

COHERENCE MEASURES IN AUTOMATIC TIME MIGRATION VELOCITY ANALYSIS

J.S. Maciel, J.C. Costa, and J. Schleicher

email: *jesse@ufpa.br*

keywords: *Migration velocity analysis, time migration, tomography*

ABSTRACT

Time-migration velocity analysis can be carried out automatically by evaluating the coherence of the migrated seismic events in the common-image gathers (CIGs). The performance of a conjugate-gradient method for automatic time-migration velocity analysis depends on the the coherence measures used in the objective function. We compare the results of four different coherence measures, being are conventional semblance, differential semblance, an extended differential semblance using extended images, and the product of the latter with conventional semblance. In our numerical experiments, the objective functions based on conventional semblance and on the product of conventional semblance with extended differential semblance provided the best velocity models, as evaluated by the flatness of the resulting common-image gathers. The method can be easily extended to anisotropic media.

INTRODUCTION

Kirchhoff time migration aims at focusing seismic events at an image point by stacking the observed data along a analytically specified diffraction-traveltime surface. This surface represents the traveltime of a diffraction at the image point. The choice of an analytic representation of the stacking surface restricts the applicability of Kirchhoff time migration to media with moderate velocity variations in the scale of the acquisition geometry.

Time migration tries to describe the full complexity of the traveltimes of the wave propagation between the image point and the earth's surface by a relatively small number of parameters at each image point. In this respect, the parameter set attributed to each image point defines an effective constant-velocity medium that defines the diffraction-traveltime surface there, but is strictly valid only at this image point.

Time migration velocity analysis consist of estimating this set of parameters at each image point. In spite of is obvious intrinsic contradictions this process has a number of advantages, the most important ones being its algorithmic simplicity, low computational cost, and high robustness in comparison to the construction of a depth migration velocity model.

Because of the underlying assumption of a smoothly varying velocity, the resulting parameter sets are also supposed to vary smoothly in the image space. Consequently, they can be represented by smooth interpolators such as, e.g., B-splines. The grid spacing between adjacent nodes can be of the order of the length of the seismic line, which greatly reduces the dimension of the parameter space.

Simplicity and robustness of time migration justify its use in the standard processing sequence of most seismic data. Even when the time migrated image is not convenient for structural interpretation, the so-constructed time-migration velocity model can still be used for the construction of an initial depth velocity model by means of time-to-depth conversion. This initial depth velocity model can then later be refined using more sophisticated depth migration-velocity analysis techniques like tomographic methods (Billette et al., 2003).

Recently, we have studied different methods of obtaining an initial model for time time migration (Schleicher et al., 2008; Schleicher and Costa, 2009). Now, we are interested in how to refine such an initial model. An efficient criterion to improve a time-migration velocity model is to require that migrated images of the same image point as obtained, e.g., from migration of neighbouring common-offset sections, must exhibit a certain degree of similarity. A common method to evaluate this similarity is the use of common-image gathers (CIGs), where all migrated traces that correspond to the same horizontal position in the image space are displayed next to each other. If the model parametrization is consistent with the data, the corresponding events must align horizontally in a CIG (Faye and Jeannot, 1986; Zhu et al., 1998). This criterion can be automatically applied by evaluating the coherence along horizontal lines through the CIG (Sattlegger, 1975; Abbad et al., 2009).

In this work, we investigate different implementations of time-migration velocity analysis that use the optimization of different types of coherence measures in the CIGs. In our numerical test, we use conventional semblance (Neidell and Taner, 1971), differential semblance (Symes and Carazzone, 1991), and two related new coherence measures based on extended images.

COHERENCE MEASURES FOR TIME-MIGRATION VELOCITY ANALYSIS

Migration velocity analysis by means of the minimization of differential semblance (Symes and Carazzone, 1991) was proposed by Chauris and Noble (2001), since this is frequently considered the most robust of the conventional coherence measures. To solve this optimization problem, the adjoint-state method (Mulder and ten Kroode, 2002; Plessix, 2006) is very convenient to analytically calculate the gradient of the objective function. We start from the general migration-velocity-analysis functional for 2D or 3D time or depth migration using differential semblance as proposed by Mulder and ten Kroode (2002),

$$\begin{aligned} \text{Minimize } J[\boldsymbol{\eta}(\mathbf{x})] &= \frac{1}{2} \int d\mathbf{x} \frac{\int dz \int dh \left[\int d\zeta S(h - \zeta) \frac{\partial}{\partial \zeta} (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right]^2}{\int dz \int dh \left[\int d\zeta S(h - \zeta) \Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta) \right]^2} \\ \text{subject to } I(\mathbf{x}, h) &= \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') \int d\xi W(\xi, h, \mathbf{x}'; \boldsymbol{\eta}) \mathcal{D}[\xi, h, t = \tau_D(\xi, h, \mathbf{x}'; \boldsymbol{\eta})], \end{aligned} \quad (1)$$

where $\mathbf{x} = (x, z)$ denotes the image point, with x representing one (2D) or two (3D) horizontal coordinates and z representing depth or vertical time. Also, $\boldsymbol{\eta}(\mathbf{x})$ represents the parameter set that describes the diffraction-traveltime surface at image point \mathbf{x} , and $I(\mathbf{x}, h)$ is the migrated section at \mathbf{x} for half-offset h , possibly filtered by a spatial filter F , e.g., a mask to pass only selected regions to the migration-velocity analysis, or some smoothing operator. Moreover, $\Gamma(\mathbf{x}, h)$ represents the muting operator applied to the CIGs, $S(h)$ is a smoothing operator, $\mathcal{D}(\xi, h, t)$ represents a single common-offset section with ξ being the CMP coordinate. Finally, W indicates a weight function for the migration operator, which includes a frequency filter (derivative in 3D, half-derivative in 2D), and τ_D represents the diffraction-traveltime surface.

Gradient computation using the adjoint-state method

The augmented Lagrangian associated with problem (1) is easily seen to be given by

$$\mathcal{L}(I, \boldsymbol{\eta}, \lambda) = \frac{1}{2} \int d\mathbf{x} \frac{D\Lambda(x)}{\Lambda(x)} + \int d\mathbf{x} \int dh \lambda(\mathbf{x}, h) \left\{ I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') M(\mathbf{x}, h, \boldsymbol{\eta}) \right\}, \quad (2)$$

where $\lambda(\mathbf{x}, h)$ are Lagrangian multipliers. Moreover, to simplify the notation in the derivation of the functional's gradient below, we have introduced the symbols

$$D\Lambda(x) \equiv \int dz \int dh \left[\int d\zeta S(h - \zeta) \frac{\partial}{\partial \zeta} (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right]^2, \quad (3)$$

$$\Lambda(x) \equiv \int dz \int dh \left[\int d\zeta S(h - \zeta) (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right]^2, \quad (4)$$

and

$$M(\mathbf{x}, h, \boldsymbol{\eta}) \equiv \int d\xi W(\xi, h, \mathbf{x}'; \boldsymbol{\eta}) \mathcal{D}[\xi, h, t = \tau_D(\xi, h, \mathbf{x}'; \boldsymbol{\eta})]. \quad (5)$$

