A PROPOSAL FOR REGULARIZED INVERSION FOR AN ILL-CONDITIONED DECONVOLUTION OPERATOR

PROPUESTA DE INVERSIÓN REGULARIZADA PARA UN OPERADOR DE DECONVOLUCIÓN MAL CONDICIONADO

Herling Gonzalez[1*], Sheryl Avendaño[2] and German Camacho[1]

[1]Ecopetrol S.A. - Instituto Colombiano del Petróleo (ICP), A.A. 4185 Bucaramanga, Santander, Colombia
[2]UTP Consultorias, Bucaramanga, Santander, Colombia

e-mail: herling.gonzalez@ecopetrol.com.co

ABSTRACT

From the inverse problem theory aspect, deconvolution can be understood as the linear inversion of an ill-posed and ill-conditioned problem. The ill-conditioned property of the deconvolution operator make the solution of inverse problem sensitive to errors in the data. Tikhonov regularization is the most commonly used method for stability and uniqueness of the solution. However, results from Tikhonov method do not provide sufficient quality when the noise in the data is strong. This work uses the conjugate gradient method applied to the Tikhonov deconvolution scheme, including a regularization parameter calculated iteratively and based on the improvement criterion of Morozov discrepancy applied on the objective function. Using seismic synthetic data and real stacked seismic data, we carried out a deconvolution process with regularization and without regularization based on a conjugated gradient algorithm. A comparison of results is also presented. Applying regularized deconvolution on synthetic data shows improved stability on the solution. Additionally, real post-stack seismic data shows a direct application for increasing the vertical resolution even with noisy data.

*To whom correspondence should be addressed

RESUMEN

Desde el punto de vista de la teoría de problemas inversos, la deconvolución puede ser entendida como una inversión lineal de un problema mal-puesto y mal-condicionado. La característica del mal-condicionamiento del operador de deconvolución hace que la solución del problema inverso sea sensitiva a errores en los datos. La regularización de Tikhonov es el método más común empleado para estabilizar la solución y obtener su unicidad. Sin embargo, los resultados del método de Tikhonov no proveen calidad suficiente cuando el ruido en los datos es fuerte. Este trabajo hace uso del método del gradiente conjugado, basado en el esquema de Tikhonov aplicado a la deconvolución, cuyo parámetro de regularización es calculado iterativamente teniendo en cuenta el criterio de discrepancia de Morozov en la función objetivo. Haciendo uso de datos sísmicos sintéticos como datos reales apilados, se llevó a cabo el proceso de deconvolución con y sin regularización basada en el algoritmo del gradiente conjugado. Se llevó a cabo una comparación del esquema planteado. Aplicando la deconvolución regularizada en los datos sintéticos muestra una mejora en la estabilidad de la solución y los datos sísmicos post-apilados mostraron un incremento de la resolución vertical aun con ruido en los datos.

Palabras clave: *Regularización de Tikhonov, Gradiente conjugado, Teoría de inversión, Procesamiento sísmico.*

RESUMO

Desde o ponto de vista da teoria de problemas inversos, a deconvolução pode ser entendida como uma inversão linear de um problema mal posto e mal condicionado. A característica do mal condicionamento do operador de deconvolução faz que a solução do problema inverso seja sensitiva a erros nos dados. A regularização de Tikhonov é o método mais comum utilizado para estabilizar a solução e obter sua unicidade. Porém, os resultados do método de Tikhonov não fornecem qualidade suficiente quando o ruído nos dados é alto. Este trabalho faz uso do método do gradiente conjugado, baseado no esquema de Tikhonov aplicado à deconvolução, cujo parâmetro de regularização é calculado iterativamente tendo em conta o critério de discrepância de Morozov na função objetivo. Fazendo uso de dados sísmicos sintéticos como dados reais empilhados, foi realizado o processo de deconvolução com e sem regularização baseado no algoritmo do gradiente conjugado. Realizou-se uma comparação do esquema proposto. Aplicando a deconvolução regularizada nos dados sintéticos mostra uma melhoria na estabilidade da solução e os dados sísmicos pós-empilhados mostraram um aumento da resolução vertical mesmo com ruído nos dados.

Palavras-chave: *Regularização de Tikhonov, Gradiente conjugado, Teoria de inversão, Processamento sísmico.*

1. INTRODUCTION

Deconvolution is an inversion process that is commonly used in many areas of science and engineering such as image and signal processing. In reflection seismic applications, deconvolution increases the vertical resolution. A basic model for a seismic trace is that it can be described by the convolution of a reflectivity series with a wavelet and added noise. Deconvolution results are generally considered as an approximation to the reflectivity series in a stratified medium, estimated from the reflected wavefield. There is extensive literature on the development of new inversion algorithms that improve deconvolution quality. Deconvolution quality in seismic data is assessed by its capacity to correctly predicting the position, amplitude and phase of the reflectivity series (Yilmaz, 1987; Zala, 1992; Leinbach, 1995). Deconvolution problems can have deterministic or non-deterministic approaches, depending on whether the wavelet is known or not (Wang, Wang & Perz, 2006). In this paper, we have assumed that the wavelet is known and band-limited, and have been estimated, for instance, from well log data or by statistical correlation of seismic data. The deconvolution process can be achieved by using regularization in the inversion (Karsli, Guney & Senkaya, 2012; Chen, Wang & Chen, 2012). By imposing an additional constraint on the estimated model, regularization can be implemented even if the convolution operates with insufficient data (Fomel, 2007) or when data contain high noise level (Hansen, 2010).

The solution to inverse problems commonly have stability issues and a multiplicity of solutions (Sen, 2006). From the mathematical standpoint, all inverse problems can be classified as well-posed or ill-posed. According to Hadamard (1923), an inverse problem is well-posed on the strict sense if:

1. There is a solution.
2. The solution is unique.
3. The solution is stable with respect to perturbations in the data.

Stability is related to the fact that minor changes introduced in the data of the problem generate minor variations in the solution. If any of the three conditions is not met, the problem is defined as ill-posed. Errors produced during measurement, as well as those introduced by numerical methods disassociate the extracted solution from the estimated information (Hansen, 2010; Montenegro, 2010). Techniques that enable recast from an ill-posed problem to a well-posed problem are also known as regularization. Deconvolution is an inverse problem, known for its sensitivity to noise in geophysical data processing (Van der Baan & Pham, 2008).

In this paper we propose an additional iterative calculation to estimate the regularization parameter ε_i within the Tikhonov regularization scheme applied to the deconvolution of noisy data. This iterative calculation uses the bi-section and drying methods to find roots in the solution and the Morozov (1984) discrepancy criteria. We have also used concepts of deblurring theory on the convolution of seismic traces. The results of the regularized deconvolution method, presented here, show an improvement in the stability of the solution in the presence of noise, as well as a direct application for increasing the vertical resolution of stacked seismic data.

