



## Rede neural artificial para picking automático e aprendizado supervisionado para diferenciação dos eventos sísmicos

Contiero, K. \*, IAG-USP, Araujo, A. S., IAG-USP, Diogo, L. A., IAG-USP

Copyright 2018, SBGf - Sociedade Brasileira de Geofísica

Este texto foi preparado para a apresentação no VIII Simpósio Brasileiro de Geofísica, Salinópolis, 18 a 20 de setembro de 2018. Seu conteúdo foi revisado pelo Comitê Técnico do VIII SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.

### Abstract

Much of the time processing a shallow refraction seismic study is spent on picking for first break identification, which, despite the technological advancement in seismic prospecting, remains dependent on the visual evaluation of the seismic interpreter. Thus, the automation of this process provides a more efficient result and allows more time for the interpretation of data and geological features. The objective of the present project was to develop a multilayer perceptron algorithm supervised in python language that performs automatic picking and, also, to develop a method to distinguish each seismic event - direct wave and refractions - through least squares method and Chi<sup>2</sup> analysis corresponding to each event. The program proved to be efficient, since the accuracy level is above 90%. The implemented methods are being evaluated in different types of real data acquired in research projects developed by the Applied Geophysics group of IAG / USP.

### Introdução

O método geofísico de refração sísmica rasa é muito utilizado por ser simples e barato, e possui várias aplicações, dentre elas: determinar a espessura da cobertura sedimentar sobre o embasamento, seja para designar a quantidade de material que é necessário se retirar antes da mineração (GOUTLY & BRABHAM, 1984), para quantificar o material não consolidado em determinados depósitos geológicos (LAWTON & HOCHSTEIN, 1993), para determinar e delinear paleocanais (PAKISER & BLACK, 1957) ou indicar a espessura da camada de regolito em terrenos não intemperizados (DENTITH et al., 1992). Porém, grande parte do tempo gasto em um processamento de um ensaio de sísmica de refração rasa é despendido com a identificação do tempo de chegada da onda direta (*first break*), que, apesar do avanço tecnológico ocorrido na prospecção sísmica, permanece, ainda, dependente da avaliação visual do intérprete sísmico. Assim, a automação deste processo proporciona um resultado mais eficiente e permite mais tempo à interpretação dos dados e das feições geológicas.

Devido a estudos sísmicos tridimensionais (3-D) e linhas sísmicas bidimensionais (2-D) de alta densidade, o volume de dados tem crescido exponencialmente, o que foi assimilado pelas melhorias nas velocidades e

capacidades dos computadores. Entretanto, três processos ainda têm suas análises feitas de forma visual: determinação de primeiras chegadas (*first-break*), edição de traços com ruído e análise de velocidades. Estes processos consomem cerca de 75% do tempo do intérprete sísmico, embora utilizem menos de 50% dos recursos computacionais existentes (McCORMACK et al., 1993). As dificuldades no processo manual de identificação do tempo de chegada da onda direta, denominado *picking*, estão relacionadas a vários efeitos, como a experiência do operador humano, sensibilidade e capacidade dos olhos do operador, a escala da imagem e dados de amplitude, a sensibilidade do cursor, e SNR - razão sinal-ruído - (SENKAYA & KARSH, 2014), sendo um processo demorado e altamente subjetivo. Muitos algoritmos existentes para automação deste processo (ERVIN, et al., 1983; NEEF & WYATT, 1986; SPAGNOLINI, 1991) apresentam uso limitado, já que precisam de muitos dados confiáveis e, quando se muda a razão sinal ruído, podem produzir uma resposta não aceitável para determinados intérpretes (MIRANDA, 2000).

Utilizamos uma técnica de redes neurais artificiais para criação de um algoritmo para *picking* automatizado, que levará a um ganho de qualidade na interpretação, uma vez que a maior parte do tempo do processamento poderá ser empregado em casos mais complexos. Outro objetivo do projeto foi determinar cada evento sísmico. Assim, identificado o *first break*, os eventos são agrupados em onda direta, primeira refração, segunda refração e assim por diante, através de um método numérico, o que pode ser de grande utilidade para agilizar a interpretação no método de refração.

### Metodologia

Para a construção de uma boa representação da subsuperfície pelo Método de Refração Sísmica, é necessário precisão nos pontos, para que seja possível a interpretação dos tempos de *first breaks* (diretas e refratadas), ou seja, um bom *picking*.