The variation of the Lagrangian functional (2) with respect to I , c e λ can then be written as

$$\begin{aligned} \delta\mathcal{L} = & \int dx \frac{\int dz \int dh \left[\int d\zeta S(\zeta - h) \frac{\partial}{\partial \zeta} (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right] \left[\int d\beta S(\beta - h) \frac{\partial}{\partial \beta} (\Gamma(\mathbf{x}, \beta) \delta I(\mathbf{x}, \beta)) \right]}{\Lambda(x)} \\ & - \int dx \frac{D\Lambda(x)}{\Lambda(x)^2} \int dz \int dh \left[\int d\zeta S(\zeta - h) \Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta) \right] \left[\int d\beta S(\beta - h) \Gamma(\mathbf{x}, \beta) \delta I(\mathbf{x}, \beta) \right] \\ & + \int d\mathbf{x} \int dh \lambda(\mathbf{x}, h) \left\{ \delta I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') \delta \boldsymbol{\eta}(\mathbf{x}') \cdot \nabla_{\boldsymbol{\eta}} M(\mathbf{x}, h, \boldsymbol{\eta}) \right\} \\ & + \int d\mathbf{x} \int dh \delta \lambda(\mathbf{x}, h) \left[I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') M(\mathbf{x}, h, \boldsymbol{\eta}) \right]. \end{aligned} \quad (6)$$

The last factor in the numerator of the fraction in the first line of equation (6) can be integrated by parts. This yields

$$\int d\beta S(\beta - h) \frac{\partial}{\partial \beta} (\Gamma(\mathbf{x}, \beta) \delta I(\mathbf{x}, \beta)) = - \int d\beta \Gamma(\mathbf{x}, \beta) \delta I(\mathbf{x}, \beta) \frac{\partial}{\partial \beta} S(\beta - h), \quad (7)$$

where the boundary term vanishes, because the muting operator Γ satisfies $\Gamma(\mathbf{x}, h_{min}) = \Gamma(\mathbf{x}, h_{max}) = 0$.

Substitution of equation (7) back in equation (6) yields, after reorganization of the integrals in the first and second lines with convenient changes of the integration variables,

$$\begin{aligned} \delta\mathcal{L} = & - \int dx \int dz \int dh \delta I(\mathbf{x}, h) \Gamma(\mathbf{x}, h) \frac{\int d\beta \frac{\partial}{\partial h} S(h - \beta) \left[\int d\zeta S(\zeta - \beta) \frac{\partial}{\partial \zeta} (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right]}{\Lambda(x)} \\ & - \int dx \int dz \int dh \delta I(\mathbf{x}, h) \Gamma(\mathbf{x}, h) \frac{D\Lambda(x)}{\Lambda(x)^2} \int d\beta S(h - \beta) \left[\int d\zeta S(\zeta - \beta) \Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta) \right] \\ & + \int d\mathbf{x} \int dh \lambda(\mathbf{x}, h) \left\{ \delta I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') \delta \boldsymbol{\eta}(\mathbf{x}') \cdot \nabla_{\boldsymbol{\eta}} M(\mathbf{x}, h, \boldsymbol{\eta}) \right\} \\ & + \int d\mathbf{x} \int dh \delta \lambda(\mathbf{x}, h) \left[I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') M(\mathbf{x}, h, \boldsymbol{\eta}) \right], \end{aligned} \quad (8)$$

which can then be regrouped into

$$\begin{aligned} \delta\mathcal{L} = & - \int dx \int dz \int dh \delta I(\mathbf{x}, h) \\ & \left\{ \Gamma(\mathbf{x}, h) \left[\frac{1}{\Lambda(x)} \int d\beta \frac{\partial}{\partial h} S(h - \beta) \left(\int d\zeta S(\zeta - \beta) \frac{\partial}{\partial \zeta} (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right) \right. \right. \\ & \left. \left. + \frac{D\Lambda(x)}{\Lambda(x)^2} \int d\beta S(h - \beta) \left(\int d\zeta S(\zeta - \beta) \Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta) \right) \right] - \lambda(\mathbf{x}, h) \right\} \\ & - \int d\mathbf{x} \int dh \lambda(\mathbf{x}, h) \left[\int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') \delta \boldsymbol{\eta}(\mathbf{x}') \cdot \nabla_{\boldsymbol{\eta}} M(\mathbf{x}, h, \boldsymbol{\eta}) \right] \\ & + \int d\mathbf{x} \int dh \delta \lambda(\mathbf{x}, h) \left[I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') M(\mathbf{x}, h, \boldsymbol{\eta}) \right]. \end{aligned} \quad (9)$$

It is now convenient to choose λ as

$$\begin{aligned} \lambda(\mathbf{x}, h) = & \Gamma(\mathbf{x}, h) \left\{ \frac{1}{\Lambda(x)} \int d\beta \frac{\partial}{\partial h} S(h - \beta) \left[\int d\zeta S(\zeta - \beta) \frac{\partial}{\partial \zeta} (\Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta)) \right] \right. \\ & \left. + \frac{D\Lambda(x)}{\Lambda(x)^2} \int d\beta S(h - \beta) \left[\int d\zeta S(\zeta - \beta) \Gamma(\mathbf{x}, \zeta) I(\mathbf{x}, \zeta) \right] \right\} \end{aligned} \quad (10)$$

and also use that

$$I(\mathbf{x}, h) = \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') M(\mathbf{x}, h, \boldsymbol{\eta}) . \quad (11)$$

Substitution in equation (9) yields then

$$\delta\mathcal{L} = - \int d\mathbf{x} \delta\boldsymbol{\eta}(\mathbf{x}) \cdot \nabla_{\boldsymbol{\eta}} \left\{ \int dh \int d\mathbf{x}' \lambda(\mathbf{x}', h) F(\mathbf{x}' - \mathbf{x}) M(\mathbf{x}, h, \boldsymbol{\eta}) \right\} , \quad (12)$$

where $\lambda(\mathbf{x}', h)$ is now given by equation (10). More explicitly, the variation to be calculated is

$$\delta\mathcal{L} = - \sum_{j=1}^N \int d\mathbf{x} \delta\eta_j \left\{ \int dh \int d\mathbf{x}' \lambda(\mathbf{x}', h) F(\mathbf{x}' - \mathbf{x}) \frac{\partial}{\partial\eta_j} M(\mathbf{x}, h, \boldsymbol{\eta}) \right\} , \quad (13)$$

where N represents the number of parameters characterizing the diffraction-traveltime surface.

In the derivative of the migrated image $M(\mathbf{x}, h, \boldsymbol{\eta})$ with respect to the traveltime parameters $\boldsymbol{\eta}$, the derivative of the weight function is of a lower order in frequency than the derivative of the phase term and can, thus, be neglected (Mulder and ten Kroode, 2002), i.e.,

$$\frac{\partial}{\partial\eta_j} M(\mathbf{x}, h, \boldsymbol{\eta}) \approx \int d\xi W(\xi, h, \mathbf{x}; \boldsymbol{\eta}) \mathcal{D}'[\xi, h, t = \tau_D(\xi, h, \mathbf{x}; \boldsymbol{\eta})] \frac{\partial\tau_D}{\partial\eta_j} . \quad (14)$$

This yields the final expression for the gradient of the objective function as

$$\delta\mathcal{L} = - \sum_{j=1}^N \int d\mathbf{x} \delta\eta_j \left\{ \int dh \int d\mathbf{x}' \lambda(\mathbf{x}', h) F(\mathbf{x}' - \mathbf{x}) \int d\xi W(\xi, h, \mathbf{x}; \boldsymbol{\eta}) \mathcal{D}'[\xi, h, t = \tau_D(\xi, h, \mathbf{x}; \boldsymbol{\eta})] \frac{\partial\tau_D}{\partial\eta_j} \right\} . \quad (15)$$

In other words, the dependence of the objective function on the medium parameters is essentially contained in the derivative $\partial\tau_D/\partial\eta_j$. Note that this derivation is valid for general medium representations, independently of the actual type and number of parameters $\boldsymbol{\eta}$. In particular, it can be applied in anisotropic media once a suitable traveltime parameterization is available.

