A cycle-skipping analysis in transformed domains for full waveform inversion using particle swarm optimization (PSO)

Jheyston O. Serrano [1], Sergio A. Abreo [1*], Ana B. Ramirez [1] and Reynam C. Pestana [2]. [1] Universidad Industrial de Santander, Colombia. [2] UFBA, Center for Research in Geophysics and Geology (CPGG) and National Institute of Science Petroleum Geophysics (INCT-GP/CNPq).

Abstract

Full waveform inversion (FWI) is a state-of-the-art method used to estimate subsurface parameters, such as the seismic velocity. FWI is an iterative method that requires an adequate starting velocity (SV) model as input, to converge to the correct solution. A SV model is considered adequate for the FWI when its low frequencies are correctly estimated or cycle-skipping events are not present. Currently, some strategies have been used to build SV models such as analytical methods, reflection tomography, and global optimization methods. In this work, we focus on the use of particle swarm optimization (PSO), which estimates a SV model by minimizing the number of cycle-skipping events can be measured in three different domains: time, frequency and complex trace domain. The computational cost of the proposed PSO method for SV estimation is reduced through the use of graphical processor units (GPUs). We show that, among the analyzed metrics and domains, the least square error metric of the cycle-skipping in the complex trace domain outperforms the others domains in the estimation of adequate SVMs.

Introduction

Full waveform inversion (FWI) is a state-of-the-art method used to estimate subsurface parameters, such as the seismic velocity model. It was first proposed by Lailly (1983) and then introduced to the seismic community by Tarantola (1984). The FWI method has many challenges, such as its high computational cost and the requirement of an adequate starting velocity (SV) model of the subsurface (R. Wu and Wu, 2013). A SV model for FWI is considered adequate when the low frequency components of the model are correctly estimated or cycle-skipping (CS) events are not present (Virieux and Operto, 2009). It has been established that CS occurs when the phase-shift between the modeled and observed traces is higher than half the period of the main wavelet (Ma, 2012). Figure 1 depicts a wavelet signal in the [top] having a phase-shift higher than $T/2$ in comparison to the wavelet signal in the [middle]. On the other hand, the wavelet signal in the [bottom] has a phase-shift that is lower than $T/2$ in comparison with the wavelet in the [middle]. Studying

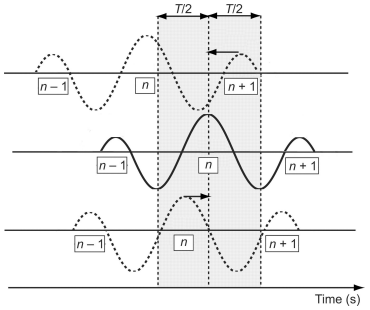

Figure 1: Cycle-skipping definition.

CS is of major interest because if CS is present, then FWI converges towards a local minimum solution (Luo and Wu, 2013). Different strategies have been proposed to build adequate SV models, such as analytical methods (Dines and Lytle, 1979), reflection tomography (Biondi and Almomin, 2012), and global optimization (Mrinal K. Sen, 2013). In this paper, we propose the use of the particle swarm optimization (PSO) (Mrinal K. Sen, 2013) technique together with a cycle-skipping measure in three different domains (time; frequency- using the Fourier transform; and complex trace- using the Hilbert transform), to estimate SV models for FWI. In particular, PSO uses a set of particles where every particle is a velocity model with dimensions $N_x \times N_z$. Each unknown velocity pixel is one dimension of the particle and it is updated independently. This independence allow us to use parallel processing and speedup the proposed PSO implementation. We show that the best SV models are obtained by minimizing the cycle-skipping of the traces in the complex trace domain using PSO.

Method

Full Waveform Inversion (FWI)

The FWI method estimates the subsurface velocity model $\mathbf{m}$ by solving a local optimization problem where the least square error between the seismic data recorded at the surface $\mathbf{d}_{obs}(\mathbf{m}_{true})$, and the modeled data $\mathbf{d}_{mod}(\mathbf{m})$ is minimized (Tarantola, 1984). The formulation is given by

$$\arg\min_{\mathbf{m}} ||\mathbf{d}_{mod}(\mathbf{m}) - \mathbf{d}_{obs}(\mathbf{m}_{true})||_2^2, \quad (1)$$

Since the inverse problem is not linear, the cost function given in Equation (1) has several local minima. The inverse problem is solved iteratively by selecting an initial model $\mathbf{m}_0$, and updating it by using Newton-like methods