2. TIKHONOV REGULARIZATION

Tikhonov regularization suggests (for the inversion problem) adding a stabilizing matrix or function (Tikhonov, 1963) such that:

$$\mathbf{Lm} = \mathbf{d} \rightarrow \mathbf{R\hat{m}} = \mathbf{d} \tag{1}$$

where $\mathbf{L}$ is the convolution operator, $\mathbf{R}$ is the regularized version of the operator, $\mathbf{m}$ is the physical model or series of reflectivity and $\mathbf{\hat{m}}$ is the approximate solution of the model, $\mathbf{d}$ is the response to the model $\mathbf{m}$ and $\mathbf{\hat{d}}$ are the data with noise from the seismic data. A regularized least-square solution l_2 of (1) is equivalent to minimizing the function

$$min\ \{\|\mathbf{Lm} - \mathbf{d}\|^2 + \varepsilon^2\|\mathbf{\Gamma m}^2\} \tag{2}$$

where ε is a parameter of regularization and $\boldsymbol{\Gamma}$ is the regularizing matrix (Fomel, 2007). This can be the identity $\mathbf{I}$ operator or a first order finite differences $\mathbf{D}$ matrix. Therefore, for the linear system with noise in (1), the best approximate solution can be obtained by minimizing the "functional smoothing" in (2) with an optimal value of ε (Tikhonov & Arsenin, 1977). $\mathbf{D}$ has a smoothing nature to ensure the linear variation of $\mathbf{Lm}$ around the proximity of $\mathbf{m}$ (Sen & Roy, 2003). In our case, the solution approximated by least squares of (2) is (Fomel, 2007):

$$\hat{\mathbf{m}} = (\mathbf{L}^{\mathrm{T}}\mathbf{L} + \varepsilon^2\,\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{d} \qquad (3)$$

This solution is regularized for a given value ε according to the following expression (Hansen, 2010):

$$min\{\;\|\mathbf{L}\hat{\mathbf{m}}_\varepsilon - \mathbf{d}\|^2\;\} \qquad (4)$$

Following a convolution model for the seismic trace, we have assumed the reflectivity data as $R(t)$ convoluted with a punctual blurring function called the Point Spread Function (*PSF*), which generates seismic data $S(t)$. This function is equivalent to the seismic wavelet and is assumed to be constant and time invariant. *PSF* in image processing is a blurring function that lowers image resolution and, for our purposes, is one-dimensional.

Any pseudo-multidimensional array of traces, can be expressed as:

$$R(t,x) * PSF = S(t,x) \qquad (5)$$

where variable t is the Two Way Time (TWT) of the trace and variable x represents the horizontal coordinate (or offset) of traces. The above expression is analogous to our linear system $\mathbf{Lm} = \mathbf{d}$ where $\mathbf{L}$ is the convolution matrix or function *PSF*, see Figure 2. By using elements p_i of matrix *PSF* with an odd magnitude $2r + 1$, we have a Toepliz matrix of the $(n + 2r) \times (n)$ dimension, for this operator (Hansen, 2002):

$$\mathbf{L} = \begin{pmatrix} p_{\boldsymbol{r}} & 0 & \dots & 0 & \\ p_{\boldsymbol{r}-1} & p_{\boldsymbol{r}} & \dots & 0 & 0 \\ p_{\boldsymbol{r}-2} & p_{\boldsymbol{r}-1} & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots \\ p_0 & p_1 & \dots & 0 & 0 \\ p_{-1} & p_0 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & 0 & 0 \\ p_{-\boldsymbol{r}} & p_{-(\boldsymbol{r}-1)} & \dots & 0 & 0 \\ & p_{-\boldsymbol{r}} & \dots & 0 & \\ 0 & 0 & \dots & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & p_{-(\boldsymbol{r}-1)} & p_{-(\boldsymbol{r}-2)} \\ 0 & 0 & \dots & p_{-\boldsymbol{r}} & p_{-(\boldsymbol{r}-1)} \\ 0 & 0 & \dots & 0 & p_{-\boldsymbol{r}} \end{pmatrix} \qquad (6)$$

so we can carry out the deconvolution process to find $\hat{\mathbf{m}}$ in accordance with expression (3).
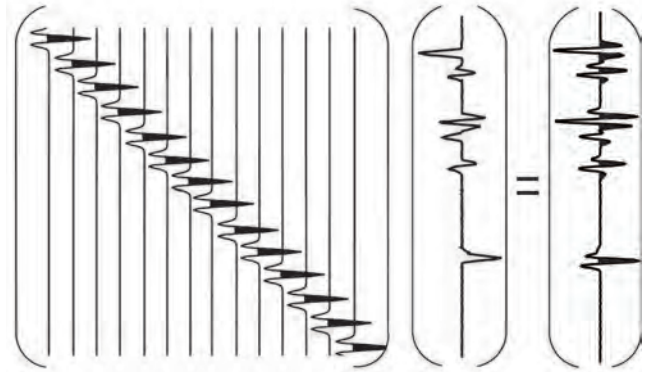

Figure 1. Stationary convolution model in the time domain. Each column of the matrix represents the blurring function or PSF, whose Ricker-type wavelet is zero phase with a central frequency of 50Hz operating on the reflectivity series to generate the seismic trace.

3. REGULARIZATION AND CONJUGATE GRADIENT

The effect of parameter ε in the solution of *Equation 14* is described as follows: When $\varepsilon > 1$, the regularization solution $\hat{\mathbf{m}}$ is stable. However, as tradeoff, the error of $\hat{\mathbf{m}}$ increases with respect to solution $\mathbf{m}$. When $\varepsilon < 1$, the error of $\hat{\mathbf{m}}$ decreases, but regularization becomes unstable

and sensitive to noise, leading to unwanted $\hat{m}$ solutions (Tikhonov & Arsenin, 1977; Hansen, 2010), see Figure 2. Taking into account the error in the data, we can go further than *Equation 1*:

$$
\begin{aligned}
\mathbf{m} - \hat{\mathbf{m}} &= \mathbf{m} - \mathbf{R}_\varepsilon\, \hat{\mathbf{d}} \\
&= \mathbf{m} - \mathbf{R}_\varepsilon\, \hat{\mathbf{d}} + \mathbf{R}_\varepsilon\, \mathbf{d} - \mathbf{R}_\varepsilon\, \mathbf{d} \\
&= \mathbf{m} - \mathbf{R}_\varepsilon\, \hat{\mathbf{d}} + \mathbf{R}_\varepsilon(\hat{\mathbf{d}} - \hat{\mathbf{d}})
\end{aligned}
\tag{7}
$$

where $\mathbf{R}_\varepsilon = (\mathbf{L}^T\mathbf{L} + \varepsilon^2\mathbf{D}^T\mathbf{D})^{-1}\,\mathbf{L}^T$ and $\hat{\mathbf{d}} = \mathbf{L}\hat{\mathbf{m}}_\varepsilon$. Taking into account the error in the approximation of the model and the error in the data, the error in the calculation of the model is:

$$
\begin{aligned}
\|\mathbf{m} - \hat{\mathbf{m}}\| &\leq \|\mathbf{m} - \mathbf{R}_\varepsilon\, \mathbf{d}\| + \|\mathbf{R}_\varepsilon\|\,\|(\mathbf{d} - \hat{\mathbf{d}})\| \\
&\leq E_{aprox} + E_{datos}
\end{aligned}
\tag{8}
$$

The error in the data is expressed as the vector $\hat{\mathbf{d}} - \mathbf{d} = \eta$, that we call "tolerance". The magnitude of vector η is related iteratively with $\|\mathbf{R}_\varepsilon\|$ through the configurate gradient algorithm.