Implementational aspects

The implementation of an algorithm of time-migration velocity analysis based on the above discussion requires a spatial representation of the velocity model. If we assume that the model parameters are represented by B-splines of the form

$$\eta_j(\mathbf{x}) = \sum_I \sum_J \eta_j^{IJ} B_I(x_1) B_J(x_3) , \quad (16)$$

we can represent the gradient $\delta\mathcal{L}$ as

$$\delta\mathcal{L} = - \sum_{j=1}^N \sum_I \sum_J \delta\eta_j^{IJ} \int d\mathbf{x} B_I(x_1) B_J(x_3) \int dh \int d\mathbf{x}' \lambda(\mathbf{x}', h) F(\mathbf{x}' - \mathbf{x}) \int d\xi W(\xi, h, \mathbf{x}; \boldsymbol{\eta}) \mathcal{D}'[\xi, h, t = \tau_D(\xi, h, \mathbf{x}; \boldsymbol{\eta})] \frac{\partial\tau_D}{\partial\eta_j} . \quad (17)$$

In other words, in terms of the B-splines coefficients, we can conclude that

$$\frac{\partial\mathcal{L}}{\partial\eta_j^{IJ}} = \frac{\partial J}{\partial\eta_j^{IJ}} = - \int d\mathbf{x} B_I(x_1) B_J(x_3) \left\{ \int dh \int d\mathbf{x}' \lambda(\mathbf{x}', h) F(\mathbf{x}' - \mathbf{x}) \int d\xi W(\xi, h, \mathbf{x}; \boldsymbol{\eta}) \mathcal{D}'[\xi, h, t = \tau_D(\xi, h, \mathbf{x}; \boldsymbol{\eta})] \frac{\partial\tau_D}{\partial\eta_j} \right\} . \quad (18)$$

Moreover, from the point of view of computational efficiency, it is important to observe that the derivative of the migrated image $M(\mathbf{x}, h, \boldsymbol{\eta})$ with respect to the medium parameters $\boldsymbol{\eta}$, given by equation (14), is easily calculated during the migration process. It represents a second sum with a slightly modified weight factor, adding only a very small extra cost to the migration procedure.

The computation of the gradient is thus given simply by projecting the product of the two images, $\lambda(\mathbf{x}, h) \in \partial M(\mathbf{x}, h, \boldsymbol{\eta})/\partial \eta_j$, into the B-splines bases that represent the model, that is,

$$\frac{\partial \mathcal{L}}{\partial \eta_j^{IJ}} = \frac{\partial J}{\partial \eta_j^{IJ}} = - \int d\mathbf{x} B_I(x_1) B_J(x_3) \left[\int dh \int d\mathbf{x}' \lambda(\mathbf{x}', h) F(\mathbf{x}' - \mathbf{x}) \frac{\partial M}{\partial \eta_j} \right]. \quad (19)$$

The main cost of the gradient evaluation is contained in the computation of the expression within brackets (Mulder and ten Kroode, 2002). To evaluate this expression, two groups of image gathers are needed, being conventional CIGs and corresponding CIGs associated with $\partial M(\mathbf{x}, h, \boldsymbol{\eta})/\partial \eta_j$. In this way, the overall cost of the methods is approximately twice the cost of a simple pre-stack time migration.

The gradient $\partial J/\partial \eta_j^{IJ}$ can be used in optimization methods like conjugate-gradient or quasi-Newton methods, to update the velocity model and minimize the differential semblance in the CIGs. In this way, a given initial model as obtained from, for instance, conventional velocity analysis, can be refined into a better time-migration velocity model.

Time migration in isotropic media

To apply the algorithm to 3D time-migration velocity analysis in isotropic media, let now the image point be explicitly denoted by $\mathbf{x} \equiv (x, y, \tau)$, where (x, y) denote the coordinates of the projection of the image point into the earth's surface, and where τ is the traveltimes along the image ray to the image point. Also, we approximate the diffraction-traveltime surface by

$$\begin{aligned} \tau_D(\xi, h, \mathbf{x}) = t_s + t_r = & \sqrt{\frac{(\xi_x - h_x - x)^2 + (\xi_y - h_y - y)^2}{c^2(x, y, \tau)} + \tau^2} \\ & + \sqrt{\frac{(\xi_x + h_x - x)^2 + (\xi_y + h_y - y)^2}{c^2(x, y, \tau)} + \tau^2}, \end{aligned} \quad (20)$$

where $c(x, y, \tau)$ represents the time migration velocity. Since this is the only medium parameter describing the diffraction-traveltime surface, we observe that $N = 1$ and $\eta_1 = c$ in equation (15). Thus, the parameter derivatives reduce to

$$\begin{aligned} \frac{\partial \tau_D(\xi, h, \mathbf{x})}{\partial c} &= \frac{\partial t_s}{\partial c} + \frac{\partial t_r}{\partial c} \\ &= -\frac{1}{c} \left\{ t_s \left[1 - \left(\frac{\tau}{t_s} \right)^2 \right] + t_r \left[1 - \left(\frac{\tau}{t_r} \right)^2 \right] \right\}. \end{aligned} \quad (21)$$

To correct for the geometrical spreading factor during migration, we choose the weight function as proposed by Peles et al. (2001), i.e.,

$$W(\xi, h, \mathbf{x}) = \frac{1}{2} c \tau \left(\frac{t_s}{t_r} + \frac{t_r}{t_s} \right) \sqrt{\frac{1}{t_s} + \frac{1}{t_r}}. \quad (22)$$

Alternatively, one can use a unit weight function (i.e., no weight and no spreading compensation) during migration. In this case, an approximate geometrical-spreading correction needs to be applied to the data after migration, before the velocity analysis.

