



Shallow Seismic Refraction: Identification and Automatic Interpretation of Arrival Times

Klaus Contiero, Liliana Alcazar Diogo

Copyright 2019, SBGf - Sociedade Brasileira de Geofísica

This paper was prepared for presentation during the 16th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 19-22 August 2019.

Contents of this paper were reviewed by the Technical Committee of the 16th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

Automation of human activities through computational methods is essential in today's world, whether in the academic world, in the corporate world or in everyday tasks, and in geophysical processing is not different.

Much of the time spent processing a shallow refraction seismic study is spent on picking for first break identification, which, despite technological advances, remains dependent on the visual evaluation of the seismic interpreter. Thus, the automation of this process provides a more efficient result and provides more time for interpretation.

In this work we used the peak detection program (Z-score). The application of the Z-score showed automatic picking efficiency, identifying 88% of the first breaks, with 92% of the detections having an error less or equal to 2%.

In addition, a method for interpreting seismic events, identifying the velocities and thicknesses of the layers, was successful: Simultaneous Inversion.

Thus, several automations in the processes of the seismic refraction method have proved to be satisfactory in real data.

Introduction

The geophysical method of shallow seismic refraction is widely used because it is simple and cheap, and has several applications, for example: determining the thickness of the sedimentary cover on the basement, to designate the amount of material that needs to be removed before mining (GOUTLY & BRABHAM, 1984), to quantify the unconsolidated material in certain geological deposits (LAWTON & HOCHSTEIN, 1993), to determine and delineate paleoceanals (PAKISER & BLACK, 1957), to indicate the thickness of the regolith layer in non-weathered soils (DENTITH et al., 1992), among others.

However, much of the time spent in processing a shallow refraction seismic study is spent with the identification of first breaks, which, despite the technological advance in seismic prospecting, still remains dependent on the visual evaluation of the seismic interpreter. Thus, the automation of this process can provide a more efficient

result and allow a faster return of data interpretation and geological features.

The improvements in computational resources have allowed to deal with the large volume of seismic data, however, three processes still have their analysis done visually: determination of first breaks, editing of traces with noise and speed analysis. These processes consume about 75% of the time of the seismic interpreter, although they use less than 50% of the existing computational resources (McCORMACK et al., 1993).

The difficulties in the manual process of identification of the first arrival times, called picking, are related to several effects, such as human operator experience, operator sensitivity and capacity, image scale, cursor sensitivity and signal-noise ratio (SENKAYA & KARSH, 2014), being a time-consuming and highly subjective process. Many algorithms exist for automation of this process (ERVIN, et al., 1983; NEAG & WYATT, 1986; SPAGNOLINI, 1991), since they require a lot of reliable data and when the signal-noise ratio changes, an answer not acceptable to certain interpreters (MIRANDA, 2000).

A technique that can be used for automatic picking is the peak detection through the dispersion principle, where we identify as a peak a value that is a number of standard deviations from a moving average, called Z-score (BRAKEL, 2017).

In this project we intend to evaluate the use of the Z-score to track the times of the first arrivals. Another goal is to automate the separation of each seismic event. Thus, identified the first breaks, the events are grouped in direct wave, first refraction, second refraction and so on, through a numerical method, which can be very useful to speed the interpretation in the refraction method. For this we use a simultaneous inversion of all seismic events.

Method

We adopted the Z-score algorithm, which calculates the moving average for a given number of points and if the next point is a certain number of standard deviations of this average, is considered a peak. In this work we use 30 points to calculate the moving average and if the next point is 5 deviations from the mean, it is considered a peak.

The mean x_m of a set of values $\{x_i\}$ is determined by $x_m = \sum_{i=1}^N \frac{x_i}{N}$. The sum of the quadratic deviations σ^2 , called variance, is defined as follows $\sigma^2 = \sum_{i=1}^N \frac{(x_i - x_m)^2}{N}$. And the standard deviation σ of the set of values is determined by the positive square root of the variance, $\sigma = +\sqrt{\sigma^2}$.

