APPLICATION OF AN AUGMENTED LAGRANGIAN METHOD FOR CONSTRAINED FULL WAVEFORM INVERSION

A. W. Camargo and L. T. Santos

email: *ra080529@ime.unicamp.br*
keywords: *FWI, Augmented Lagrangian, Constrained Optimization*

ABSTRACT

Most Full Waveform Inversion (FWI) applications use quasi-Newton based optimization algorithms to perform the inversion for the velocity model. In this report, we will discuss the application of an Augmented Lagrangian method, which can easily incorporate constraints to the FWI problem. We will also consider that the velocity model can be written as a linear combination of some base functions, in order to reduce the computational cost. In the presented numerical examples, such base functions are chosen trivially, due to the fact that the tested velocity models are very simple. Nevertheless, they show the importance of considering the constraints to the optimization problem.

INTRODUCTION

The seismic Full Waveform Inversion (FWI) is a powerful tool to seismic imaging because it tries to estimate subsurface medium parameters from seismic data. However, accurate reconstruction of interfaces is yet a big challenge for current techniques. The fundamental basis for FWI unites two important worlds of computational geophysics that are seismic modeling (direct problem) and seismic inversion (inverse problem). A good review on the classic FWI theory cab be found in Virieux and Operto (2009). A big difficulty is how to avoid local minima but there are many algorithms based on FWI that try to improve the final results. See, e.g., Symes (2008); Leeuwen and Herrmann (2016); Bunks et al. (1995). Most of the applications mentioned above use algorithms for unconstrained optimization, such as the a limited memory quasi-newton method – LBFGS (see, e.g., Nocedal and Wright (1999)). In general, regularization terms are added to the objective function to help accelerate the convergence of the procedure.

In this work, we analyse the FWI as a constrained optimization problem, where we use an Augmented Lagrangian algorithm, the software Algencan developed by Birgin and Martínez (2014a). To reduce the cost of optimization we will consider that the bidimensional velocity model can be written as a combination of some base functions, $\phi_i : I\!R^2 \to I\!R$ $(i = 0, 1, \ldots, n)$, i.e., the acoustic propagation velocity c is given by

$$c(\mathbf{w}; \mathbf{x}) = \phi_0(\mathbf{x}) + \sum_{i=1}^{n} w_i \phi_i(\mathbf{x}), \tag{1}$$

where $\mathbf{x} = (x, z) \in I\!R^2$ are the spatial variables and $\mathbf{w} \in I\!R^n$ (weights) are the parameters that will be inverted by the algorithm. In the experiments done in this work, the base functions ϕ_i were chosen according to the desired complexity of the velocity model. For example, for a constant gradient velocity model we can chose $\phi_0(\mathbf{x}) = c_0$ (constant), $\phi_1(\mathbf{x}) = x$ and $\phi_2(\mathbf{x}) = z$.

SEISMIC MODELING

The seismic modeling plays a key role for FWI. In the two-dimensional case of a acoustic velocity medium, the wave equation with constant density and considering the velocity given by equation (1), we have

$$\frac{1}{c(\mathbf{w};\mathbf{x})^2}u_{tt}(\mathbf{w};\mathbf{x},t) - \Delta u(\mathbf{w};\mathbf{x},t) = F(\mathbf{x},t), \tag{2}$$

where Δ is laplacian operator, F is the function that describes the source and u is the wavefield. Notice that we are considering an implicit dependency on the variable $\mathbf{w}$ in the wavefield.

To solve equation (2) we use a finite-difference method of second-order in time and $2M^{th}$-order in space (Liu and Sen, 2009). Typical values for M in the literature are $1 \leq M \leq 5$. For the numerical method to be well defined, we must include the initial and boundary condition, which are

$$u(\mathbf{w};\mathbf{x},t) = 0, \quad \text{for} \quad t \leq 0 \quad \text{and} \quad u_t(\mathbf{w};\mathbf{x},0) = 0. \tag{3}$$

To avoid boundary reflections, we use an absorbing boundary condition based on the work of Kosloff and Kosloff (1986). The numerical scheme for solving (2) is given by

$$\widetilde{\gamma}_{j,k}u_{j,k}^{\ell+1} = \bar{\gamma}_{j,k}u_{j,k}^{\ell} - \widehat{\gamma}_{j,k}u_{j,k}^{\ell-1} + r_{j,k}^2\Delta u_{j,k}^{\ell} + c_{j,k}^2\Delta t^2 F_{j,k}^{\ell}, \tag{4}$$

where the indices j, k, ℓ refer to discretization $x_j = x_0 + j\Delta$, $z_k = z_0 + k\Delta$, $t_\ell = \ell\Delta t$, with $j = 0, 1, \ldots, J$, $k = 0, 1, \ldots, K$, $\ell = 0, 1, \ldots, L$, where

$$u_{j,k}^{\ell} \approx u(\mathbf{w}; x_j, z_k, t_\ell), \quad r_{j,k} = c_{j,k}\Delta t/\Delta, \quad \text{and} \quad c_{j,k} = c(\mathbf{w}; x_j, z_k). \tag{5}$$

The laplacian operator is approximated by

$$\Delta u_{j,k}^{\ell} = 2a_0 u_{j,k}^{\ell} + \sum_{m=1}^{\mathrm{M}} a_m \left(u_{j-m,k}^{\ell} + u_{j+m,k}^{\ell} + u_{j,k-m}^{\ell} + u_{j,k+m}^{\ell}\right), \tag{6}$$

where the a_m's are the finite difference coefficients given by (Liu and Sen, 2009)

$$\begin{cases} a_m = \dfrac{(-1)^{m+1}}{m^2} \displaystyle\prod_{n=1, n\neq m}^{\mathrm{M}} \left|\frac{n^2}{n^2 - m^2}\right|, \; m = 1, \ldots, M, \\ a_0 = -2\displaystyle\sum_{m=1}^{\mathrm{M}} a_m, \end{cases} \tag{7}$$

Parameters $\widetilde{\gamma}_{j,k}$, $\widehat{\gamma}_{j,k}$ and $\bar{\gamma}_{j,k}$ are related to the absorbing boundary conditions and are given by (Kosloff and Kosloff, 1986)

$$\tilde{\gamma}_{j,k} = 1 + \Delta t\gamma_{j,k}, \quad \bar{\gamma}_{j,k} = 2 - \Delta t^2\gamma_{j,k}^2, \quad \hat{\gamma}_{j,k} = 1 - \Delta t\gamma_{j,k}, \tag{8}$$

where

$$\gamma_{j,k} = \begin{cases} \dfrac{\alpha_0}{\cosh^2(\mu j\Delta)}, & j = 0, 1, \ldots, J_b, \text{ and} \\ & k = 0, \ldots, K, \\ \dfrac{\alpha_0}{\cosh^2(\mu(J-j)\Delta)} + \dfrac{\alpha_0}{\cosh^2(\mu(K-k)\Delta)}, & j = J - J_b, \ldots, J, \text{and} \\ & k = K - K_b, \ldots, K, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Morever, α_0 and μ are dimensional constants, $0 < J_b < J$ and $0 < K_b < K$ indicate how many indices will be considered for the absorbing boundary. In this work we have used $\alpha_0 = 140\ \mathrm{s}^{-1}$ and $\mu = 8.3\ \mathrm{km}^{-1}$ as recommended in Seki and Nishikawa (1988), and $J_b = K_b = 50$.