(Goldstein, 1965), as follows

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \alpha_k \Delta \mathbf{m}^k, \quad (2)$$

where α_k is the step size and $\Delta \mathbf{m}^k$ of the k^{th} iteration is given by

$$\Delta \mathbf{m}^k = -[\mathbf{H}(\mathbf{m}^k)]^{-1} \mathbf{g}(\mathbf{m}^k), \quad (3)$$

where $[\mathbf{H}(\mathbf{m}^k)]^{-1}$ is the inverse of the Hessian matrix and $\mathbf{g}(\mathbf{m}^k)$ is the gradient of the Equation (1) both evaluated in $\mathbf{m}^k$. In the first iteration the inverse of the Hessian matrix is the identity matrix and for the following iterations the product of $[\mathbf{H}(\mathbf{m}^k)]^{-1}\mathbf{g}(\mathbf{m}^k)$ is given by the L-BGFS algorithm (Liu and Nocedal (1989)). Additionally, the gradient function $\mathbf{g}(\mathbf{m}^k)$ is computed using the formulation in Plessix (2006). An inadequate SV model ($\mathbf{m}_0$) produces convergence to a local minimum which means a wrong final velocity model.

In the next section, we describe a methodology to estimate the initial model $\mathbf{m}_0$, using global optimization together with three different metrics that minimizes the CS between the observed and modeled traces.

SV Model Estimation

Particle swarm optimization (PSO) is a global optimization technique that uses a group of particles to explore the entire space of possible solutions by moving randomly through it (Ranjita S. and. Shalicahan, 2007). The formulation used in PSO to find the position of the j^{th}-dimension of the i^{th} particle is given by

$$\mathbf{x}_{i,j}^{k+1} = \mathbf{x}_{i,j}^k + \mathbf{v}_{i,j}^{k+1} \cdot dt, \quad (4)$$

where $\mathbf{x}_{i,j}^{k+1}$ is the new position vector associated with the i^{th} subsurface velocity model, $\mathbf{x}_{i,j}^k$ is the present position vector of the i^{th} particle, $\mathbf{v}_{i,j}^{k+1}$ is the new velocity vector and dt is the step size used by the particle to move inside the searching space. An update in the velocity vector is obtained as

$$\mathbf{v}_{i,j}^{k+1} = (w \cdot \mathbf{v}_{i,j}^k) + c_1 \cdot rand() \cdot \frac{(\mathbf{p}_i^k - \mathbf{x}_{i,j}^k)}{dt} + c_2 \cdot rand() \cdot \frac{(\mathbf{g}^k - \mathbf{x}_{i,j}^k)}{dt}, \quad (5)$$

where w is an inertial weight that controls the particle movement; c_1 and c_2 are the weights associated with the local and global behavior of the swarm, respectively; $rand()$ is a random number between $[0-1]$ with normal distribution; $\mathbf{p}_i^k$ is the best local position; and $\mathbf{g}^k$ is the best global position. Each particle is moving in a j-dimensional space, which for the geophysical problem is the number of unknown velocity parameters. The PSO solution at each iteration is accepted when a minimum error measure between the observed and the modeled data is obtained. The following sections describe the error metrics, in three different transform domains, that are used in the PSO method to update the estimated SV model.

CS in Time Domain

In time domain, we use the correlation coefficient as measure of CS between modeled and observed traces. However an amplitude correction is performed to both modeled and observed traces, before computing the correlation coefficient. Thus, the correlation takes into account the effects of multiples and reflections of deeper areas. The amplitude correction used in this paper is given by

$$\mathbf{d}_{obs}^{s,r}(t) = \mathbf{d}_{obs}^{s,r}(t) \times q^2(t), \quad (6)$$

where s is the source number, r is the receiver number, t is time and $q^2(t)$ is a polynomial function that correct the exponential decay.

