 3.12 and 3.13, used in the first cycle.

By iterating the inversion process for all the picked reflections in all the recorded traces in a CO GPR data set, the algorithm can recover the stratigraphy and EM velocity distribution within the analyzed subsurface. These quantities can then be used in glaciological GPR surveys to recover the total volume, density distribution, and water content of glaciers and snowfields.

3.2.1 Amplitude recovery

An important assumption made by the inversion algorithm is that the input reflection amplitudes depend only on the subsurface EM impedance contrasts, while the amplitude decay caused by processes like intrinsic attenuation, signal scattering, and wavefront expansion is either corrected or disregarded. One of the advantages of the auto-picking algorithm described in chapter 2 is that it does not depend on the signal amplitudes, and therefore it can be implemented without the need of any pre-applied amplitude recovery. The algorithm can thus pick the original reflection amplitudes, without any subjective distortion caused by the assumptions made by the interpreter with regards to the EM properties of the subsurface, in order to counteract the amplitude decay. For example, spreading losses caused by wavefront expansion can be corrected by using the subsurface EM velocity distribution to calculate the distance traveled by the radar signal. However, this information is not usually known in advance and it is the very aim of the described inversion process. Any subjective alteration of the reflection amplitudes would significantly affect the calculated EM impedance contrasts (Eqs. 3.11 and 3.22), and influence the estimated subsurface stratigraphy and EM velocity distribution.

On the other hand, the spreading losses can be objectively corrected during the inversion process itself, by using the EM velocities which are iteratively calculated for each layer. In the general case of the n -th cycle, the amplitude A_n of the n -th reflection can be automatically corrected for the wavefront

expansion by using the arrival times twt_k of the first n reflections and the estimated EM velocities v_k of the first n layers to recover the distance r traveled by the wavelet. Specifically, assuming a $1/r$ amplitude decay with distance, the amplitude A_n in Eq. 3.24 would be substituted with the corrected amplitude A_{c_n} (Dossi et al., 2015a; 2016):

$$A_{c_n} = A_n (1 + r) \quad (3.27)$$

with

$$r = \sum_{k=1}^n v_k (\text{twt}_k - \text{twt}_{k-1})$$

$$\text{twt}_0 = 0$$

In case intrinsic attenuation or signal scattering within the subsurface cannot be disregarded, it is necessary to apply an appropriate gain function to the data set. This correction does introduce a subjectivity factor, since it would depend on the interpreter's experience, as well as the available information regarding the EM properties of the subsurface. If the dependance of the attenuation on the traveled distance was known, the gain function could be inserted into the inversion process as well.

3.2.2 Radiation Pattern

The reference amplitude is the peak amplitude of the incident wavelet at the first interface, and it is necessary in the inversion algorithm in order to reconstruct the subsurface EM impedance contrasts by comparing it with the reflection amplitudes (Eqs. 3.11 and 3.22). Assuming a negligible intrinsic attenuation within the subsurface, the reference amplitude corresponds to the peak amplitude of the wavelet initially transmitted into the ground by the GPR system.

In this study, the reference amplitude was separately measured by recording the peak amplitude of a single reflection from a vertical 3×3 m metallic surface, propagating in air. The signal was recorded using a ProEx GPR system equipped with 250 MHz bistatic Malå Geoscience shielded antennas, with a 31 cm offset parallel to the metallic surface. The same system was used to acquire the GPR profiles analyzed in section 3.4 and in chapter 5, therefore the reference amplitude estimated hereafter was subsequently used in the inversion of the analyzed field data sets.

In this analysis, four different data sets were recorded, each one with the GPR system being rotated by 90° , in order to study possible variations in the radiation pattern and to remove artifacts. In each series, the GPR system was gradually moved closer to the metallic surface along its normal axis, starting at a distance of about 9 m. The transmitter was automatically triggered every 0.5 s, resulting in four GPR profiles containing a single reflection from the air-metal interface, which can be automatically picked in order to analyze the amplitude decay with distance. The four profiles are shown in Fig. 3.2, each with superimposed the picked main reflection phase used for the analysis.

Assuming a total reflection from the metallic surface, the peak reflected amplitudes were used to recover the reference amplitude A_0 by fitting the following function, which assumes a $1/r$ amplitude decay with distance, similarly to Eq. 3.27:

$$A(r) = \frac{A_0}{1 + r} + B \quad (3.28)$$

Ideally, the value of B in Eq. 3.28 should be either null or negligible, so that the theoretical amplitude $A(0)$ of the wavelet first emitted by the transmitter is indeed equal to A_0 . The picked

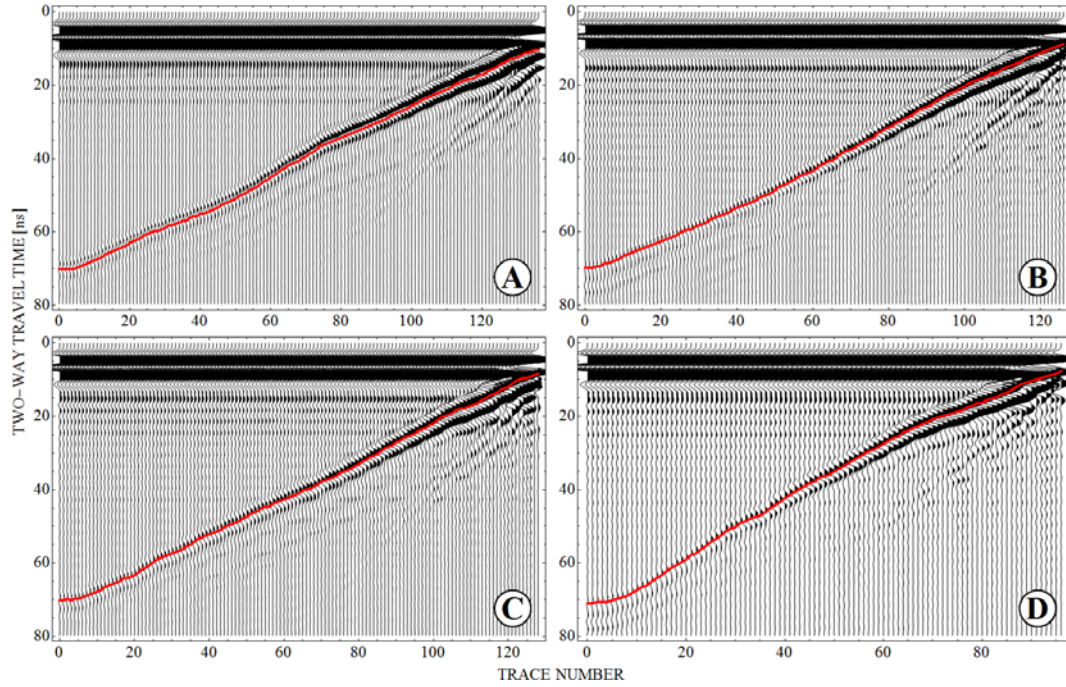


Figure 3.2: Reflections from a single metallic surface recorded with a 250 MHz GPR system located at varying distances in air. The antennas were kept parallel to the surface, while being rotated by 90° in each graph, in order to analyze possible variations in the radiation pattern and to remove artifacts. The peak amplitudes of the main phase of each reflection are marked in each graph, and they are used in Fig. 3.3 to study the signal amplitude decay with the distance traveled from the transmitter.

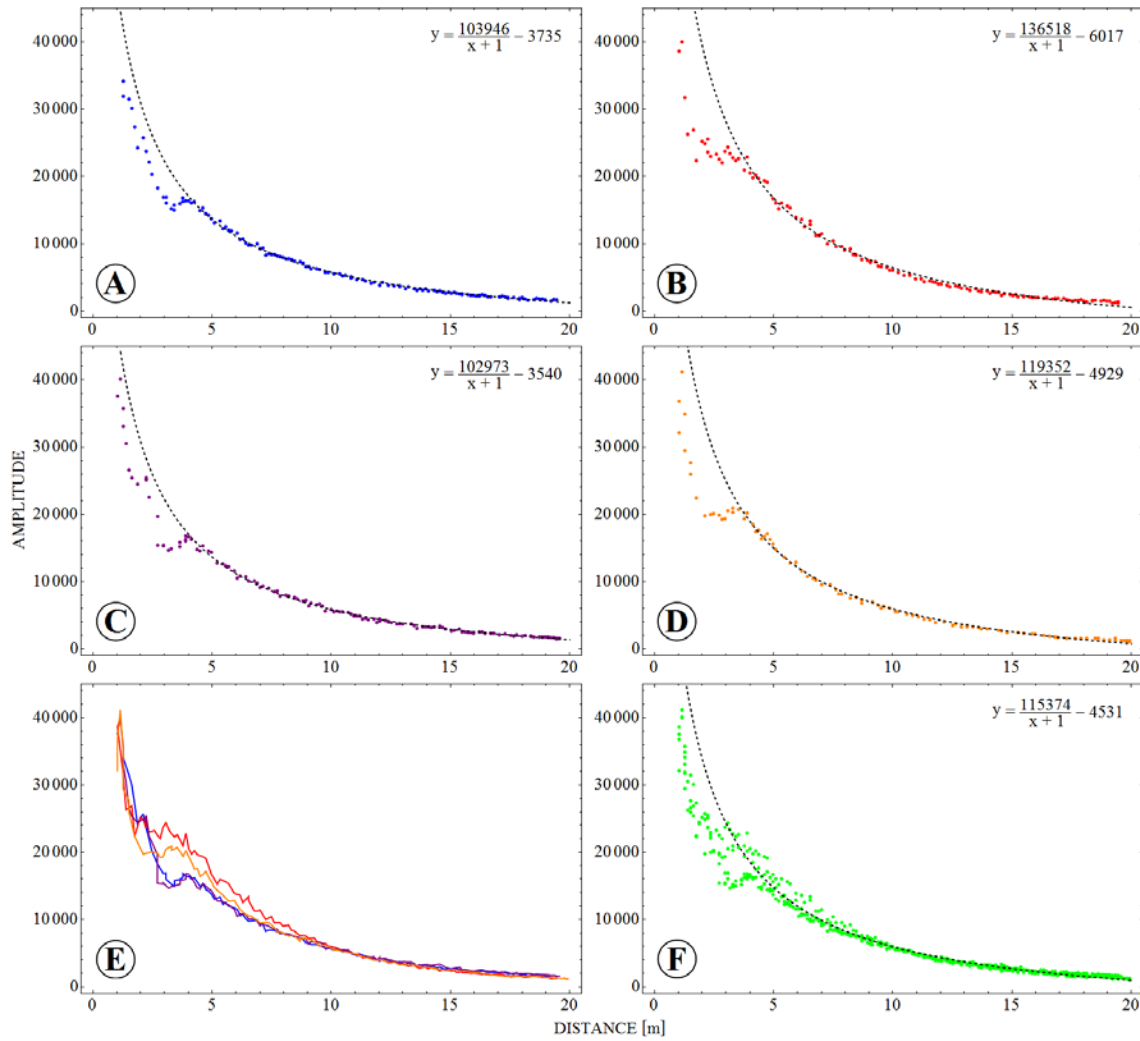


Figure 3.3: Analysis of the amplitude decay of a GPR signal with respect to the distance traveled from the transmitter. The amplitudes picked in Fig. 3.2 and the resulting fitted functions are plotted in modulus in each graph, showing the dependance of the peak wavelet amplitude on the traveled distance, assuming a total reflection from the metallic surface. The antennas are kept parallel to the surface, while being rotated by 90° degrees in each graph from (A) to (D), and the four series are compared in (E) in order to highlight possible variations in the radiation pattern. The four data sets are combined in (F), showing the corresponding fitted function. In all cases the functions are fitted to the amplitudes sampled at distances greater than 4 m, in order to disregard the erratic behavior at smaller distances, caused by the interference between the reflected wavelet and the airwave (Fig. 3.2).

FITTING PARAMETERS	SERIES				
	1	2	3	4	All
A_0	103946	136518	102973	119352	115374
B	-3735	-6017	-3540	-4929	-4531
$ B/A_0 $ [%]	3.6	4.4	3.4	4.1	3.9
$ \Delta A_0/A_{0All} $ [%]	9.9	18.3	10.7	3.4	0

Table 3.3: Analysis of the fitting parameters obtained from the amplitude decays shown in Fig. 3.3. The function given in Eq. 3.28 show an overall good fit to the picked amplitudes, with the estimated values of B being consistently lower than 5% the values of A_0 in all cases. The analysis also shows an average 10% variability of the values of A_0 estimated in each series, with respect to the value A_{0All} obtained when combining all the data sets (Fig. 3.3F).

amplitudes and the resulting fitted equations are plotted in Fig. 3.3 as functions of the distance traveled by the reflected signal from the transmitter, which is calculated by combining the two-way travel time of the picked reflection with the EM velocity in air (i.e. about 30 cm/ns). Each series shows an asymptotic $1/r$ decay with distance, which accurately fits Eq. 3.28, while the estimated values of B are consistently lower than 5% the values of A_0 in all cases (Table 3.3).

The picked amplitudes in Fig. 3.3 show an erratic behavior at distances lower than 4 m, due to the interference between the reflected wavelets and the airwave, which can be either constructive or destructive, depending on the arrival time of the reflection in each recorded trace. This interference can also be noticed in Fig. 3.2, where the 4 m distance threshold corresponds to a two-way travel time equal to about 13.3 ns, with the distance between the GPR system and the metallic surface being equal to about 2 m. Therefore, the fitting analysis in Fig. 3.3 was limited to the amplitudes picked at distances greater than 4 m. The subsequent lack of viable amplitude data at smaller distances, combined with residual noise and possible minor changes in the radiation pattern between each series, results in an average 10% variability of the value A_0 calculated in each series, with respect to the value A_{0All} obtained when combining all the data sets (Table 3.3).

The uncertainty of the reference amplitude has to be taken into account when it is being used as input of the inversion procedure. Nevertheless, the analyzed series do not show significant departures from the approximated $1/r$ amplitude decay with distance, nor do they show major changes in the radiation pattern when the antennas are rotated (Fig. 3.3E). The resulting reference amplitude to be used in the inversion of the field GPR data sets in section 3.4 and in chapter 5 was therefore set equal to 120000, taking into account the average estimated values of A_0 and B (Table 3.3). However, before being able to compare the measured reference amplitude with the peak reflection amplitudes acquired in the ground-based GPR surveys, it is necessary to take into account the differences in radiation pattern between the two signals, which respectively travel through the air and the ground, before calculating the subsurface EM impedance contrasts. The change in directivity between the signal transmitted by an air-coupled GPR system and the signal transmitted by a ground-coupled one is caused by the refractive focusing associated with the air-ground interface (Jol, 2009).

The generic radiation pattern of a GPR signal transmitted into the ground is divided into a main central lobe and several minor lateral lobes, with most of the energy contained within a cone defined by the critical angle θ_C of the radiation pattern, which is given by the following equation (Annan and Cosway, 1992; Jiao et al., 2000)

$$\theta_C = \arcsin \left[\frac{1}{\sqrt{\epsilon_r}} \right] \quad (3.29)$$

The signal wavefront can therefore be approximated with an expanding spherical cap defined by θ_C , and characterized by a surface area S equal to

$$S = 2\pi r^2 (1 - \cos[\theta_C]) \quad (3.30)$$

When measuring the reference amplitude (Fig. 3.3), the wavefront of the transmitted signal can be approximated with an expanding hemisphere. In fact, the relative permittivity ϵ_r of air is equal to 1, thus resulting in a value of θ_C equal to 90° (Eq. 3.29) and a wavefront surface area equal to $2\pi r^2$ (Eq. 3.30). Assuming that the same amount of energy within the hemispherical wavefront of the reference signal is concentrated within the spherical cap transmitted into the ground during the surveys, the reference amplitude is obtained by multiplying the estimated amplitude A_0 by a factor $1/\sqrt{1 - \cos[\theta_C]}$, with A_0 already being corrected for the spreading losses of the signal traveling through the air (Eq. 3.28) during the previously described measurements (Fig. 3.3). Although approximate, this correction is necessary in order to account for the differences in radiation pattern between the reference signal recorded in air and the signal transmitted into the ground, and to avoid overcompensating for the spreading losses of the reflected wavelets with respect to those of the reference wavelet.

3.3 Synthetic data analysis

The automated picking and inversion algorithm was applied to a synthetic CO GPR data set, which was constructed using the GPRMax 2-D software (Giannopoulos, 2005). The original model consists of seven layers, with both horizontal and dipping reflectors, above a semi-infinite base layer. Each layer is homogeneous, lossless, non-magnetic, and non-dispersive, with the stratigraphy and EM velocity distribution given in Fig. 3.4A. The velocity distribution shows values typical of a glacier above a rocky basement (Table 3.1), and it is characterized by both increasing and decreasing velocity contrasts.

The simulated survey consists of a ground-coupled 400 MHz bistatic GPR system, which transmits Ricker wavelets with a constant 70 cm offset and a 27.8 cm trace interval. The constructed GPR data set is shown in Fig. 3.4B, with superimposed the horizons selected for the amplitude inversion among those automatically picked, each marking the main phase of their respective reflections (i.e. the second phase for a Ricker wavelet). The simulated velocity contrasts result in both positive and negative polarities in the horizons to be used in the inversion, which constitutes a further test for the inversion algorithm. The GPR profile also contains isolated hyperbolic diffractions caused by the abrupt polygonal boundaries of the designed layers (Fig. 3.4A). These diffractions can interfere with the recorded reflections and alter the picked amplitudes and travel times in a few traces of the picked horizons. This is useful to study the effect of interference on the inversion results, similarly to what can be expected in real data sets. For continuity reasons, possible small gaps within the automatically picked horizons resulting from signal distortions caused by either noise or interference, such as the 1 trace-wide gap in the deepest reflection in Fig. 3.4B, can be automatically filled before the inversion by interpolating the amplitudes and travel times picked at both ends of each gap.