To automatically identify the separation of seismic events, we proposed to perform a simultaneous inversion of all events, direct wave and refractions. The inversion unknowns are the linear and angular coefficients that adjust to the lines, thus determining the velocity of the layers and their thicknesses.

The angular coefficients, b , of the lines are equal to the inverse of the velocities of the layers and the linear coefficient, a_2 , is known as the intersection time, the crossing point on the vertical axis.

Thus, for the first layer the space equation we have $y_1 = b_1 \cdot x$, where y_1 represents the time and $b_1 = \frac{1}{V_1}$. And for the second layer the equation will be $y_2 = a_2 + b_2 \cdot x$, where $b_2 = \frac{1}{V_2}$ and a_2 the intersection time. We wish to find the x point of encounter of the lines according to the unknowns of the model, therefore: $x = a_2 \left(\frac{V_1 V_2}{V_2 - V_1} \right)$.

The inversion problem was solved by the least squares method, implemented in Python through the Least-square fitting (leastsq) of the Scipy library (JONES et al., 2001). For an goal function $y = f(x, p)$ we want to fit a set of data $\{x_i, y_i\}$, minimizing the sum of the errors $E(p)$ defined by $E(p) = \sum_{i=0}^{N-1} e_i^2(p)$, where $e_i = |y_i - f(x_i, p)|$.

Initially, we used a synthetic model to validate the proposed inversion. The algorithm proved to be efficient and robust for use with real data.

Results

We used a seismogram for testing with real data from a study in the municipality of Termas de Ibirá-SP (RUIZ, 2014) and the Z-score method for detection of peaks presented good results for automatic detection, as shown in Figure 1.

Table 1 shows some detection results and error in relation to manual picking, and the proximity of the marked points automatically and manually can be visualized by plotting the times together, as shown in Figure 2.

Initially, we used a synthetic model to test the validity of the hypothesis and the robustness of the algorithm. For a model with velocities $V_1 = 500$ m / s, $V_2 = 1200$ m / s, thickness of the first layer $h_1 = 10$ m and intersection time of $t_i = 0,0363623737$ s, we fed the algorithm with synthetic arrival times, as shown in Figure 3, and a priori different values V_1 , V_2 and, obtaining precisely, after the inversion, the values $V_1 = 500$ m / s, $V_2 = 1200$ m / s and $t_i = 0,0363623737$ s.

After testing the robustness of the method, we use the same principle in the real time arrival data, shown in Figure 4.

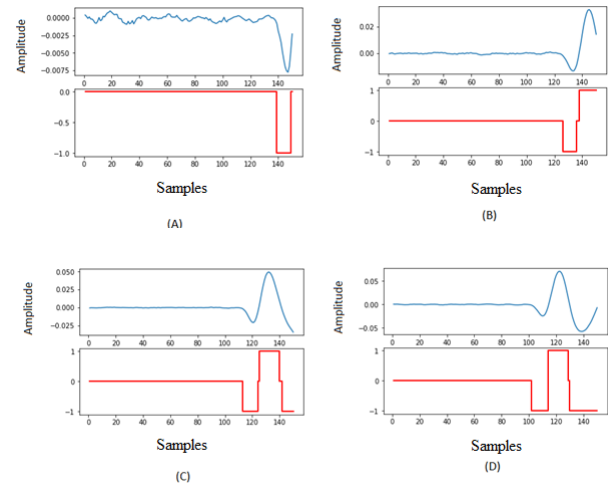


Figure 1 – Z-score results for some traces, with trait and sample detected as peak, respectively: (A) 20 and 140; (B) 35 and 127; (C) 50 and 114; (D) 60 and 103. In red we have the numerical result of the detection, where the samples detected as peaks have a value of 1, positive peak, or -1, negative peak.

Table 1 – Difference between automatic and manual picks. The first traits have a large error, which decreases with the increase of the offset to the source.

