

DISEÑO DE UNA ESTRATEGIA PARA LA ADQUISICIÓN DE DATOS SÍSMICOS USANDO MUESTREO COMPRESIVO

JULIAN GONZALO MANTILLA ARIAS



CPS | RESEARCH GROUP

Universidad Industrial de Santander
Facultad de Ingenierías Fisicomecánicas
Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones
Maestría en Ingeniería Electrónica
Bucaramanga
2017

DISEÑO DE UNA ESTRATEGIA PARA LA ADQUISICIÓN DE
DATOS SÍSMICOS USANDO MUESTREO COMPRESIVO

Autor:
JULIAN GONZALO MANTILLA ARIAS

Trabajo de investigación presentado como requisito para optar al título de
Magíster en Ingeniería Electrónica

Directora:
PhD. Ana Beatriz Ramírez Silva

Codirectora:
Geóloga. Nancy Espíndola Lopez

Universidad Industrial de Santander
Facultad de Ingenierías Fisicomecánicas
Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones
Maestría en Ingeniería Electrónica
Bucaramanga
2017

AGRADECIMIENTOS

Gracias al grupo de investigación CPS por el apoyo prestado durante la elaboración de este proyecto. Al ICP por los conocimientos adquiridos en procesamiento de datos sísmicos que contribuyeron a la interdisciplinariedad de la investigación.

Este trabajo fue desarrollado con el apoyo de ECOPETROL y COLCIENCIAS como parte del proyecto de investigación No. 0266-2013.

CONTENIDO

	Pág.
INTRODUCCIÓN	13
1 MUESTREO COMPRESIVO: TEORÍA	15
1.1. MÉTODOS DE RECONSTRUCCIÓN DE DATOS COMPRIMIDOS .	17
1.2. INVERSIÓN DE ONDA COMPLETA	20
2 ADQUISICIÓN DE DATOS SÍSMICOS	22
2.1. DEFINICIONES GEOFÍSICAS	22
2.2. EXPLORACIÓN SÍSMICA	23
2.3. ADQUISICIÓN SÍSMICA 2D CONVENCIONAL	23
2.4. ADQUISICIÓN SÍSMICA 2D USANDO MUESTREO COMPRESIVO	24
3 PRUEBAS Y RESULTADOS	27
3.1. SELECCIÓN DEL DOMINIO DE REPRESENTACIÓN <i>SPARSE</i> . . .	27
3.2. DESCRIPCIÓN Y ACONDICIONAMIENTO DE LOS DATOS DE PRUE- BA	27
3.3. MEDICIÓN DEL ERROR DE APROXIMACIÓN	33
3.4. MATRIZ DE MUESTREO	35
3.5. SELECCIÓN DEL PARÁMETRO DE RECONSTRUCCIÓN λ	36
3.6. VALIDACIÓN DEL DISEÑO DE ADQUISICIÓN USANDO <i>CS</i> . . .	40
3.7. RECONSTRUCCIÓN DE DATOS SÍSMICOS REALES	42
3.8. RECONSTRUCCIÓN DE DATOS SÍSMICOS REALES CON ANCHO DE BANDA LIMITADO	51
3.9. RECONSTRUCCIÓN DE DATOS SÍSMICOS SINTÉTICOS PARA <i>FWI</i>	60
4 CONCLUSIONES	67
REFERENCIAS	69
BIBLIOGRAFÍA	72
ANEXOS	75