The correlation metric used to measure CS between the modeled and the observed data, per source and receiver, is defined as

$$C_{s,r} = \frac{\Sigma_t(\mathbf{d}_{obs}^{s,r}(t) - \hat{\mathbf{d}}_{obs}^{s,r})(\mathbf{d}_{mod}^{s,r}(t) - \hat{\mathbf{d}}_{mod}^{s,r})}{\sqrt{(\Sigma_t(\mathbf{d}_{obs}^{s,r}(t) - \hat{\mathbf{d}}_{obs}^{s,r})^2)(\Sigma_t(\mathbf{d}_{mod}^{s,r}(t) - \hat{\mathbf{d}}_{mod}^{s,r})^2)}}, \quad (7)$$

where $\hat{\mathbf{d}}_{obs}^{s,r}$ or $\hat{\mathbf{d}}_{mod}^{s,r}$ represents the average in the time direction. We compute the average of all the correlation coefficients for all receivers $\hat{C}_s$ which corresponds to the correlation coefficient per shot. Finally, we compute the average of the correlation coefficients for all shots $\hat{C}$ such that we obtain the correlation coefficient per velocity model. The proposed PSO implementation accepts a new seismic velocity model as SV model if its correlation coefficient is greater than the correlation coefficient for the velocity model in the previous iteration.

CS in Frequency Domain

In order to compute the CS in the frequency domain, first we compute the Fourier transform of the observed and modeled data $\mathbf{f}_{obs}^{s,r}(\omega)$ and $\mathbf{f}_{mod}^{s,r}(\omega)$. Particularly, we determine both the amplitude and phase of each frequency component of all traces. $\theta_{obs}^{s,r}(\omega)$ is the phase of $\mathbf{d}_{obs}^{s,r}(\mathbf{m}_{true})$, and $\theta_{mod}^{s,r}(\omega)$ is the phase of $\mathbf{d}_{mod}^{s,r}(\mathbf{m}^k)$. The CS-measure in frequency $CS_{FT}^{s,r}$ is obtained by counting the number of frequencies whose difference between $\theta_{obs}^{s,r}(\omega)$ and $\theta_{mod}^{s,r}(\omega)$ is larger than π. Figure 2 and Algorithm 1 summarize the strategy to compute CS measure in frequency $CS_{FT}^{s,r}$.

From $CS_{FT}^{s,r}$, we compute the L^2-norm in the receivers direction $\hat{CS}_{FT}^{s}$ such that a CS-Fourier transform coefficient per shot is obtained. Then, we average the coefficient measure for all sources $\hat{CS}_{FT}$ to obtain the CS-Fourier coefficient per velocity model. The PSO implementation strategy in frequency accepts a new seismic velocity model as SV model if its CS-Fourier coefficient is smaller than the CS-Fourier coefficient for the velocity model in the previous iteration.

Only the frequency band between 0 and $10\ Hz$ is used in the proposed strategy only the low frequency components are necessary to estimate adequate SV models.

CS in the Complex Trace Domain

The complex trace allows the extraction of seismic attributes, such as the reflector's intensity, the wrapping of

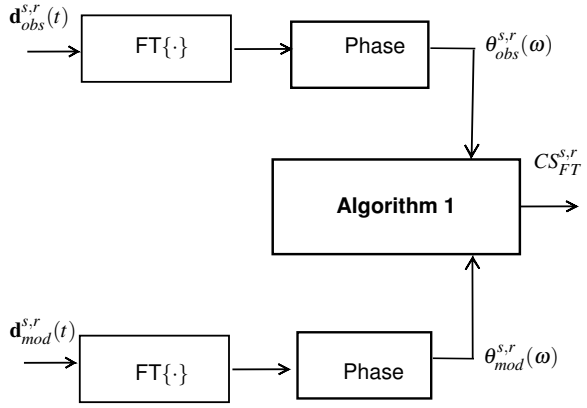

Figure 2: Proposed method to compute the cycle-skipping (CS) in the frequency domain.

Algorithm 1 - CS in frequency domain

for each source $s = 1, \cdots, S$ **do** ▷ S, number of sources
  for each receiver $r = 1, \cdots, R$ **do** ▷ R, receivers
    $Counter = 0$ ▷ Clear the counter
    for each frequency $\omega = 0, \cdots, W$ **do**
      ▷ W, frequencies fewer than 10 Hz
      $dif = |\theta^{s,r}_{mod}(\omega) - \theta^{s,r}_{obs}(\omega)|$ ▷ Phase delay
      if $dif > \pi$ **then**
        $Counter = Counter + 1$ ▷ CS counter
      end if
    end for
    $C^{s,r}_{FT} = Counter$ ▷ Saving the CS count
  end for
end for

the phase or the instant frequency of the seismic traces. The complex trace is defined as

$$\tilde{\mathrm{g}}(t) = g(t) + iH\{g(t)\}, \tag{8}$$

where $\tilde{\mathrm{g}}(t)$ is known as the analytical signal, and $H\{\cdot\}$ is the Hilbert transform of the trace $g(t)$. We estimate the phase of the analytic signal as