Alternative objective functions

A point of criticism to differential semblance as an annihilator to measure the absence of moveout of events in a CIGs relates to its rather local character (Abbad et al., 2009). Differential semblance can

result in small values even if the event still exhibits some residual moveout, particularly for acquisitions with a large offset variation. This feature of differential semblance is a disadvantage for velocity analysis, particularly for anisotropic media. Abbad et al. (2009) propose an alternative objective function that has the disadvantage of not being analytically differentiable. To extend the active region of the annihilator and avoid this problem, we test the performance of the method using the following alternative functionals

$$J_1[\boldsymbol{\eta}(\mathbf{x})] = \frac{1}{2} \int dx \left\{ 1 - \frac{\int dz \left[\int_0^H dh \mathcal{I}(\mathbf{x}, h) \right]^2}{H \int dz \int_0^H dh \mathcal{I}^2(\mathbf{x}, h)} \right\} = \frac{1}{2} \int dx \mathcal{S}(x), \quad (23)$$

$$J_2[\boldsymbol{\eta}(\mathbf{x})] = \frac{1}{2} \int dx \frac{\int dz \int_0^L d\gamma W(\gamma) \int_\gamma^{H-\gamma} dh W_h(h) [\Delta \mathcal{I}(\mathbf{x}, h, \gamma)]^2}{\int dz \int dh \mathcal{I}^2(\mathbf{x}, h)} = \frac{1}{2} \int dx \mathcal{E}(x), \quad (24)$$

$$J_3[\boldsymbol{\eta}(\mathbf{x})] = \frac{1}{2} \int dx \mathcal{S}(x) \mathcal{E}(x), \quad (25)$$

where J_1 is based on conventional semblance $\mathcal{S}(x)$, J_2 uses an extended version $\mathcal{E}(x)$ of the differential semblance, and J_3 uses the product of the latter two coherence measures. In the above functionals, $\mathcal{I}(\mathbf{x}, h)$ is a smoothed version of the migrated image, represented as

$$\mathcal{I}(\mathbf{x}, h) = \int d\zeta S_z(z - \zeta) \int_0^H d\sigma S_h(h - \sigma) \Gamma(x, \zeta, \sigma) I(x, \zeta, \sigma), \quad (26)$$

with $I(x, z, h) = I(\mathbf{x}, h)$ given by the second line of problem (1). Moreover, equations (23) and (24) make use of extended images given by

$$\Delta \mathcal{I}(\mathbf{x}, h, \gamma) = \mathcal{I}(\mathbf{x}, h + \gamma) - \mathcal{I}(\mathbf{x}, h - \gamma). \quad (27)$$

The objective functions (24) and (25) using the extended differential semblance $\mathcal{E}(x)$ retain the desirable properties of differential semblance, but extend the domain of the annihilator.

Gradient of the functionals To use the new functionals in an optimization method, we need their gradients. Since the derivations are rather similar, we determine, as an example, the gradient of functional (24). Its Lagrangian is given by

$$\mathcal{L}(\boldsymbol{\eta}) = \frac{1}{2} \int dx \mathcal{E}(x) + \int d\mathbf{x} \int dh \lambda(\mathbf{x}, h) \left\{ I(\mathbf{x}, h) - \int d\mathbf{x}' F(\mathbf{x} - \mathbf{x}') M(\mathbf{x}, h, \boldsymbol{\eta}) \right\}. \quad (28)$$

The variation of this Lagrangian functional is

$$\begin{aligned} \delta \mathcal{L}(\boldsymbol{\eta}) &= \int dx \frac{\int dz \int_0^L d\gamma W(\gamma) \int_\gamma^{H-\gamma} dh W_h(h) \Delta \mathcal{I}(x, z, h, \gamma) \delta \Delta \mathcal{I}(x, z, h, \gamma)}{\int dz \int dh \mathcal{I}^2(\mathbf{x}, h)} \\ &\quad - \int dx \frac{\mathcal{E}(x)}{\int dz \int d\xi \mathcal{I}^2(x, z, \xi)} \int dz \int d\xi \mathcal{I}(x, z, \xi) \delta \mathcal{I}(x, z, \xi) \\ &\quad + \text{other terms,} \end{aligned} \quad (29)$$

where the other terms arise from the variation of the second term of equation (28) and are identical to the last two lines of equation (6).

Because of equation (26), the variation of the smoothed image \mathcal{I} is given by

$$\delta \mathcal{I}(\mathbf{x}, h) = \int d\zeta S_z(z - \zeta) \int_0^H d\sigma S_h(h - \sigma) \Gamma(x, \zeta, \sigma) \delta I(x, \zeta, \sigma) \quad (30)$$

and that of the extended images by

$$\begin{aligned} \delta \Delta \mathcal{I}(\mathbf{x}, h, \gamma) &= \delta \mathcal{I}(\mathbf{x}, h + \gamma) - \delta \mathcal{I}(\mathbf{x}, h - \gamma) \\ &= \int d\zeta S_z(z - \zeta) \int_0^H d\sigma [S(h + \gamma - \sigma) - S(h - \gamma - \sigma)] \Gamma(x, \zeta, \sigma) \delta I(x, \zeta, \sigma). \end{aligned} \quad (31)$$

Substituting these expressions, as well as the definition

$$\mathcal{J}(\mathbf{x}, h, \gamma) \equiv \int_{\gamma}^{H-\gamma} d\sigma S_h(h-\sigma)W_h(\sigma)\Delta\mathcal{I}(\mathbf{x}, \sigma, \gamma), \quad (32)$$

in equation (29) yields, after a convenient reorganization of the integrals,

$$\begin{aligned} \delta\mathcal{L}(\boldsymbol{\eta}) = & - \int d\mathbf{x} \int dh \delta I(\mathbf{x}, h)\Gamma(\mathbf{x}, h) \\ & \left\{ \frac{\int d\zeta S_z(z-\zeta) \int_0^L d\gamma W(\gamma) [\mathcal{J}(x, \zeta, h+\gamma, \gamma) - \mathcal{J}(x, \zeta, h-\gamma, \gamma)]}{\int d\zeta \int d\xi \mathcal{I}^2(x, \zeta, \xi)} \right. \\ & \left. + \frac{\mathcal{E}(x)}{\int dz \int d\xi \mathcal{I}^2(x, z, \xi)} \int d\zeta S_z(z-\zeta) \int d\xi S(h-\xi)W_h(\xi)\mathcal{I}(x, \zeta, \xi) - \lambda(\mathbf{x}, h) \right\} \\ & + \text{other terms}, \end{aligned} \quad (33)$$

where the other terms now equal the last two lines of equation (9). From equation (33), we recognize that with the following choice for $\lambda(\mathbf{x}, h)$,

$$\begin{aligned} \lambda(\mathbf{x}, h) = \lambda_2(\mathbf{x}, h) = \Gamma(\mathbf{x}, h) & \left\{ \frac{\int d\zeta S_z(z-\zeta) \int_0^L d\gamma W(\gamma) [\mathcal{J}(x, \zeta, h+\gamma, \gamma) - \mathcal{J}(x, \zeta, h-\gamma, \gamma)]}{\int d\zeta \int d\xi \mathcal{I}^2(x, \zeta, \xi)} \right. \\ & \left. + \frac{\mathcal{E}(x)}{\int dz \int d\xi \mathcal{I}^2(x, z, \xi)} \int d\zeta S_z(z-\zeta) \int d\xi S(h-\xi)W_h(\xi)\mathcal{I}(x, \zeta, \xi) \right\}. \end{aligned} \quad (34)$$

the final expression for the variation of the gradient of the objective function is again given by equation (15).