SEISMIC INVERSION

The inverse problem tries to find an approximation for the velocity model that best fits the seismic data $d_s(\mathbf{x}_r, t)$, where $\mathbf{x}_r = (x_r, z_r)$ are the receiver positions for each shot s. For conventional FWI, the inverse problem gives an approximation of $\mathbf{m}(\mathbf{x}) = 1/c(\mathbf{x})^2$ for $\mathbf{x}$ in a certain pre-defined grid. For example, for a 100×100 grid in xz-plane, there are exactly 10^4 values to be determined. Therefore, to reduce the number of variables to be inverted, we will consider that the velocity is given by a the linear combination given by (1) and thus the inverse problem can be formulated as

$$\min_{\mathbf{w}} f(\mathbf{w}), \tag{10}$$

where $f(\mathbf{w})$ is the objective function and can be given by the classical least-squares functional,

$$f(\mathbf{w}) = \frac{1}{2} \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} \int_0^T \Big[u_s(\mathbf{w}; \mathbf{x}_r, t) - d_s(\mathbf{x}_r, t) \Big]^2 \, \mathrm{d}t, \tag{11}$$

where $[0, T]$ is the time window of the sampling, and N_s and N_r are the number of shots and receivers, respectively. Moreover, u_s is the simulated data for shot s, using velocity $c(\mathbf{w}; \mathbf{x})$.

Equation (10) is an unconstrained optimization problem and, in general, regularization terms are added to the objective function. Some previous works analyse different objective functions for the inverse problem (see, e.g., (Shin and Ha, 2008; Wu et al., 2014; Liu et al., 2016)). We can rewrite equation (11) as,

$$f(\mathbf{w}) = \frac{1}{2} \sum_{s=1}^{N_s} \int_0^T \int_{I\!R^2} E(u_s(\mathbf{w}; \mathbf{x}, t)) \, \mathrm{d}\mathbf{x} \, \mathrm{d}t, \tag{12}$$

where

$$E(u_s(\mathbf{w}; \mathbf{x}, t)) = \sum_{i=1}^{N_r} \delta(\mathbf{x} - \mathbf{x}_{r_i}) \Big[u_s(\mathbf{w}; \mathbf{x}, t) - d_s(\mathbf{x}, t) \Big]^2, \tag{13}$$

with δ being the two-dimensional Dirac's delta fuction. If we are interested in using some opimization algorithm that uses derivatives, it is necessary to compute the gradient of the objective function. Applying the chain rule on equation (12), and writing the result in jacobian notation, we obtain

$$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{w}} = \frac{1}{2} \sum_{s=1}^{N_s} \int_0^T \int_{I\!R^2} \frac{\mathrm{d}E}{\mathrm{d}u_s} \cdot \frac{\mathrm{d}u_s}{\mathrm{d}c} \cdot \frac{\mathrm{d}c}{\mathrm{d}\mathbf{w}} \, \mathrm{d}\mathbf{x} \, \mathrm{d}t. \tag{14}$$

The difficulty in using equation (14) is the high cost to compute the jacobian $\mathrm{d}u_s/\mathrm{d}c$. However, through the use of the adjoint method (Plessix (2006); Fichtner (2011)) it is possible to avoid the calculus of such term. The gradient can be computed as

$$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{w}} = -2 \int_{I\!R^2} \left[\int_0^T p(\mathbf{x}, t) \, \frac{\partial^2 u_s(\mathbf{w}; \mathbf{x}, t)}{\partial t^2} \, \mathrm{d}t \right] \frac{\phi(\mathbf{x})}{c(\mathbf{w}; \mathbf{x})^3} \, \mathrm{d}\mathbf{x}, \tag{15}$$

where $\phi(\mathbf{x}) = [\phi_1(\mathbf{x}), \ldots, \phi_n(\mathbf{x})]$, $p(\mathbf{x}, t)$ is the adjoint wavefield given by the solution of

$$\frac{1}{c(\mathbf{w}; \mathbf{x})^2} p_{tt}(\mathbf{x}, t) - \Delta p(\mathbf{x}, t) = -\widetilde{E}(\mathbf{x}, t), \tag{16}$$

with

$$\widetilde{E}(\mathbf{x}, t) = \sum_{i=1}^{N_r} \delta(\mathbf{x} - \mathbf{x}_{r_i}) \Big[u_s(\mathbf{w}; \mathbf{x}, t) - d_s(\mathbf{x}, t) \Big], \tag{17}$$

and the following "final" conditions,

$$p(\mathbf{x}, T) = 0, \quad \text{and} \quad p_t(\mathbf{x}, T) = 0. \tag{18}$$

The adjoint equation (16) is the wave equation (2) with a different source term. It can be physically interpreted as a backpropagation of the residual value. Thus, we can use the numerical scheme used for the modeling, with the following transformation in order to have initial conditions, $\tilde{p}(\mathbf{x}, t) = p(\mathbf{x}, T - t)$.

AN AUGMENTED LAGRANGIAN METHOD

Let us consider the following constrained optimization problem

$$\text{Minimize} \quad f(\mathbf{w}) \quad \text{subject to} \quad h(\mathbf{w}) = 0, \quad \mathbf{w} \in \Omega, \tag{19}$$

where $f : I\!R^n \to I\!R$ and $h : I\!R^n \to I\!R^m$ are differentiable functions, and $\Omega \subset I\!R^n$. The Augmented Lagragian function is given by

$$L_\rho(\mathbf{w}, \lambda) = f(x) + \frac{\rho}{2} \sum_{i=1}^{m} \left[h_i(\mathbf{w}) + \frac{\lambda_i}{\rho} \right]^2, \tag{20}$$

where $\rho > 0$, $\mathbf{w} \in \Omega$ and $\lambda \in I\!R^m$ is an approximation for the Lagrange multiplier's vector. Notice that $L_\rho(\mathbf{w}, 0)$ is the classical L_2 penalty function. The Augmented Lagrangian function (20)be seen as a penalty function in which the "punishment" for infeasibility points is not applied with respect to the original constraints $h(\mathbf{w}) = 0$ but with a respect to the shifted constraints $h(\mathbf{w}) + \lambda/\rho = 0$.

The Augmented Lagrangian algorithm for solving (19) proceeds by minimizing function (20) with respect to $\mathbf{w}$ at each iteration and updating the Lagrange multipliers vector (λ) and the penalty parameter (ρ). In other words, at each iteration the algorithm solves the subproblem

$$\text{Minimize} \quad L_\rho(\mathbf{w}, \lambda) \quad \text{subject to} \quad \mathbf{w} \in \Omega, \tag{21}$$

for a given ρ and λ. Ideally, if enough progress has been obtained in terms of improvement of feasibility, the same penalty parameter may be used at the next "outer iteration", but the penalty parameter must be increased if progress is not satisfactory. The vector λ/ρ can be interpreted as a "shift", because in (21) the penalty parameter, will not punish only the violation of infeasibilities (this would be the case if the shift is null), but the infeasibilities modified by the shift. The idea is that, even with a penalty parameter of moderate value, a reasonable choice of the shift makes possible the coincidence, or near coincidence, of the solution to problem (19). As there are updates of both Lagrange multipliers and the penalty parameter, we must take care to ensure the boundedness of λ and this is done by a safeguard: when ρ tends to infinity, common sense dictates that the shifts should be tend to zero.