Assim, um método computacional que simule o comportamento visual e tomada de decisões de um intérprete humano é encontrado na tecnologia neurocomputacional, inicialmente desenvolvida por neurobiologistas como modelos de entendimento do sistema nervoso humano, e destaca-se das técnicas computacionais convencionais por sua habilidade de se adaptar e aprender através de repetida exposição a exemplos, interpolação no caso de falta de dados e robustez no processamento de dados ruidosos (MIRANDA, 2000).

Neste projeto, foi desenvolvido um algoritmo em linguagem python do tipo perceptron multicamadas com retropropagação de erro, aprendizado supervisionado e alimentação à frente, determinando os melhores parâmetros para a rede – número de camadas, neurônios por camada, função de ativação, número de dados para treinamento – e sua aplicabilidade para *picking* automatizado em dados ruidosos. A saída da unidade de processamento é a combinação linear das entradas, armazenada num vetor, multiplicadas por pesos pré-existentes.

O traço é dividido em janelas temporais normalizadas, com as amplitudes da janela divididas pela de maior módulo dentro do intervalo. Para  $n$  dados de entrada, a rede percorrerá um traço sísmico e deve retornar um valor próximo de 1 (saída positiva) quando o *first break* estiver na  $m$ -ésima amostra e retornar um valor próximo de 0 (saída negativa) para todas as outras janelas. A posição da  $m$ -ésima amostra é escolhida dentro da janela, conforme Figura 1.

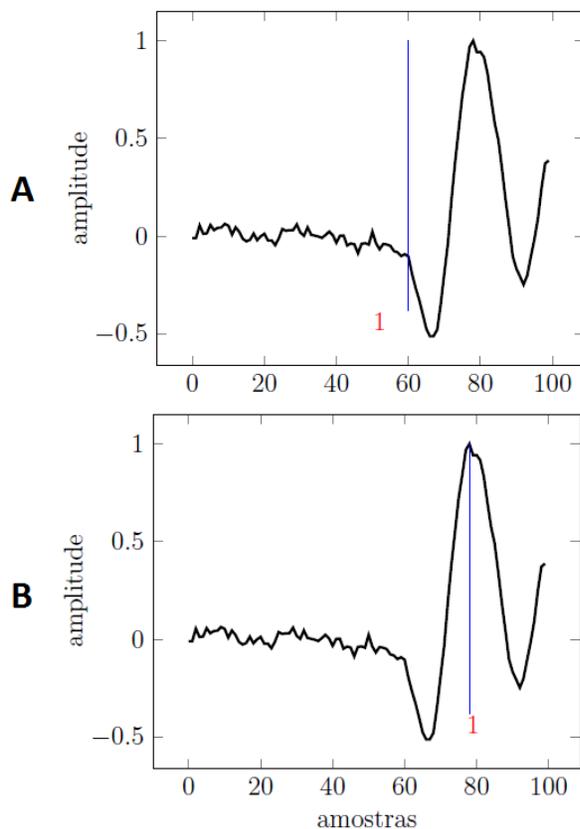


Figura 1 – Exemplo de traço sísmico com  $n = 100$  entradas e saída positiva (traço azul) na  $m$ -ésima amostra (A)  $m = 60$ , primeira chegada (B)  $m = 75$ , maior amplitude do pulso da primeira chegada (Adaptado de ARAUJO, 2015).

Inicialmente, os pesos são aleatórios, sendo ajustados após cada iteração pelo algoritmo de retropropagação do erro, até que o erro esteja dentro do aceitável (no caso, entre  $10^{-1}$  e  $10^{-2}$ ). O número de dados de entrada,  $n$ , é

modificado até se obter o melhor parâmetro, assim como o número de camadas ocultas e o número de neurônios em cada camada, como mostra a Figura 2.

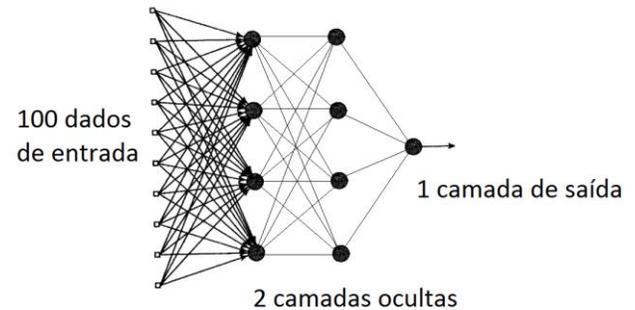


Figura 2 – Exemplo de estrutura da rede para 100 dados de entrada e duas camadas ocultas (Adaptado de MIRANDA, 2000).