The corresponding derivations for functionals (23) and (25) leads to the following expressions for $\lambda(\mathbf{x}, h)$,

$$\begin{aligned} \lambda(\mathbf{x}, h) = \lambda_1(\mathbf{x}, h) = \Gamma(\mathbf{x}, h) & \int d\zeta S_z(z-\zeta) \int d\sigma S_h(\sigma-h) \left\{ \frac{\int d\xi \mathcal{I}(x, \zeta, \xi)}{H \int d\alpha \int d\beta \mathcal{I}^2(x, \alpha, \beta)} \right. \\ & \left. + \frac{\int d\alpha [\int d\beta \mathcal{I}(x, \alpha, \beta)]^2 \mathcal{I}(x, \zeta, \sigma)}{H [\int d\alpha \int d\beta \mathcal{I}^2(x, \alpha, \beta)]^2} \right\}, \end{aligned} \quad (35)$$

$$\lambda(\mathbf{x}, h) = \lambda_3(\mathbf{x}, h) = \mathcal{E}(x)\lambda_1(\mathbf{x}, h) + \mathcal{S}(x)\lambda_2(\mathbf{x}, h), \quad (36)$$

where the latter is an immediate consequence of the product rule of differentiation. In both cases, the variation of the Lagrangian is still described by the same expression (15).

NUMERICAL EXAMPLES

We have applied the above technique of velocity model refinement to the Marmousoft data (Billette et al., 2003). These data were constructed by Born modeling in a smoothed version of the Marmousoft model (Versteeg and Grau, 1990). The true Marmousoft velocity model is depicted in Figure 1.

Model representation

As detailed above, we chose to represent the velocity model by B-splines, using a grid with 15×31 nodes, with grid spacing intervals of 0.1 s in vertical time and 250 m in the horizontal direction. The starting model was a homogeneous model with a constant velocity of 1500 m/s.

For the extended differential semblance, we chose a window length of four traces and in all functionals, we applied the smoothing operator five times along the vertical and offset axes.

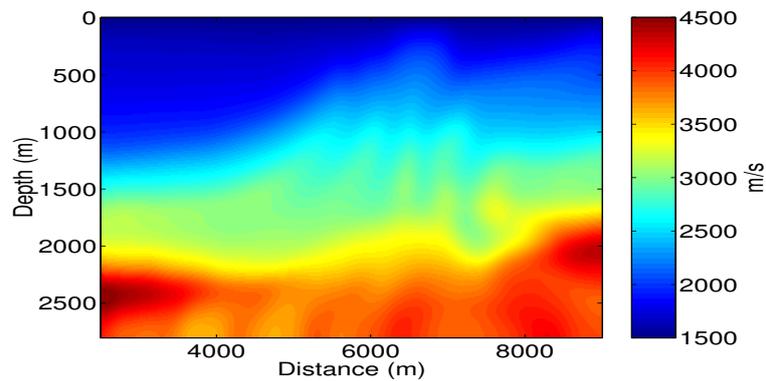


Figure 1: Marmousoft velocity model.

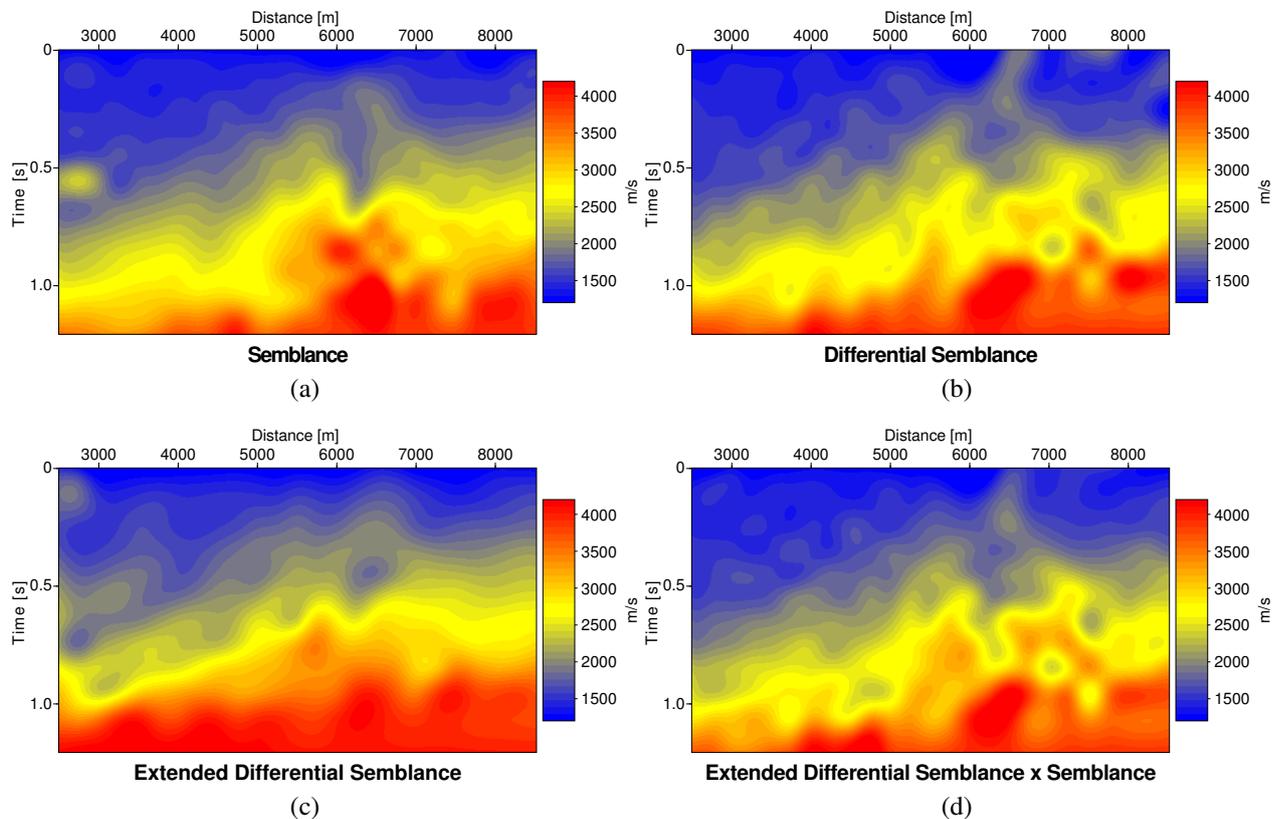


Figure 2: Velocity models obtained from a constant starting model using a conjugate gradient method with the objective functions as detailed in the text. (a) Semblance; (b) Differential Semblance; (c) Extended Differential Semblance; (d) Extended Differential Semblance \times Semblance.

Optimization method

To estimate the B-splines coefficients, we employed an implementation of the conjugate-gradient method (Byrd et al., 1993, 1995; Zhu et al., 1994, 1997), with at most 100 evaluations of the objective function and a maximum 40 iterations. The convergence criterion was a precision tolerance of 10^{-10} .