LISTA DE FIGURAS

	Pág.
Figura 1. Señal de audio representada en su dominio de representación natural y en el dominio de la frecuencia	16
Figura 2. Experimento sísmico terrestre	24
Figura 3. <i>Common shot gather</i> sintético	25
Figura 4. Líneas de adquisición sísmica con una distribución de fuentes uniforme y aleatoria	26
Figura 5. <i>Common receiver gather</i> número 1214 con diferentes tipos de visualización.	28
Figura 6. Representación del <i>CRG</i> 1214 con el 1% de los coeficientes ($\rho = 0.001$) en el dominio curvelet.	30
Figura 7. Representación del <i>CRG</i> número 1214 con el 1% de los coeficientes ($\rho = 0.001$) en el dominio de Fourier.	31
Figura 8. Representación del <i>CRG</i> número 1214 con el 1% de los coeficientes ($\rho = 0.001$) en el dominio wavelet.	32
Figura 9. Nivel <i>sparse</i> de los datos sísmicos agrupados por receptor común.	34
Figura 10. Matriz de muestreo aleatorio para el dato real, eliminado el 10%, 20% y 30% de las fuentes	36
Figura 11. Vista superior del cubo de submatrices A_k , derivadas de la matriz de muestreo aleatorio de la Figura 10 (b).	37
Figura 12. Porcentaje de trazas eliminadas en cada <i>CRG</i> , al aplicar las submatrices de muestreo de la Figura 11.	38
Figura 13. Curvas de correlación y SNR promedio de las trazas sísmicas de prueba reconstruidas con diferentes valores de λ	39
Figura 14. Sección del <i>CRG</i> número 1214 reconstruido con diferentes valores λ	41
Figura 15. <i>Common source gather</i> número 159, reconstruido tras eliminar un 10% ($sr=0.1$) de las fuentes.	44
Figura 16. <i>Common source gather</i> número 159, reconstruido tras eliminar un 20% ($sr=0.2$) de las fuentes.	45
Figura 17. <i>Common source gather</i> número 159, reconstruido tras eliminar un 30% ($sr=0.3$) de las fuentes.	46
Figura 18. <i>Common receiver gather</i> reconstruidos donde se evidencia la aparición de eventos anómalos en ángulos opuestos a los eventos sísmicos.	47
Figura 19. Espectro en frecuencia del <i>CSG</i> 159 original y los <i>CSG</i> reconstruidos	48
Figura 20. Sección sísmica apilada utilizando los datos reconstruidos a partir de la adquisición aleatoria con $sr = 0.3$	49

Figura 21. Sección sísmica apilada utilizando los datos adquiridos aleatoriamente con $sr = 0.3$	50
Figura 22. Curvas de correlación y SNR promedio de las trazas sísmicas con ancho de banda limitada al ser reconstruidas con diferentes valores de λ	52
Figura 23. Sección de un <i>CRG</i> con ancho de banda limitada reconstruido con diferentes valores λ	53
Figura 24. <i>Common source gather</i> número 159 con ancho de banda limitado que fue reconstruido tras eliminar un 10% ($sr=0.1$) de las fuentes.	54
Figura 25. <i>Common source gather</i> número 159 con ancho de banda limitado que fue reconstruido tras eliminar un 20% ($sr=0.2$) de las fuentes.	55
Figura 26. <i>Common source gather</i> número 159 con ancho de banda limitado que fue reconstruido tras eliminar un 30% ($sr=0.3$) de las fuentes.	56
Figura 27. Espectro en frecuencia del <i>CSG</i> original y los <i>CSG</i> reconstruidos con ancho de banda limitado	57
Figura 28. Sección sísmica apilada utilizando los datos reconstruidos con ancho de banda limitado a partir de la adquisición aleatoria con $sr = 0.3$	58
Figura 29. Sección sísmica apilada utilizando los datos con ancho de banda limitada adquiridos aleatoriamente con $sr = 0.3$	59
Figura 30. <i>Common source gather</i> sintético reconstruido tras eliminar un 30% ($sr=0.3$) de las fuentes.	61
Figura 31. Espectro en frecuencia de un <i>CSG</i> sintético y sus respectivos <i>CSG</i> reconstruidos	62
Figura 32. Modelo de velocidad original, inicial y de control obtenido con la <i>FWI</i> para los datos uniformes	64
Figura 33. Modelo de velocidad obtenido al aplicar <i>FWI</i> sobre los datos reconstruidos de una adquisición aleatoria con $sr = 0.3$	65
Figura 34. [Modelo de velocidad obtenido con la <i>FWI</i> de los datos aleatorios con $sr = 0.3$ sin aplicar el proceso de reconstrucción.	66
Figura 1. Características geométricas para la agrupación de datos sísmicos	75
Figura 2. Traza sísmica con la amplitud balanceada y normalizada utilizando <i>AGC</i> y su respectiva función de ganancia.	77