When $\varepsilon \to 0$, $\|\mathbf{R}_\varepsilon\|$ get closer to $\|(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\| = \|(\mathbf{L})^{-1}\|$ To understand the effect of $\varepsilon \to 0$, we took in consideration the condition cond($\mathbf{L}$) of operator-matrix $\mathbf{L}$

$$
\text{cond}(\mathbf{L}) = \|\mathbf{L}\|\,\|\mathbf{L}^{-1}\| \gg 1
\tag{9}
$$

From this last expression, it can be proposed that there is a set of $\{\varepsilon_i\}$ that optimize in a comprehensive manner the trade-off between error in the data and the error in approximation of the model (Figure 3).
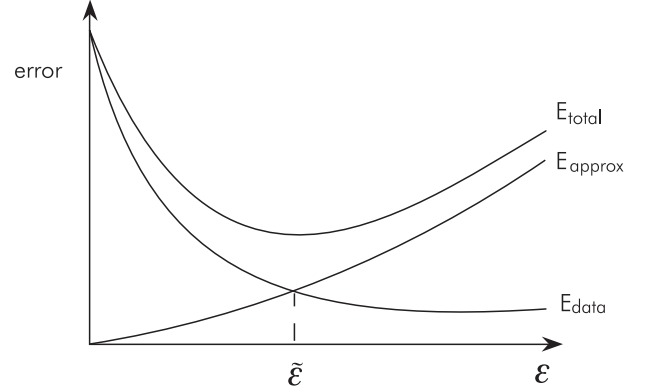

Figure 3. Error curves for the solution of the model in the inversion process. Error in data, error in approximation and total error, modified Tikhonov and Arsenin, 1977.

The technique chosen to minimize *Equation 4* is the conjugate gradient technique, which consists of quadratically expressing linear relationship $\mathbf{A}\mathbf{m} = \mathbf{d}$ as:

$$
f(\mathbf{m}) = \frac{1}{2}\mathbf{m}^T\mathbf{A}\mathbf{m} - \mathbf{d}^T\mathbf{m} + c
\tag{10}
$$

Space model parameters
(acceptable models)

Space model responses
(with and without noise)

$\mathbf{m}$

$\mathbf{d} = \mathbf{L}\mathbf{m}$

$\hat{\mathbf{m}} = \left(\mathbf{L}^T\mathbf{L} + \varepsilon\,\mathbf{D}^T\mathbf{D}\right)^{-1}\mathbf{L}^T\,\hat{\mathbf{d}}$

$\hat{\mathbf{d}} = \mathbf{d} + \eta$

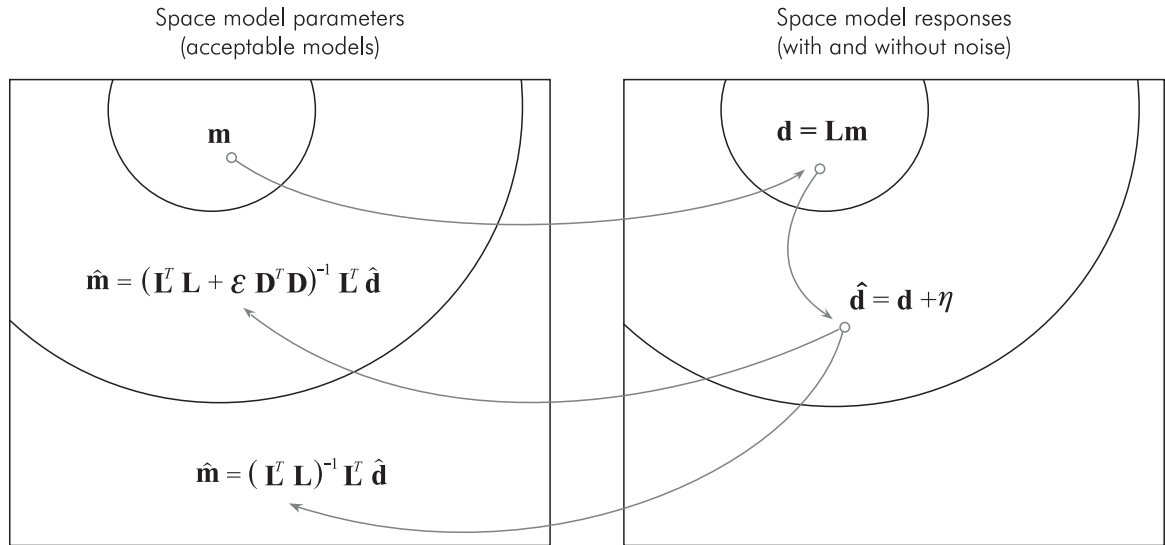$\hat{\mathbf{m}} = \left(\mathbf{L}^T\mathbf{L}\right)^{-1}\mathbf{L}^T\,\hat{\mathbf{d}}$

Figure 2. This figure illustrates the advantages of regularization (modified from Hansen, 2010). Due to the ill-conditioning of operator $\mathbf{L}$, the least squares solution $\hat{\mathbf{m}} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\,\hat{\mathbf{d}}$ in the space model, may be quite far from the exact solution $\mathbf{m}$ even with minor noise or perturbations in the data, that is, when $\|\eta\| \ll \|\mathbf{d}\|$. Besides, the regularized operator solution $\hat{\mathbf{m}} = (\mathbf{L}^T\mathbf{L} + \varepsilon^2\mathbf{D}^T\mathbf{D})^{-1}\mathbf{L}^T\mathbf{d}$ tends to get closer to the exact solution $\mathbf{m}$ under noise conditions.

where $\mathbf{A}$ is a direct modeling operator with a defined positive square (Hestenes & Stiefel, 1952), $\langle \mathbf{m}, \mathbf{Am} \rangle = \mathbf{m}^T \mathbf{Am} = \langle \mathbf{Am}, \mathbf{m} \rangle > \mathbf{0}$, and symmetrical matrix $\mathbf{A}^T = \mathbf{A}$, here $\langle,\rangle$ represents the internal product. This quadratic vector equation is in terms of $\mathbf{m}$ whose gradient is equal to:

$$\nabla f(\mathbf{m}) = \mathbf{Am} - \mathbf{d} \qquad (11)$$

where the direction of maximum decrease is $-\nabla f = \mathbf{d} - \mathbf{Am}$. The conjugate gradient method is a method of descent that begins to iterate in an initial model $\mathbf{m}_0$ and continues in the direction of maximum descent of the paraboloid represented by quadratic function (10), resulting in a succession of models $\mathbf{m}_k$ that converge at a solution $\hat{\mathbf{m}}$ close to $\mathbf{m}$.