Para identificar os eventos sísmicos a partir dos tempos, a cada conjunto de 5 pontos, deslocados de um em um (1 a 5, 2 a 6, 3 a 7, ...), ajustamos uma reta pelo método dos mínimos quadrados e calculado seu desvio quadrático ( $\chi^2$ ). Quando um ajuste é bom, o  $\chi^2$  é pequeno (VUOLO, 1996), então, quando os dados forem de um mesmo evento, espera-se que isto seja satisfeito. Quando o  $\chi^2$  não mais seguir esta regra, significará que se trata de um outro evento sísmico. Portanto, um outro script será criado para delimitar os eventos, além da RNA que identificará o *first break*.

### Resultados

Utilizamos dados reais provenientes de um estudo no município de Termas de Ibirá-SP (Ruiz, 2014). Para os  $n$  dados de entrada, janelas muito pequenas ( $n < 50$ ) ou muito grandes ( $n > 200$ ) não produzem a resposta esperada, e as diferenças para valores entre 75 e 125 não são consideráveis, portanto a janela escolhida para a quantidade de valores de entrada foi  $n = 100$ . Observamos que quanto mais central está a saída positiva, melhor é o ajuste, além disso, estar associada ao primeiro máximo é mais facilmente identificada do que no início do pulso. Por fim, duas camadas ocultas são o necessário para identificar bem funções contínuas (CYBENKO, 1988).

Determinados os melhores parâmetros e a melhor arquitetura para a rede neural, treinou-se a rede com 25 traços e foi possível obter um *picking* compatível com o de um operador humano, como mostra a Figura 3.

Com os valores dos tempos realizamos a separação dos eventos, através da análise do  $\chi^2$ , variando de 5 em 5 os pontos analisados, conforme Figura 4.

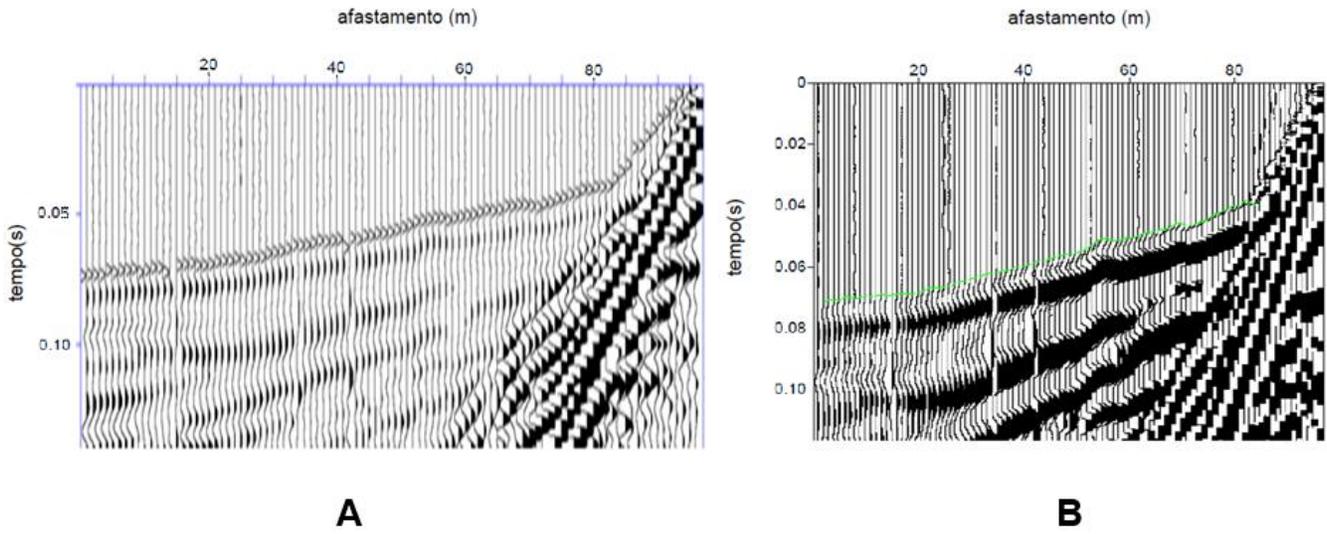


Figura 3 – Comparação entre o *picking* manual feito no *suxpicker* em (A) e o realizado pela rede neural em (B).