LISTA DE TABLAS

Tabla 1.	Valor promedio de λ teórico para los diferentes porcentajes de muestreo aleatorio. Utilizando la transformada curvelet como base de representación <i>sparse</i>	38
Tabla 2.	Valores de λ experimental, con los cuales se obtiene el mejor promedio de correlación y SNR.	39
Tabla 3.	Valor promedio de λ teórico para los diferentes porcentajes de muestreo aleatorio, del dato sísmico con ancho de banda limitado.	51
Tabla 4.	Valores de λ experimental, con los cuales se obtiene el mejor promedio de correlación y <i>SNR</i> , en la reconstrucción del dato sísmico con ancho de banda limitada.	52
Tabla 5.	Norma l_2 de la diferencia entre el modelo de velocidad real y los modelos obtenidos con <i>FWI</i>	63

LISTA DE ANEXOS

	Pág.
Anexo A. Agrupación de datos sísmicos.....	75
Anexo B. <i>Automatic gain control (AGC)</i>	76

RESUMEN

TÍTULO: Diseño de una Estrategia para la Adquisición de Datos Sísmicos Usando Muestreo Compresivo*

AUTOR: Julian Gonzalo Mantilla Arias**

PALABRAS CLAVE: Muestreo compresivo, datos sísmicos, señal *sparse*, reconstrucción de datos, transformada curvelet, *GPSR*, inversión de onda completa.

DESCRIPCIÓN:

La adquisición de datos sísmicos es una de las etapas en el proceso de exploración de hidrocarburos, que tradicionalmente se basa en un arreglo uniforme de fuentes y sensores sobre la superficie terrestre, para generar ondas de presión y medir las ondas reflejadas en las diferentes capas del subsuelo terrestre. Sin embargo, esta distribución no siempre es posible debido a obstáculos en la superficie terrestre que impide la ubicación exacta de las fuentes.

El muestreo compresivo es una nueva alternativa en la adquisición de datos sísmicos la cual permite cambiar la distribución uniforme de fuentes por una aleatoria, flexibilizando la ubicación de las fuentes y reduciendo su número en comparación con la técnica tradicional. El muestreo compresivo se basa en una representación poco densa de los datos sísmicos para lo cual se utilizó la transformada curvelet.

Las pruebas se realizaron en un conjunto de datos reales 2D, eliminando un 10%, 20% y 30% de las fuentes de acuerdo a una matriz de muestreo aleatorio previamente diseñada. Como resultado de la reconstrucción de las fuentes eliminadas se obtuvieron *common shot gathers* visualmente similares al original y que conservan el mismo contenido frecuencial, en especial para las bajas frecuencias donde se pierde menos información. Para complementar los resultados, se planteó un segundo escenario en el cual se reconstruyen datos sísmicos sintéticos para utilizarlos en un proceso de inversión de onda completa, obteniendo como resultado un modelo de velocidad similar al que se obtendría con una adquisición uniforme sin crear artefactos extraños que perjudiquen la calidad del modelo de velocidades.

* Trabajo de investigación.

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Maestría en Ingeniería Electrónica. Director: PhD. Ana Beatriz Ramírez Silva. Codirector: Geóloga. Nancy Espíndola Lopez.

ABSTRACT

TITLE: Design of a Strategy for Seismic Data Acquisition Applying Compressive Sensing*

AUTHOR: Julian Gonzalo Mantilla Arias**

KEYWORDS: Compressive sensing, seismic data, sparse signal, data reconstruction, curvelet transform, GPSR, full wave inversion.

DESCRIPTION:

Seismic data acquisition is one of the stages in the process of oil and gas exploration, that traditionally is based in uniform arrays of sources and sensors located at the earth surface, to generate pressure waves and measure the reflected waves in the earth subsurface. However, the uniform distribution is not always possible due to the presence of obstacles on the earth surface that obstruct the exact location of the sources.