Input: $\mathbf{m}_0, l, \mathrm{tol}$, Output: $\hat{\mathbf{m}}$

1. $\mathbf{p}_0 = \mathbf{r}_0 = -\nabla \phi = \mathbf{d} - \mathbf{Am}_0$

2. **for** $k = 0, 1, \ldots l$ **do**

3. $\quad \alpha_k = \dfrac{\langle \mathbf{r}_k, \mathbf{r}_k \rangle}{\langle \mathbf{p}_k, \mathbf{Ap}_k \rangle}$

4. $\quad \mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k \mathbf{p}_k$

5. $\quad \mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{Ap}_k$

6. $\quad$ **if** $\|\mathbf{r}_{k+1}\| < tol$ **then**

7. $\qquad$ **return** $\mathbf{m}_{k+1}$

8. $\quad$ **else**

9. $\qquad \beta_k = \dfrac{\langle \mathbf{r}_{k+1}, \mathbf{r}_{k+1} \rangle}{\langle \mathbf{r}_k, \mathbf{r}_k \rangle}$

10. $\qquad \mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k$

11. $\quad$ **end if**

12. **end for**

13. **return** $\mathbf{m}_{k+1}$

Algorithm 1. Conjugate gradient with a matrix-vector operation, two point products, two scalar-vector multiplications, one sum operation and one combined scalar multiplication and vector addition operation per iteration.

By using the conjugate gradient method, we calculate a value of $\hat{\mathbf{m}}_\varepsilon$ for any given ε_i, such that the residual $\|\mathbf{r}_{\varepsilon_i}\| = \|\hat{\mathbf{d}} - \mathbf{L}\hat{\mathbf{m}}_\varepsilon$ is minimized. As we have outlined in Figure 4, for a given value of ε between $0 < \varepsilon_i \leqq 1$ and

an initial model $\mathbf{m}_0$ (i.e. initial residual is $\mathbf{r}_0 = \hat{\mathbf{d}} - \mathbf{Lm}_0$), $\|\mathbf{r}_{\varepsilon_i}\|$ decreases until it reaches a critical value. Further iterations $\{k\}$ of conjugate gradient, increase the residual $\|\mathbf{r}_{\varepsilon_i}\| = \|\hat{\mathbf{d}} - \mathbf{L}\hat{\mathbf{m}}_{\varepsilon_i}$ (see Figure 4).
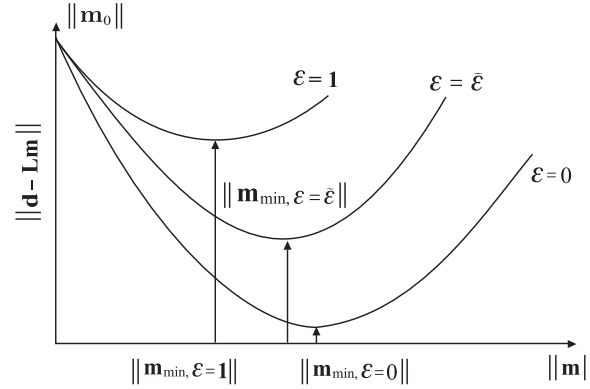

Figure 4. Curves of residuals $\|\mathbf{r} = \mathbf{d} - \mathbf{Lm}\|$ vs model approximation $\|\mathbf{m}\|$, for different values of ε using the conjugate gradient method.

In Figure 4, notice that $\varepsilon = 0$, is the solution of problem $\mathbf{Lm} = \mathbf{d}$, whose solution is unstable and sensitive to perturbations in data. Again, this is due to the ill-conditioned operator $\mathbf{L}$ (Shewchuk, 1994). In Figure 4, the curve with value $\varepsilon = 1$, has a lower condition number for operator $\mathbf{A} = \mathbf{L}^T \mathbf{L} - \mathbf{D}^T \mathbf{D}$, but the solution is not as close to the solution of the system. In the third curve in Figure 4 with optimal value $\varepsilon = \tilde{\varepsilon}$, the solution is stable and close to $\mathbf{m}_{\varepsilon = 0}$, with a lower calculated residual value $\mathbf{r}$. This behavior raises the question of how to choose this optimal parameter value.

To choose the optimal value $\varepsilon = \tilde{\varepsilon}$ we have used the post-election strategy based on the discrepancy criteria (Morozov, 1984). This criteria establish that the norm of the difference between the observed data and the data generated by the regularized solution cannot be lower than the noise level δ:

$$\tau_1 \delta \leqslant \|\hat{\mathbf{d}} - \mathbf{L}\hat{\mathbf{m}}_\varepsilon\| \leqslant \tau_2 \delta \qquad (12)$$

where $1 \leqslant \tau_1 \leqslant \tau_2$ (Bonesky, 2009). For our deconvolution problem we have taken $\tau_1 = \tau_2 = 1$. That means that the norm for the data generated by the regularized solution cannot be more accurate than the norm of the observed data (Engl, Hanke & Neubauer, 1996). In other words, the noise level δ should be proportional to $\|\boldsymbol{\eta}\|$,

where $\hat{\mathbf{d}} - \mathbf{d} = \boldsymbol{\eta}$, and proportional to $\|\mathbf{r}\| = \|\hat{\mathbf{d}} - \mathbf{L}\hat{\mathbf{m}}\|$. To estimate the relative noise level $\boldsymbol{\eta}$ in synthetic data, we have used normalization according to the following expression

$$\delta_{\mathrm{relat}} = \frac{\|\boldsymbol{\eta}\|}{\|\mathbf{d}\|} \qquad (13)$$

Now, to estimate the relative noise level in seismic data, we have applied the following expression

$$\delta_{\mathrm{relat}} = \frac{\|\hat{\mathbf{d}} - \mathbf{L}\hat{\mathbf{m}}\|}{\|\hat{\mathbf{d}}\|} \qquad (14)$$

Now, we can assume that in order to obtain a proper $\varepsilon = \tilde{\varepsilon}$, we have to find the values of the roots of equation $\|\mathbf{r}_{\varepsilon_{\mathbf{i}}}\|^2 - \delta^2$ (Montenegro, 2010). That is to say:

$$f(\varepsilon_i) = \|\mathbf{r}_{\varepsilon_{\mathbf{i}}}\|^2 - \delta^2 \qquad (15)$$