For the FWI problem, let us assume that we know the velocity somewhere in the analysed region (this information may come from a vertical well profile, for example). Therefore, the constraints can be written as

$$h(\mathbf{w}; x_q, z_q) = c(\mathbf{w}; x_q, z_q) - v_q(x_q, z_q), \quad q = 1, 2, \ldots, Q, \tag{22}$$

where each pair (x_q, z_q) indicates the position of a known velocity value $v_q(x_q, z_q)$. For example, for a known vertical well profile, we can use x_q fixed for different values of z_q. In addition, the set Ω will be a box, i.e., $\Omega = \{\mathbf{w} \in I\!R^n \mid \mathbf{a} \leq \mathbf{w} \leq \mathbf{b}\}$, with $\mathbf{a}, \mathbf{b} \in I\!R^n$. The choice for $\mathbf{a}$ and $\mathbf{b}$ must be made to ensure that $c(\mathbf{w}; \mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega$.

NUMERICAL EXPERIMENTS

We have used the software ALGENCAN (Birgin and Martínez (2014b)) which contains the computational implementation of the Augmented Lagrangian method discussed in the previous section. The book Birgin and Martínez (2014a) presents a detailed description of ALGENCAN, with both theoretical and practical aspects.

We will consider teo different velocity models. The seismic data was obtained by the finite-difference algorithm previously presented, with $M = 3$ and the source term given by $F(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_s)S(t)$, where $\mathbf{x}_s = (x_s, z_s)$ is the source position and S is a Ricker wavelet, with peak frequency of $f_p = 20$ Hz,

$$S(t) = (1 - 2(f_p \pi t)^2) e^{-(f_p \pi t)^2} \tag{23}$$

For the discretization, we have

$$F_{j,k}^{\ell} = \frac{1}{\Delta^2} \delta_{j,j_s} \delta_{k,k_s} S(t_\ell), \tag{24}$$

where j_s, k_s are the indices for the source position and $\delta_{m,n}$ denotes the Kronecker's delta. Altough we are useing absorbing boundary conditions, the seismic survey will be simulated considering a free-surface at $z = -40$ m.

Problem	Estimated $\mathbf{w}$	CPU Time (h)	Flag
COp	$(0.5000, 3.0000)$	6.4	Solution found
COp-vp5	$(0.4585, 3.1142)$	3.7	Infeasible stationary point found
UOp	$(0.8214, 0.5654)$	2.7	Solution found
COp30	$(0.5000, 3.0000)$	10.1	Solution found
COp30-vp5	$(0.4585, 3.1142)$	3.8	Infeasible stationary point found
UOp30	$(0.8229, 0.5579)$	1.8	Solution found

Table 1: Numerical results for Experiment 1.

Experiment 1: Constant Gradient Velocity

For the first experiment, the velocity model is given by,

$$v(x, z) = 2.0 + 0.5x + 3.0z, \tag{25}$$

with x and z measured in kilometers and t in seconds, so that velocities are measured in km/s. For convenience, we will denote by v the true velocity used to create the observed data and by c the velocity that will be estimated by Algencan. For $-4 \leq x \leq 4$ and $0 \leq z \leq 2.5$ we have $v_{min} = 2$ km/s $v_{max} = 5.5$ km/s as illustrated in Figure 1. The finite difference parameters are $\Delta = 10$ m, $\Delta t = 1$ ms and $T = 2$ s. The seismic survey consists of three shots, with $\mathbf{x}_s = -3.5, -2.5, -1.5$ and $z_s = 0$. For each shot the first receiver is at a distance of 200 m from the source with 41 receivers with a distance between them of 125 m. Figure 1 also shows, for the second shot, the modeled data with the exact velocity model, without noise and with a 30% random noise added.

For the inverse problem, we want to estimate the weights $\mathbf{w} = (w_1, w_2)$, such that

$$c(\mathbf{w}; x, z) = \phi_0(\mathbf{x}) + w_1\phi_1(x, z) + w_2\phi_2(x, z), \tag{26}$$

where $\phi_0(\mathbf{x}) = 2$, $\phi_1(x, z) = x$ and $\phi_2(x, z) = z$. We choose $\Omega = \{\mathbf{w} \in I\!R^2 \mid 0 \leq \mathbf{w}_1, \mathbf{w}_2 \leq 5\}$, and for the equality constraints given by equation (22) we have considered a vertical velocity profile at $x_q = 2$ km and $Q = 49$ equally spaced values for z_q between 60 m 2.46 km (black $\times$'s in Figure 1a).

In order to analyze whether the addition of the constraints contribute positively to the problem, we also solve the inverse problem without the constraints, considering the set Ω only. In this case, the subproblem (21) is solved only once.

The initial guesses were $\mathbf{w}^0 = (0, 0)$ and $\lambda_q^0 = 0$, and we have used $\varepsilon_{opt} = 10^{-5}$ and $\varepsilon_{feas} = 10^{-6}$, as the precisions for the stopping criteria for optimality and feasibility, respectively. The maximum number of iterations for the method is 15, with a maximum number of inner iteration equals to 20 (subproblems).

The results for the inverted weights using the ALGENCAN package are shown in Table 1, where we have used the following code:

- COp means that the problem was solved considering the constraints and observed data without noise.
- COp-vp5 means that the problem was solved considering the constraints with a 5% random noise added to the known velocity profile and observed data without noise.
- UOp means that the problem was solved without constraints and observed data without noise.
- COp30, COp30-vp5 and UOp30 indicates that we have added a 30% noise to the observed data.

From Table 1 we see that when the problem has constraints we have more success to find the best velocity model approximation, even in the case where the constraint was contaminated with random noise. For the unconstrained case, even though the solution found is a minimizer, it is not a good enough. Figure 2 shows the modeled data for the second shot using the estimates found and Figure 3 shows the difference between the observed and modeled data. As expected for COp, the difference between the data is very close to zero. In Figure 4 we show the velocity models obtained with the optimized weights and Figure 5 depicts the percentage errors. We can observe that for the COp cases the percentage errors are very small (a maximum of 2%) whereas for the UOp cases the percentage errors can reach 80%. To look in more detail, in Figure 6 we plot the velocity profile used for the constraints compared to the estimated one.

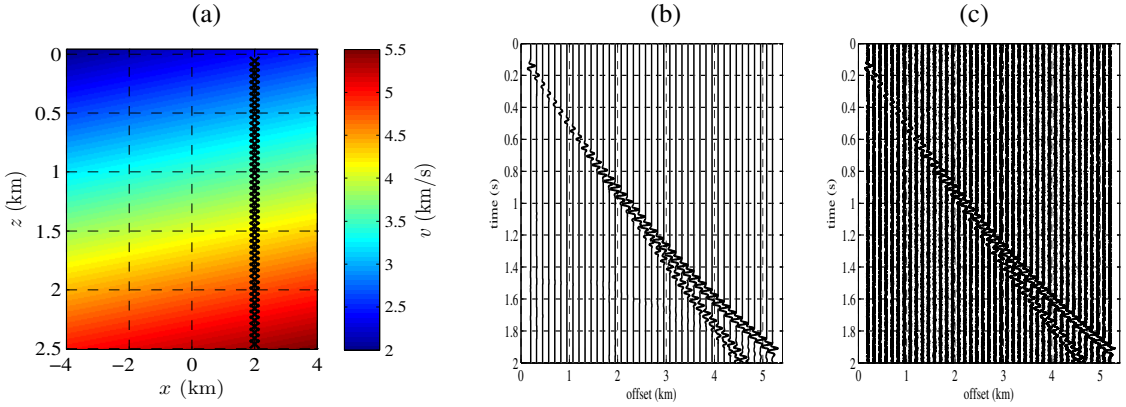

Figure 1: Data for Experiment 1: (a) Exact velocity model; (b) Second shot without noise; (c) Second shot with the addition of 30% random noise.