With regards to the input inversion parameters, the EM velocity in the shallowest layer and the transmitter-receiver offset are known by construction, the reflected amplitudes and travel times were automatically picked using the previously described algorithm, while the reference amplitude was recovered by simulating the same GPR system with both transmitter and receiver located in the same point in void space. In order to study the effects that the input EM velocity has on the inversion results, the uncertainty associated with the EM velocity in the shallowest layer is set equal to 0.1 cm/ns, while all the other input quantities have uncertainties equal to the last significant digit of

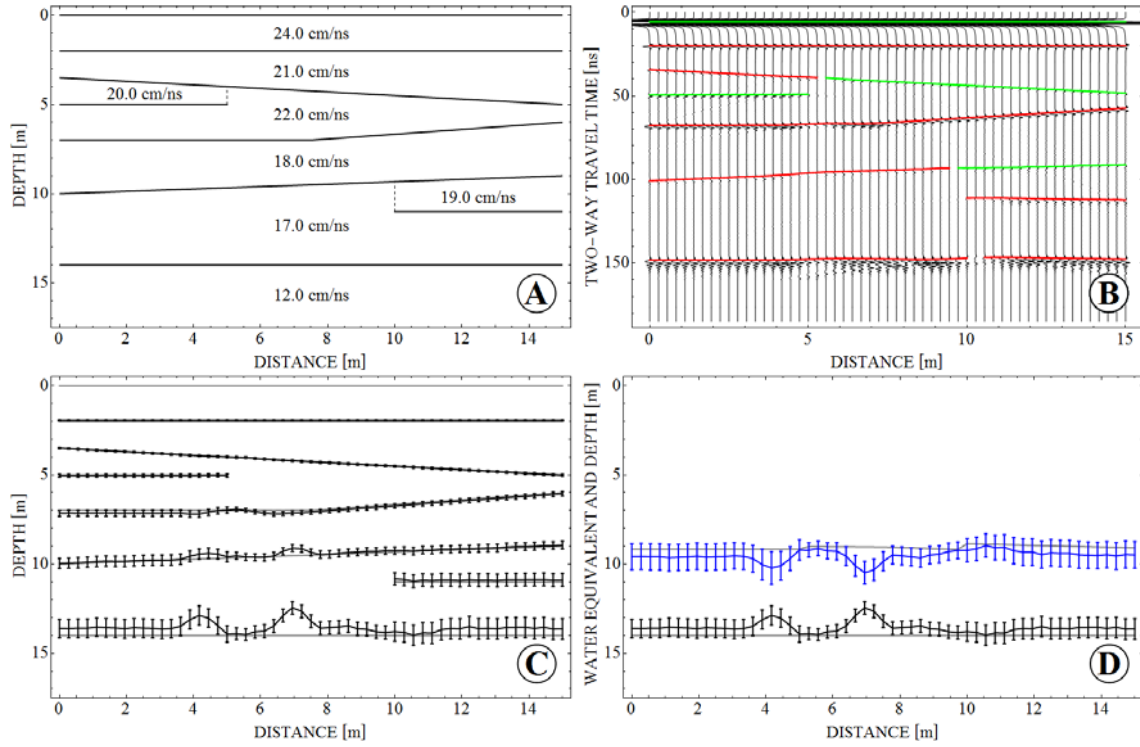


Figure 3.4: Analysis of a synthetic CO GPR data set, showing (A) the stratigraphy and the EM velocity distribution of the model; (B) the constructed GPR profile, with superimposed the horizons used for the analysis; (C) the inverted stratigraphy (black), compared with the initial model (gray); and (D) the recovered total thickness (black) and water equivalent (blue), compared with the corresponding theoretical values (gray). Positive amplitudes in (B) are marked in green, negative amplitudes in red. The synthetic profile was constructed using the GPRMax 2-D software (Giannopoulos, 2005).

their respective values. The effects that the input reflected amplitudes have on the inversion results are instead discussed in chapter 4, which analyzes the relation between peak amplitude error and the sampling rate selected during data acquisition.

The inverted stratigraphy is compared to the stratigraphy of the initial model in Fig. 3.4C, showing an overall good consistency for the first six interfaces, while also slightly different results for the deepest one. Specifically, a few instances can be noticed in which the aforementioned hyperbolic diffractions cause evident distortions in the inverted stratigraphy already in the fourth interface, with the effects being iteratively propagated to the deeper layers and also being added to other distortions caused by deeper interfering diffractions. These propagating distortions created the two main bulges in the deepest interface, respectively located at 4 and 7 m (Fig. 3.4C).

Another important discrepancy between the simulated and reconstructed deepest interface is an evident constant difference, which makes the reconstructed deeper layer about 30 cm shorter on average. While this difference is clearly visible in the deepest interface, the other interfaces show an overall good correspondence with the inversion results (Fig. 3.4C). Therefore, the discrepancy in the deepest interface is most likely caused by a general underestimation of the EM velocities within the deepest layer, combined with the larger time interval in which the deepest layer is defined (Fig. 3.4B) with respect to the shallower layers, which increases the resulting error. The main uncertainty factors during the inversion are the input EM velocity and the reference and reflected amplitudes, with the initial errors iteratively propagating to the deeper layers. However, in this case the EM velocity of the shallowest layer is exactly known by construction, which suggests the main cause of the velocity underestimation within the deepest layer to be the input amplitudes producing an overestimation of the EM impedance contrast in the upper interface of the layer. Such overestimation can be caused by an underestimation of the reference amplitude, as well as by an overestimation of the reflected amplitudes, possibly caused by either diffraction-reflection interference or by an overcompensating amplitude recovery function (Eq. 3.27). In the latter case, possible differences between the theoretical and simulated amplitude decay would have a greater effect in the deeper layers. Moreover, the large number of layers makes the deeper layers more susceptible to error propagation, so that slight errors in the EM velocity estimated in the shallower layers can build up and combine with the overestimated EM impedance contrast to produce the observed results.

The estimated water equivalent and total thickness of the simulated glacier are shown in Fig. 3.4D, and compared to their respective theoretical values. Besides the previously described causes of errors in the calculation of the total glacier thickness, the uncertainty of the water equivalent also greatly depends on the inverted EM velocities, since small changes in the EM velocity correspond to relatively larger changes in the density of the simulated air-ice mixture (Table 3.1). The described underestimation of the EM velocity in the deepest layer, corresponding to an overestimation of its density, combined with the underestimated thickness of the layer, causes a slightly larger difference between the theoretical and calculated water contents (Fig. 3.4D), with respect to the difference between the theoretical and calculated thicknesses.

3.4 Field data analysis

In this section, the presented inversion algorithm is applied to the glaciological CO GPR profile shown in Fig. 2.3, which was used in chapter 2 to describe the proposed auto-picking algorithm. The profile was acquired in May 15th 2014 on an alpine glacieret as part of a 3-D GPR survey to study its internal stratigraphy, density distribution, total volume, and water content, and the results are presented in greater details in chapter 5. The data set was recorded using a ProEx GPR system equipped with

250 MHz ground-coupled bistatic Malå Geoscience shielded antennas, with an average 20 cm trace interval and a transmitter-receiver offset equal to 31 cm. The processing sequence consisted of DC removal, drift correction, band pass filtering, and background removal, while the amplitude recovery was applied during the inversion process, as previously discussed. The applied signal processing does not include migration, since it requires as input the EM velocity distribution, which is not usually known in advance and it is the very aim of the inversion process. Moreover, the EM velocity can show significant vertical and lateral variation in small glaciers and glacierets (Colucci et al., 2015), and an inaccurate migration velocity can cause significant signal distortions, as explained in section 2.2.

The EM velocity in the shallowest layer is assumed constant at 21.0 cm/ns, as calculated by applying Looyenga’s empirical formula (Eqs. 3.1 and 3.5) to direct density measurements sampled at different depths in a 4 m deep snow pit. In case a more detailed EM velocity distribution for the shallower layer was available, it would certainly improve the accuracy of the inversion results, both in terms of their absolute values and their lateral variations. Nevertheless, a constant EM velocity can be considered as a normalization velocity, while the estimated subsurface EM impedance contrasts mainly depend on the input reference and reflected amplitudes. The uncertainties associated with the EM velocity of the shallowest layer is set equal to of 0.1 cm/ns, while the uncertainties of the reference and reflected amplitudes is set at 5% their absolute values. These input quantities are the ones that contribute the most to the uncertainty of the inversion results, which is calculated by applying the maximum error propagation formula to each equation used during the inversion process.

The horizons used for the inversion, shown in Fig. 3.5A, were selected among those automatically picked by the algorithm (Fig. 2.4), after the automated grouping process was able to identify the main recorded reflections (Figs. 2.9 and 2.11). Shallower coherent reflections located within the 0-20 ns time interval (Fig. 3.5A) were not considered for the inversion due to possible amplitude distortions caused by interference within the more heterogeneous fresh snow, as well as residual coherent noise from the airwave-groundwave interference, which was removed using the background removal. Moreover, the EM velocity of the shallowest layer was calculated from direct density measurements averaged along the 4 m deep snow pit, which approximately covers the considered 0-20 ns time interval.

The reconstructed internal stratigraphy of the glacieret is shown in Fig. 3.5B, together with the associated uncertainties. For continuity reasons, gaps shorter than 15 traces (i.e. about 3 m) in the selected horizons (Fig. 3.5A) were automatically filled before the inversion by interpolating the amplitudes and travel times picked at both ends of each gap. Other gaps in the deepest interface, which marks the rocky base of the glacieret, were also filled regardless of their width, just to ensure the correct calculation of the total thickness of the glacieret. The reconstructed stratigraphy in Fig. 3.5B highlights the iterative nature of the inversion process, with the uncertainties of the results increasing with depth and being larger in areas with a higher number of layers. Since the inversion process is applied to each trace independently, sudden changes in the stratigraphy or even local spikes can be expected, while the global lateral trends are preserved, which means that the inversion results can be used to describe the subsurface on a global (i.e. tens of meters) rather than a local (i.e. tens of centimeters) scale, while the large number of recorded traces makes the analysis statistically sound.

The estimated density distribution is presented in Fig. 3.6A, showing a general increase of the density with depth, as expected, as well as significant lateral density variations along the profile. Due to the aforementioned independence of the inversion results in each trace, the densities were smoothed for visual clarity using a 41 traces (i.e. 8 m) wide moving average window, in order to better highlight the lateral changes and to filter out unavoidable outliers. The presence of these outliers is still noticeable in a few darker areas in Fig. 3.6A, which show densities close to, and in some cases exceeding, the density of pure ice (i.e. 917 kg/m³). In the latter case, the excessively high densities

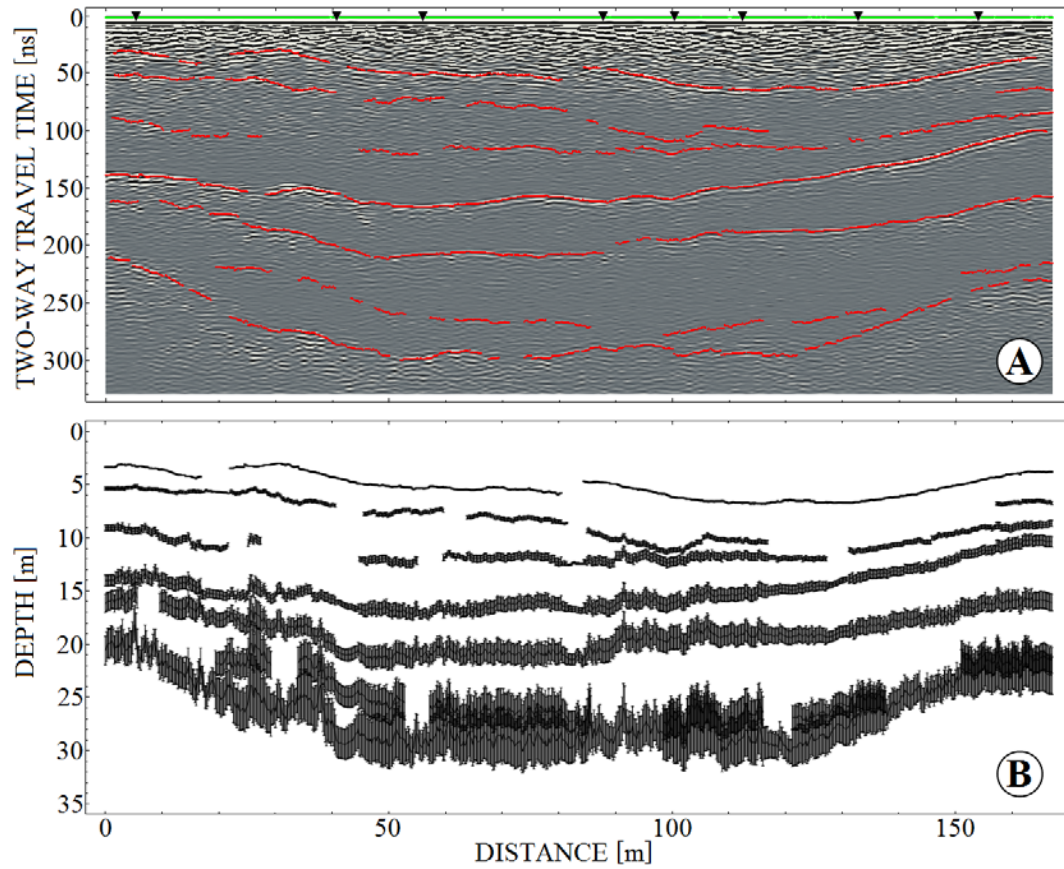


Figure 3.5: Example of amplitude inversion applied to the GPR profile shown in Fig. 2.3. The figure shows the horizons (A) selected for the analysis among those automatically picked (Fig. 2.4), as well as the inverted stratigraphy (B). Positive amplitudes in (A) are marked in green, negative amplitudes in red. The black triangles mark points of intersection with other GPR profiles in the 3-D survey.

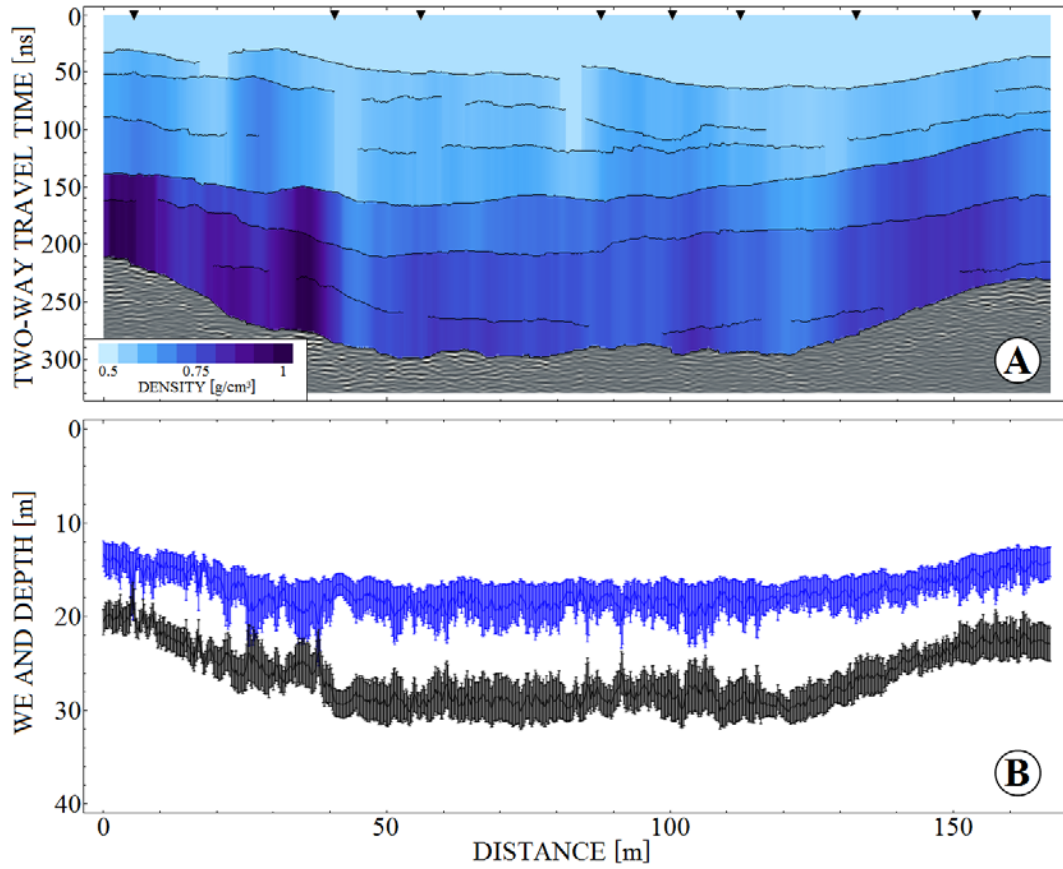


Figure 3.6: Example of amplitude inversion applied to the GPR profile shown in Fig. 2.3. The figure shows the inverted density distribution (A), and the total thickness (black) and water equivalent (blue) calculated at each trace of the profile (B). For visual clarity, a 41 traces (i.e. 8 m) wide smoothing window was applied to the results in (A), in order to better highlight the lateral density variations, and also to remove unavoidable outliers. No smoothing operator was applied to the results in (B). The black triangles mark points of intersection with other GPR profiles in the 3-D survey.

calculated in these few traces indicate a clear underestimation of the EM velocities below the minimum threshold of 16.8 cm/ns for pure ice, as calculated from Eqs. 3.1 and 3.5.