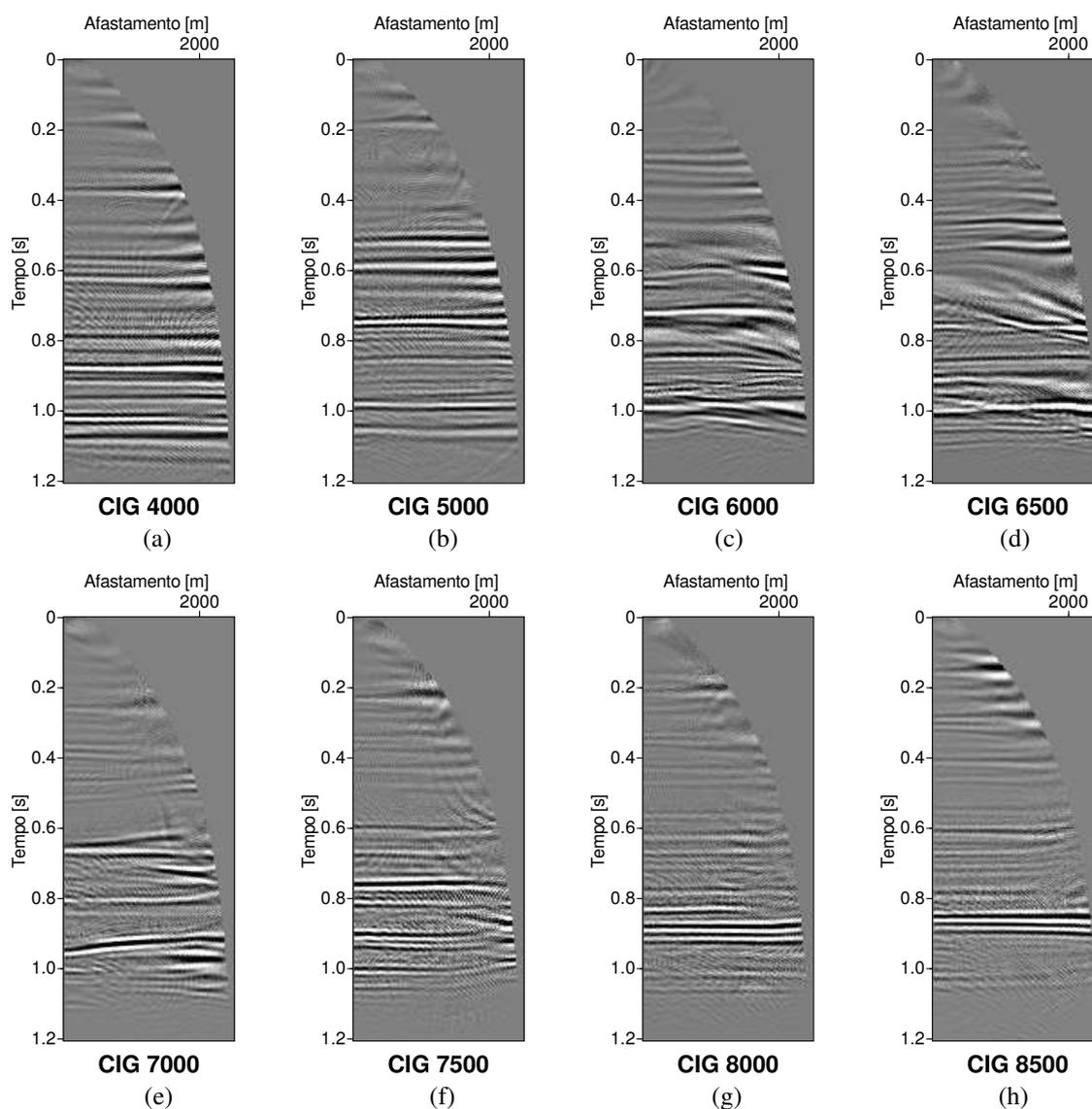


Figure 3: Selected CIGs along the Marmousoft model after automatic time migration velocity analysis using the semblance functional (24).

Velocity results

Figure 2 shows the resulting velocity models using the four tested functionals. We observe that the main general features are correctly recovered by all four objective functions. The functionals using differential semblance and the product of extended differential semblance times conventional semblance yield the most detailed models. Since the original Marmousoft model of Figure 1 is in depth, it is hard to judge which of the time-migration models of Figure 2 is actually the best one.

Image gathers

The best test for a time-migration model is, of course, a time migration. Figures 3 to 6 show selected common-image gathers after time migration using the four different velocity models of Figure 2. The CIGs corresponding to the objective functions using semblance (model in Figure 2a), differential semblance (model in Figure 2b), extended differential semblance (model in Figure 2c), and extended differential

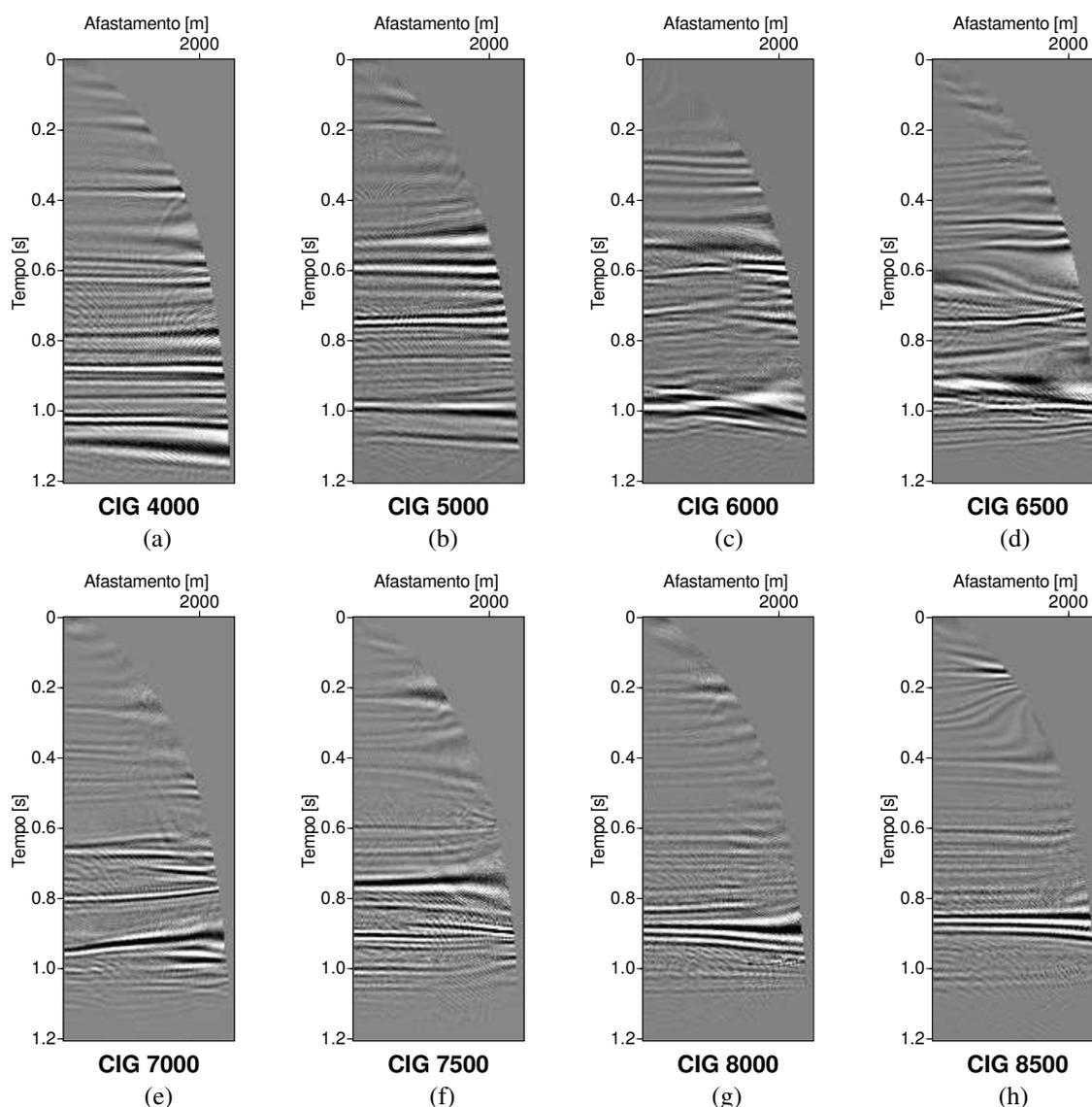


Figure 4: Selected CIGs along the Marmousoft model after automatic time migration velocity analysis using the differential semblance functional (1).