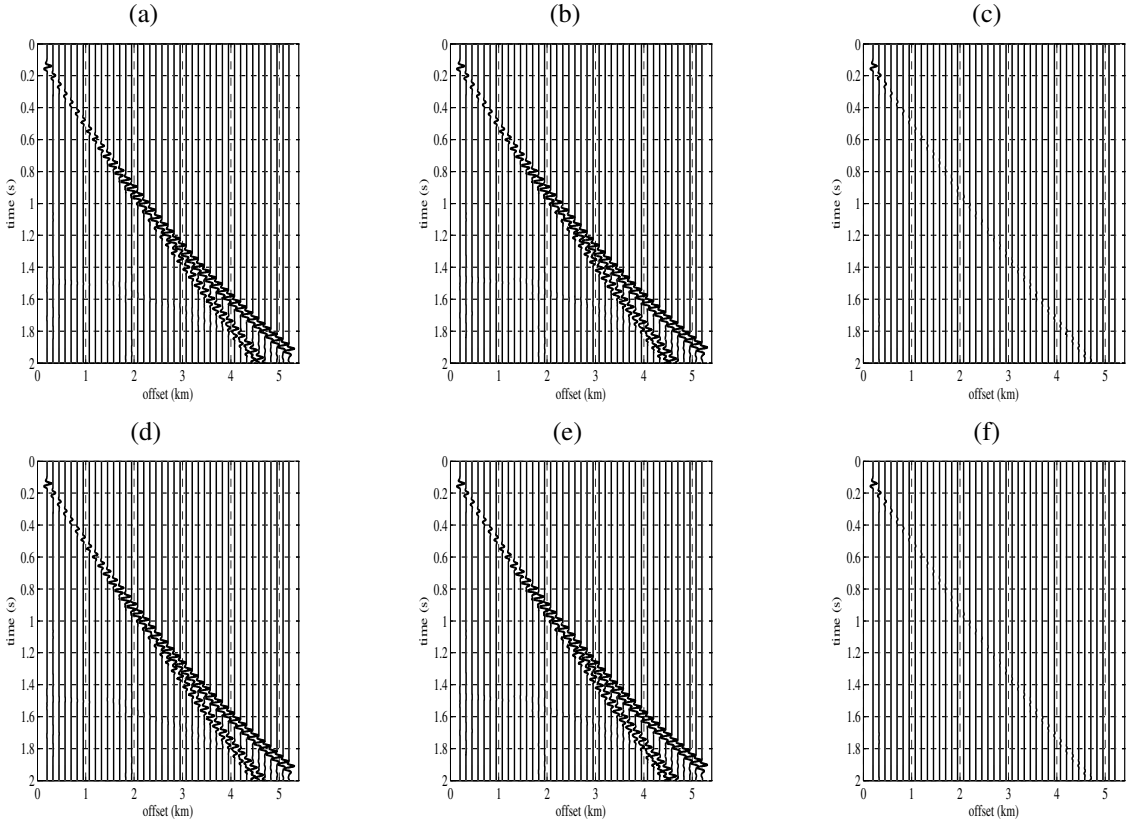

Figure 2: Modeled seismic data for the second shot in Experiment 1, after the optimization procedure: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

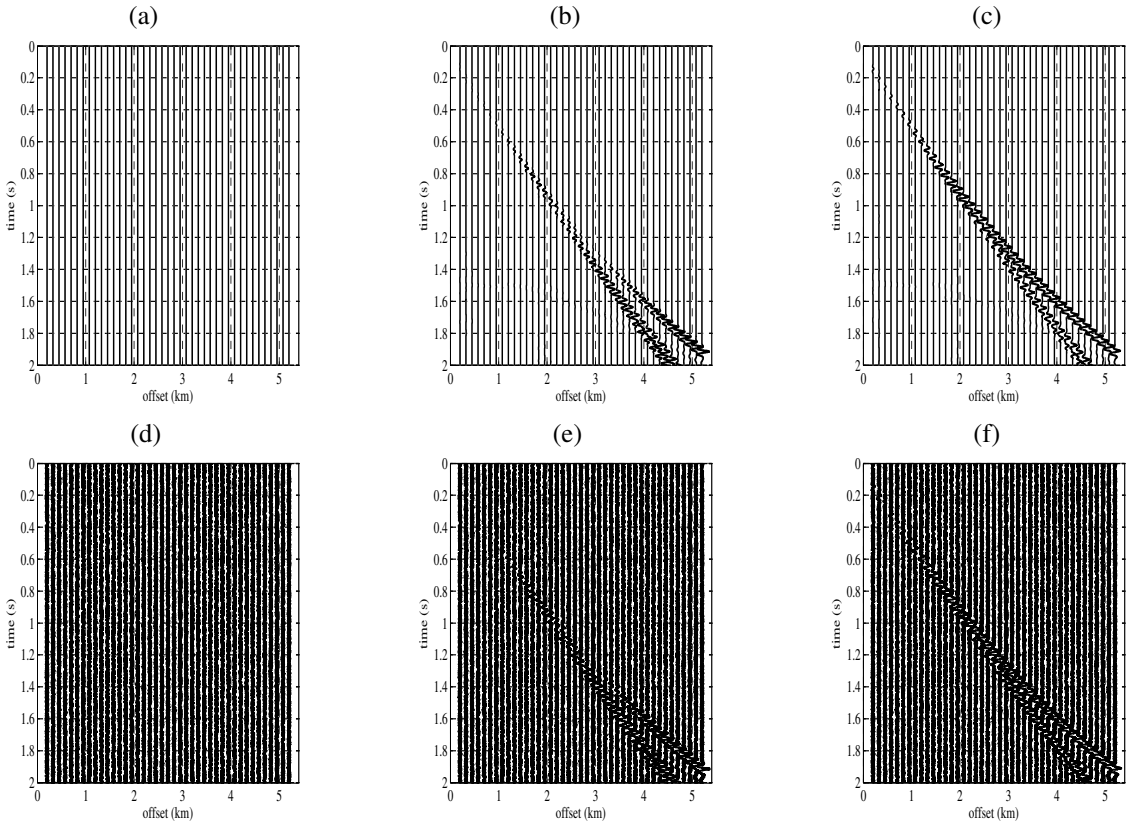

Figure 3: Difference between modeled and observed data for the second shot in Experiment 1: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