As previously discussed, the main uncertainty factors during the inversion process are the input EM velocity and the reference and reflected amplitudes. The underestimation of the EM velocities can be the result of an underestimation of the EM velocity in the shallower layer, which in this case was set constant along the entire GPR profile. This implies that in some areas the actual EM velocity can be either higher or lower, and this initial error would then propagate to the deeper layers, thus leading to the observed results. Another possible cause of the underestimated EM velocities can be an overestimation of the subsurface EM impedance contrasts, which could indicate an underestimation of the measured reference amplitude, as well as an overestimation of the reflected amplitudes due to factors like isolated ice lenses, internal debris, constructive signal interference, or overcompensation for amplitude decay. The deeper layers are the ones more likely to show outliers in the inversion results, especially in areas with higher stratification, due to the iterative nature of the inversion process, which can combine all these factors and propagate them to the deeper layers. Moreover, possible underestimations of the EM velocity in the shallower layers could still be acceptable for air-ice mixtures (Table 3.1), and therefore they may not be noticeable when compared with the EM velocities in neighboring traces, as opposed to clearly unrealistic inversion results such as the ones discussed.

Besides local outliers, the reconstructed density distribution is useful when analyzing the glacieret on a global scale. For example, the distribution shows a clear divide between the densities estimated above and below the central interface, located at about 140 ns in Fig. 3.6A, while the internal stratification of the two main glaciological units is characterized by smaller density changes due to weaker EM impedance contrasts. The deeper denser part, composed of compact firn and ice, represents the section of the glacieret that remained after the 2013 summer melting period, while the shallower part represents the materials accumulated during the 2013-2014 winter period.

The reconstructed density distribution can be combined with the stratigraphy in order to estimate the water equivalent (Eq. 3.6), which is shown in Fig. 3.6B compared to the total thickness of the glacieret. The average thickness along the analyzed GPR profile is equal to 26.5 m, while the average water equivalent is equal to 17.4 m. It is interesting to notice the opposite lateral trends of the two quantities, in which an increase in the thickness correspond to a decrease in the water equivalent, which is caused by the fact that both quantities are derived from the reconstructed EM velocity distribution. Considering a set of adjacent traces with the same number of reflections at equal travel times, a lateral increase in the average EM velocity correspond to an increase in the total thickness, but also to a decrease in the average density (Table 3.1). For small EM velocity variations, the opposite changes in density and thickness could compensate each other in the calculation of the water equivalent (Eq. 3.6), resulting in smaller changes (Table 3.2). However, for large EM velocity variations, the density changes become dominant with respect to changes in the thickness, resulting in the opposite lateral trends of the thickness and water equivalent in Fig. 3.6B.

In chapter 5, the auto-picking and inversion procedure applied to the analyzed glaciological GPR profile is extended with the same parameters to all the 25 profiles in a 3-D survey covering the entire surface of the glacieret, in order to estimate its total volume and water content, while the stability of the procedure is verified by comparing the inversion results along different profiles at each intersection, such as those marked with black triangles in Figs. 3.5A and 3.6A.

Chapter 4

Amplitude sampling

The main uncertainty factors during the proposed inversion procedure are the input EM velocity and the reference and reflected amplitudes, with the initial errors iteratively propagating to the deeper layers. In glaciological GPR surveys, the EM velocity of the shallower layer can be estimated by applying well-known empirical formulas to direct density measurements, and the accuracy of the input EM velocity distribution mainly depends on the spatial density of these direct measurements. As previously discussed, while a constant EM velocity in the shallower layer can be considered as a normalization velocity, the accurate reconstruction of the subsurface EM impedance contrasts strongly depends on the input reference and reflected amplitudes, which have to accurately sample the peak amplitudes of their respective original analog signals.

Sampling is a non-linear process used during data acquisition in order to convert analog signals, which contain an infinite number of values, into a smaller and numerically manageable discrete series, resulting in an inevitable and unrecoverable loss of the data between the sampled values. Two different kinds of sampling error can be identified:

1. The quantization error is the round-off error resulting from the analog-to-digital conversion, and it is defined by the difference between the theoretical analog signal and the resulting sampled value at a given time instant.
2. The peak amplitude error is defined as the difference between the theoretical analog peak amplitude and the actual digitized peak amplitude, which can be shifted in time by at most half the sampling interval.

This chapter analyzes the peak amplitude error, specifically the effects that the sampling rate selected during data acquisition has on the picked reflection amplitudes, and consequently on the inversion results, when such amplitudes are used as input of the inversion algorithm.

Consider an analog signal defined by a continuous function $A(t)$, the sampling interval Δt specified during data acquisition, and a local maximum of the function located at time $t_{\max}$, whose amplitude $A(t_{\max})$ has to be accurately sampled. The worst sampling scenario, in which the peak amplitude error is the largest possible, is obtained when $t_{\max}$ is between two adjacent samples located at time instants t_k and $t_k + \Delta t$ so that the sampled amplitudes $A(t_k)$ and $A(t_k + \Delta t)$ are equal (Kuffel et al., 2000; Schon, 2013). In this case, both adjacent samples would be considered the digitized peak amplitude $A_{\max}$, however in the more general case one of the two sampled amplitudes would be lower than such value, while the other would be higher, and therefore the latter would be considered the

digitized peak amplitude. The maximum possible time shift of the digitized peak amplitude with respect to the theoretical time instant $t_{\max}$ is equal to half the sampling interval Δt , in which case $t_{\max}$ would be located exactly in the midpoint between t_k and $t_k + \Delta t$, like for example in the case of a symmetric peak sampled in the worst case scenario, while in the more general case $t_{\max}$ would be closer to one of the two adjacent samples.

The choice of the sampling rate to be used during data acquisition is usually based on the well-known Nyquist-Shannon theorem, which gives practical lower limits to the sampling rate in order to accurately preserve the spectral content of the original analog signal. However, while the sampling rate may accurately reconstruct a signal in terms of its frequency components, it may still cause significant peak amplitude errors in the recorded data sets, which would then negatively affect any further quantitative amplitude analysis and cause unnecessary and preventable uncertainties in the inversion results. The peak amplitude error has a random behavior and it depends on the sampling interval, the arrival time, and the shape of the analyzed peak. Nevertheless, the sampling rate can be related to the central frequency of the analyzed signal, in order to establish a minimum threshold above which the amplitude uncertainty is contained below a specified percentage of the actual peak amplitude value.

4.1 State of the art

The Nyquist-Shannon theorem states that a continuous signal characterized by a band-limited frequency spectrum, with a maximum frequency $f_{\max}$, can be fully specified within a time interval of duration T by a set of $2Tf_{\max}$ discrete values (Nyquist, 1928; Shannon, 1949). The theorem is commonly used as a rule of thumb during data acquisition, suggesting that the sampling rate f_s should be equal to at least twice the highest frequency component in the spectrum of the analog signal to be digitized. In other words, in order to avoid aliasing effects and correctly reconstruct a band-limited signal, at least in terms of its spectral content, the value of $f_{\max}$ has to be lower than the Nyquist frequency f_{Ny} , which is defined as half the sampling rate f_s :

$$f_{Ny} = \frac{1}{2\Delta t} \quad (4.1)$$

where Δt is the sampling interval. For clarity, the following text will always refer to the Nyquist frequency for quantitative analysis, since it can be more easily compared with the highest frequency component of the main energy band of a signal when analyzing its amplitude spectrum, while the selected sampling rate is twice the Nyquist frequency:

$$f_s = 2f_{Ny} \quad (4.2)$$

with

$$f_s = \frac{1}{\Delta t}$$

The Nyquist-Shannon theorem refers to a band-limited signal, and therefore it cannot be strictly applied to a recorded data set, since a signal limited in time is characterized by an infinite number of spectral components. Nevertheless, it is always possible to define a maximum frequency $f_{\max}$ in the signal spectrum above which the signal energy can be considered negligible, and this frequency is referred to in the following as the Nyquist-Shannon threshold. The theorem thus offers a useful guideline when choosing a value for f_{Ny} during data acquisition, which has to be higher than $f_{\max}$ in order to accurately reconstruct the spectral content of the original analog signal. For example, when

sampling an analog signal characterized by a central frequency f_C , Jol (2009) recommends using a value of f_{Ny} equal to at least $3f_C$ for GPR surveys, while Drijkoningen (2003) suggests using a value of f_{Ny} equal to at least $4f_C$ as a rule of thumb for reflection seismics.

While accurately describing the behavior of the reconstructed frequency spectrum with respect to the selected sampling rate, the Nyquist-Shannon theorem does not take into account neither the signal distortions caused by the significant and unrecoverable amplitude data loss in the time domain following the analog-to-digital conversion, nor the implications that such distortions have on any further quantitative analysis of the digitized signal. This is because the theorem was initially developed in the context of telegraph transmissions, in which the signal must be properly recorded in terms of its spectral content, while the specific amplitude values were not essential for the correct recovery of the transmitted information, as long as the signal-to-noise ratio of the recorded signal was acceptable. Nevertheless, it is necessary to assess the effects that the sampling rate selected during data acquisition has on the recorded amplitudes, and consequently on both the qualitative and quantitative analysis and the subsequent interpretation of geophysical data sets. In fact, the recorded peak amplitudes are important input for several signal processing and inversion algorithms, such as the Amplitude-vs-Offset (AVO) analysis and the inversion procedure described in the chapter 3.

The interpolation of the recorded data sets and the subsequent re-sampling of the reconstructed analog signal can theoretically be considered as a possible way to reduce sampling-related signal distortions. However, the accuracy of the reconstructed signal depends on the applied interpolation method, which may require systematic assumptions on the signal behavior between samples. These assumptions may be accurate for synthetic data sets, for which the wavelet shape is known, but they are not necessarily acceptable nor verifiable for real data sets, in which the actual data are limited to the sampled values, while any information between samples is unrecoverably lost. Interpolation and re-sampling are analyzed by Strange (2013) as a possible way to increase temporal resolution, therefore his focus is to reduce the time uncertainty $\Delta t/2$ of the signal peak, rather than the peak amplitude error. His conclusion is that the interpolation of a discrete data set in order to increase resolution is less accurate than directly sampling the original analog signal at higher rates, while interpolating data sets acquired at low sampling rate is always preferable with respect to using the raw sampled data (Strange, 2013). Moreover, accurate interpolation techniques tend to be more computationally intensive and time consuming, especially when applied to large data sets, while computationally effective interpolation methods tend to be less accurate (Strange, 2013).

More importantly, even if the analog signal is accurately reconstructed through interpolation, the subsequent re-sampling process simply re-introduces the initial sampling-related distortion problem, since a new value of f_{Ny} has to be selected in order to re-sample the reconstructed signal. Therefore, even if the analyzed data set is re-sampled in order to accurately measure a specific peak amplitude, other recorded signal peaks may still show considerable peak amplitude errors, depending on their shape and arrival time, as well as the newly selected sampling interval.

4.2 Synthetic data analysis

The following sections analyze the effects that different sampling rates have on the peak amplitudes sampled from generic analog wavelets. The analysis is also applied to both CMP and CO synthetic data sets, in order to study the behavior of the peak amplitude error, as well as the effects of noise and interference. In the CO case, which is more strictly related to the previously described inversion procedure, the effects that the sampling rates has on the reconstructed stratigraphy, velocity distribution, and water equivalent of the model shown in Fig. 3.4A are further discussed.

NYQUIST FREQUENCY f_{Ny} [GHz]	SAMPLING INTERVAL Δt [ns]	SINE FUNCTION A [%]	RICKER FUNCTION A [%]	GENERIC WAVELET A [%]
1.0	0.50	70.7	84.1	69.1
1.5	0.33	86.6	92.7	85.7
2.0	0.25	92.4	95.9	91.9
2.5	0.20	95.1	97.4	94.8
3.0	0.17	96.6	98.2	96.3
3.5	0.14	97.5	98.6	97.3
4.0	0.13	98.1	99.0	97.9

Table 4.1: Analysis of the peak amplitudes sampled in Figs. 4.1A-C. The table lists the Nyquist frequency f_{Ny} , the sampling interval Δt , and the peak amplitude A of the main signal peak in each graph, sampled in the worst case scenario and shown as a percentage of the actual peak amplitudes.

Wavelet analysis

Consider three different analog signals, specifically a sine function, a Ricker function, and a generic wavelet simulating both GPR and seismic sources, respectively given by the following equations

$$S(t) = \sin[2\pi f_S t] \quad (4.3)$$

$$R(t) = \left(1 - \frac{1}{2}(2\pi f_R(t - \delta))^2\right) e^{-\frac{1}{4}(2\pi f_R(t - \delta))^2} \quad (4.4)$$

$$G(t) = -A \sin[2\pi f_G(\alpha t - \beta)] e^{-2\pi f_G(\alpha t - \beta)^2} \quad (4.5)$$

where A is a normalizing amplitude, the central frequencies f_S , f_R , and f_G are respectively equal to 0.5, 0.3, and 0.5 GHz, the time coefficient α is equal to 0.1, and the temporal parameters δ and β are equal to 3 and 0.3 ns, respectively.

The analyzed analog signals are respectively shown in Figs. 4.1A-C, together with the discrete series sampled in the worst case scenario with respect to the main signal peak in each graph, using values of f_{Ny} equal to 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, and 4.0 GHz in all three cases. The chosen Nyquist frequencies are all well above the main energy bands of the analyzed signals, which are shown in Figs. 4.1D-F, and therefore they all satisfy the Nyquist-Shannon theorem, so that the amplitude spectrum is the same for all the analyzed discrete series in each case. The peak amplitudes sampled in Figs. 4.1A-C are shown in Table 4.1 as a percentage of the actual peak amplitudes of the original analog signals. As expected, the results show that the higher the Nyquist frequency is, the lower the maximum peak amplitude error becomes in all three cases. This behavior can also be observed in Figs. 4.1G-I, which show the peak amplitudes sampled in the worst case scenario for values of f_{Ny} between 0.8 and 5.2 GHz, which are still above the main energy bands in all three cases (Figs. 4.1D-F). More specifically, Figs. 4.1G-I show that the higher the Nyquist-to-central frequency is, the lower the maximum peak amplitude error becomes. In particular, both the sine function and the generic wavelet have a central frequency equal to 0.5 GHz (Figs. 4.1D and 4.1F), and the corresponding sampling errors show a very similar dependence on f_{Ny} (Figs. 4.1G and 4.1I). On the other hand, the Ricker function is characterized by a lower central frequency (i.e. 0.3 GHz, Fig. 4.1E), which correspond to a higher

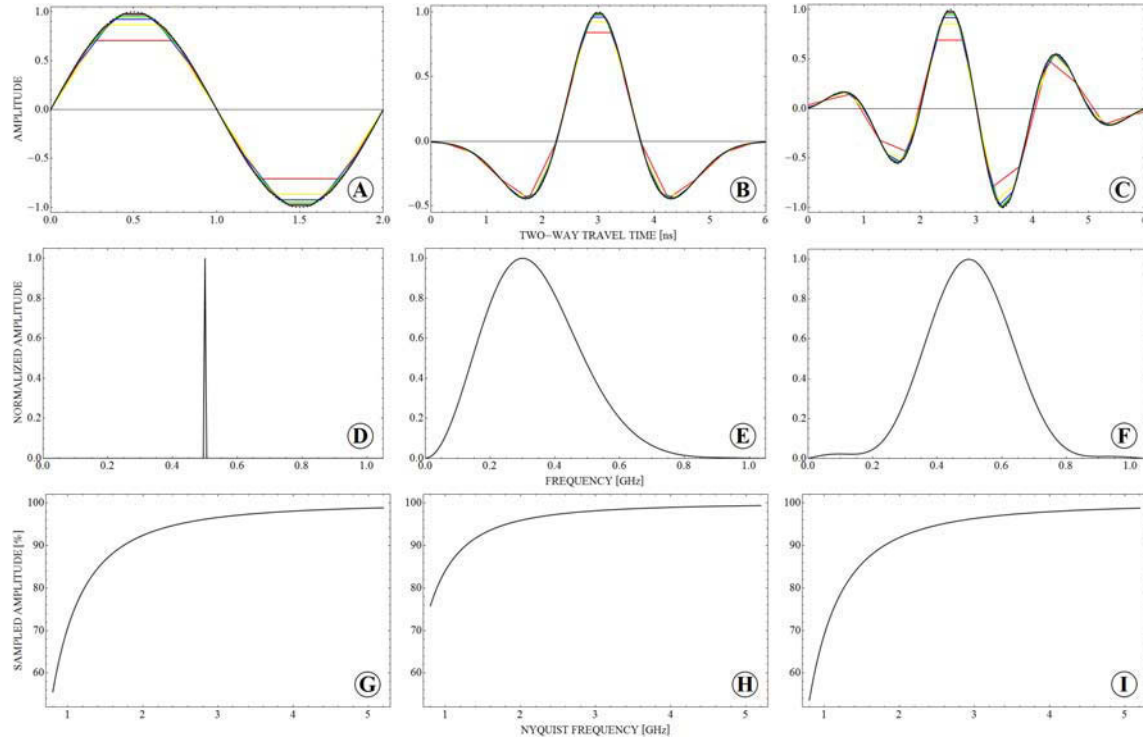


Figure 4.1: Analysis of the peak amplitude error for a sinusoid signal, a Ricker function, and a generic wavelet. The figure shows the analog signals (dashed lines) and seven digital series (A-C) sampled in the worst case scenario with respect to the main signal peak in each graph, their amplitude spectra (D-F), and the resulting peak amplitudes for values of f_{Ny} between 0.8 and 5.2 GHz (G-I), also sampled in the worst case scenario and shown as a percentage of the actual peak amplitudes. The values of f_{Ny} used for sampling in (A-C) are equal to 1.0 (red), 1.5 (yellow), 2.0 (blue), 2.5 (green), 3.0 (brown), 3.5 (gray), and 4.0 (black) GHz, all well above the main energy bands of the respective signals (D-F), and the resulting sampled peak amplitudes are shown in Table 4.1. The maximum peak amplitude error is lower than 5% for values of f_{Ny} above 2.5 GHz in (G), 1.82 GHz in (H), and 2.56 GHz in (I), which are equal to about 5, 6, and 5 times the respective central frequencies (i.e. 0.5, 0.3, and 0.5 GHz).