If $\varepsilon > \tilde{\varepsilon}$ the residual is larger than δ and *Equation 15* is positive, while for $\varepsilon < \tilde{\varepsilon}$, the residual is smaller than and *Equation 15* becomes negative. That suggest the iterative search $\varepsilon_i + 1 = \varepsilon_i + \frac{\varepsilon_i - \varepsilon_{i-1}}{f(\varepsilon_i) - f(\varepsilon_{i-1})}$, illustrated in Figure 5.
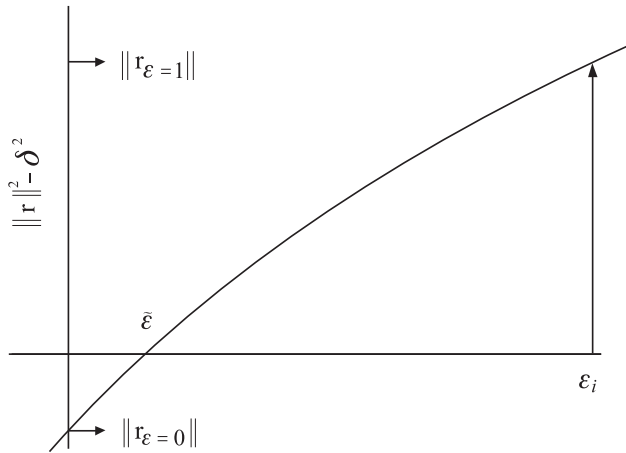

Figure 5. Curve of ($\|\mathbf{r}\|^2 - \delta^2$) with respect to ($\varepsilon_i$). The optimal value $\varepsilon = \tilde{\varepsilon}$ corresponds to the root $\|\mathbf{r}_{\tilde{\varepsilon}}\|^2 - \delta^2 = 0$

The following algorithm was developed to carry out the deconvolution process by regularization:

Input: $\mathbf{L}, \mathbf{d}, \mathbf{m}_0, \delta, \varepsilon_0, n$, tol, Output: $\hat{\mathbf{m}}_{\tilde{\varepsilon}}$

1. $\mathbf{r}_0 = \mathbf{d} - \mathbf{L}\mathbf{m}_0$
2. for $i = 0, 1, \ldots n$ do
3.   Call CG with $\mathbf{A} = \mathbf{L}^T\mathbf{L} + \varepsilon_i^2\mathbf{D}^T\mathbf{D}$
4.   if $\mathbf{r}_{\varepsilon_i} < tol$ then
5.   return $\mathbf{m}_{\varepsilon_i}$
6.   else
7. $\mathbf{m}_0 = \mathbf{m}_{\varepsilon_i}$
8.   if $i = 0$ then
9. $\varepsilon_{i+1} = \varepsilon_0/2$
10.   else
11. $\varepsilon_{i+1} = \varepsilon_i + \dfrac{\varepsilon_i - \varepsilon_{i-1}}{f(\varepsilon_i) - f(\varepsilon_{i-1})},$
                where $f(\varepsilon_i) = \|\mathbf{r}_{\varepsilon_i}\|^2 - \delta^2$
12.   end if
13.   end if
14. end for
15. return $\mathbf{m}_{\tilde{\varepsilon}}$

Algorithm 2. Regularized inversion combined with Conjugate Gradient for the iterative calculation of ε_i based on the search for the roots of equation ($\|\mathbf{r}\|^2 - \delta^2$) with regard to ($\varepsilon_i$) (line 11).

4. SYNTHETIC DATA

Tests with synthetic data enabled us to evaluate the response of the deconvolution process under the effects of noise in a controlled manner, as well as the results of the inversion strategy. We started with a simple case of deconvolution $\hat{\mathbf{m}} = \mathbf{L}^{-1}\mathbf{d}$ without noise. For the direct problem we use a zero phase Ricker wavelet with a central frequency of 50Hz (see Figure1) convoluted with a known reflectivity (see Figure 1 and 6a). The inversion result is presented in Figure 6b, which shows edge effects due to the sudden end of data at border (Gibbs oscillation). Changing the criteria of the residual value in the inversion process from 1x10⁻⁵ to 1x10⁻³, this edge effect is mitigated; however, the solution is not recovered (see Figure 6c). Then, using the regularization method suggested in Section 3 (Algorithm 2) and taking
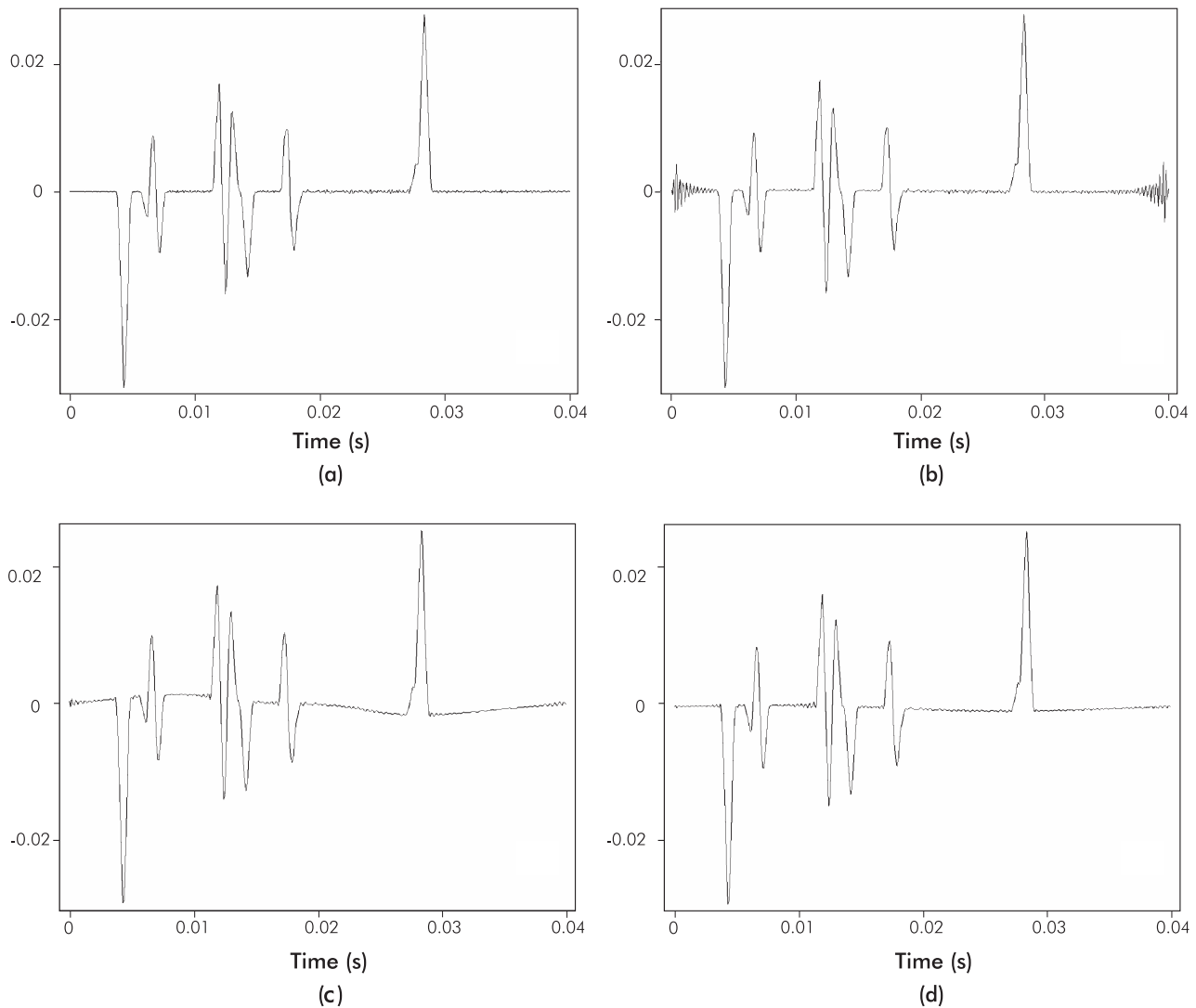
Figure 6. Results of the reflectivity inversion without noise, (a) original reflectivity used to generate the data (exact solution), (b) inversion by CG with residual $\|\mathbf{r}_{k+1}\| < 1e^{-5}$, (c) inversion by CG with residual $\|\mathbf{r}_{k+1}\| < 1e^{-3}$, (d) inversion regularized by CG with residual $\|\mathbf{r}_{k+1}\| < 1e^{-4}$, delta of 0.01 and $\varepsilon = 0.0107147$.