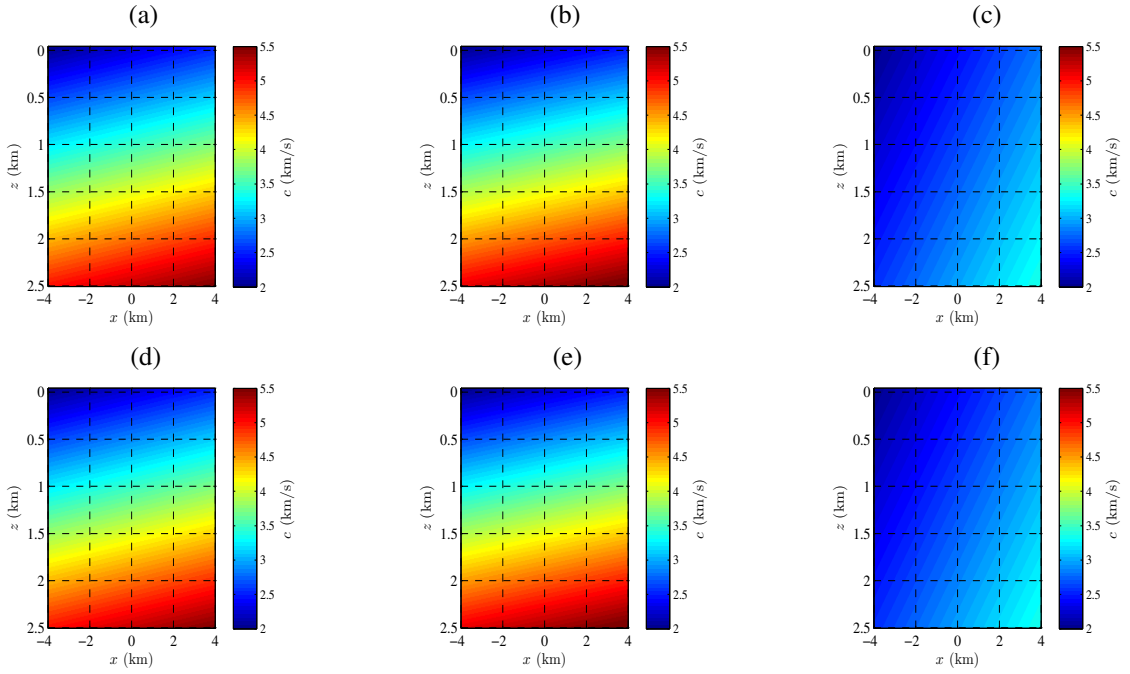

Figure 4: Velocity models estimated by inversion in Experiment 1: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

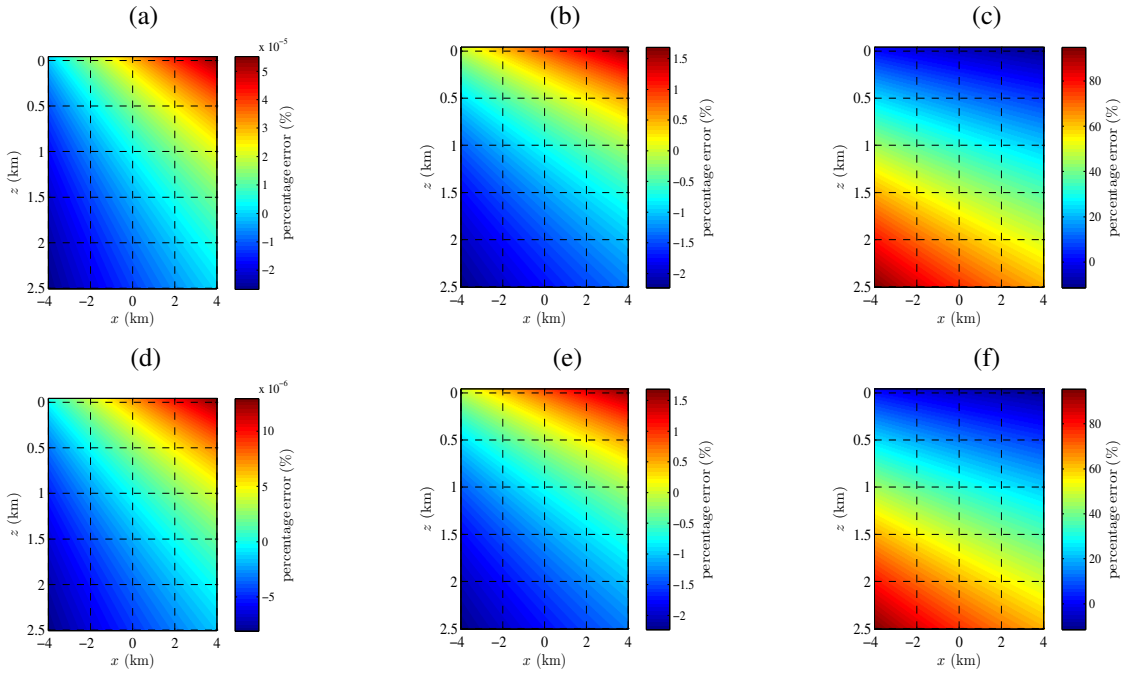

Figure 5: Percentual error for the velocity in Experiment 1: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

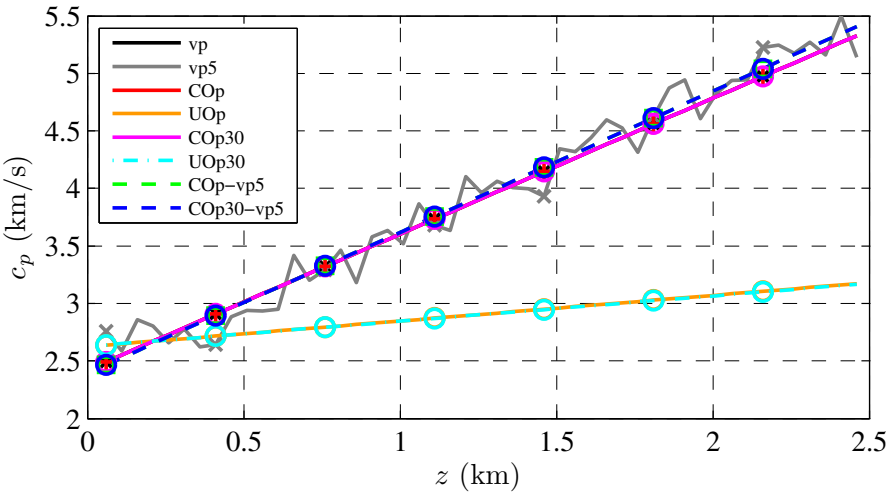

Figure 6: Comparison of the estimates with the exact value in the vertical velocity profile at $x_p = 2$ km in Experiment 1.

Problem	Estimated $\mathbf{w}$	CPU Time (h)	Flag
COp	(2.5000, 2.5000, 2.5000, 2.5000 3.5000, 3.5000, 3.5000)	5.7	Solution found
COp-vp5	(2.6115, 2.4142, 2.4637, 2.4945 3.5179, 3.5302, 3.5529)	4.2	Infeasible stationary point found
UOp	(5.0000, 5.0000, 5.0000, 5.0000 5.0000, 5.0000, 4.7123)	2.9	Solution found
COp30	(2.5000, 2.5000, 2.5000, 2.5000 3.5000, 3.5000, 3.5000)	6.6	Solution found
COp30-vp5	(2.6119, 2.4142, 2.4637, 2.4945 3.5179, 3.5178,3.5302)	4.3	Infeasible stationary point found
UOp30	(4.9775, 3.8496, 2.9918, 2.5100 2.3027, 2.5535, 2.5448)	4.7	Failure

Table 2: Numerical results for Experiment 2.