$$\theta^{s,r}_{obs}(t) = \arctan\left(\frac{H\{g^{s,r}_{obs}(t)\}}{g^{s,r}_{obs}(t)}\right), \tag{9}$$

where $\theta^{s,r}_{obs}(t)$ is the analytical phase obtained by $\mathbf{d}^{s,r}_{obs}(\mathbf{m}_{true})$ and $\theta^{s,r}_{mod}(t)$ is the analytical phase obtained by $\mathbf{d}^{s,r}_{mod}(\mathbf{m}^k)$. The CS-complex trace coefficient $CS^{s,r}_{CT}$ is 1 when the difference between $\theta^{s,r}_{obs}(t)$ and $\theta^{s,r}_{mod}(t)$ is higher than π. Figure 3 and Algorithm 2 summarize the strategy to compute $CS^{s,r}_{TC}$.

Note that all the frequency components are involved in the phase shift computation in the complex trace domain. Similar to the CS frequency domain, the L^2-norm in the receivers direction $\hat{CS}^{s}_{TC}$ is computed, which retrieves the CS-complex trace coefficient per shot. Then, we compute the average of all sources such that we obtain the CS-complex trace coefficient per velocity model. At the end, a new velocity model is accepted as SV model when $\hat{CS}_{TC}$ is smaller than for the velocity model in the previous iteration.

Algorithm 2 CS in complex trace domain

for each source $s = 1, \cdots, S$ **do** ▷ S, number of sources
  for each receiver $r = 1, \cdots, R$ **do** ▷ R, receivers
    $Counter = 0$ ▷ Clear the counter
    for $t = 0, \cdots, N_t$ **do**
      ▷ N_t, adquisition time
      $dif = |\theta^{s,r}_{mod}(t) - \theta^{s,r}_{obs}(t)|$ ▷ Phase delay
      if $dif > \pi$ **then**
        $Counter = Counter + 1$ ▷ CS counter
      end if
    end for
    $C^{s,r}_{TC} = Counter$ ▷ Saving the CS count
  end for
end for

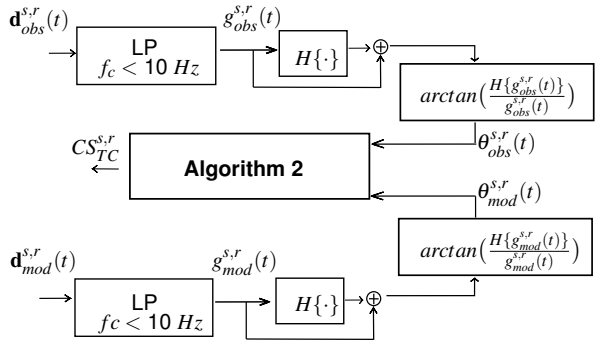

Figure 3: Proposed method to compute the cycle-skipping (CS) in the complex trace domain.

Experiment description

In all the experiments, we use as true velocity model the image shown in Figure 4. The observed traces are obtained by using the finite difference in time domain (FDTD) implementation of the acoustic wave equation with constant density given by

$$\frac{1}{\mathbf{m}_{true}(x,z)^2}\frac{\partial^2 P(x,z)}{\partial t^2} = \frac{\partial^2 P(x,z)}{\partial x^2} + \frac{\partial^2 P(x,z)}{\partial z^2} + src(x,z), \tag{10}$$

where $P(x,z)$ is the pressure field, $\mathbf{m}_{true}(x,z)$ is the true subsurface velocity, x and z are the spatial coordinates, t the time variable, and $src(x,z)$ is the seismic source. For all the results described in the next section, only 5 shots located at 1200 m, 1950 m, 2700 m, 3450 m, and 4200 m are used, at a depth of 75 m. Also, 175 receptors located from 500 m to 4875 m, separated from 25 m, at a depth of 75 m. The time step for the FDTD was selected to be $dt = 4$ ms and the total propagation time was 2.5 s. The source was modeled as

$$g_s = -2(\pi\hat{f})^2(t-t_0)e^{(-(\pi f)^2(t-t_0)^2)}, \tag{11}$$

where $\hat{f} = 3\ Hz$ is the central frequency, and $t_0 = 0.5$ s is the time delay. The modeling of the input wavelet propagating

through the medium was implemented using a second order stencil of finite differences for the temporal derivative and an eighth order stencil for the spatial derivatives. The modeled traces are also obtained using Equation (10) but using $\mathbf{m}^k$ as seismic velocity model. $\mathbf{m}^k$ is the estimated velocity model obtained by the best global particle, at each iteration of the PSO method. When the new velocity model gives a larger CS measure than for the current velocity model, then the estimated model is not stored and a new random search direction is performed. This process will be repeated until the stopping criterion have been reached.