1×10^{-4} as the residual value, $\delta = 0.01$ and an epsilon ε_i, of 0.0010715, we get the results displayed in Figure 6d. This illustrates in data the effect of trade-off between error in the data and error in approximation of the model. In a second numeric experiment, we added Gaussian noise to the data and the inversion was carried out in the same way as the previous experiment, i.e. taking a residual value of 1×10^{-4}, with and without regularization. The value of the Gaussian noise applied was lower than the residual of 1×10^{-4}. The effects of the noise on the inversion data are illustrated in Figures 7a and b.

When noise is included in the data, the solution becomes unstable (Figure 7b) due to the ill-conditioning of operator **L**. Additionally, the CG method in the iterative process takes a larger number of steps for convergence (more than 300 iterative steps), sacrificing accuracy for stability. The difference in the results is due to the fact that the noise acts as a perturbation that is amplified in the iterative process due to the instability of the inversion. These results show the advantages of regularization when looking for stable solutions for an inversion problem in the presence of noise, as the deconvolution of real seismic data.
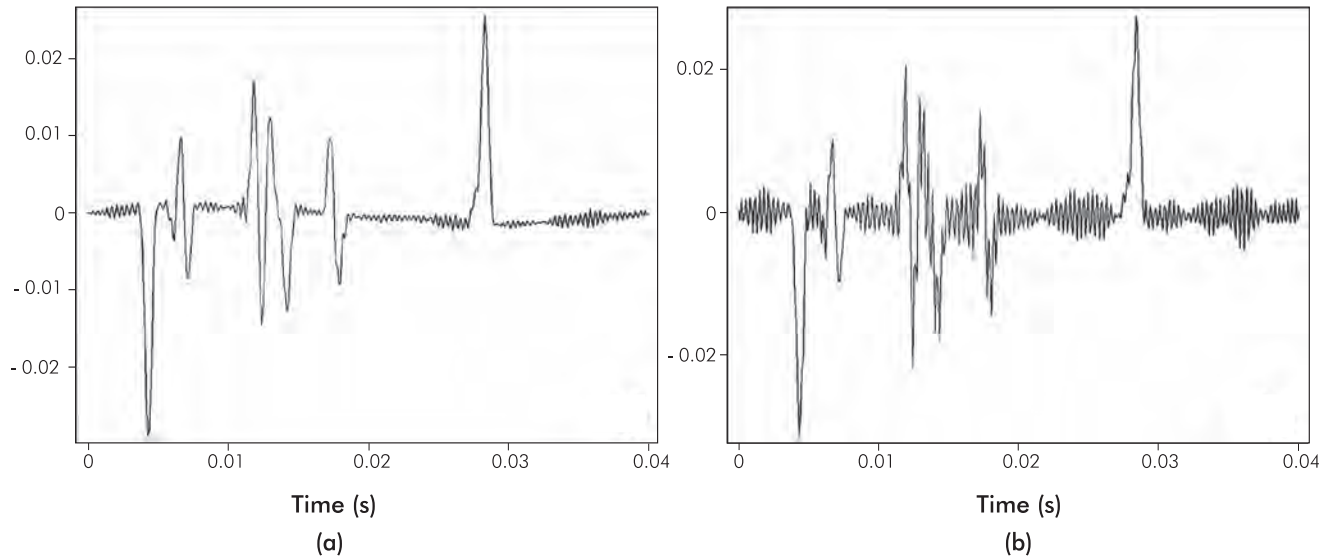
Figure 7. Result of deconvolution with Gaussian noise in data. (a) Inversion by regularization using values of δ=0.01 relative were used with a tolerance in residual of 1.0×10^{-4} for the conjugate gradient and the estimated value was $\varepsilon = 0.12882$. (b) non-regularized inversion.

5. STACKED SEISMIC DATA

The method was applied on real seismic data extracted from a section of an area of interest. For the inversion-deconvolution process, a wavelet was statistically extracted in the target zone, where a preprocessing phase zero correction in seismic data was applied. The extraction of the wavelet is a common routine process in the oil industry for seismic characterization and seismic well calibration. The section analyzed was CMP 25160 to 25378 (218 meters) from 0 to 4 seconds in TWT. The target zone is between 1 and 2 seconds (see Figure 8a). The wavelet that was calculated is a Ricker (Figure 8b) estimated by statistical correlation.