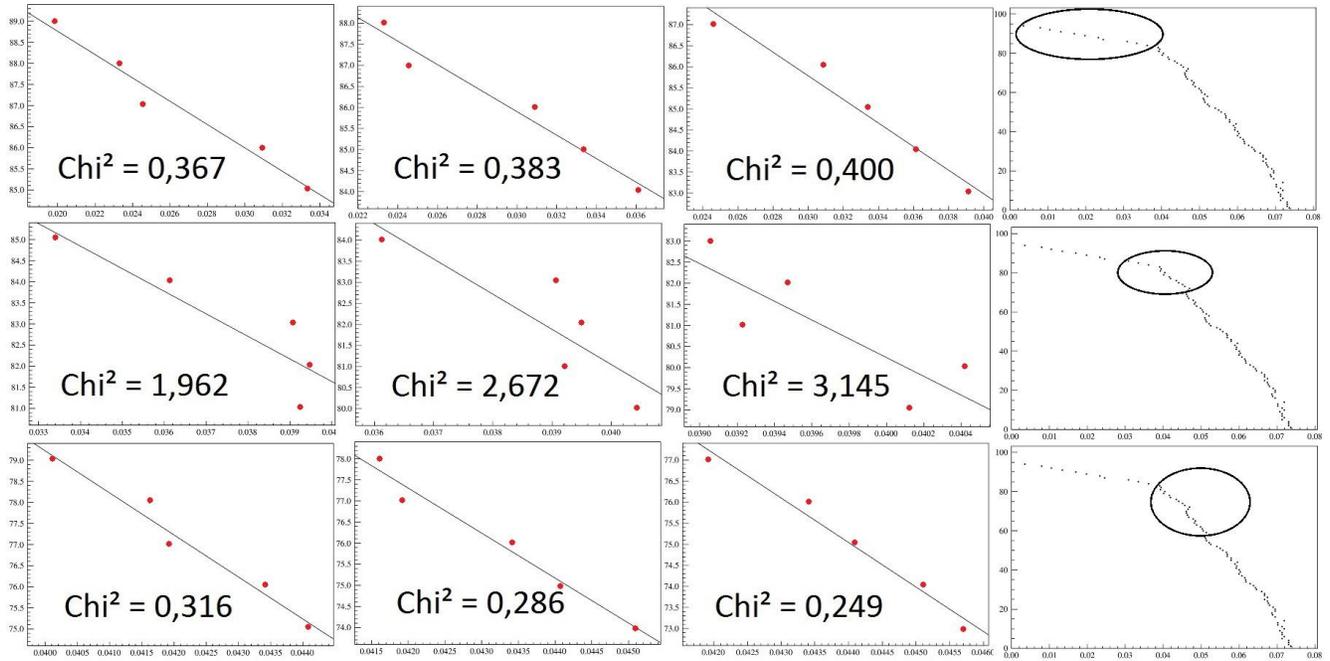


Figura 4 – Análise do Chi<sup>2</sup> em detalhe para conjuntos de 5 pontos em regiões diferentes, inicialmente na região da onda direta, em seguida o limiar entre direta e refração e, por último, a região da onda refratada.

Delimitados os eventos, foram identificados por retas de cores diferentes, como na Figura 5.

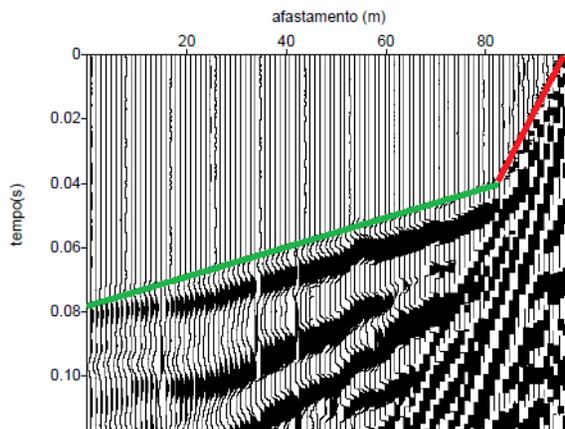


Figura 5 – Diferenciação dos eventos sísmicos através da análise do  $\chi^2$ , em vermelho a onda direta e em verde a primeira refração.