semblance multiplied by semblance (model in Figure 2d), are depicted in Figure 3, 4, 5, and 6, respectively. We see that the general aspect of all CIGs is that most events are reasonably flat. Independently of the objective function, the central region between CIG 6000 and CIG 7500 presents the most difficulties. This problem should be attributed to the general limitations of time migration in geologically complex areas rather than taking it as an indication of a poor velocity model. Visual inspection seems to indicate that the best flattening is achieved by the semblance (Figure 3) and product (Figure 6) objective functions.

For a more quantitative analysis, we have compiled Table 1. It shows the mean values of all four objective functions after optimization of each of them. We see that minimum semblance objective function is obtained when minimizing itself. On the other hand, all other three objective functions are minimized when minimizing the product objective function. This results confirms in a more quantitative manner what our visual inspection of the CIGs already indicated, namely that the resulting velocity models of the semblance and product objective functions produce the best flattened image gathers.

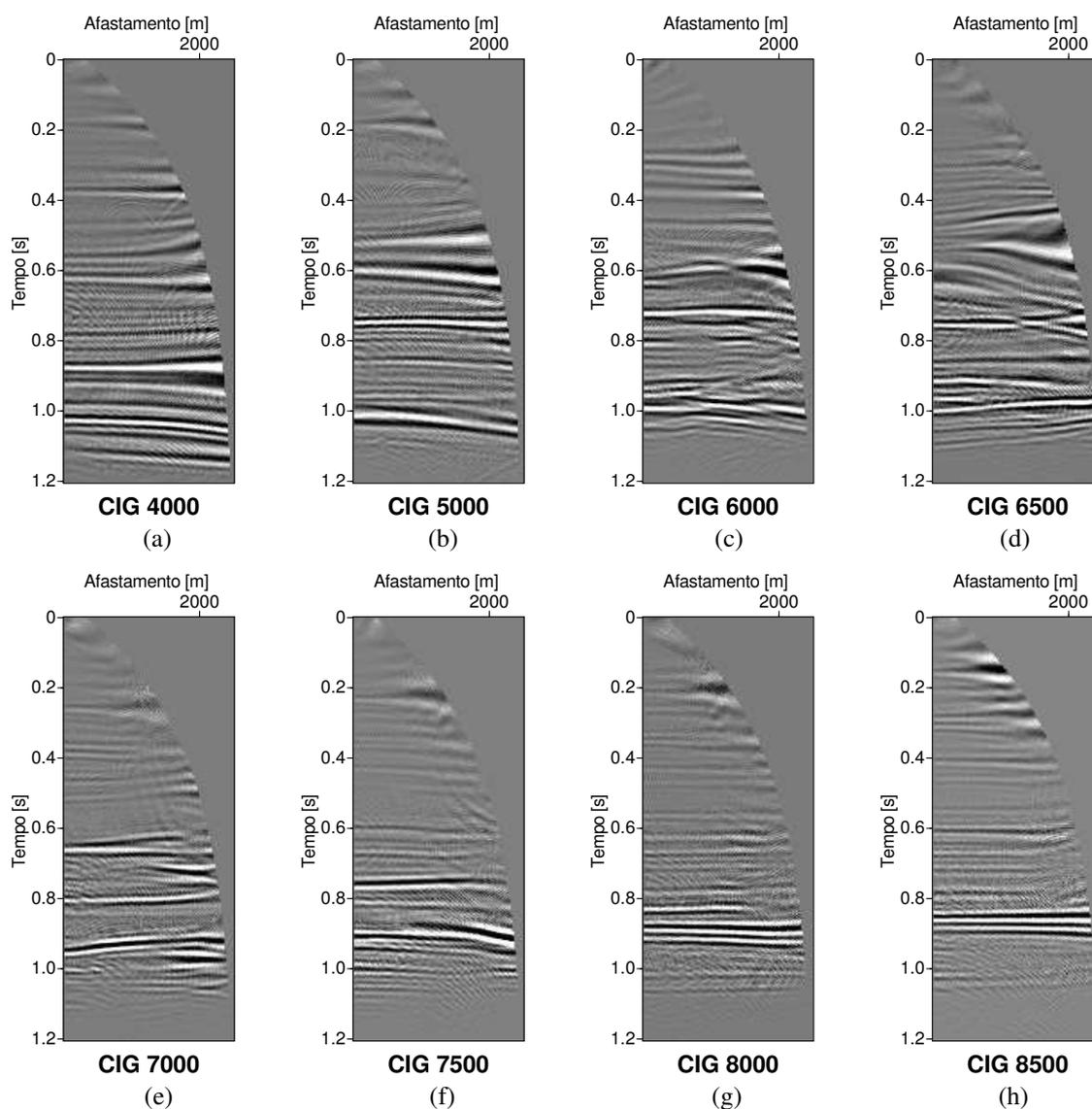


Figure 5: Selected CIGs along the Marmousfit model after automatic time migration velocity analysis using the extended differential semblance functional (23).

CONCLUSIONS

Time-migration velocity analysis can be carried out automatically by evaluating the coherence of the migrated seismic events in the common-image gathers (CIGs). In this work, we have studied the performance of an adjoint-state implementation of a conjugate-gradient method for automatic time-migration velocity analysis using four objective functions based on different coherence measures. The four coherence measures tested are conventional semblance, differential semblance, an extended differential semblance using extended images, and the product of the latter with conventional semblance. In our numerical experiments, the objective functions based on conventional semblance and on the product of conventional semblance with extended differential semblance provided the best velocity models, as evaluated by the flatness of the resulting common-image gathers.

The present approach to time-migration velocity analysis can be extended to arbitrary anisotropic media. Although analytic expressions to describe the diffraction-traveltime surfaces are no longer available