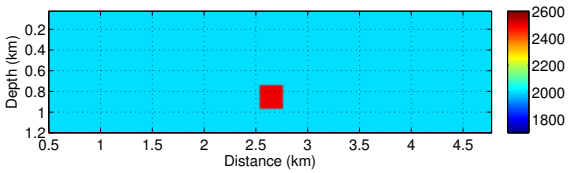

Figure 4: True velocity model ($\mathbf{m}_{true}$).

Results

The particles dimension is set to 175 pixels (25×7), and an example of one particle is shown in the random velocity model of Figure 5-a). The dimension is reduced by using a coarse resolution due to computational cost. Also, once a velocity model is estimated by the proposed PSO implementations, a low pass filter of 5×5 dimensions is applied to the model in order to avoid the strong reflections coming from the high velocity contrasts. The smoothed velocity model of Figure 5-a) is shown in Figure 5-b). Figure 6-a), 6-b), and 6-c) depict the estimated SV models obtained with the PSO method using three different domains: time, frequency and complex trace, respectively. The images in Figure 6 are the average of 100 resulting models obtained with the proposed PSO methods. In all cases, 100 iterations and 500 particles were used. These models were used as starting points for the multi-scale FWI method with frequencies 5, 10 and 20 Hz, (Carey Bunks and Chavent., 1995). The final velocity models are presented in Figure 7-a), 7-b), and 7-c) for time, frequency and complex trace, respectively. Qualitatively, the best model was obtained when PSO uses the CS measure in the complex trace.

Additionally, we evaluate quantitatively the results obtained with the three different domains. Table 1 shows the $L2$-error norm of the models obtained with PSO in the three different domain compared with the true velocity model, as well as the $L2$-error norm of the velocity models obtained after FWI compared with the true velocity model. The best SV models and FWI resulting models were obtained when the CS is measured in the complex trace domain. This occurs because the CS measure in the complex trace allows correcting for the phase changes produced by the wave field reflections over the square diffracting area.

Table 2 and Table 3 show the execution times to obtained the subsampled SVM (SSVM) and full SVM (FSVM) using C and CUDA-C languages, respectively. Note that a speedup a factor of 6 is approximately obtained when

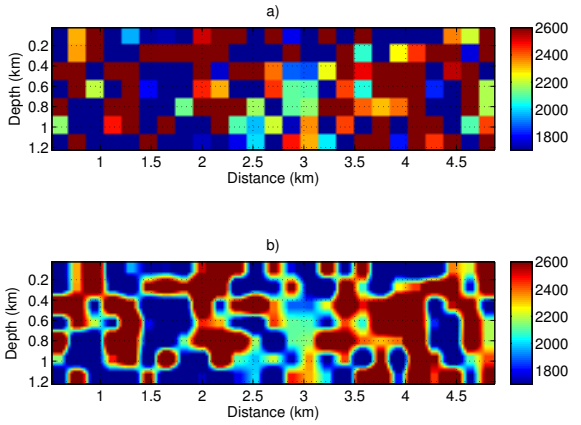

Figure 5: a) Initial velocity model for the PSO. b) Smoothed version of the initial velocity model for the PSO.

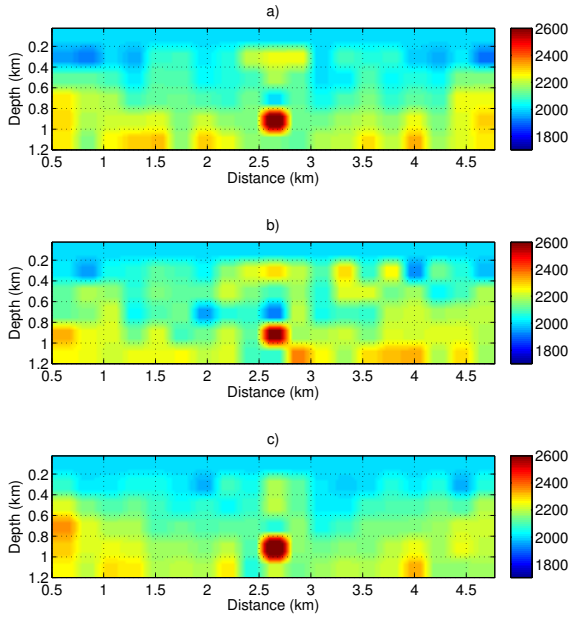

Figure 6: Velocity models obtained with PSO using a) correlation in time, b) CS in frequency, and c) CS in the complex trace.