Nyquist-to-central frequency ratio for the analyzed f_{Ny} interval, resulting in lower sampling errors (Fig. 4.1H) with respect to the other wavelets (Figs. 4.1G and 4.1I).

This relation between the Nyquist frequency, the central signal frequency, and the peak amplitude error can be used to define a minimum threshold for the value of f_{Ny} , which can be used to limit sampling-related signal distortions during data acquisition. The sampling analysis in Figs. 4.1G-I shows that the maximum peak amplitude error is lower than 5% for values of f_{Ny} above 2.48 GHz for the sine function, which is about 5 times the signal central frequency (i.e. 0.5 GHz, Fig. 4.1D); for values of f_{Ny} above 1.82 GHz for the Ricker function, which is around 6 times the signal central frequency (i.e. 0.3 GHz, Fig. 4.1E); and for values of f_{Ny} above 2.56 GHz for the generic wavelet, which is more than 5 times the signal central frequency (i.e. 0.5 GHz, Fig. 4.1F).

Common Mid-Point data set

The model shown in Fig. 4.2A consists of four layers, with both horizontal and dipping reflectors, above a semi-infinite base layer, and it was used as input in the GPRMax 2-D software (Giannopoulos, 2005) to create a synthetic CMP GPR data set. The simulated layers are homogeneous, lossless, non-magnetic, and non-dispersive, with EM velocities typical of a glacier above a rocky basement (Table 3.1). The simulated survey consists of a ground-coupled 400 MHz bistatic GPR system, which transmits Ricker wavelets with an offset that varies from 0 m, i.e. with the transmitter and the receiver both located at the mid-point in Fig. 4.2A, to a maximum value of 10 m, increasing by about 14.9 cm at each step. The constructed GPR data set is shown in Fig. 4.2B, with superimposed the picked horizons used for the sampling analysis, each marking the main phase of their respective reflections (i.e. the second phase for a Ricker wavelet). Specifically, the graphs in Figs. 4.3A-E show the peak amplitudes of the groundwave and of the four constructed reflections, sampled using values of f_{Ny} equal to 8.5, 4.0, 2.5, 1.5, and 1.2 GHz. The chosen Nyquist frequencies are all well above the main energy band of the analyzed signal, shown in Fig. 4.3F, and therefore they all satisfy the Nyquist-Shannon theorem, so that the amplitude spectrum is the same for all the analyzed discrete signals. Moreover, the values of f_{Ny} used in the analysis also satisfy the threshold recommended by Jol (2009) for GPR of three times the central frequency of the recorded signal (i.e. three times 0.4 GHz, Fig. 4.3F).

The picked amplitudes in Fig. 4.3 are shown as a percentage in each trace of the amplitudes picked with f_{Ny} equal to 8.5 GHz, which are used as reference. The general trend shows a decrease in the variability of the sampled peak amplitudes, and therefore of the maximum peak amplitude error, with increasing values of f_{Ny} . This behavior is also highlighted in Table 4.2, which shows the minimum sampled amplitudes in each graph for each sampling rate, given as a percentage of the amplitudes picked with the highest f_{Ny} . The results show maximum peak amplitude errors as high as 6.5% (4 GHz, blue line), 14.5% (2.5 GHz, red line), 12.4% (1.5 GHz, green line), and 18.9% (1.2 GHz, grey line). There are a few cases in Figs. 4.3A-E where the sampled peak amplitude is actually higher at lower values of f_{Ny} , indicating that even the highest sampling rate does not necessarily guarantee the recovery of the actual peak amplitude, although the resulting sampling errors are negligible.

A quasi-periodic behavior of the peak amplitude error can be noticed in Fig. 4.3A, caused by the linear dependence of the groundwave arrival time with respect to the offset (Fig. 4.2B), combined with the regular sampling grid, which uses a constant sampling interval. It can also be noticed in Fig. 4.2B that the groundwave goes in interference with the first reflection for offsets greater than about 7 m, causing the picked horizon to be cut at about 9 m. The interference also causes the picked signal phase of the groundwave to shorten its time interval, which corresponds to a higher central frequency for the signal phase, that in turn causes a visible increase in the peak amplitude error above 7 m in Fig. 4.3A.

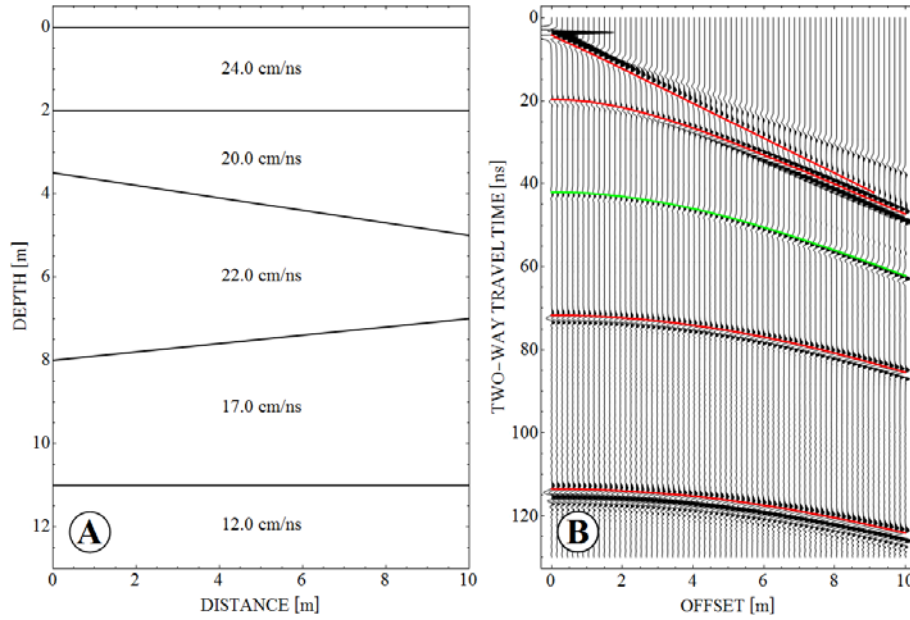


Figure 4.2: Analysis of a synthetic CMP GPR data set, showing (A) the stratigraphy and EM velocity distribution of the model, and (B) the resulting GPR profile, with superimposed the horizons used for the analysis. Positive amplitudes are marked in green, negative amplitudes in red. For visual clarity, a linear gain function was applied to the data set. The synthetic profile was constructed using the GPRMax 2-D software (Giannopoulos, 2005).

NYQUIST FREQUENCY f_{Ny} [GHz]	SAMPLING INTERVAL Δt [ns]	GROUND WAVE A_{min} [%]	REFLECTION			
			1 A_{min} [%]	2 A_{min} [%]	3 A_{min} [%]	4 A_{min} [%]
1.2	0.42	81.1	81.5	84.3	84.6	89.0
1.5	0.33	87.6	88.1	88.5	89.9	92.0
2.5	0.20	85.5	96.0	95.7	96.3	97.1
4.0	0.13	93.5	98.2	98.4	98.5	98.9
8.5	0.06	100.0	100.0	100.0	100.0	100.0

Table 4.2: Analysis of the maximum peak amplitude error for the picked amplitudes shown in Figs. 4.3A-E. The table lists the Nyquist frequency f_{Ny} , the sampling interval Δt , and the minimum sampled peak amplitude A_{min} , given as a percentage of the peak amplitude sampled with the highest f_{Ny} .

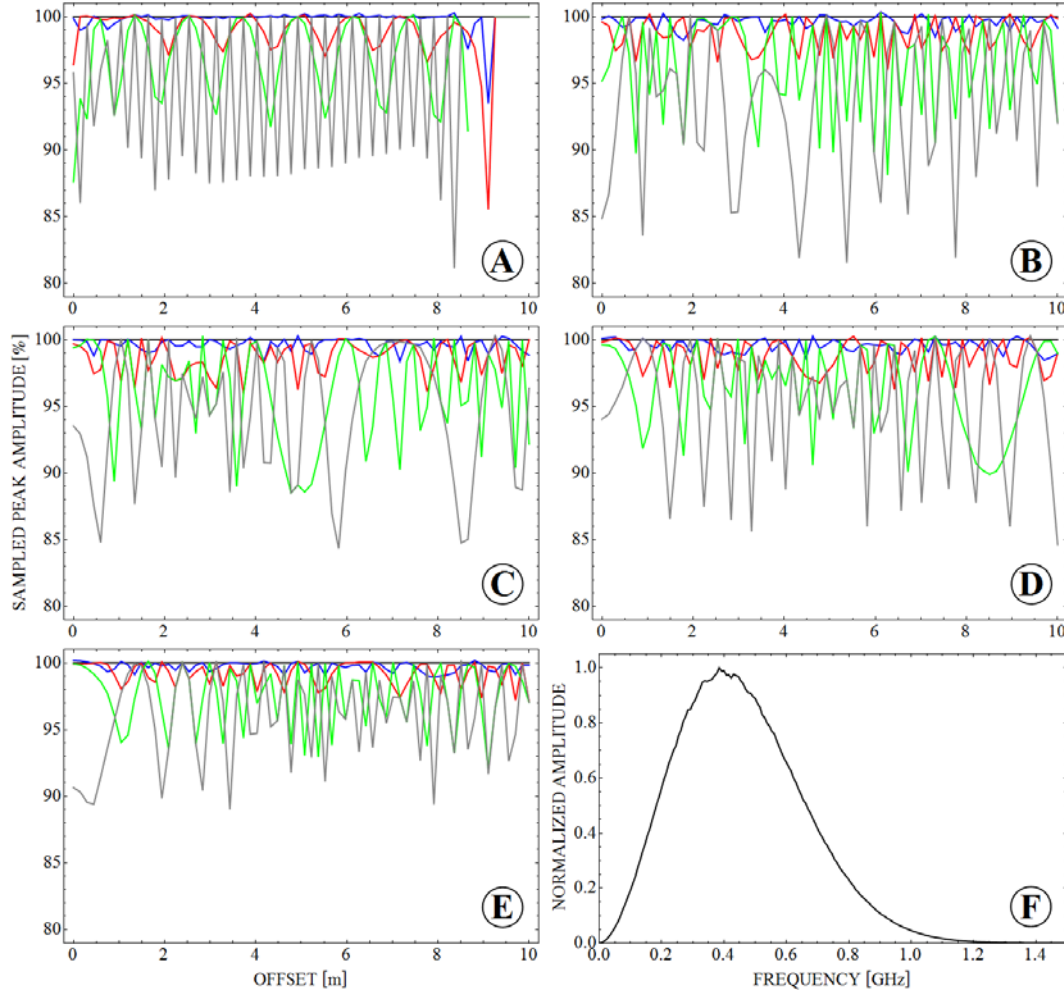


Figure 4.3: Analysis of the peak amplitudes sampled along the groundwave and the four reflections in the synthetic CMP GPR profile shown in Fig. 4.2B. The figure highlights the variations of the peak amplitudes sampled with different sampling rates along the groundwave (A), as well as the first (B), second (C), third (D), and fourth (E) reflections, given as a percentage of the amplitudes picked with the highest f_{Ny} . The normalized frequency spectrum of the zero-offset trace sampled with the highest f_{Ny} is also shown (F). The values of f_{Ny} used for sampling are equal to 8.5 (black, used as reference), 4.0 (blue), 2.5 (red), 1.5 (green), and 1.2 (grey) GHz, all well above the main energy band (F).

This behavior is consistent with the fact that a lower Nyquist-to-central frequency ratio increases the maximum peak amplitude error, and it also causes the picked horizon to be cut even shorter at lower values of the Nyquist frequency (Fig. 4.3A) due to signal distortions.

Besides the quasi-periodic behavior for perfectly planar events recorded in a regular sampling grid, the peak amplitude error is generally random in nature, and it depends on the sampling interval, the arrival time, and the shape of the analyzed signal peak. This random behavior is clearly shown along the hyperbolic-shaped reflections (Figs. 4.3B-E), and it prevents any possible prediction or estimation of the sampling error. On the other hand, besides the interfering groundwave, the results in Table 4.2 show that the maximum peak amplitude error is lower than 5% in all the reflections for values of f_{Ny} equal to at least 2.5 GHz, which is more than 6 times the signal central frequency (i.e. 0.4 GHz, Fig. 4.3F). Therefore, it is possible to limit the peak amplitude error before data acquisition by setting a minimum threshold for the sampling rate, with respect to the central frequency of the recorded signal.

Common Offset data set

In section 3.3, the previously described auto-picking and amplitude inversion procedure was applied to a synthetic CO GPR data set, and the results were shown in Fig. 3.4. The simulated survey consists of a ground-coupled 400 MHz bistatic GPR system, which transmits Ricker wavelets with a constant 70 cm offset and a 27.8 cm trace interval. Similarly to the synthetic CMP data set shown in Fig. 4.2B, which simulates the same GPR system, the CO GPR profile shown in Fig. 3.4B is significantly oversampled, using a value of f_{Ny} equal to 8.5 GHz. This value is greatly above the main energy band of the simulated GPR signal (Fig. 4.3F), which ensures the accurate sampling of the peak reflection amplitudes used in the inversion. In this section, the same sampling analysis applied in Fig. 4.3 to the synthetic CMP survey is also applied to the CO GPR data set, in order to study the effects that the sampling rate has not only on the picked amplitudes, but also on the inversion results.

The analysis is limited to the first six reflections in the GPR profile (Fig. 3.4B), since the amplitudes of neither the airwave-groundwave interference nor the base reflection are necessary to recover the internal EM impedance contrasts of the simulated glacier. The values of f_{Ny} used in the analysis are equal to 8.5, 4.0, 2.5, 1.5, and 1.2 GHz, all well above the main energy band of the signal (Fig. 4.3F), and the resulting picked amplitudes are shown in Fig. 4.4 as a percentage in each trace of the amplitudes picked with f_{Ny} equal to 8.5 GHz, which are used as reference. Similarly to the CMP case, the graphs in Fig. 4.4 show a mostly random behavior of the peak amplitude error, depending on how much the recorded reflections differ from perfectly planar events. As previously discussed, a mostly planar reflection (Fig. 4.4B) in a regular sampling grid can result in quasi-periodic peak amplitude errors, and in the extreme case of a perfectly flat reflection (Fig. 4.4A), it can result in a laterally constant sampling error. However, the deeper reflections (Fig. 4.4C-F) can be deformed by either interference or lateral changes in the shallower stratigraphy and EM velocity distribution, resulting in more erratic changes in the peak amplitude error.

The sampled amplitudes show a decrease in the variability of the peak amplitude error with increasing values of f_{Ny} , similarly to the previous case, and this behavior is also highlighted in Table 4.3, which shows the minimum sampled amplitudes in each graph for each sampling rate, given as a percentage of the amplitudes picked with the highest f_{Ny} . The results show maximum peak amplitude errors as high as 1.5% (4 GHz, blue line), 3.6% (2.5 GHz, red line), 12.6% (1.5 GHz, green line), and 15.5% (1.2 GHz, grey line). Similarly to the previous case, the maximum peak amplitude error in Table 4.3 is lower than 5% in all the reflections for values of f_{Ny} equal to at least 2.5 GHz, which is more than 6 times the signal central frequency (i.e. 0.4 GHz, Fig. 4.3F). Therefore, it is possible to set before

NYQUIST FREQUENCY	SAMPLING INTERVAL	REFLECTION					
f_{Ny} [GHz]	Δt [ns]	1 A_{min} [%]	2 A_{min} [%]	3 A_{min} [%]	4 A_{min} [%]	5 A_{min} [%]	6 A_{min} [%]
1.2	0.42	98.4	85.5	87.6	84.5	86.5	88.7
1.5	0.33	90.8	88.8	93.3	90.0	87.4	92.8
2.5	0.20	97.4	96.4	96.5	96.7	96.5	97.5
4.0	0.13	99.9	98.5	98.6	98.7	98.8	99.0
8.5	0.06	100.0	100.0	100.0	100.0	100.0	100.0

Table 4.3: Analysis of the maximum peak amplitude error for the picked amplitudes shown in Fig. 4.4. The table lists the Nyquist frequency f_{Ny} used, the sampling interval Δt , and the minimum sampled peak amplitude A_{min} , given as a percentage of the peak amplitude sampled with the highest f_{Ny} .

data acquisition a minimum threshold for the sampling rate, with respect to the central frequency of the recorded signal, in order to limit the peak amplitude error, which can have a significant impact on any quantitative analysis applied to the recorded data set.

The effects that the sampling rate has on the amplitude inversion results for the analyzed synthetic GPR data set are shown in Fig. 4.5, which highlights the changes in the calculated total glacier thickness (Fig. 4.5A) and water equivalent (Fig. 4.5B) with respect to the values obtained with f_{Ny} equal to 8.5 GHz, which are used as reference. Due to the fact that lower sampling rates generally correspond to higher maximum peak amplitude errors (Fig. 4.4), a similar behavior can be noticed in the inversion results, which show greater differences at lower sampling rates with respect to the reference values. Specifically, we can notice in Fig. 4.5 a maximum difference of about +3.0 (4 GHz, blue line), +10.3 (2.5, red line), +31.3 (1.5 GHz, green line), and +37.1 (1.2 GHz, grey line) cm for the total glacier thickness, and about -3.5 (4 GHz, blue line), -11.7 (2.5, red line), -31.4 (1.5 GHz, green line), and -34.7 (1.2 GHz, grey line) cm for the water equivalent.