Trace	Manual Pick (s)	Automatic Pick (s)	Diference (s)	Diference (%)
94	0,003507	0,0185	-0,014993	81%
93	0,007971	0,0180	-0,010029	56%
92	0,010522	0,0155	-0,004978	32%
91	0,013391	0,0170	-0,003609	21%
90	0,016899	0,0205	-0,003601	18%
89	0,019768	0,0225	-0,002732	12%
88	0,023276	0,0260	-0,002724	10%
87	0,024551	0,0270	-0,002449	9%
86	0,030866	0,0305	0,000366	1%
85	0,033385	0,0345	-0,001115	3%
84	0,036115	0,0375	-0,001385	4%
83	0,039054	0,0405	-0,001446	4%
82	0,039474	0,0410	-0,001526	4%
81	0,039222	0,0420	-0,002778	7%
80	0,040419	0,0420	-0,001581	4%
79	0,04012	0,0420	-0,00188	4%

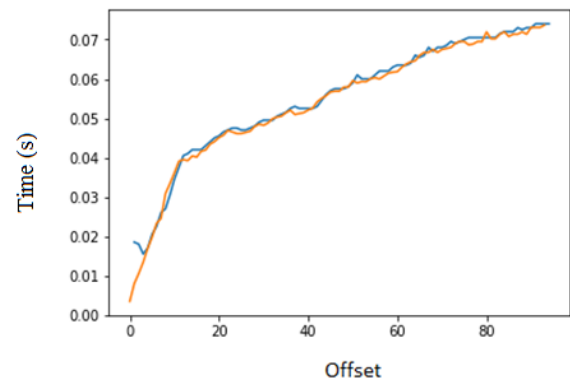


Figure 2 – Manual and Automatic Picks plotted together. In orange we have the manual pick and in blue, the automatic.

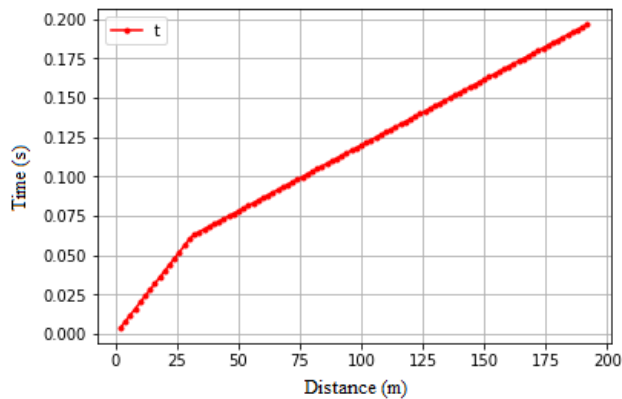


Figure 3 – Synthetic model used in the inversion, the points represent the first breaks and the time values are used in the inversion to obtain $V_1 = 500$ m/s, $V_2 = 1200$ m/s and $t_i = 0.0363623737$ s.

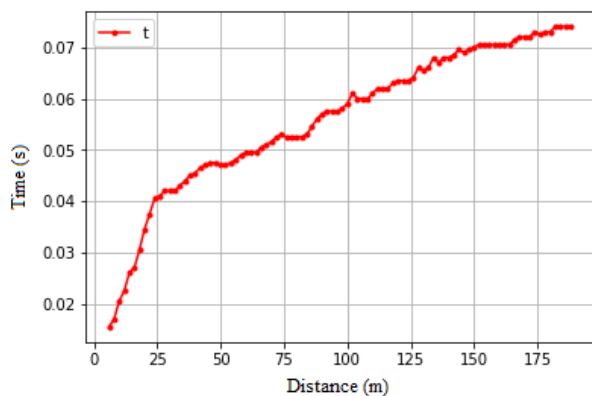


Figure 4 – Real data used in the inversion, the points represent the first breaks obtained by the automatic detection, the time values are used in the inversion to obtain $V_1 = 596$ m/s, $V_2 = 4735$ m/s and $t_i = 0.037037154$ s. From these values we estimate the thickness of the first layer $h_1 = 11,12555795$ m.