### Discussão e Conclusões

A determinação dos parâmetros se deu baseada em informações prévias, como a escolha de duas camadas ocultas, mas também em fatores práticos, como a janela de amostras ( $n = 100$ ) e a posição da saída positiva da janela ( $m = 75$ ). Isso se dá, provavelmente, porque com aproximadamente 100 amostras podemos observar o ruído antes da primeira chegada e a forma da onda (*wavelet*), a escolha de uma posição intermediária para a saída positiva minimiza o efeito do ruído e minimiza a sensibilidade do processo quanto à variação da *wavelet*. O número de neurônios em cada camada se mostrou de menor importância, já que para valores maiores que 5 neurônios por camada não se nota diferença significativa. Portanto, fixou-se cada camada com 5 neurônios.

Esta arquitetura de rede se mostrou eficiente, com mais de 90% de acerto com o *picking* manual, sendo rara a necessidade de ajustes manuais após o picking automatizado pela rede neural, o que mostra a validade do modelo e sua aplicabilidade, seja em outros estudos, seja nas disciplinas de graduação dos institutos de bacharelado em Geofísica. A diferenciação dos eventos também foi satisfatória e, uma vez que o  $\chi^2$  aumenta de ordem de grandeza, percebemos a mudança de evento e podemos caracterizá-los de forma bem visual (Figura 4).

Dois tipos de redes neurais são comumente usadas para identificação do *first break*: modelos baseados em parâmetros, que usa tradicionais conceitos de processamento de sinais, e modelos baseados em reconhecimentos de padrões visuais, que usa processamento de imagens (MCCORMACK, 2003). O presente trabalho apresenta uma ideia diferente de ambas.

Todavia, novas abordagens serão empregadas nos meses restantes da pesquisa até a finalização do Trabalho de Conclusão de Curso (TCC).

### Agradecimentos

À Sociedade Brasileira de Geofísica (SBGf) pela bolsa de iniciação científica do primeiro autor.

### Referências

- APPENDIX 6 SEISMIC REFRACTION METHOD, Cambridge University. Acesso em 24 out 2017. <[http://www.cambridge.org/ee/download\\_file/825719//>](http://www.cambridge.org/ee/download_file/825719//>).
- ARAUJO AS. 2015. Rastreamento automático das primeiras quebras em dados sísmicos. Trabalho de Graduação, Universidade de São Paulo, São Paulo.
- CYBENKO G. 1988. Continuous Valued Neural Networks with Two Hidden Layers are Sufficient, Technical Report, Department of Computer Science, Tufts University.
- DENTITH MC, EVANS BJ, PAISH KF & TRENCH A. 1992. Mapping the regolith using seismic refraction and magnetic data: Results from the Southern Cross Greenstone Belt, Western Australia. *Exploration Geophysics*, 23: 97–104.
- ERVIN CP, MCGINNIS LD, OTIS RM & HALL ML. 1983. Automated analysis of marine refraction data: A computer algorithm. *Geophysics*, 48(5): 582-589.
- GOULTY NR, BRABHAM PJ. 1984. Seismic refraction profiling in opencast coal exploration, *First Break*, 2(5): 26–34.
- LAWTON DC & HOCHSTEIN MP. 1993. Physical properties of titanomagnetite sands, *Geophysics*, 45: 394–402.
- MCCORMACK MD. A review of automated first-break picking and seismic trace editing techniques. In: *Geophysical Applications of Artificial Neural Networks and Fuzzy Logic*. Springer, Dordrecht, 2003. p. 1-12.
- MCCORMACK, MD, ZAUCHA DE, DUSHEK DW. 1993. First-break refraction event picking and seismic data trace editing using neural networks. *Geophysics*, 58: 67-78.
- MIRANDA AIF. 2000. Identificação automática das primeiras quebras em traços sísmicos por meio de uma rede neural, Dissertação de mestrado, Universidade Federal do Pará, Belém.
- NEEF DB, WYATT SB. 1986. Noise suppression by the radial amplitude-slope rejection method. *Geophysics Prospecting*, 40: 587 – 604.

RUIZ DMG. 2014. Ensaio de sísmica rasa multicomponente em Termas de Ibirá – SP, Trabalho de Graduação, Universidade de São Paulo, São Paulo.

PAKISER LC & BLACK RA. 1957. Exploring for ancient channels with the refraction seismograph. *Geophysics*, 22: 32–47.

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

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

VUOLO JH. 1996. Qualidade de ajuste, in *Fundamentos da Teoria dos Erros*, 189-20.