The opposite trends of the two quantities with a decreasing sampling rate is explained by the fact that the general decrease in the picked reflection amplitudes (Fig. 4.4) causes an underestimation of the EM impedance contrasts, with a fixed reference amplitude, resulting in a general increase in the estimated average EM velocities with respect to the initial model (Fig. 3.4A). The increase in the average EM velocities causes an increase in the estimated thickness, with fixed travel times, as well as a decrease in the estimated densities of the analyzed air-ice mixtures (Table 3.1), which become dominant in the calculation of the water equivalent (Eq. 3.6).

An interesting effect of the peak amplitude error can be noticed in the 1.5 GHz inversion results (i.e. the green line in Fig. 4.5), which show a constant component that prevent the calculated total thickness and water equivalent from ever reaching the reference values (black line). A similar smaller effect can be noticed for the 2.5 GHz results (red line) and this behavior is caused by the laterally constant peak amplitude error for the first reflection (Fig. 4.4A), which is respectively equal to 2.6% and 9.2% in the 2.5 and 1.5 GHz results (Table 4.3). This causes a significant error in the estimated EM velocities in the second layer, which then propagates to the deeper layers due to the iterative nature of the inversion algorithm.

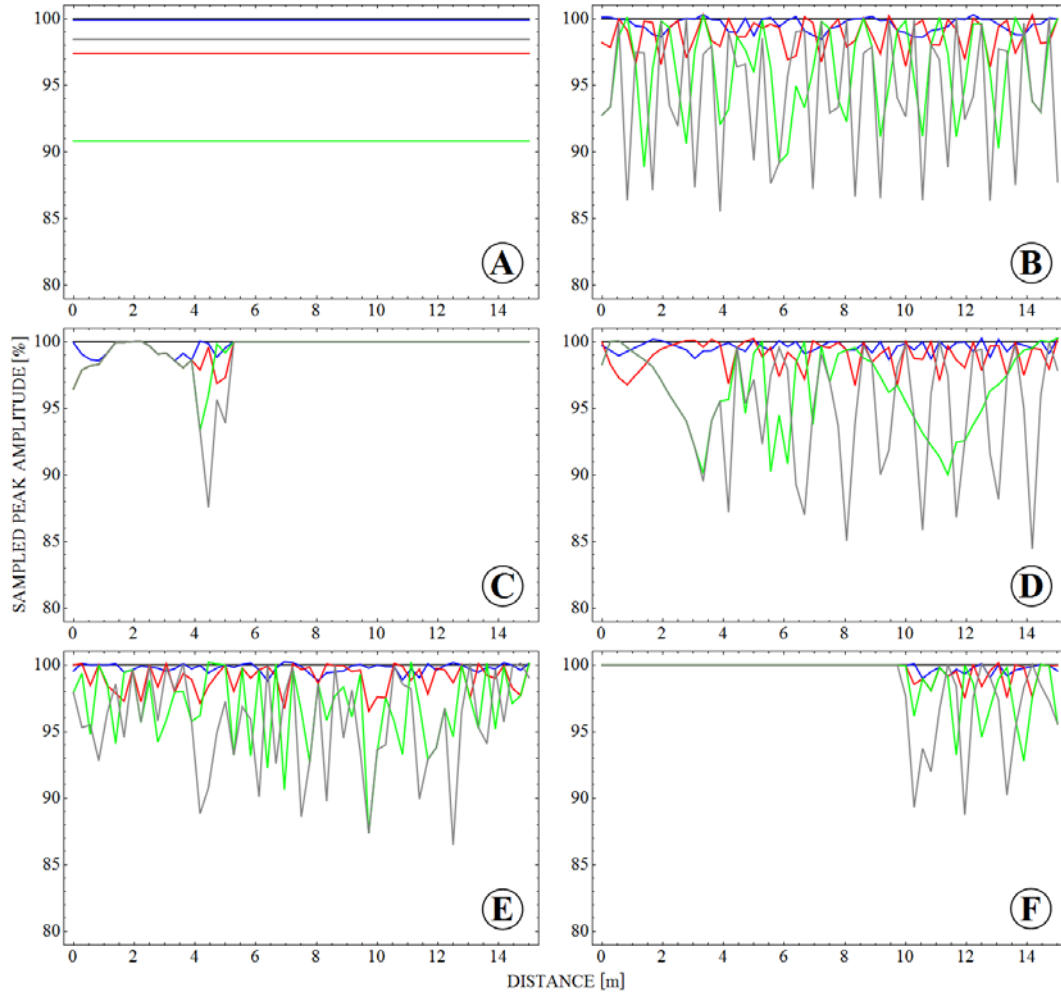


Figure 4.4: Analysis of the peak amplitudes sampled along the first six reflections in the synthetic CO GPR profile shown in Fig. 3.4B. The figure highlights the variations of the peak reflected amplitudes sampled with different sampling rates along each reflection (A-F), given as a percentage of the amplitudes picked with the highest f_{Ny} . The third and sixth reflections are spatially limited (Fig. 3.4B), but their respective graphs (C, F) are plotted for clarity over the entire length of the GPR profile. The values of f_{Ny} used for sampling are equal to 8.5 (black, used as reference), 4.0 (blue), 2.5 (red), 1.5 (green), and 1.2 (grey) GHz, all well above the main energy band (Fig. 4.3F). The effects that the sampling rate has on the amplitude inversion results are shown in Fig. 4.5.

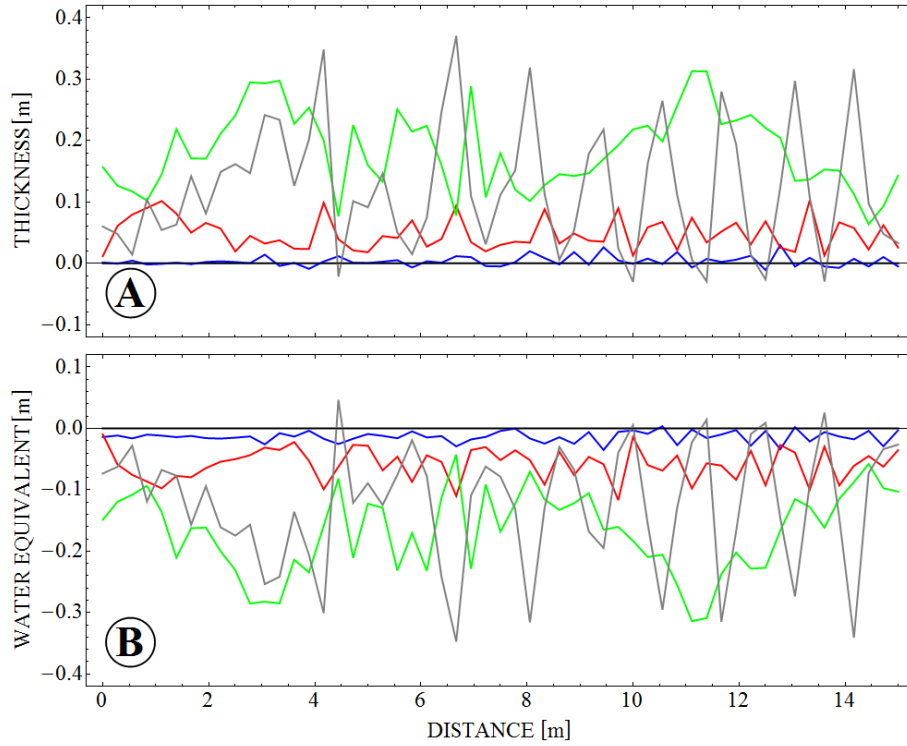


Figure 4.5: Analysis of the variations of the total thickness (A) and water equivalent (B) estimated at different sampling rates for the model shown in Fig. 3.4. The quantities are given as a difference between the values estimated at the various sampling rates with those calculated with the highest f_{Ny} . The values of f_{Ny} used for sampling are equal to 8.5 (black, used as reference), 4.0 (blue), 2.5 (red), 1.5 (green), and 1.2 (grey) GHz, all well above the main energy band (Fig. 4.3F).

4.3 Field data analysis

The signal shown in Fig. 4.6A was acquired using a ground-based 500 MHz GPR system, with a value of f_{Ny} equal to 18.8 GHz, which is greatly above the main energy band of the signal, shown in Fig. 4.6B. The signal was oversampled in order to study changes in the maximum peak amplitude error at different sampling rates without the use of interpolation, so that actual sampled amplitudes can be analyzed. The various discrete time series are therefore obtained by selecting samples from the oversampled trace which are separated by a specified number of time intervals, with such number increasing by one for each series. The worst sampling scenarios for a signal peak are then obtained in each series by selecting two consecutive samples around the maximum value in the oversampled trace, such that the amplitude difference between the two selected samples is the smallest possible.

Two examples of discrete time series obtained with increasing time intervals are shown in Figs. 4.6C-D, respectively sampled in the worst case scenario with respect to the third and fourth peaks marked in Fig. 4.6A. Since the construction of the discrete series does not involve interpolation, the two samples selected in each series on the two sides of the original peaks do not necessarily have the same value like they would have in the actual worst case scenario, and the line joining them is not necessarily flat like in the previously describes synthetic cases (Figs. 4.1A-C). Nevertheless, the selected discrete series contain actual amplitude samples from the initial oversampled trace and therefore they constitute realistic GPR traces recorded at different sampling rates. Twenty discrete time series have been constructed in the worst case scenario for all five peaks marked in Fig. 4.6A, with the time interval increasing by one sample for each series, while the Nyquist frequency remains well above the main energy band of the recorded signal (Fig. 4.6B). The sampled peak amplitudes are shown in Table 4.4 as a percentage of the actual peak amplitudes recorded in the original oversampled trace.

As expected, the results show a general increase in the peak amplitude errors with decreasing sampling rates, while also showing in some cases greatly different errors for the same value of f_{Ny} at the different peaks, mainly due to the different frequency content in each signal phase and, to a lesser extent, the aforementioned constraints in the construction of the discrete series, which used the original samples without interpolation. A few examples are given by the flat peaks of the second and third signal phases, which cause either small or null errors for a large interval of f_{Ny} , while the larger half-height width causes a slow increase in the peak amplitude error with decreasing values of f_{Ny} . This behavior is particularly noticeable in the second peak, for which the error is within 5% for all cases shown in Table 4.4, while in the third peak the error reaches 10.2% with the lowest sampling rate. On the other hand, the fourth signal peak has a higher frequency content which causes a smaller half-height width and consequently a greater increase in the peak amplitude error, reaching values of 27.4% in the worst case in Table 4.4. Similar results were found in the synthetic CMP data set analyzed in Fig. 4.2, in which the groundwave goes in interference with the first reflection at larger values of the transmitter-receiver offset, causing the frequency content of its signal phase to increase and thus resulting in larger peak amplitude errors for the same sampling rates (Fig. 4.3A).

This behavior highlights the difficulty in the estimation or prediction of the peak amplitude error in a GPR data set, with different signal phases showing various levels of uncertainty due to different frequency contents. Moreover, not all the signal phases in a recorded trace are sampled in the worst case scenario (Figs. 4.6C-D), and the peak amplitude error shows a quasi-random behavior which depends on the selected sampling interval, as well as the arrival time and the shape of the analyzed signal peak. As previously discussed, interpolation and re-sampling may be considered in order to reduce the maximum peak amplitude error, however the re-sampling process simply re-introduces the initial sampling-related distortion problem, when selecting a new value of f_{Ny} with which to re-sample

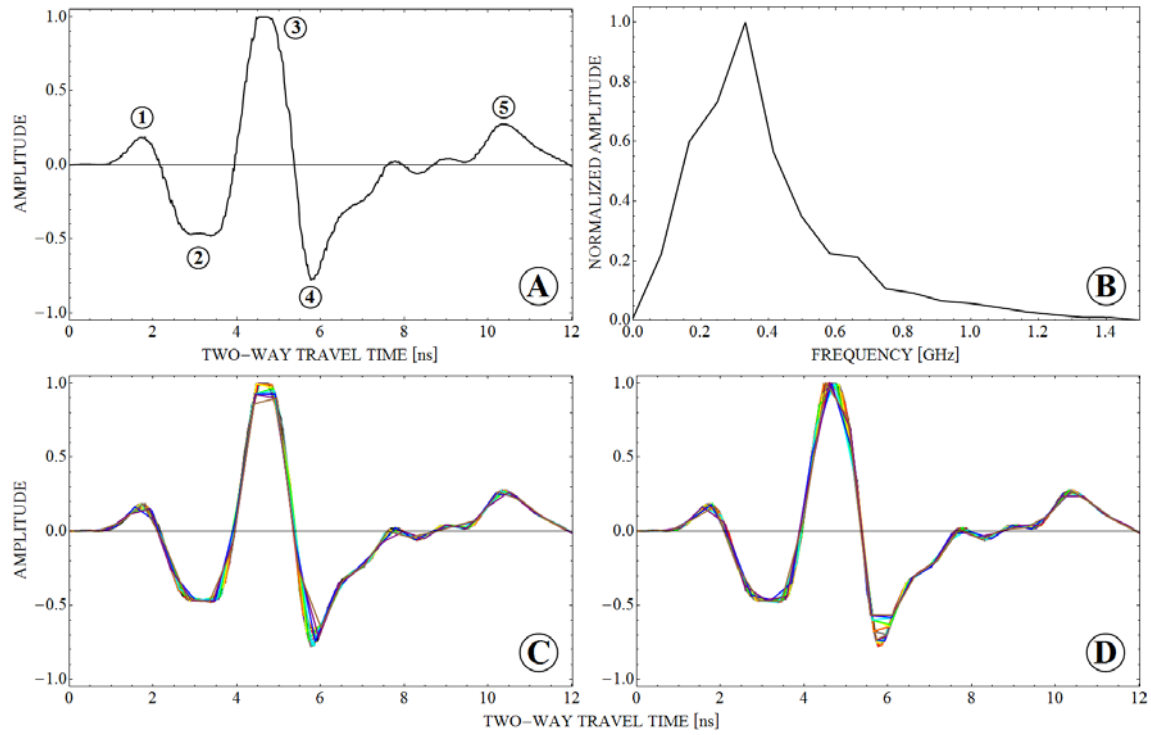


Figure 4.6: Analysis of the peak amplitudes sampled at different peaks of a real GPR signal. The figure shows the analyzed signal (A), the frequency spectrum (B), and the digitized series sampled in the worst case scenario for the third (C) and fourth (D) peaks, chosen as examples. The discrete series were acquired without interpolation, with the sampling interval increasing by one sample for each series. The peak amplitudes sampled in the worst case scenario for all five peaks marked in (A) are analyzed in Table 4.4.

NYQUIST FREQUENCY	SAMPLING INTERVAL	PEAK				
f_{Ny} [GHz]	Δt [ns]	1 A [%]	2 A [%]	3 A [%]	4 A [%]	5 A [%]
18.81	0.03	100.0	99.1	100.0	99.3	100.0
9.41	0.05	100.0	98.6	100.0	99.3	100.0
6.27	0.08	100.0	98.6	100.0	99.3	100.0
4.70	0.11	100.0	98.6	100.0	97.1	100.0
3.76	0.13	100.0	98.7	100.0	96.0	99.2
3.14	0.16	100.0	98.7	100.0	94.5	98.2
2.69	0.19	97.5	98.1	100.0	94.5	98.3
2.35	0.21	90.7	97.4	100.0	94.5	98.3
2.09	0.24	90.7	97.4	99.9	92.0	97.1
1.88	0.27	89.1	96.8	100.0	90.1	97.1
1.71	0.29	90.1	97.0	99.7	87.1	96.2
1.57	0.32	88.8	96.9	99.5	87.1	95.6
1.45	0.35	82.0	96.1	98.9	83.6	94.9
1.34	0.37	81.6	95.6	99.5	82.6	94.9
1.25	0.40	81.6	96.8	99.5	77.3	93.0
1.76	0.43	78.6	96.5	96.7	77.3	92.2
1.11	0.45	78.6	96.7	94.8	75.1	90.0
1.05	0.48	75.6	97.0	92.7	73.4	87.5
0.99	0.50	75.6	96.9	92.5	72.9	85.7
0.94	0.53	67.4	97.1	89.8	72.6	85.7

Table 4.4: Analysis of the peak amplitudes sampled in the worst case scenario at the five signal peaks marked in Fig. 4.6A. The table lists the Nyquist frequencies f_{Ny} , the sampling interval Δt , and peak amplitude A of each signal peak, sampled in the worst case scenario and shown as a percentage of the actual peak amplitude, given by the maximum value in the original oversampled trace.

the reconstructed analog signal. Moreover, the accuracy of the interpolation is not verifiable for real data sets, since the actual data are limited to the sampled values, while any information between samples is unrecoverably lost.