Experiment 2: Plane Interface

The velocity model for this experiment is illustrated in Figure 7 and is given by

$$v(x,z) = \begin{cases} 2.5, & \text{for} \quad z \leq 1.0, \\ 3.5, & \text{for} \quad z > 1.0. \end{cases} \tag{27}$$

The seismic survey consists of five shots with for $x_s = -2.0, -1.5, -1.0, -0.5, 0.0$ and $z_s = 0$. For each shot the first receiver is at a distance of 200 m from the source with 41 receivers spaced by 50 m. The finite difference parameters are $\Delta = 10$ m, $\Delta t = 10^{-3}$ s, and the maximum time registrated is $T = 2$ s. Figure 7 also shows the modeled data, for the third shot located at -1.0 km, using the exact velocity model and without noise and with 30% random noise added.

Since we are considering that the velocity model can be described by a combination of base functions, we take $\phi_0(\mathbf{x}) = 0$ and

$$\phi_i(x,z) = \begin{cases} 1, & \widehat{z}_{i-1} \leq z \leq \widehat{z}_i, \\ 0, & \text{otherwise,} \end{cases} \quad , \quad i = 1, 2, \ldots, 7, \tag{28}$$

where $\widehat{z}_i = 0.25i$ $(i = 1, 2, \ldots, 5)$, $\widehat{z}_6 = 2$ and $\widehat{z}_7 = 4$. Therfore, the optimal solution is $\mathbf{w}^* = (2.5, 2.5, 2.5, 2.5, 3.5, 3.5, 3.5)$. We choose $\Omega = \{\mathbf{w} \in I\!R^7 \mid 1.5 \leq \mathbf{w}_i \leq 5,\ i = 1, 2, \ldots, 7\}$, and for the equality constraints given by equation (22) we have considered a vertical velocity profile at $x_q = 1.5$ km and $Q = 49$ equally spaced values for z_q between 60 m 2.46 km (black $\times$'s in Figure 7a.

For the optimization procedure, the initial guesses were $\mathbf{w}^0 = (1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5)$ and $\lambda_q^0 = 0$, and we have used $\varepsilon_{opt} = 10^{-5}$ and $\varepsilon_{feas} = 10^{-6}$, as the precisions for the stopping criteria for optimality and feasibility, respectively. The maximum number of iterations for the method is 15, with a maximum number of inner iteration equals to 20 (subproblems). The corresponding velocity model for $\mathbf{w}^0$ is depicted in Figure (8).

Table 2 shows the numerical results, which are very similar to the results for Experiment 1. The estimated values for the weights from the optimization procedure with constraints are much better than the ones that came from the unconstrained procedure. In particular, for the UOp problem, the solution found is almost the upper limit for the box defined by Ω. We also have a new flag for UoOp30, indicating a failure to obtain the solution.

Figure 9 shows the modeled data for the third shot at $x_s = -1.0$km and Figure 10 shows the differences with the observed data. In Figure 11 we show the velocity models obtained with the optimized weights and Figure 12 depicts the percentage errors.

We can observe that for the COp cases, the percentage errors are very small (a maximum of 4%) whereas for the UOp cases the percentage errors can reach 50%. In Figure 13 we plot the velocity profile used for the constraints compared to the estimated one.

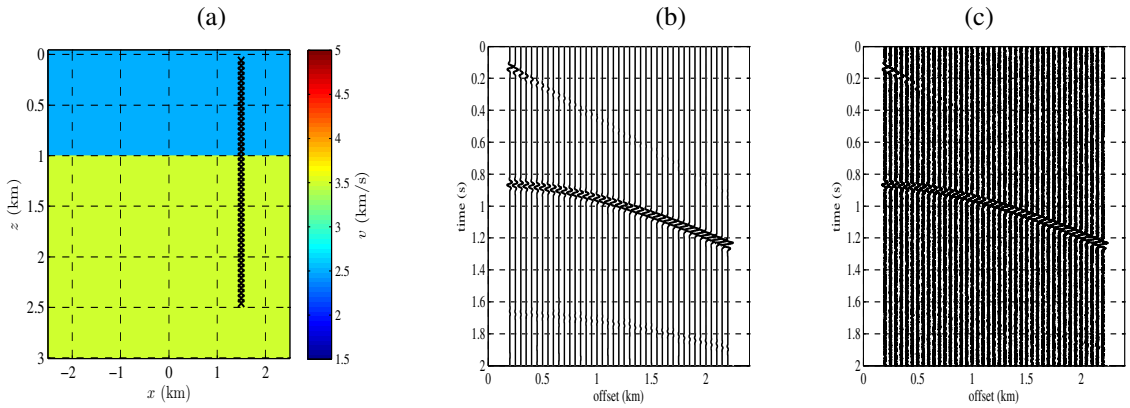

Figure 7: Data for Experiment 2: (a) Exact velocity model; (b) Third shot without noise; (c) Third shot with the addition of 30% random noise.

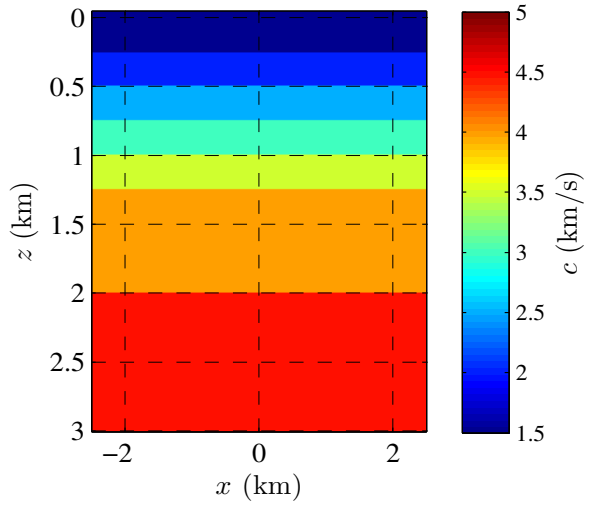

Figure 8: Initial velocity model for Experiment 2.