The deconvolution of this seismic section was carried out only with regularization (Figure 9) It is important to remember that the success of any inversion process with regularization is subject to the level of noise in data. If we have a seismic with poor image quality and high noise in the traces we cannot expect good results.
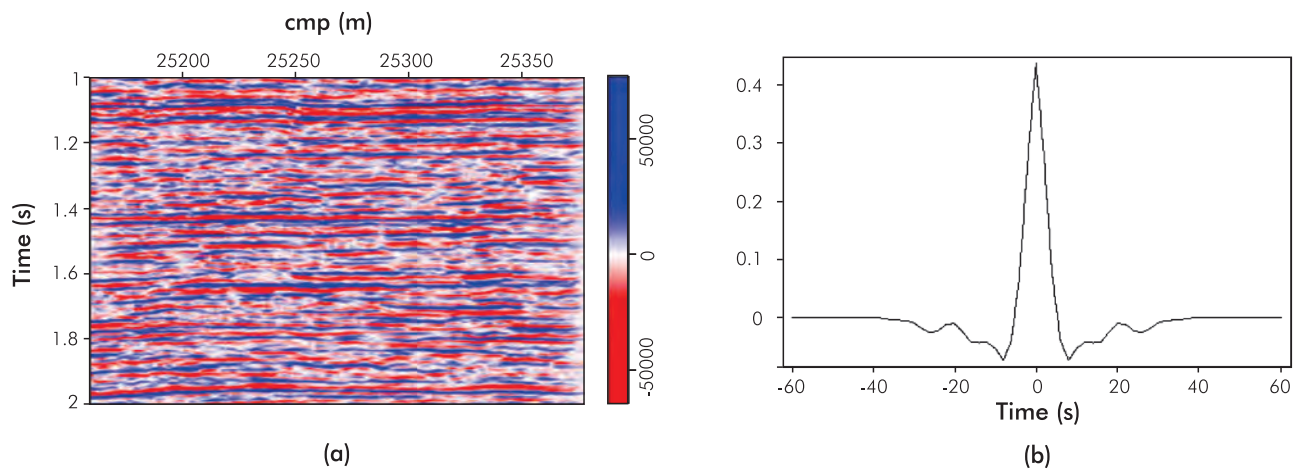

Figure 8. (a) Seismic section of interest in TWT. (b) Wavelet extracted from the seismic section by statistical correlation.
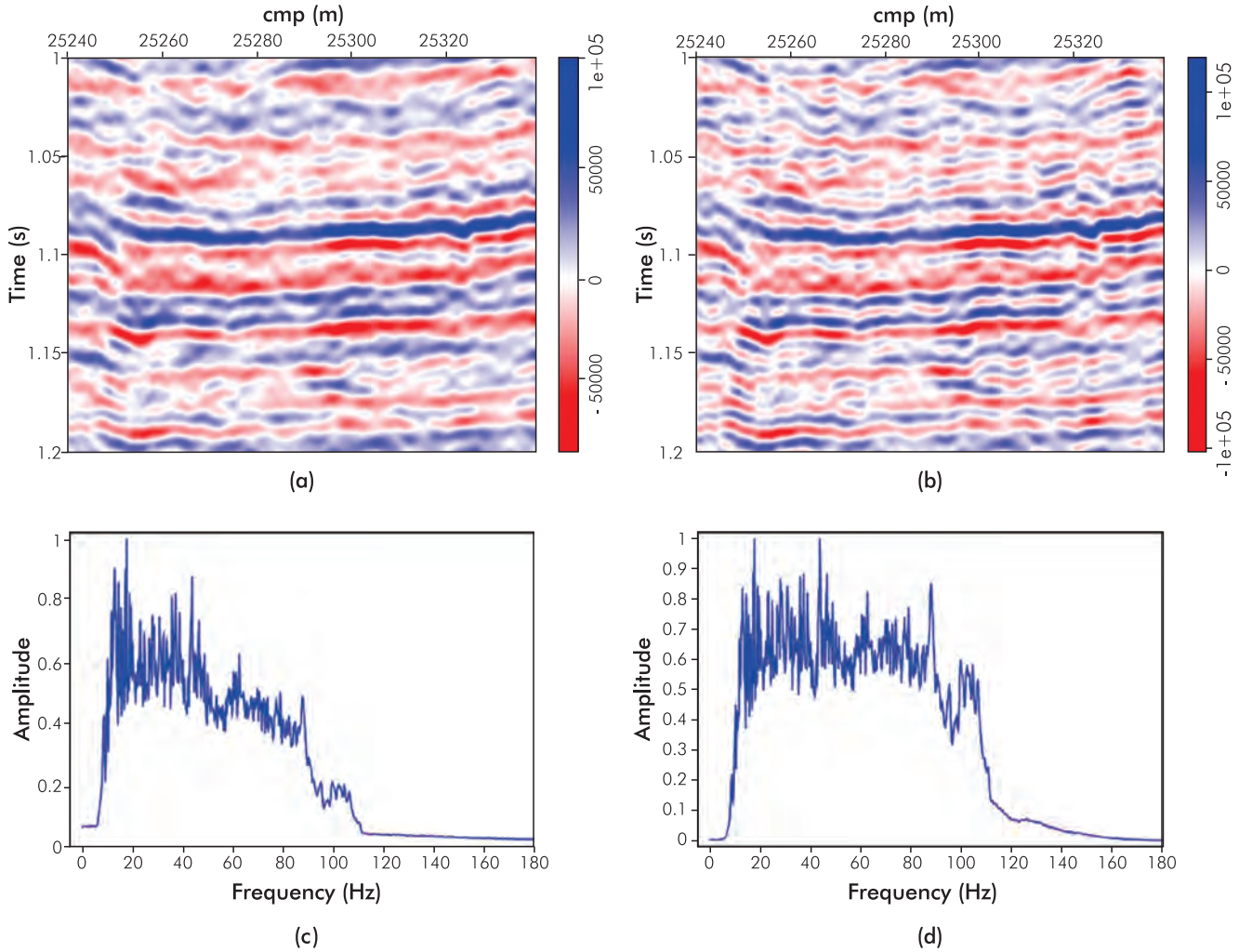
Figure 9. (a) Seismic section of interest (b) Section after regularized deconvolution based on the method described in section 3. (c) spectral amplitude before deconvolution (d) spectral amplitude after regularized deconvolution. $\delta_{rela} = 0.07263$, $\varepsilon = 0.4791$ and $Tol = 1 \times 10^{-6}$.

Different relative δ values were used between 0.1 and 0.01 for this regularized inversion in accordance with the discrepancy criteria and tolerance in the residual 1×10^{-6}. We have assumed that the rounding errors in calculations are less than the noise in the data. The inversion process applied is based on post-inversion as follows:

1. Estimating δ_{rela} with the CG method without regularization: The objective is to test the behavior of the error or residual $\mathbf{r} = \hat{\mathbf{d}} - \mathbf{L}\hat{\mathbf{m}}_\varepsilon$ with 100 or more iterations on the non-regularized solution, that is $\varepsilon = 0$. If the residual does not vary or the deconvolution process declines in successive iterations for a given value of relative noise, it is taken as the limit to make the deconvolution invertible. Based on the Morozov criteria,

nothing can be more accurate than the noise level. An exploration can be carried out with relative values of δ between 0.1 and 0.01 based on expression 14. If you know the uncertainty or standard deviation of the data you are working with, calculation δ_{rela} is direct and you would use *Equation 13*.