On the other hand, the peak amplitude error can be limited before data acquisition by setting a minimum threshold for the sampling rate, with respect to the central frequency of the recorded signal. The synthetic data analysis in section 4.2 suggested, as a rule of thumb, using during data acquisition a value of f_{Ny} equal to at least 6 times the central frequency of the recorded signal, in order to limit the peak amplitude error within 5%. This rule is also acceptable for the sampled amplitudes in Table 4.4, in which case the minimum threshold would be equal to 3.0 GHz, with the nominal central frequency of the GPR system being equal to 500 MHz. Moreover, the rule can still be applied when considering the actual central frequency of about 350 MHz (Fig. 4.6B) calculated for the analyzed signal (Fig. 4.6A), in which case the minimum threshold would be equal to about 2.1 GHz. For values of f_{Ny} above this threshold, the peak amplitude error in Table 4.4 is within 5% for most of the analyzed signal peaks, except for the fourth one, whose 5.5% error could still be acceptable, and for the first one, whose 9.3% error is most likely due to a lower Nyquist-to-central frequency ratio, combined with the aforementioned sampling constraints, as previously discussed.

Chapter 5

Field application

5.1 Prevala glacieret

In this chapter the automated picking and inversion procedure is applied to a pseudo 3-D glaciological GPR survey, which was constructed by combining a set of 2-D profiles acquired on the Prevala glacieret (Fig. 5.1), which is located in the Julian Alps (Italy). The surface area of the glacieret shown in Fig. 5.1B was measured during an airborne LiDAR survey in October 2013, as part of a study on the 2013 summer mass balance of the analyzed glacieret (Forte et al., 2014b). The LiDAR survey period corresponds with the beginning of the winter season at the glacieret elevation (i.e. about 1830-1960 m a.s.l.), while the analyzed GPR survey was acquired on May 15th 2014, which is after the winter accumulation period and before the start of the summer melting period.

The data set was recorded using a ProEx GPR system equipped with 250 MHz ground-coupled bistatic Malå Geoscience shielded antennas, resulting in 25 intersecting profiles covering the entire surface of the glacieret, with a total length of about 3.1 km. An odometer was used as trigger for the transmitter, with an average 20 cm trace interval. A RTK GPS device was used to locate each trace of the recorded GPR profiles, with a centimetric average accuracy for the measured latitude and longitude and a decimetric average accuracy for the elevation. The sampling interval was set equal to about 0.4 ns, corresponding to a Nyquist frequency of 1.25 GHz, which is 5 times the nominal central frequency of the GPR system. The processing sequence consisted of DC removal, drift correction, band pass filtering within the 100-125-600-750 MHz frequency interval, and background removal using a 51 traces (i.e. about 10 m) wide window within the 5-100 ns and 320-387 ns time intervals.

The picking procedure presented in chapter 2 and applied to profile P1 (Fig. 2.3) was extended with the same parameters to all the other GPR profiles in the 3-D survey. The picking algorithm was used to automatically track every laterally coherent event at least 10 traces (i.e. about 2 m) in length, while the patching process was used to connect consecutive horizons separated by gaps at most 5 traces (i.e. about 1 m) wide. The grouping process was used to automatically identify the main reflections and to select the specific reflection phases to be used in the inversion procedure. The processing sequence applied to the data set did not include any amplitude recovery, therefore the algorithm was able to extract the original amplitudes for the inversion, thus avoiding any possible subjective alteration caused by the assumptions made by the interpreter with regards to the EM properties of the subsurface. Spreading losses caused by wavefront expansion were instead iteratively corrected during the inversion process itself, while intrinsic attenuation was considered negligible in all GPR profiles,

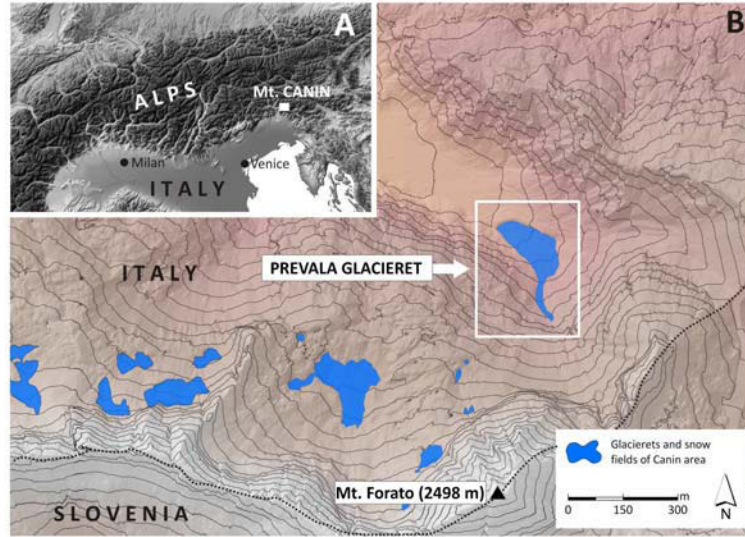


Figure 5.1: Location map of the Prevala glacieret in the Julian Alps (Italy). The figure shows the position of Mt. Canin in North-East Italy (A), and the glacierets and snowfields in the Mt. Canin area (B). The surface areas shown in (B) were measured during an airborne LiDAR survey in October 2013, as part of a study on the summer mass balance of the Prevala glacieret (Forte et al., 2014b).

assuming that the liquid water content could be disregarded during the data acquisition in May.

The auto-picking and inversion analysis was applied to the entire volume of the glacieret, from the surface down to its rocky basement, thus including areas with clearly defined reflections that can be easily tracked, as well as areas characterized by low signal-to-noise ratio, noise, or interference, that prevent the construction of long and clearly defined horizons with undistorted amplitudes. While amplitude distortions due to local events, such as interfering reflections or hyperbolic diffractions caused by internal debris, could be found over the entire volume of the glacieret, the deeper layers are more susceptible to such distortions due to the lower signal-to-noise ratio caused by amplitude decay. For example, the rocky basement of a glacier is usually characterized by rough topography and debris, which cause several interfering reflections and diffractions that may prevent the accurate tracking of a clearly defined basal reflection, thus contributing to the uncertainty of the inversion results, particularly the estimated total thickness of the glacieret.

The EM velocity within the shallowest layer was estimated using direct density measurements sampled at different depths in a 4 m deep snow pit. The measured densities are shown in Fig. 5.3A, while the resulting EM velocities in Fig. 5.3B were estimated using Looyenga's empirical formula (Eqs. 3.1 and 3.5). The interpolated functions (blue lines) were used to calculate the average values through integration, which are respectively equal to 543 kg/m^3 and 20.8 cm/ns . The input EM velocity for the shallowest layer was therefore set equal to 21.0 cm/ns in all GPR profiles during the inversion process. As previously discussed, a more detailed input velocity distribution would certainly improve the inversion results, nevertheless a constant EM velocity can be considered as a normalization velocity, while the internal impedance contrasts mainly depend on the input reference and reflected amplitudes.

The inversion algorithm was applied to each GPR trace independently and the estimated values

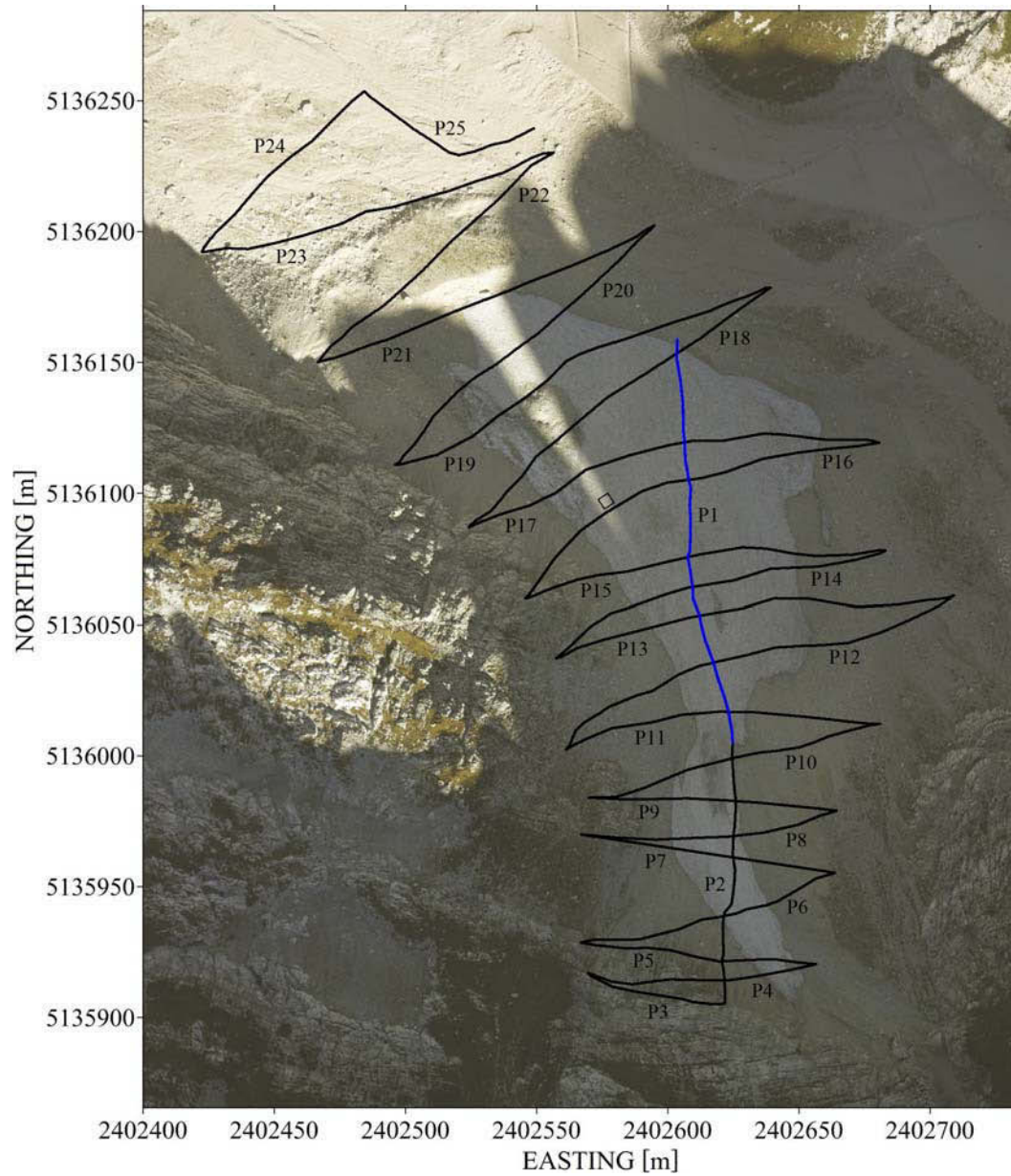


Figure 5.2: Orthophoto of the Prevala glacieret acquired in October 2013, with superimposed the positions of the snow pit (black square) and the GPR profiles (black lines), marked P1-P25, acquired in May 2014. Profile P1 was used in chapters 2 and 3 to illustrate the auto-picking and inversion procedure, and it is marked in blue.

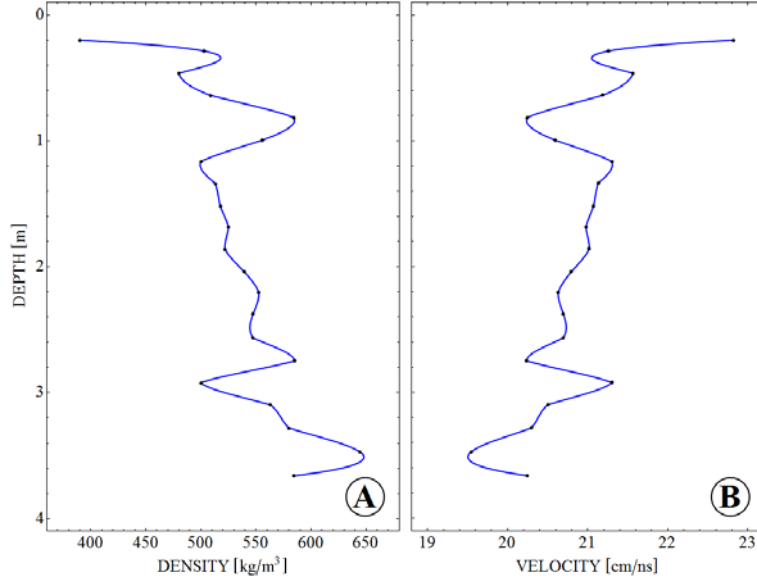


Figure 5.3: Direct measurements from a snow pit. The figure shows the densities measured at different depths (A), and the resulting EM velocities (B) estimated using Looyenga’s empirical formula (Eqs. 3.1 and 3.5). The direct measurements (black dots) are interpolated (blue line) in order to obtain an average EM velocity for the shallowest layer to be used as input for the inversion algorithm.

of the total thickness and water equivalent of the glacieret are respectively shown in Figs. 5.4 and 5.6, interpolated between the various profiles in the 3-D survey. Each figure shows the raw data independently calculated in each GPR trace, as well as the same data smoothed in each GPR profile using a 41 traces (i.e. 8 m) wide moving average window, which removes small scale (i.e. tens of centimeters) lateral changes while preserving large scale (i.e. tens of meters) variations.

The inversion results show an overall good consistency with the topography surrounding the glacieret (Fig. 5.5), with the shallowest areas highlighting the mountain sides narrowing toward the higher southern tip, as well as the broader moraine enclosing the glacieret on the northern side. The deeper central area is also compared in Figs. 5.4B and 5.6B with the surface area of the glacieret as measured in October 2013, and it shows a good correspondence, considering possible differences in the surface topography of the glacieret and the surrounding sediments and bedrock, which were covered by snow during the 2014 GPR survey, but not during the 2013 LiDAR survey.

Nevertheless, the central area highlights a deep denser level consisting of the glacieret that remained after the 2013 summer melting period, buried under the snow fallen during the 2013-2014 winter accumulation period, and subsequently altered by internal thermomechanical processes. The interface between these two main glacial bodies was identified in section 3.4 as the central reflection in profile P1, located at a time-depth of around 140 ns in Fig. 3.5A. The estimated density distribution in Fig. 3.6A shows a clear divide between the densities estimated above and below such interface, while the internal stratification of the two main glaciological units is characterized by smaller density changes due to weaker impedance contrasts. As previously discussed, the darker areas in Fig. 3.6A, which originate from this interface and then propagate to the deeper layers, can be attributed to internal

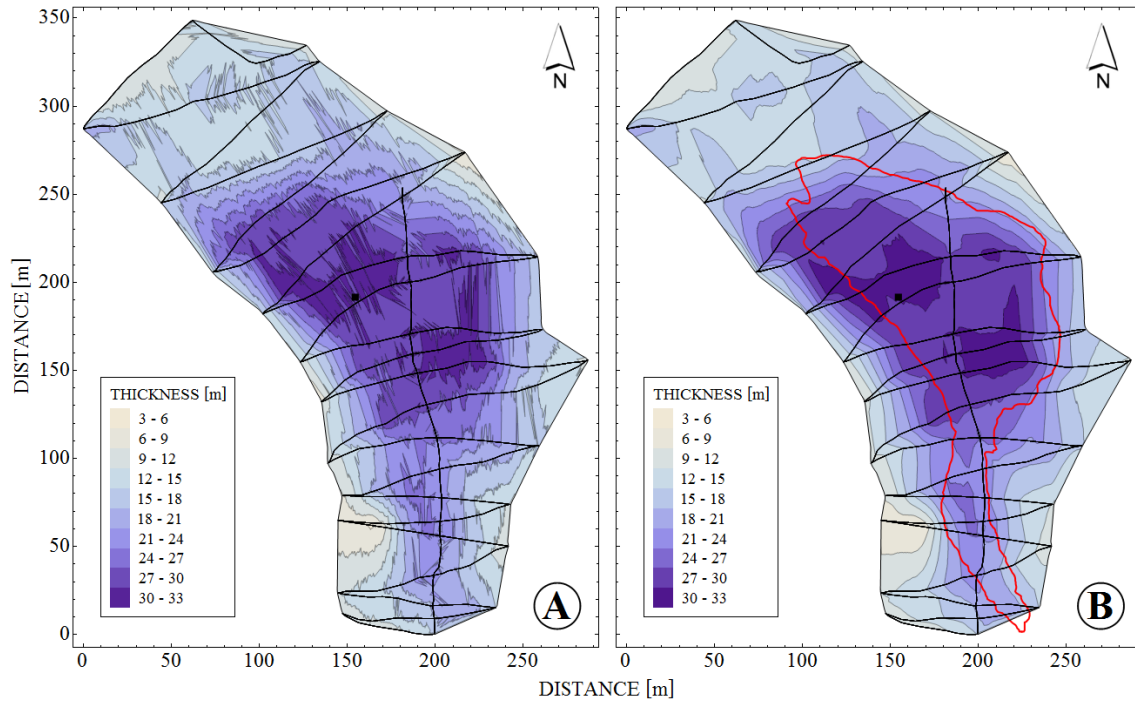


Figure 5.4: Contour plot of the total thickness of the Prevala glacieret as estimated from the 3-D GPR survey acquired in May 2014. The figure interpolates the thicknesses calculated in each GPR trace (A), and the thicknesses smoothed in each trace using a 8 m wide averaging window moving across each profile (B). The positions of the snow pit (black square) and of the GPR profiles (thick black lines) are superimposed to each graph. The boundaries of the glacieret measured in October 2013 (red line) are also superimposed in (B) for comparison.

debris caused by rocks rolling down from the mountain sides and then settling at lower elevations (i.e. the left side of profile P1 in Fig. 3.6A) before the winter accumulation period.

The consistency of the inversion result was also tested at the fifteen points where the GPR profiles cross each other, and the results are shown in Tables 5.1 and 5.2 for the densities estimated in each interpreted layer, and in Table 5.3 for the calculated water equivalent. Possible differences in the reconstructed stratigraphy at a crossing point, caused by weaker impedance contrasts that prevent an interface from being automatically recognized on both the crossing profiles, are corrected by splitting single layers into different layers with the same density values.