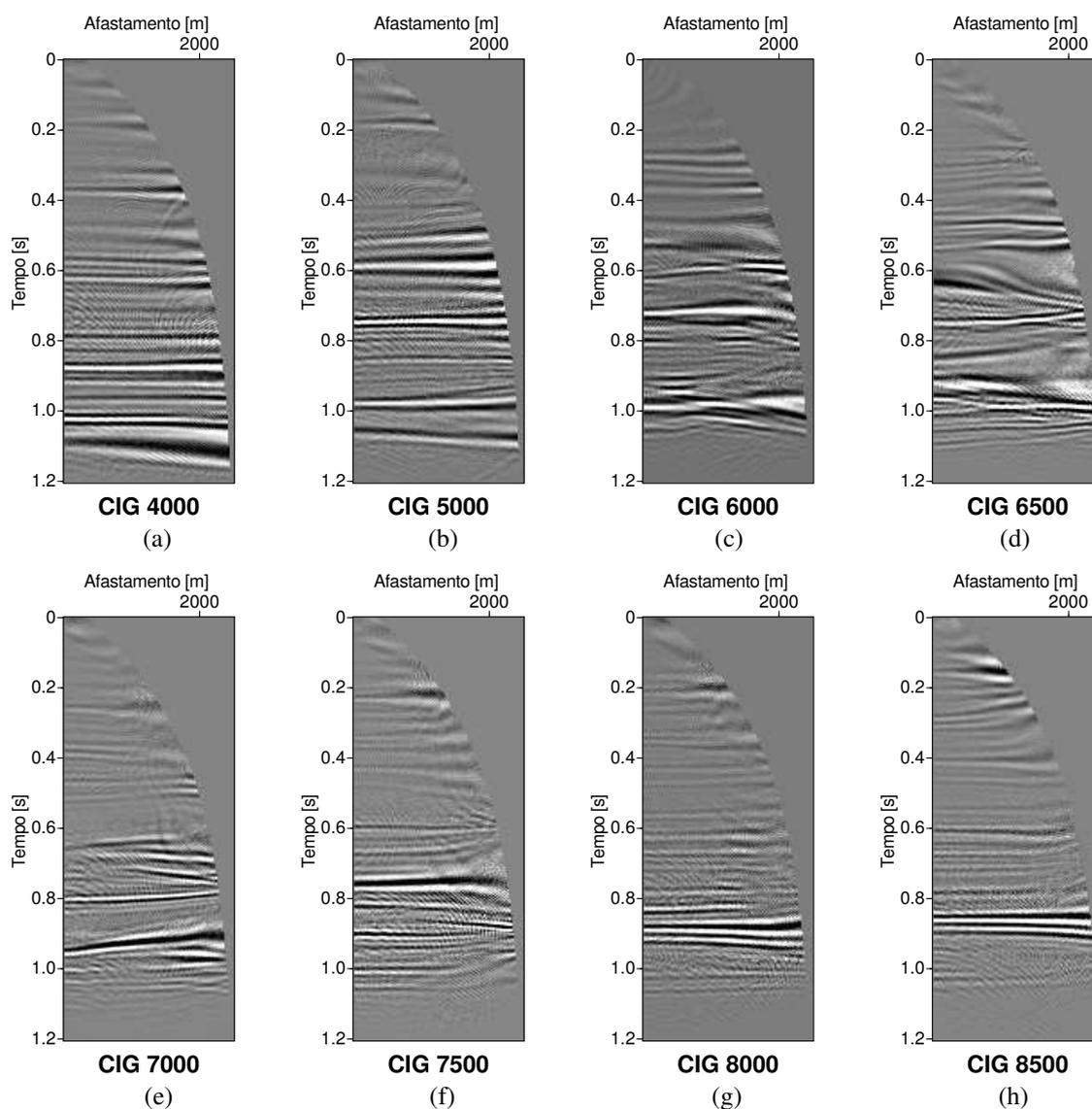


Figure 6: Selected CIGs along the Marmousfit model after automatic time migration velocity analysis using the product functional (25) of extended differential semblance and semblance.

for the more general cases of anisotropy, the numerical evaluation of these surfaces is still simple and computationally efficient. An example for nonhyperbolic velocity analysis was given by Abbad et al. (2009). For VTI media, analytic approximate traveltimes representations are available in the literature (see, e.g., Alkhalifah and Tsvankin, 1995; Fomel, 2004). Aleixo and Schleicher (2010) presented highly accurate two-parameter traveltimes approximations for VTI media. Therefore, an extension of the ideas presented in this work to VTI media is straightforward.

ACKNOWLEDGEMENTS

This work was kindly supported by the Brazilian agencies CAPES, FINEP, and CNPq, as well as Petrobras and the sponsors of the *Wave Inversion Technology (WIT) Consortium*.

Optimized objective function	Objective function values			
	Semblance	Differential Semblance	Extended Differential Semblance	Extended Differential Semblance \times Semblance
Semblance	0.0602710	25.83693	2.829875	1.425242
Differential Semblance	0.0894132	26.01136	2.659655	1.346235
Extended Differential Semblance	0.0648145	34.74438	4.180604	2.116063
Extended Differential Semblance \times Semblance	0.1037669	24.07932	2.617682	1.307488

Table 1: Values of the objective functions after migration velocity analysis. Left column: Optimized objective function. Other columns: Resulting values of all four objective functions.

REFERENCES

- Abbad, B., Ursin, B., and Rappin, D. (2009). Automatic nonhyperbolic velocity analysis. *Geophysics*, 74(2):U1–U12.
- Aleixo, R. and Schleicher, J. (2010). Traveltime approximations for qP waves in vertical transversely isotropy media. *Geophysical Prospecting*, 58:191–201.
- Alkhalifah, T. and Tsvankin, I. (1995). Velocity analysis for transversely isotropic media. *Geophysics*, 60(05):1550–1566.
- Billette, F., Le Begat, S., Podvin, P., and Lambare, G. (2003). Practical aspects and applications of 2D stereotomography. *Geophysics*, 68(3):1008–1021.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1993). A limited memory algorithm for bound constrained optimization. *Tech. Report, EECS Department, Northwestern University, NAM-08*.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16(5):1190–1208.
- Chauris, H. and Noble, M. (2001). Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: application to synthetic and real data sets. *Geophys. J. Int.*, 144:14–26.
- Faye, J. P. and Jeannot, J. P. (1986). Prestack migration velocities from focusing depth analysis. pages 438–440. SEG.
- Fomel, S. (2004). On anelliptic approximations for qP velocities in VTI media. *Geophysical Prospecting*, 52:247–259.
- Mulder, W. and ten Kroode, A. (2002). Automatic velocity analysis by differential semblance optimization. *Geophysics*, 67:1184–1191.
- Neidell, N. S. and Taner, M. T. (1971). Semblance and other coherency measures for multichannel data. *Geophysics*, 36:482–497.
- Peles, O., Tygel, M., and Kosloff, D. (2001). A practical approach to true-amplitude migration. *Journal of Seismic Exploration*, 10:183–203.
- Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys. J. Int.*, 167:495–503.

- Sattlegger, J. W. (1975). Migration velocity determination: Part i. philosophy. *Geophysics*, 40(1):1–5.
- Schleicher, J., Costa, J. C., and Novais, A. (2008). Time-migration velocity analysis by image-wave propagation of common-image gathers. *Geophysics*, 73(5):VE161–VE171.
- Schleicher, J. and Costa, J. C. (2009). Migration velocity analysis by double path-integral migration. *Geophysics*, 74(6):WCA225–WCA231.
- Symes, W. W. and Carazzone, J. J. (1991). Velocity inversion by differential semblance optimization. *Geophysics*, 56(5):654–663.
- Versteeg, R. and Grau, G., editors (1990). *The Marmousi Experience*. EAEG.
- Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1994). L-BFGS-B - Fortran subroutines for large-scale bound constrained optimization. *Tech. Report, EECS Department, Northwestern University*, NAM-11.
- Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). L-BFGS-B - Fortran subroutines for large-scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560.
- Zhu, J., Lines, L., and Gray, S. (1998). Smiles and frowns in migration/velocity analysis. *Geophysics*, 63(04):1200–1209.