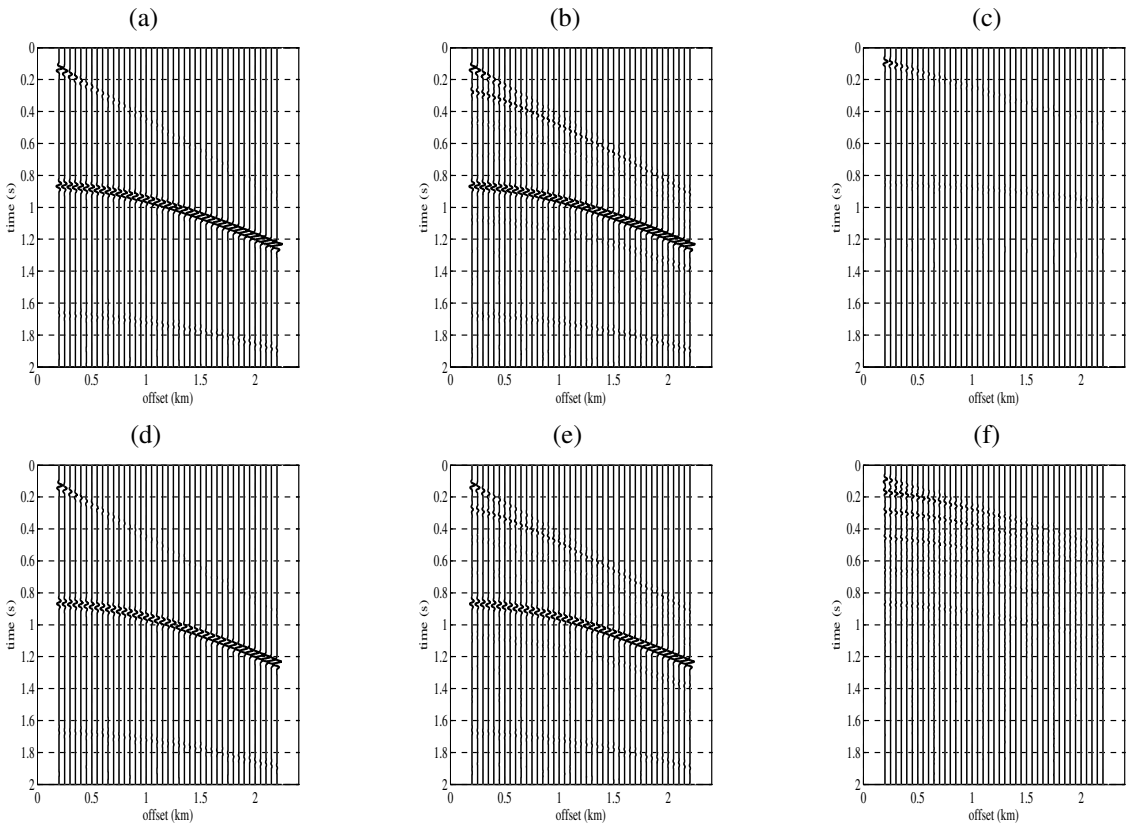

Figure 9: Modeled seismic data for the third shot in Experiment 2, after the optimization procedure: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

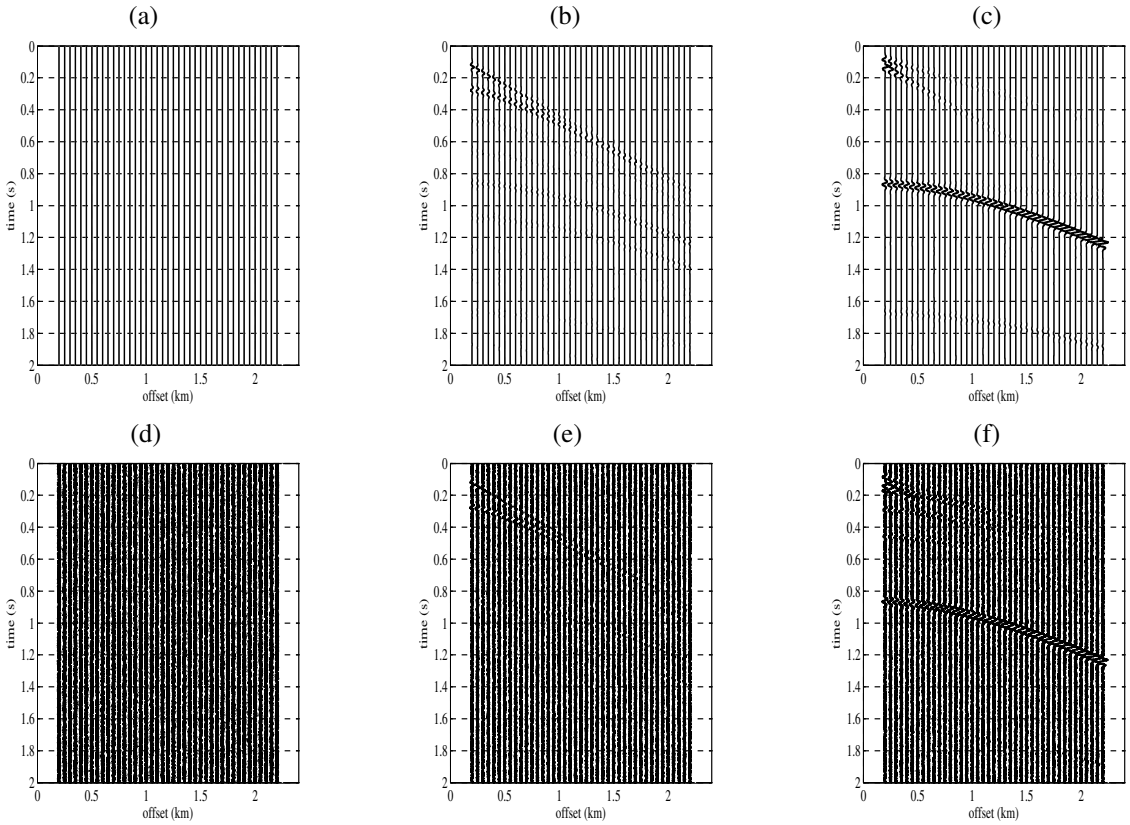

Figure 10: Difference between modeled and observed data for the third shot in Experiment 2: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

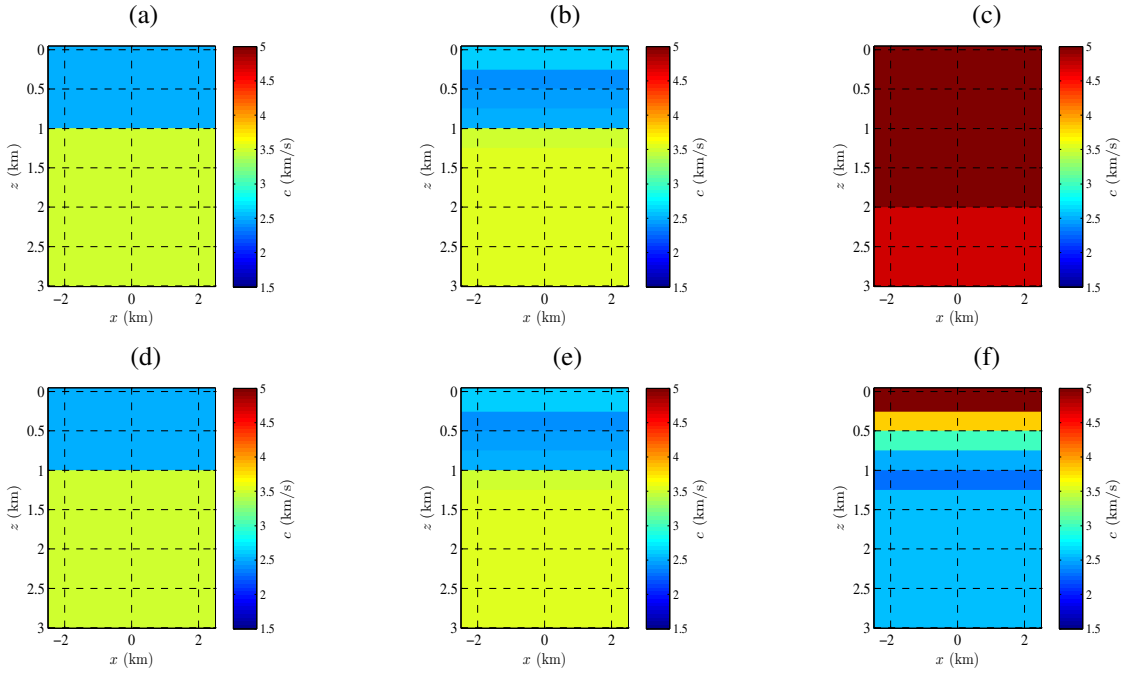

Figure 11: Velocity models estimated by inversion in Experiment 2: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