The presented tables compare the average value and the variability of each quantity, which are calculated using a 11 traces (i.e. 2 m) wide window centered at each crossing point in each GPR profile of the 3-D survey. The density results in Tables 5.1 and 5.2 that exceed the density of pure ice (i.e. 917 kg/m^3), either in terms of their average values or their variability range, are marked in bold. As expected, these excessively high densities tend to be contained within the deeper layers due to the error propagation during the iterative inversion process, while possible overestimated densities within the shallower layers, caused by signal distortions due to interfering reflections or hyperbolic diffractions, could still be acceptable for air-ice mixtures (Table 3.1).

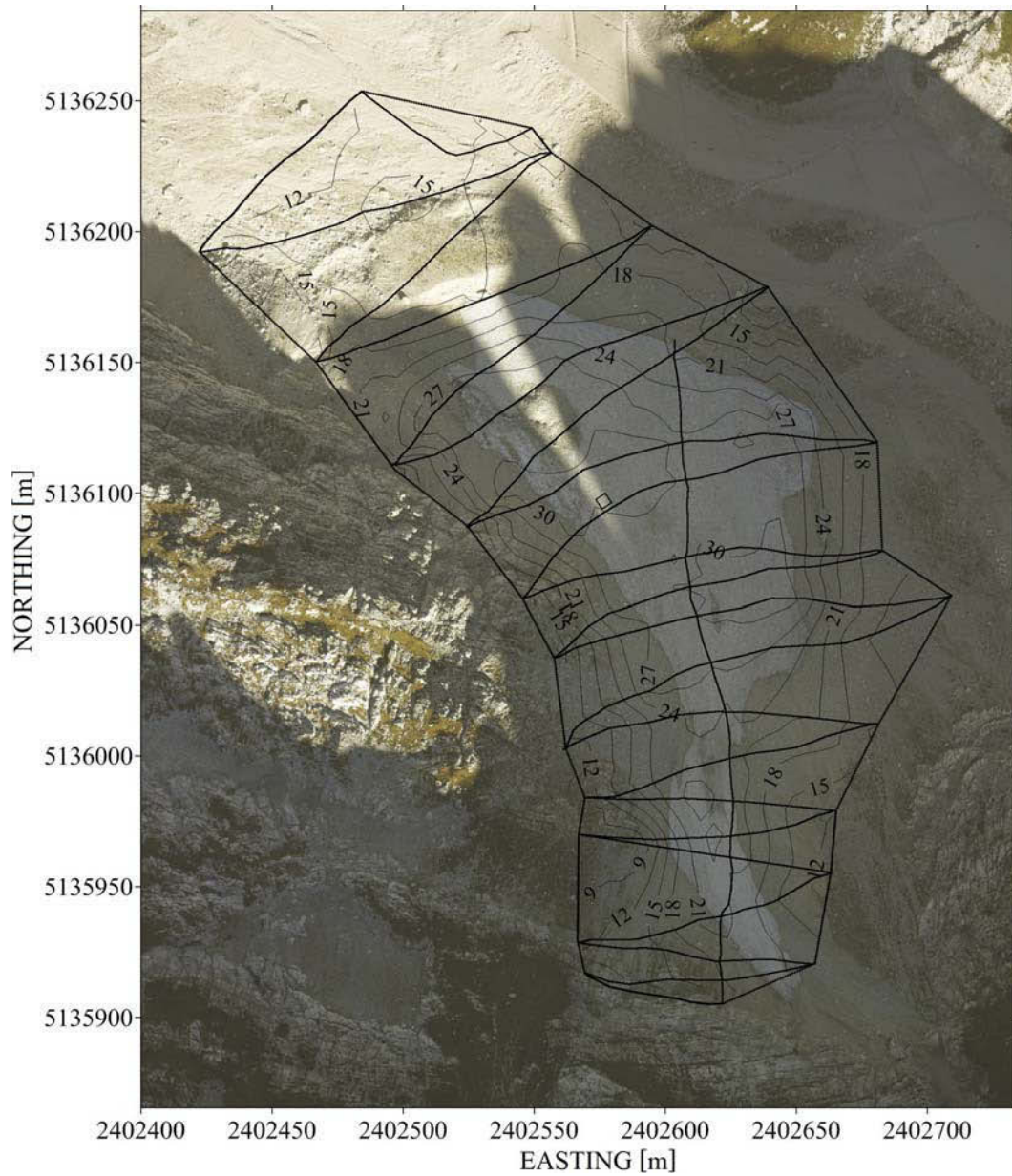


Figure 5.5: Orthophoto of the Prevala glacieret acquired in October 2013, with superimposed the positions of the snow pit (black square), the GPR profiles and boundaries (thick black lines) of the 3-D survey acquired in May 2014, and the thickness values (thin black lines) presented in Fig. 5.4B. The inversion results show an overall good consistency with the imaged topography, considering possible differences in the surface topography of the glacieret and the surrounding sediments and bedrock, which were covered by snow during the 2014 GPR survey, but not during the 2013 LiDAR survey.

Pn	DENSITY $\left[\frac{\text{kg}}{\text{m}^3}\right]$			Pn	DENSITY $\left[\frac{\text{kg}}{\text{m}^3}\right]$		
	Pn	P1	Pn - P1		Pn	P1	Pn - P1
P11	523 $\pm$ 0	523 $\pm$ 0	0	P15	523 $\pm$ 0	523 $\pm$ 0	0
	603 $\pm$ 137	604 $\pm$ 49	1		523 $\pm$ 0	581 $\pm$ 53	58
	613 $\pm$ 134	627 $\pm$ 41	14		545 $\pm$ 30	599 $\pm$ 53	54
	644 $\pm$ 166	738 $\pm$ 74	94		563 $\pm$ 43	624 $\pm$ 65	61
	769 $\pm$ 300	776 $\pm$ 135	7		696 $\pm$ 143	708 $\pm$ 121	12
	900 $\pm$ 304	785 $\pm$ 135	115		745 $\pm$ 145	728 $\pm$ 152	17
P12	523 $\pm$ 0	523 $\pm$ 0	0	P16	523 $\pm$ 0	523 $\pm$ 0	0
	607 $\pm$ 65	574 $\pm$ 57	33		553 $\pm$ 22	573 $\pm$ 79	20
	630 $\pm$ 84	599 $\pm$ 60	31		568 $\pm$ 35	633 $\pm$ 17	65
	630 $\pm$ 84	731 $\pm$ 154	101		602 $\pm$ 45	594 $\pm$ 105	8
	768 $\pm$ 243	797 $\pm$ 138	29		727 $\pm$ 144	735 $\pm$ 115	8
	793 $\pm$ 232	823 $\pm$ 138	30		785 $\pm$ 147	786 $\pm$ 109	1
P13	523 $\pm$ 0	523 $\pm$ 0	0	P17	523 $\pm$ 0	523 $\pm$ 0	0
	584 $\pm$ 71	569 $\pm$ 61	15		523 $\pm$ 0	547 $\pm$ 0	24
	597 $\pm$ 71	589 $\pm$ 69	8		573 $\pm$ 105	571 $\pm$ 72	2
	623 $\pm$ 88	631 $\pm$ 85	8		588 $\pm$ 122	667 $\pm$ 245	79
	690 $\pm$ 107	698 $\pm$ 123	8		832 $\pm$ 260	705 $\pm$ 289	127
	733 $\pm$ 91	731 $\pm$ 145	2		870 $\pm$ 251	724 $\pm$ 299	146
	733 $\pm$ 91	745 $\pm$ 134	12				
P14	523 $\pm$ 0	523 $\pm$ 0	0	P18	523 $\pm$ 0	523 $\pm$ 0	0
	578 $\pm$ 74	606 $\pm$ 52	28		523 $\pm$ 0	604 $\pm$ 74	81
	595 $\pm$ 95	629 $\pm$ 61	34		548 $\pm$ 23	631 $\pm$ 78	83
	665 $\pm$ 127	665 $\pm$ 66	0		548 $\pm$ 23	684 $\pm$ 113	136
	665 $\pm$ 127	722 $\pm$ 56	57		650 $\pm$ 144	1210 $\pm$ 1104	560
	701 $\pm$ 159	740 $\pm$ 65	39		692 $\pm$ 161	1132 $\pm$ 1252	440
	701 $\pm$ 159	771 $\pm$ 67	70				

Table 5.1: Comparison between the density values estimated in each layer at the eight crossing points between profile P1 and the profiles P11-P18. The table shows the mean density values and their variability range, calculated in each profile using a 11 traces-wide window centered at each crossing point. Since the input density of the shallowest layer is fixed as constant, its variability is zero. The table also shows the difference between the mean densities estimated in each pair of crossing profiles. The density results that exceed the density of pure ice (i.e. 917 kg/m³), either in terms of their average values or their variability range, are marked in bold. At the P1-P18 crossing point, clearly overestimated densities are obtained in the deepest two layers of profile P1, due to outliers in the reconstructed density distribution (Fig. 3.6A), as discussed in section 3.4.

P _n	DENSITY $\left[\frac{\text{kg}}{\text{m}^3}\right]$			P _n	DENSITY $\left[\frac{\text{kg}}{\text{m}^3}\right]$		
	P _n	P2	P _n - P2		P _n	P2	P _n - P2
P4	523 ± 0	523 ± 0	0	P8	523 ± 0	523 ± 0	0
	552 ± 16	633 ± 111	81		619 ± 118	523 ± 31	96
	582 ± 13	684 ± 115	102		629 ± 125	544 ± 40	85
	662 ± 107	727 ± 91	65		647 ± 121	556 ± 39	91
	686 ± 121	771 ± 113	85		737 ± 180	662 ± 148	75
P5	523 ± 0	523 ± 0	0	P9	523 ± 0	523 ± 0	0
	670 ± 85	661 ± 0	9		612 ± 62	570 ± 30	42
	686 ± 89	658 ± 118	28		638 ± 62	580 ± 39	58
	759 ± 132	728 ± 44	31		780 ± 174	738 ± 81	42
	799 ± 167	694 ± 141	105		815 ± 149	753 ± 80	62
P6	523 ± 0	523 ± 0	0	P10	523 ± 0	523 ± 0	0
	564 ± 71	604 ± 181	40		583 ± 57	623 ± 59	40
	584 ± 84	631 ± 182	47		601 ± 53	636 ± 65	35
	657 ± 95	677 ± 181	20		618 ± 50	652 ± 75	34
	706 ± 115	675 ± 176	31		697 ± 70	779 ± 93	82
P7	523 ± 0	523 ± 0	0		788 ± 146	902 ± 80	114
	632 ± 128	558 ± 40	74		788 ± 146	918 ± 181	130
	673 ± 174	575 ± 51	98				
	776 ± 253	607 ± 75	169				
	801 ± 248	672 ± 167	129				
	841 ± 272	726 ± 197	115				

Table 5.2: Comparison between the density values estimated in each layer at the seven crossing points between profile P2 and the profiles P4-P10. The table shows the mean density values and their variability range, calculated in each profile using a 11 traces-wide window centered at each crossing point. Since the input density of the shallowest layer is fixed as constant, its variability is zero. The table also shows the difference between the mean densities estimated in each pair of crossing profiles. The density results that exceed the density of pure ice (i.e. 917 kg/m³), either in terms of their average values or their variability range, are marked in bold.

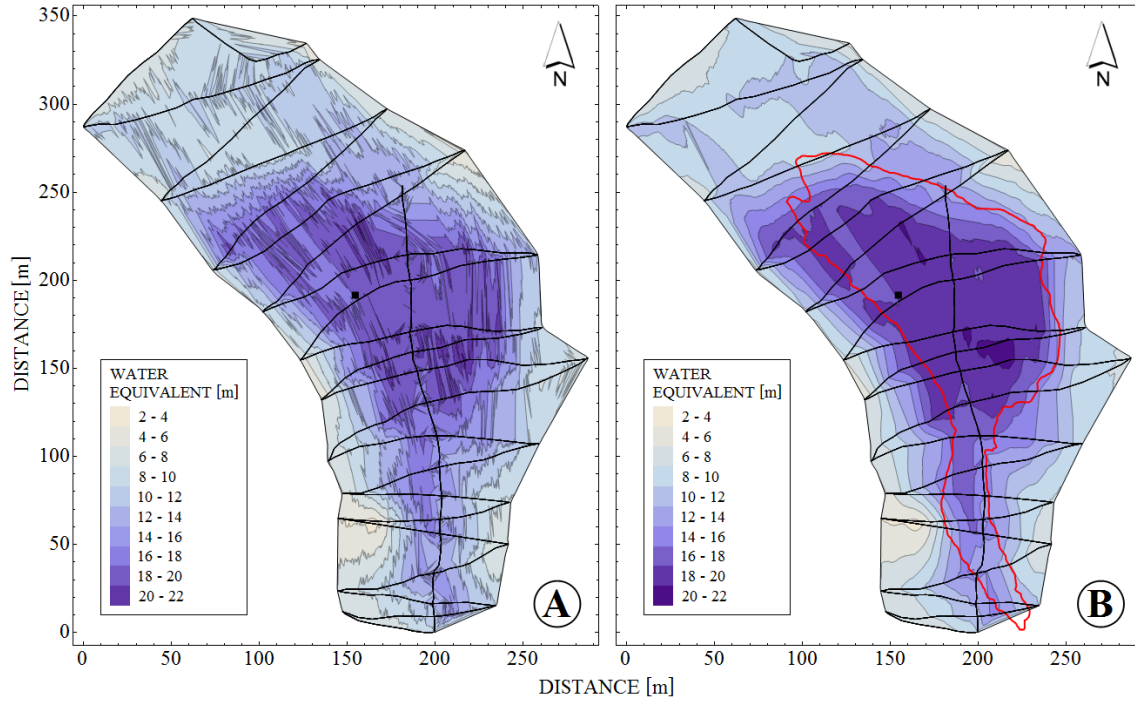


Figure 5.6: Contour plot of the water content of the Prevala glacieret as estimated from the 3-D GPR survey acquired in May 2014. The figure interpolates the water equivalent calculated in each GPR trace (A), and the water equivalent smoothed in each trace using a 8 m wide averaging window moving across each profile (B). The positions of the snow pit (black square) and of the GPR profiles (thick black lines) are superimposed to each graph. The boundaries of the glacieret measured in October 2013 (red line) are also superimposed in (B) for comparison.

P _n	WATER EQUIVALENT [m]			P _n	WATER EQUIVALENT [m]		
	P _n	P1	P _n - P1		P _n	P2	P _n - P2
P11	17.7 ± 2.5	15.3 ± 0.8	2.4	P4	10.8 ± 0.6	11.1 ± 0.9	0.3
P12	19.2 ± 1.8	17.8 ± 1.1	1.4	P5	12.5 ± 0.6	11.8 ± 1.2	0.7
P13	19.2 ± 1.1	18.4 ± 1.4	0.8	P6	13.7 ± 1.1	13.6 ± 1.7	0.1
P14	19.0 ± 1.5	18.6 ± 1.0	0.4	P7	16.9 ± 2.1	12.9 ± 1.2	4.0
P15	19.3 ± 1.2	18.1 ± 1.1	1.2	P8	15.7 ± 2.0	12.8 ± 1.0	2.9
P16	19.3 ± 1.0	18.9 ± 1.3	0.4	P9	15.1 ± 1.2	14.1 ± 0.6	1.0
P17	19.7 ± 2.5	17.4 ± 2.3	2.3	P10	14.9 ± 0.5	16.4 ± 1.0	1.5
P18	12.9 ± 0.9	14.9 ± 3.2	2.0				

Table 5.3: Comparison between the water equivalent values estimated at the fifteen crossing points between profile P1 and the profiles P11-P18 and between profile P2 and the profiles P4-P10. The table shows the mean water equivalent values and their variability range, calculated in each profile using a 11 traces-wide window centered at each crossing point. The table also shows the difference between the mean water equivalent values estimated in each pair of crossing profiles.

Similarly, the difference between the average density values estimated in different profiles at each crossing point tend to increase with depth, with the largest differences in Tables 5.1 and 5.2 respectively equal to 560 kg/m^3 in the fifth layer at the P1-P18 crossing point, and 169 kg/m^3 in the fourth layer at the P2-P7 crossing point. These examples highlight the two main possible causes of discrepancy between density values estimated in crossing GPR profiles, namely density outliers and erroneous trace positioning. In the P1-P18 crossing point, the difference in the inversion results is mainly caused by the overestimated densities in profile P1. In fact, the P1-P18 crossing point is located directly above one of the two darker areas visible in Fig. 3.6 (i.e. the leftmost black triangle), which were discussed in section 3.4, while the other one is slightly to the left of the P1-P17 crossing point (i.e. the second black triangle from the left), which also shows a considerable although lower difference in the estimated densities in Table 5.1 (i.e. 146 kg/m^3 in the sixth layer). In the P2-P7 crossing point, the difference in the estimated densities is instead mainly caused by erroneous positioning due to lacking GPS data for profile P7, whose trail is therefore substituted with a straight line (Fig. 5.2) connecting the end and starting points of the profiles P6 and P8, respectively. Nevertheless, this artificial line does not significantly distort the interpolated results shown in Figs. 5.4 and 5.6.

The discrepancy in the inversion results at the two analyzed crossing points is also visible in Table 5.3, with the average water equivalent varying by 2.0 and 4.0 m at the P1-P18 and P2-P7 crossing points, respectively. However, noticeable differences in the average water equivalent are also visible at the P1-P11, P1-P17, and P2-P8 crossing points, due to the large variability ranges of the densities estimated at these locations (Tables 5.1 and 5.2). Despite possible discrepancies due to local outliers or erroneous positioning, the inversion results in Tables 5.1, 5.2, and 5.3 show an overall good consistency at each crossing points, thus highlighting the stability of the inversion algorithm for different GPR profiles, while the main advantage of the procedure is the large number of recorded traces (i.e. 14142 traces in total for the analyzed 3-D survey) covering the entire surface of the glacieret, with the high data density making the analysis statistically sound.