2. Defining tolerance and the initial value of ε_i: Considering the number of significant figures for precision and the adjustment of the minimum residual in the conjugate gradient algorithm (see Algorithm 2), the tolerance value should be consistent with the estimate of δ_{rela} in the previous step. To choose the (initial) ε, it is important to consider the appropriate weighting between what you want to get and tolerance in the accuracy in the inversion. If we

start with an $\varepsilon = 0$, we are at the noise limit allowed and we get a poor quality image deconvolution. If we take a value of $\varepsilon > 1$, we are smoothing the result and moving away from the real solution. For practical proposals, it is a good idea to explore with initial values of $\varepsilon \geq 1$. This gives Expression 15 a root and leads to an optimal $\varepsilon = \tilde{\varepsilon}$.

3. *Assessing the result of the inversion process:* Despite the application of the inversion theory and numerical implementation, the effectiveness of the regularized inversion process should be assessed with respect to the qualitative and quantitative analyses. A quantitative criterion independent from the vertical resolution that may be a comparison of the spectral amplitude is illustrated in Figures 9a and 9d. These spectra show that no abnormal frequencies were added and there was an improvement in the spectral amplitude to recover high frequencies. The above parameters change depending on these results. Therefore, it is post-inversion to preserve the balance between what you want and the tolerance in precision.

6. CONCLUSIONS

• This paper has shown how the inverse solution of an ill-conditioned operator (deconvolution) can be solved with a regularization model based on a conjugate gradient algorithm and stabilizing the solution to noise in data. We have illustrated the trade-off between error in the data and error in approximation of the model, that is, the balance between what you want and the tolerance allowed in the precision of solution.

• The optimal ε_i was found based on the Morozov discrepancy criteria, as a root of residual. The advantage of using the inversion model presented here was to find a satisfactory response for the deconvolution operator under the effect of Gaussian noise in data. An exact solution is not assumed in the inversion process obtained due to the level of noise in the data; the ill-conditioning of the deconvolution operator, the problems with errors in the floating point representation and the propagation of perturbations in the data.

Numerical results have shown us how the solution can be stable under the influence of noise based on the Morozov discrepancy criteria, whose meaning suggests that nothing can be more accurate than the noise level. A successful application of the proposed regularization model was performed in real data.

REFERENCES

Bonesky, T. (2009). Morozov's discrepancy principle and Tikhonov-type functionals. *Inverse Problems*, 25(1), 1-11.

Chen, Z., Wang, Y. & Chen, X. (2012). Gabor deconvolution using regularized smoothing. *SEG Technical Program Expanded Abstracts*, 2012: 1-5.

Engl, H., Hanke, M. & Neubauer, A. (1996). *Regularization of inverse problems.* Dordrecht: Kluwer Academic Publishers.

Fomel, S. (2007). Shaping regularization in geophysical-estimation problems. *Geophysics*, 72(2), R29-R36.

Hadamard, J. (1923). *Lectures on Cauchy's problem in linear differential equations*. New Haven: Yale University Press.

Hansen, C. (2002). Deconvolution and regularization with Toeplitz matrices. *Numerical Algorithms*, 29(4), 323-378.

Hansen, C. (2010). *Discrete inverse problems: Insight and Algorithms*. Philadelphia: SIAM, Society for Industrial and Applied Mathematics.

Hestenes, M. R. & Stiefel, E. L. (1952). Methods of conjugate gradients for solving linear systems. *J. Research Nat. Bur. Standards*, 49(6), 409-436.

Karsli, H., Guney, R. & Senkaya, M. (2012). High resolution deconvolution by combining Fx filtering and cauchy regularization. *5th Saint Petersburg International Conference & Exhibition*, EAGE.

Leinbach, J. (1995). Wiener spiking deconvolution and minimum-phase wavelets: A tutorial. *The Leading Edge,* 14(3), 189-192.

Montenegro, A. F. (2010). Regularización de problemas inversos e imágenes borrosas. *Tesis de Maestría, Departamento de Geofísica*, Universidad Nacional de Bogotá, Bogotá, Colombia, 87pp.

Morozov, V. A. (1984). *Methods for solving incorrectly posed problems.* New York: Springer.

Sen, M. & Roy, I. (2003). Computation of differential seismograms and iteration adaptive regularization in prestack waveform inversion. *Geophysics*, 68(6), 2026-2039 .

Sen, M. K. (2006). *Seismic inversion*. Austin: SPE.

Shewchuk, R. J. (1994). An introduction to the Conjugate Gradient method without the agonizing pain. *Report, School of Computer Science*, Carnegie Mellon University, Pittsburgh, USA.

Tikhonov, A. N. (1963). Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4: 1035-1038.

Tikhonov, A. N. & Arsenin, V. Y. (1977). *Solutions of ill-posed problems*. Michigan: W. H. Winston.

Van der Baan, M. & Pham, D. T. (2008). Robust wavelet estimation and blind deconvolution of noisy surface seismic. *Geophysics*, 73(5), 37-46.

Wang, J., Wang, X. & Perz, M. (2006). Structure preserving regularization for sparse deconvolution. *SEG Annual Meeting*, New Orleans, Louisiana. Conference Paper 2006-2072.

Yilmaz, Ö. (1987). *Seismic data processing in geophysics.* Tulsa: Society of Exploration Geophysicists.

Zala, C. A. (1992). High-resolution inversion of ultrasonic traces. *IEEE Transactions on Ultrasonics*, 39(4), 438-463.

AUTHORS

Herling Gonzalez
Affiliation: *Ecopetrol S. A. - Instituto Colombiano del Petróleo.*
MSc. Physics, *Universidad Industrial del Santander.*
e-mail: herling.gonzalez@ecopetrol.com.co

Sheryl Avendaño
Affiliation:*UT UP Consultorias.*
Physics, *Universidad Industrial del Santander.*
e-mail: sherphys@gmail.com

German Camacho
Affiliation: *Ecopetrol S. A. - Instituto Colombiano del Petróleo.*
MSc. Geophysics, *Universidad Nacional de Colombia.*
e-mail: german.camacho@ecopetrol.com.co

NOTATION

$\varepsilon, \varepsilon_i$	Regularization parameter
$\mathbf{d}, \hat{\mathbf{d}}$	Right-hand side, noise data
$\mathbf{R}, \Gamma$	Regularization matrix
$\mathbf{m}, \hat{\mathbf{m}}$	Regularized solution
$\mathbf{A}$	Coefficient matrix
PSF	Point Spread Function
$\mathbf{I}$	Identity matrix
$\mathbf{D}$	Finite difference matrix
R_i	Seismic reflectivity
$\mathbf{L}$	Deconvolution operator
l_p	General p-norm
S	Seismic data
p_i	Point spread coefficients
$\mathbf{p}_i$	Search direction on Conjugated Gradient algorithm
f	Misfit function
$\mathbf{r}, \mathbf{r}_i$	Residual solution
E	Error
k	Iterative index
δ, η	Noise level, and vector noise $\hat{\mathbf{d}} - \mathbf{d} = \eta$
$\|\cdot\|_2, \|\cdot\|$	Euclidean norm
$\langle \cdot, \cdot \rangle$	Inner product
$Cond(\cdot)$	Condition number
∇	Nabla operator, gradient