Conclusions

Some proposals for automatic picking through neural networks (ARAUJO, 2015) are effective, but they produce many false positives, that is, identification of peaks where they did not exist, which require more processing to provide results. The Z-score method, besides being easy to implement, showed an advantage over other proposals, as there are no false positives. Also, since we always choose the first peak of the trace, there is no chance of more than one value per trace, which would also be a false positive. Another advantage of this method is that we could identify the first effective arrival, rather than the greater positive amplitude of the first arrival, so no correction needs to be performed.

Early traces are difficult to detect, even manually. In these traces there was a big difference between automatic and manual detections. This is due to the great influence of surface waves (ground roll) in the direct wave, which have lower velocity and do not influence offsets further away

from the source. The data that were not identified are the noisy ones, where there is a high amplitude variation before the first break, so the number of standard deviations of the desired peak for the moving average is different from the standard of the data as a whole, or where the noise has amplitude close to the first break signal.

The simultaneous inversion model provided a good fit and, as shown to be valid and robust for synthetic data, we assume that it is a reliable method. From the inversion we find in an automated way the velocity of the layers and their thicknesses.

Simulating the visual behavior for human decision taking proved to be an arduous task, but despite the difficulties, an automation in a real data was shown to be viable, providing time for the seismic interpreter and the proposal was reached.

Acknowledgments

Sociedade Brasileira de Geofísica (SBGf).

References

- ARAUJO, A.S. Rastreamento automático das primeiras quebras em dados sísmicos. Trabalho de Graduação, Universidade de São Paulo, São Paulo, 2015.
- BRAKEL, J.P. "Peak signal detection in realtime timeseries data," 2017. Recuperado em 24 de novembro de 2018, de <https://stackoverflow.com/a/22640362>.
- DENTITH, M.C., EVANS, B.J., PAISH, K.F. & TRENCH, A. Mapping the regolith using seismic refraction and magnetic data: Results from the Southern Cross Greenstone Belt, Western Australia. *Exploration Geophysics*, 23, 97–104, 1992.
- ERVIN, C.P., MCGINNIS, L.D., OTIS, R.M. & HALLI, M.L. Automated analysis of marine refraction data: A computer algorithm. *Geophysics*, 48 (5), 582–589, 1983.
- GOULTY, N.R., & BRABHAM, P. J. Seismic refraction profiling in opencast coal exploration. *First Break*, 2 (5), 26–33, 1984.
- JONES, E., OLIPHANT, E., PETERSON, P., et al. SciPy: Open Source Scientific Tools for Python. Recuperado em 24 de novembro de 2018, de <http://www.scipy.org/>, , 2001.
- LAWTON, D.C. & HOCHSTEIN, M.P. Physical properties of titanomagnetite sands, *Geophysics*, 45, 394–402, 1993.
- McCORMACK, M.D., ZAUCHA, D.E., DUSHEK, D.W. First-break refraction event picking and seismic data trace editing using neural networks. *Geophysics*, 58, 67–78, 1993.
- MIRANDA, A.I.F. Identificação automática das primeiras quebras em traços sísmicos por meio de uma rede neural, Dissertação de mestrado, Universidade Federal do Pará, Belém, 2000.

RUIZ, D.M.G. Ensaios de sísmica rasa multicomponente em Termas de Ibirá – SP, Trabalho de Graduação, Universidade de São Paulo, São Paulo, 2014.

PAKISER, L.C. & BLACK, R.A. Exploring for ancient channels with the refraction seismograph. *Geophysics*, 22, 32–47, 1957.

SENKAYA, A.M. & KARSLY, Y.H. A semi-automatic approach to identify first arrival time: the cross-correlation technique (CCT). *Earth Science Research Journal*, 18 (2), 107–113, 2014.

SPAGNOLINI, U. Adaptive picking of refracted first arrivals. *Geophysics Prospecting*, 39, 293–312, 1991.