The glacieret surface considered in the 3-D analysis is limited to the area covered by the GPR survey (Figs. 5.4 and 5.6), which is equal to about $45.3 \cdot 10^3 \text{ m}^2$. In the October 2013 LiDAR survey (Fig. 5.1B) the surface boundaries of the glacieret were clearly identified by the difference in reflectivity between the glacieret surface and the surrounding sediments and bedrock. On the other hand, the entire area was covered by snow during the GPR survey in May 2014, and therefore the surface boundaries of the glacieret were not as clearly marked. The smoothed inversion results shown in Figs. 5.4B and 5.6B were interpolated using a regular grid with a 5 m cell size, in order to estimate the total volume of the respective quantities. The volume of the glacieret was estimated equal to about $97.0 \cdot 10^4 \text{ m}^3$, corresponding to an average 21.4 m thickness, while the volume of the water content was estimated equal to about $62.0 \cdot 10^4 \text{ m}^3$, corresponding to an average 13.7 m water equivalent.

Chapter 6

Conclusion

The proposed auto-picking and inversion algorithm, which was developed and tested during the Ph.D. research, was applied to a 3-D glaciological GPR data set, in order to study the internal stratigraphy, density distribution, and water content of an alpine glacieret. The analysis demonstrated the applicability of GPR as a valuable tool to study the internal structures and properties of the subsurface from both a qualitative and a quantitative point of view. In particular, CO GPR surveys showed several advantages with respect to more traditional direct measurements, in terms of both time and effort spent on data acquisition, as well as data density, with the large number of recorded traces making any quantitative analysis statistically sound. Repeated GPR measurements can also be used to monitor temporal changes in the analyzed system and, in the case of glaciological surveys, study the evolution of a glacier, its internal processes, and estimate its mass balance (Forte et al, 2014b).

The presented auto-picking algorithm uses attribute analysis to quickly and objectively identify the main reflections within the recorded data set, and to define them in terms of their arrival time, peak amplitudes, and polarities. The algorithm is able to track every recorded event characterized by lateral phase continuity, regardless of its reflection strength or lateral amplitude variations; to connect independently picked horizons marking different parts of the same reflection; and to select specific reflection phases to be used for further analysis and interpretation, while disregarding isolated horizons not marking any automatically recognizable structure. During the Ph.D. research, the picking algorithm was shown to perform equally well on various GPR profiles acquired in different environments (Dossi et al., 2015a-b), as well as on reflection seismic data sets (Forte et al., 2016). The algorithm is mostly independent from the interpreter, except for a few required input parameters and thresholds, and it shows several advantages in terms of computing time and accuracy of the results with respect to other commonly used picking techniques. Similarly to other auto-picking algorithms, the procedure can be sensitive to low signal-to-noise ratio as well as signal interference, therefore appropriate processing may be required in order to improve the picking results. Nevertheless, the proposed algorithm can be applied without the need of amplitude recovery, given its independence from the signal amplitudes. Therefore, the original recorded amplitudes can be picked without any subjective alterations caused by the assumption made by the interpreter with regards to the EM properties of the subsurface, which would affect any further quantitative analysis. Furthermore, the algorithm does not rely on any manually selected control point, which makes the picking results more objective.

The presented inversion algorithm uses the amplitudes and travel times picked along the recorded reflections in a GPR profile to iteratively recover the EM velocity and thickness of each interpreted

layer, by reconstructing the travel path of each reflected signal. In glaciological surveys, the calculated EM velocities can then be used to estimate the densities of the air-ice mixtures within these layers using well-known empirical relations. The accuracy of the algorithm was initially tested on synthetic GPR data sets and the inversion results showed an overall good consistency with the initial model. The inversion algorithm can be sensitive to amplitude distortions due to noise or interference, which can affect the estimated impedance contrasts and possibly lead to local unrealistic results. Moreover, possible errors in the input quantities tend to propagate to the deeper layers due to the iterative nature of the inversion algorithm. However, the large number of traces in a GPR profile contributes to the stability of the final model, and while sudden changes or even local spikes in the reconstructed stratigraphy, density distribution, and water content can be expected, since the inversion is applied to each GPR trace independently, the global trends are nevertheless preserved. The inversion algorithm can be applied to any GPR data set that satisfies the assumptions made by the procedure, therefore intrinsic signal attenuation caused by electric conductivity or dielectric relaxation must be taken into consideration in environments where high-loss materials can be expected, while signal distortions due to complex stratigraphy, scattering objects, and noise sources can also affect the inversion results.

In case all the assumptions of the algorithm are satisfied, the two main uncertainty factors for the inversion results are the input EM velocity for the shallowest layer, and the reference and reflected amplitudes. The accuracy of the input EM velocity mainly depends on the availability and spatial distribution of locally sampled values, which can be estimated through either CMP analysis or, in the case of glaciological GPR surveys, direct density measurements from snow pits. In the analyzed 3-D survey a detailed EM velocity distribution for the shallowest layer was not available, and therefore a constant EM velocity was used, the implication of which have been discussed. The accuracy of the input amplitudes mainly depends on the accuracy of the picking algorithm in selecting the main phase in each reflection, the quality of the analyzed GPR data sets in terms of signal-to-noise ratio or possible interference, as well as the amplitude recovery method used during the inversion process itself. Nevertheless, another important parameter is the sampling rate used during data acquisition, which can cause significant signal distortions due to unrecoverable data loss, thus altering the peak amplitudes used during the inversion. The sampling analysis was applied to both synthetic and real GPR data sets and the results suggest, as a rule of thumb, that the sampling rate has to be equal to at least 12 times the signal central frequency, which is higher than the commonly adopted standards, in order to limit the peak amplitude error within 5%. These results can be useful for any quantitative amplitude analysis, and they can be applied to both GPR and reflection seismic data sets.

The analyzed 3-D glaciological data set was acquired in May 2015, and it comprises 25 GPR profiles covering the entire surface of the Prevala glacieret. The auto-picking algorithm was able to identify and isolate the main reflections within each profile, and recover their amplitude and travel times to be used as input for the inversion process. The inversion procedure was then independently applied to each GPR trace in the 3-D survey, and the inversion results were interpolated between profiles in order to recover the internal stratigraphy and density distribution of the glacieret, as well as its total volume and water content. The picking and inversion results were able to highlight the main interface separating the deeper part consisting of the glacieret that remained after the previous summer melting period, and the shallower part made of the materials that accumulated during the latest winter period. The results were also able to highlight the internal layering of these two glaciological units, which is characterized by weaker impedance contrasts, corresponding to smaller density changes. The stability of the inversion procedure was tested at the points of intersection between profiles, showing an overall good consistency between the density values estimated in each interpreted layer, as well as the calculated water equivalent. The interpolated 3-D results were also consistent with available

topography data, with the reconstructed thickness model being able to identify the mountain sides narrowing toward the higher southern tip, as well as the broader moraine enclosing the glacieret on the northern side. These results show the validity of GPR survey as a tool for glaciological monitoring, since it is able to provide statistically sound models for the total volume and water content of glaciers and snowfields, which provides useful applications in several fields.

Acknowledgment

The Ph.D. research was partly supported by the “Finanziamento di Ateneo per progetti di ricerca scientifica - FRA 2014” of the University of Trieste. We kindly thank Jury Agostinis, Marco Basso Bondini, Giacomo Cigoi, Roberto R. Colucci, Costanza del Gobbo, Daniele Fontana, Veronica Franco, Nicholas Gregori, Daniel Markovic, Davide Martinucci, Laura Miele, Simone Pillon, Michele Potleca, Giacomo Toffolutti, and Carlotta Zanettini for assisting in the field operations. Significant logistical support has been provided by the Parco Naturale Regionale delle Prealpi Giulie, Promotur FVG, and the Ufficio Neve e Valanghe FVG, which are gratefully acknowledged. We also thank the professors Maksim Bano and Luigi Zanzi for their external review of this thesis.

Bibliography

- [1] Annan A.P., Cosway S.W., and Sigurdsson T., “GPR for snow pack water content”, in Proc. 5th International Conference on Ground Penetrating Radar (GPR 1994), Kitchener, Ontario, Canada, June 12-16, 1994, pp. 465-475
- [2] Annan A.P. and Cosway S.W., “Simplified GPR beam model for survey design”, in Proc. 62nd SEG Annual Meeting, New Orleans, Louisiana, USA, October 25-29, 1992, pp. 356-359
- [3] Barnes A.E., “A tutorial on complex seismic trace analysis”, *Geophysics*, vol. 72, no. 6, 2007, pp. W33-W43
- [4] Barnes A.E., “Theory of 2-D complex seismic trace analysis”, *Geophysics*, vol. 61, no. 1, 1996, pp. 264-272
- [5] Birchak J.R., Gardner C.G., Hipp J.E., and Victor J.M., “High dielectric constant microwave probes for sensing soil moisture”, in Proc. IEEE, vol. 62, no. 1, 1974, pp. 93-98
- [6] Bradford J.H., Harper J.T., and Brown J., “Complex dielectric permittivity measurements from Ground-Penetrating Radar data to estimate snow liquid water content in the pendular regime”, *Water Resources Research*, vol. 45, no. 8, 2009, pp. 1-12
- [7] Bradford J.H. and Harper J.T., “Wave field migration as a tool for estimating spatially continuous radar velocity and water content in glaciers”, *Geophysical Research Letters*, vol. 32, no. 8, 2005, pp. 1-4
- [8] Chopra S. and Marfurt K.J., “Emerging and future trends in seismic attributes”, *The Leading Edge*, vol. 27, no. 3, 2008, pp. 298-318
- [9] Chopra S. and Marfurt K.J., “Seismic attributes – A historical perspective”, *Geophysics*, vol. 70, no. 5, 2005, pp. 3SO-28SO
- [10] Colucci R.R., Forte E., Boccali C., Dossi M., Lanza L., Pipan M., and Guglielmin M., “Evaluation of internal structure, volume and mass of glacial bodies by integrated LiDAR and Ground Penetrating Radar surveys: the case study of Canin Eastern Glacieret (Julian Alps, Italy)”, *Surveys in Geophysics*, vol. 36, no. 2, 2015, pp. 231-252
- [11] Davis J.L. and Annan A.P., “Ground-Penetrating Radar for high resolution mapping of soil and rock stratigraphy”, *Geophysical Prospecting*, vol. 37, no. 5, 1989, pp. 531-551
- [12] Dorn G.A., “Modern 3-D seismic interpretation”, *The Leading Edge*, vol. 17, no. 9, 1998, pp. 1262-1269

- [13] Dossi M., Forte E., Pipan M., Colucci R.R., and Bortoletto A., "Automated reflection picking and inversion: application to ground and airborne GPR surveys", in Proc. IEEE 16th International Conference on Ground Penetrating Radar, Hong Kong, June 13-16, 2016, pp. 1-6, ISBN 978-1-5090-5181-6
- [14] Dossi M., Forte E., and Pipan M., "Automated reflection picking and polarity assessment through attribute analysis: Theory and application to synthetic and real ground-penetrating radar data", *Geophysics*, vol. 80, no. 5, 2015a, pp. H23-H35
- [15] Dossi M., Forte E., and Pipan M., "Auto-picking and phase assessment by means of attribute analysis applied to GPR pavement inspection", in Proc. IEEE 8th International Workshop on Advanced Ground Penetrating Radar, Florence, Italy, July 7-10, 2015b, pp. 1-4, ISBN 978-1-4799-6495-6
- [16] Drijkoningen G.G., "Seismic data acquisition", TA3600, Delft University of Technology, Section Applied Geophysics & Petrophysics, 2003, 115 pp.
- [17] Forte E., Dossi M., Pipan M., and Del Ben A., "Automated phase attribute-based picking applied to reflection seismics", *Geophysics*, vol. 81, no. 2, 2016, pp. V141-V150
- [18] Forte E., Dossi M., Pipan M., and Colucci R.R., "Velocity analysis from common offset GPR data inversion: theory and application to synthetic and real data", *Geophysical Journal International*, vol. 197, no. 3, 2014a, pp. 1471-1483
- [19] Forte E., Dossi M., Colle Fontana M., and Colucci R.R., "4-D quantitative GPR analyses to study the summer mass balance of a glacier: a case history", in Proc. IEEE 15th International Conference on Ground Penetrating Radar, Brussels, Belgium, June 30-July 4, 2014b, pp. 352-356, ISBN 978-1-4799-6789-6
- [20] Forte E., Dossi M., Colucci R.R., and Pipan M., "A new fast methodology to estimate the density of frozen materials by means of common offset GPR data", *Journal of Applied Geophysics*, vol. 99, 2013, pp. 135-145
- [21] Forte E. and Pipan M., "Review of multi-offset GPR applications: Data acquisition, processing and analysis", *Signal Processing*, vol. 132, 2017, pp. 210-220
- [22] Giannopoulos A., "Modelling ground penetrating radar by GprMax", *Construction and Building Materials*, vol. 19, no. 10, 2005, pp. 755-762
- [23] Godio A., "Georadar measurements for the snow cover density", *American Journal of Applied Sciences*, vol. 6, no. 3, 2009, pp. 414-423
- [24] Harper J.T. and Bradford J.H., "Snow stratigraphy over a uniform depositional surface: spatial variability and measurement tools", *Cold Regions Science and Technology*, vol. 37, no. 3, 2003, pp. 289-298
- [25] Herron D.A., "Horizon autopicking", *The Leading Edge*, vol. 19, no. 5, 2000, pp. 491-492
- [26] Jiao Y., McMechan G.A., and Pettinelli E., "In situ 2-D and 3-D measurements of radiation patterns of half-wave dipole GPR antennas", *Journal of Applied Geophysics*, vol. 43, no. 1, 2000, pp. 69-89

- [27] Jol H.M., "Ground penetrating radar: Theory and applications", Elsevier Science Ltd., 2009, 524 pp., ISBN 978-0-444-53348-7
- [28] Kovacs A., Gow A.J., and Morey R.M., "The in-situ dielectric constant of polar firn revisited", *Cold Regions Science and Technology*, vol. 23, no. 3, 1995, pp. 245-256
- [29] Kuffel E., Zaengl W.S., and Kuffel J., "High voltage engineering: Fundamentals", Newnes, II ed., 2000, 534 pp, ISBN 0-7506-3634-3
- [30] Looyenga H., "Dielectric constants of heterogeneous mixtures", *Physica*, vol. 31, no. 3, 1965, pp. 401-406
- [31] Lundberg A., Richardson-Näslund C., and Andersson C., "Snow density variations: consequences for ground-penetrating radar", *Hydrological Processes*, vol. 20, no. 7, 2006, pp. 1483-1495
- [32] Moore J.C., Pälli A., Ludwig F., Blatter H., Jania J., Gadek B., Glowacki P., Mochnacki D., and Isaksson E., "High-resolution hydrothermal structure of Hansbreen, Spitsbergen, mapped by Ground-Penetrating Radar", *Journal of Glaciology*, vol. 45, no. 151, 1999, pp. 524-532
- [33] Murray T., Booth A., and Rippin D.M., "Water-content of glacier-ice: limitations on estimates from velocity analysis of surface ground-penetrating radar surveys", *Journal of Environmental and Engineering Geophysics*, vol. 12, no. 1, 2007, pp. 87-99.
- [34] Nyquist H., "Certain topics in telegraph transmission theory", *Transactions of the AIEE*, 1928, pp. 617-644
- [35] Previati M., Godio A., and Ferraris S., "Validation of spatial variability of snowpack thickness and density obtained with GPR and TDR methods", *Journal of Applied Geophysics*, vol. 75, no. 2, 2011, pp. 284-293
- [36] Richardson S.D. and Reynolds J.M., "An overview of glacial hazards in the Himalayas", *Quaternary International*, vol. 65-66, 2000, pp. 31-47
- [37] Robin G. de Q., "Velocity of radio waves in ice by means of a borehole interferometric technique", *Journal of Glaciology*, vol. 15, 1975, pp. 151-159
- [38] Schon K., "High impulse voltage and current measurement techniques: Fundamentals - Measuring instruments - Measuring methods", Springer Publishing, 2013, 280 pp, ISBN 978-3-319-00377-1
- [39] Seybold J.S., "Introduction to RF Propagation", Wiley-Interscience, 2005, 330 pp., ISBN 978-0-471-65596-1
- [40] Shannon C.E., "Communication in the presence of noise", *Proceedings of the IRE*, vol. 37, no. 1, 1949, pp. 10-21
- [41] Stiles W.H. and Ulaby F.T., "Dielectric Properties of Snow", Technical Report, Center for Research Inc., University of Kansas, 1981, pp. 91-103
- [42] Strange A.D., "Analysis of time interpolation for enhanced resolution GPR data", in *Proc. IEEE 7th International Workshop on Advanced Ground Penetrating Radar*, Nantes, France, July 2-5, 2013, pp. 1-5, ISBN 978-1-4799-0940-7

- [43] Taner M.T., Koehler F., and Sheriff R.E., "Complex seismic trace analysis", *Geophysics*, vol. 44, no. 6, 1979, pp. 1041-1063
- [44] Tillard S. and Dubois J.C., "Analysis of GPR data: wave propagation velocity determination", vol. 33, no. 1-3, 1995, pp. 77-91
- [45] Zemp M., Hoelzle M., and Haeberli W., "Six decades of glacier mass-balance observations: a review of the worldwide monitoring network", *Annals of Glaciology*, vol. 50, no. 50, 2009, pp. 101-111