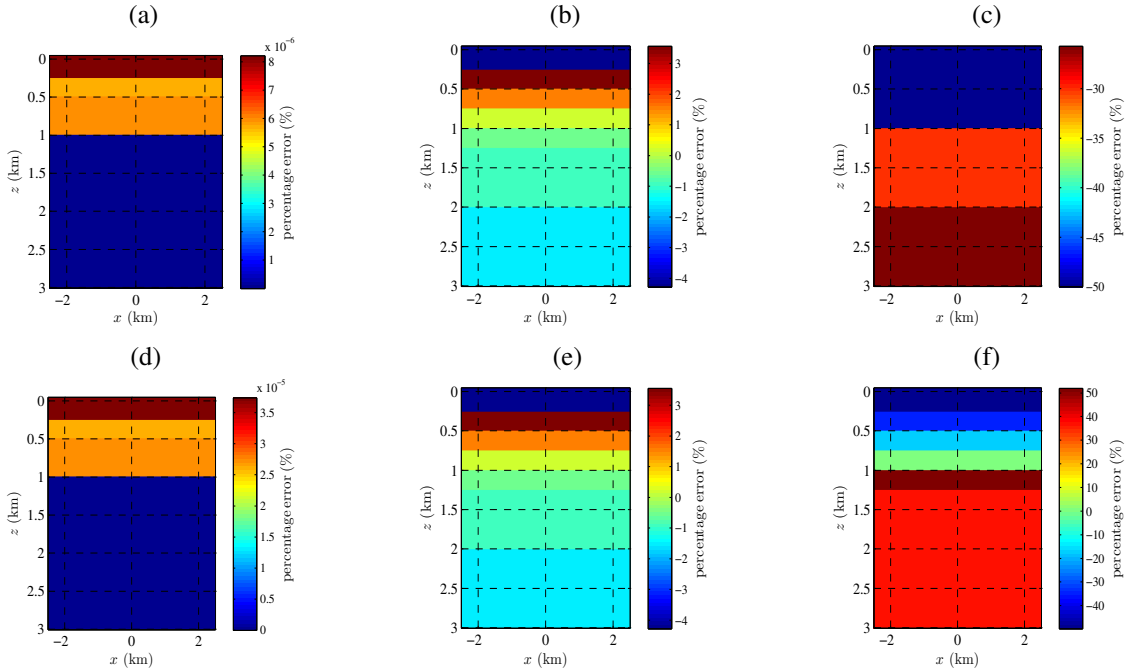

Figure 12: Percentual error for the velocity in Experiment 2: (a) COp; (b) COp-vp5; (c) UOp; (d) COp30; (e) COp30-vp5; (f) UOp30.

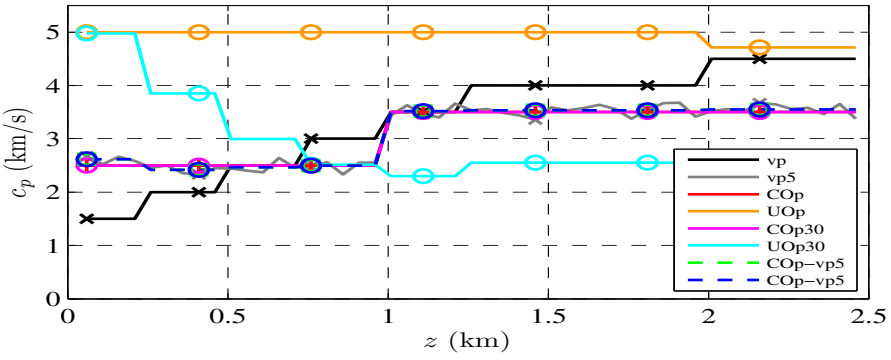

Figure 13: Comparison of the estimates with the exact value in the vertical velocity profile at $x_p = 1.5$ km in Experiment 2.

CONCLUSIONS

We tested the problem of seismic inversion from the point of view of a constrained optimization problem through an Augmented Lagragian method using the software ALGENCAN. The advantage of considering constraints in the experiments performed show good benefits for the results when compared to the unconstrained problem. Further studies are necessary, mainly to apply the same approach to more a complex velocity model and alanlyse diferent type of possible constraints.

ACKNOWLEDGMENTS

This work was kindly supported by the research agencies CAPES and CNPq, Brazil, and the sponsors of the *Wave Inversion Technology (WIT) Consortium*, Karlsruhe, Germany.

REFERENCES

Birgin, E. and Martínez, J. (2014a). *Practical Augmented Lagragian Methods for Constrained Optimization*. Society for Industrial and Applied Mathematics.

Birgin, E. and Martínez, J. (2014b). *Trustable Algorithms for Nonlinear General Optimization*. Software available online at www.ime.usp.br/ egbirgin/tango/publications.php.

Bunks, C., Saleck, F., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60:1457–1473.

Fichtner, A. (2011). *Full seismic waveform modelling and inversion*. Springer.

Kosloff, R. and Kosloff, D. (1986). Absorbing boundaries for wave propagation problems. *Journal of Computational Physics*, 63(2):363–373.

Leeuwen, T. V. and Herrmann, F. J. (2016). A penalty method for pde−constrained optimization in inverse problems. *Inverse Problems*, 32(1):1–27.

Liu, Y. and Sen, M. K. (2009). A new timespace domain high−order finite difference method for the acoustic wave equation. *Journal of Computational Physics*, 228(23):8779–8806.

Liu, Y., Teng, J., Xu, T., Wang, Y., Liu, Q., and Badal, J. (2016). Robust time−domain full waveform inversion with normalized zero−lag cross−correlation objective function. *Geophysical journal international*, 209(1):106–122.

Nocedal, J. and Wright, S. W. (1999). *Numerical optimization*. Springer.

Plessix, R. (2006). A review of the adjoint state method for computing the gradient of a functional with geophysical a applications. *Geophysical journal international*, 167(2):495–503.

Seki, T. and Nishikawa, T. (1988). Absorbing boundaries for wave propagation problems (using kosloff's method in absorbing region). In *Proceeding 9th World Conference on Earthquake Engineering*, volume 2, pages 629–634.

Shin, C. and Ha, W. (2008). A comparison between the behavior of objective function for waveform inversion in the frequency and laplace domains. *Geophysics*, 73(5):119–133.

Symes, W. (2008). Migration velocity analysis and waveform inversion. *Geophysical Prospecting*, 56(6):765–790.

Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):127–152.

Wu, R., Luo, J., and Wu, B. (2014). Seismic envelope inversion and modulation signal model. *Geophysics*, 79(3):13–24.