parallel architectures are used in the implementations. The execution times were calculated using a server with a GPU Tesla K40c of 12 Gb of RAM and a processor Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40 GHz with 256 Gb of RAM.

Finally, we present results of the SV models obtained with the different metrics at a pixel level resolution. The SV models results of PSO at pixel resolution for the time, frequency and complex trace domain are given in Figure 8-a), 8-b) and 8-c),respectively. Also, the FWI results using the SV models are presented in Figure 9.

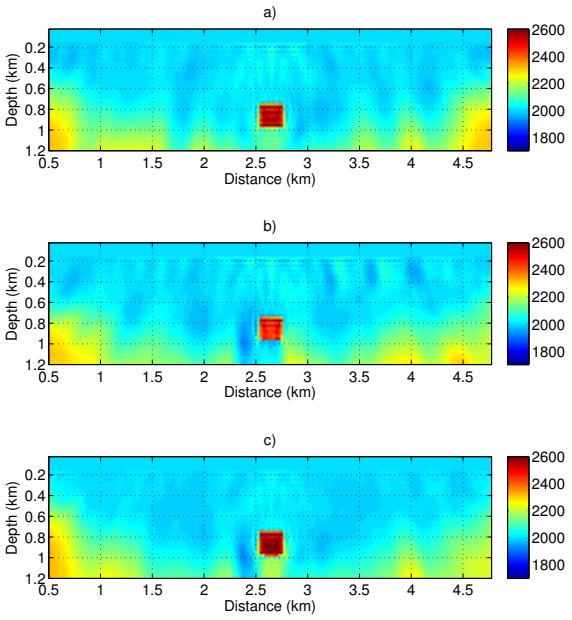

Figure 7: Velocity models obtained with FWI method in a multi-scale approach using 5,10 y 20 Hz for the initial models obtained with a) correlation in time, b) CS in frequency, and c) CS in the complex trace.

Table 1: The $L2$-error norm in PSO and FWI models with each metric.

$L2-error\ norm$	PSO (x1000)	FWI (x1000)
Correlation in time	13.663	8.2423
CS in frequency	14.788	8.1411
CS in complex trace	13.559	8.0141

Table 4 presents the $L2$-error norm of the models obtained with PSO in the three different domain compared with the true velocity model, as well as the $L2$-error norm of the velocity models obtained after FWI compared with the true velocity model, obtained at pixel level resolution.

Conclusions

In this work, we have presented a global optimization that uses three different strategies for the random search of a starting velocity model for FWI: the correlation in time, the CS in frequency and the CS in the complex trace. The proposed global optimization methods were quantitatively evaluated in terms of the model with lower $L2$- error norm when under-sampled model is used (see Table 1), and when the model has a pixel level resolution (see Table 2). The results show that, on average, the best SV model is

Table 2: The execution time for generate SVM in C.

Metric	SSVM (hours)	FSVM (hours)
Correlation in time	22.34	134.04
CS in frequency	24.5	147.25
CS in complex trace	84.4	506.43

Table 3: The execution time for generate SVM in CUDA-C.

Metric	SSVM (hours)	Time (hours)
Correlation in time	3.53	22.34
CS in frequency	4.08	24.52
CS in complex trace	14.06	84.40

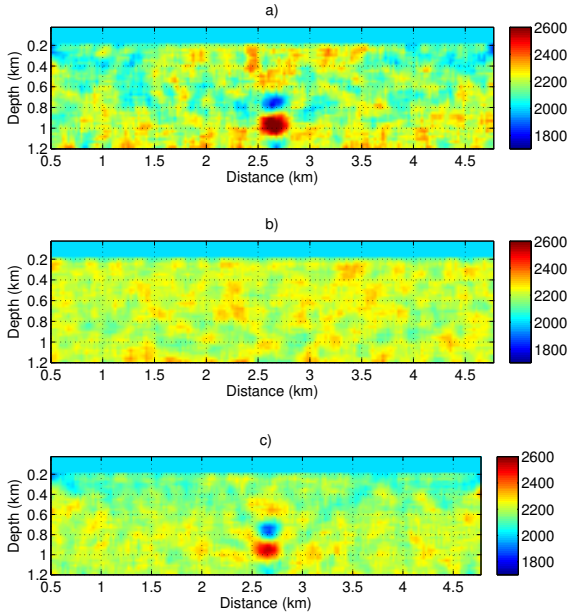

Figure 8: Velocity models obtained with PSO using a) correlation in time, b) CS in frequency, and c) CS in the complex trace.

obtained with the CS measure in the complex trace. In the complex trace domain we can find the phase changes produced by the reflections, particularly in the square diffraction zone.

The results shown in this work encourages the use of global optimization techniques to find adequate SV models for the seismic FWI, instead of relying only on *a-priori* information given by an expert. A methodology that combines both, tomography and global optimization methods can also be explored.

To mitigate the computational cost of the dimensionality in real data, we propose to search velocity models in parallel by taking advantage of their independence in the search. In this way, this methodology becomes more feasible to generate 3D SV models.

Table 4: The $L2$-error norm in PSO and FWI models with pixel resolution for each metric.

$L2-error\ norm$	PSO (x1000)	FWI (x1000)
Correlation in time	16.780	7.8208
CS in frequency	18.391	7.9594
CS in complex trace	16.150	7.2823

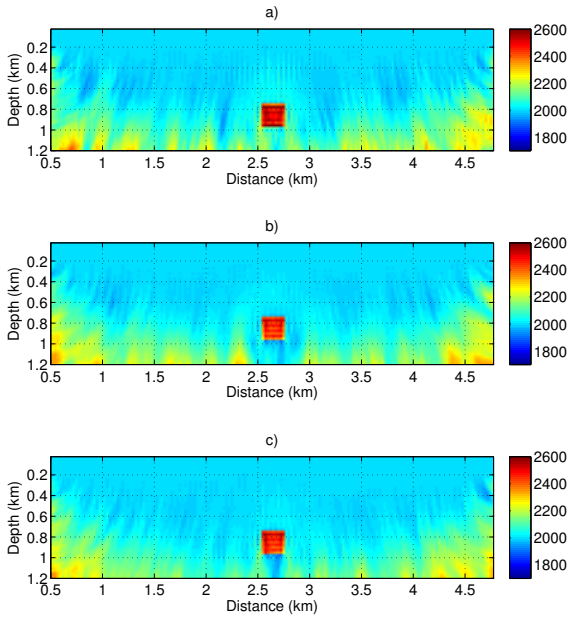

Figure 9: Velocity models obtained with FWI using a) correlation in time, b) CS in frequency, and c) CS in the complex trace.

Acknowledgements

This work is supported by the Colombian Oil Company ECOPETROL and COLCIENCIAS as a part of the research project grant No. 0266-2013.

References

Biondi, B., and A. Almomin, 2012, Tomographic full waveform inversion (tfwi) by combining full waveform inversion with wave- equation migration velocity anaylisis: SEG., **74**, 127–152.

Carey Bunks, Fatimetou Saleck, S. Z., and G. Chavent., 1995, Multiscale seismic waveform inversion: GEOPHYSICS, **60**, 1457–1473,.

Dines, K. A., and R. J. Lytle, 1979, Computerized geophysical tomography: IEEE., **67**, 1065–1072.

Goldstein, A. A., 1965, On newton's method: Numerische Mathematik, **7**, 391–393.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations.

Liu, D. C., and J. Nocedal, 1989, On the limited memory bfgs method for large scale optimization: Mathematical programming, **45**, 503–528.

Luo, J., and R. Wu, 2013, Seismic envelope inversion and modulation signal model: GEOPHYSICS, **79**.

Ma, Y., 2012, Waveform-based velocity estimation from reflection seismic data: Colorado School of Mines.

Mrinal K. Sen, University of Texas, A., 2013, Global optimization methods in geophysical inversion: Cambridge.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

R. Wu, J. L., and B. Wu, 2013, Ultra-low-frequency information in seismic data and envelope inversion: SEG, **6**, 3078.

Ranjita S. and. Shalicahan, S., 2007, Particle swarm optimization: A new tool to invert geophysical data: Society of Exploration Geophysicists, **54**, 772–775.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, 6.