Compressive sensing (CS) is a new alternative in the seismic data acquisition that allows changing the uniform distribution of sources by a random distribution, flexibilizing the source location and reducing its number in comparison with the traditional technique. Compressive sensing is based on a sparse representation of the seismic data, which it was used the curvelet transform.

The CS technique was tested in real 2-dimensional shot-gathers, removing 10 %, 20 % and 30 % of the sources, according to a previously designed matrix sensing. As reconstruction result, it was obtained visually similar common shot gathers to the original and also it conserved the same frequency spectrum, specially for the low frequencies where the lost information was less. To complement the results, a second stage was proposed where synthetic data was reconstructed to be used in full wave inversion process, getting as results velocity models similar to the obtained with an uniform acquisition, without creating strange artifacts that could damage the quality of the velocity model.

* Trabajo de investigación.

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Maestría en Ingeniería Electrónica. Director: PhD. Ana Beatriz Ramírez Silva. Codirector: Geóloga. Nancy Espíndola Lopez.

INTRODUCCIÓN

La adquisición de datos es una de las etapas en el proceso de exploración sísmica. Tradicionalmente, se utilizan esquemas de muestreo con una distribución regular o uniforme de fuentes y sensores para generar ondas de presión y medir las ondas reflejadas en las diferentes capas o estratos del subsuelo. Sin embargo, la ubicación regular de los instrumentos no siempre es posible debido a diversas causas tales como ambientales, sociales, geográficas, entre otras. Lo cual obliga a cambiar la posición de fuentes y sensores o incluso a eliminarlas de la adquisición. En cualquiera de estos casos se crean irregularidades en la información dejando huecos o trazas faltantes, siendo el caso de mayor gravedad cuando se elimina una fuente ya que se pierde todo un conjunto de trazas.

La reducción de costos es otro factor de interés en el proceso de adquisición de datos sísmicos y la forma más directa de hacerlo es reduciendo el número de fuentes a utilizar, pero reducir la cantidad de fuentes en una adquisición uniforme implica reducir el área de estudio o incrementar la distancia entre las fuentes perdiendo resolución en la información. Una nueva alternativa para la adquisición de datos sísmicos consiste en utilizar el muestreo compresivo para cambiar el esquema de muestreo tradicional por una distribución aleatoria de fuentes, dándole mayor flexibilidad a la ubicación de las fuentes y permitiendo reducir su número, en comparación con la técnica tradicional.

Como antecedentes en la investigación de este tema, se tiene principalmente los aportes realizados por Zwartjes y Herrmann. Zwartjes [1] abordó el tema como una regularización de datos faltantes en adquisiciones marinas 3D VSP utilizando el dominio de la frecuencia junto con una inversión poco densa para encontrar los coeficientes de Fourier que sinteticen la señal completamente reconstruida. Herrmann realizó el trabajos más notorio en esta área [2] utilizando directamente la teoría del muestreo compresivo, con la transformada curvelet como base de representación *sparse* y una reconstrucción con el software *Spectral Projected Gradient for l_1* (SPGL1). Los datos de prueba fueron un subconjunto de datos sísmicos con fuentes simultaneas, organizados por receptor común. Durante el desarrollo de este proyecto de investigación se realizaron pruebas que usan la técnica de muestreo compresivo para la adquisición de datos sísmicos 2D, diseñando una distribución aleatoria de fuentes y un esquema de reconstrucción para datos reales. El muestreo aleatorio se implementó por medio de simulación eliminando un 10 %, 20 % y 30 % de las fuentes utilizadas en una adquisición sísmica real, las fuentes eliminadas se reconstruyeron posteriormente siguiendo las indicaciones propuestas, como el balance de amplitudes y la selección del parámetro de reconstrucción λ . En el proceso de reconstrucción se tuvieron en cuenta dos enfoques,

uno para propósitos generales utilizando un dato sísmico real y otro para inversión de onda completa (*FWI*) utilizando datos con ancho de banda limitado. Para analizar la viabilidad de la técnica planteada se realizó un procesamiento de los datos reconstruidos dependiendo su enfoque, aplicando una *FWI* a los datos de ancho de banda limitado, y un apilado bruto a los datos de ancho de banda completo.

La selección de una base de representación *sparse* adecuada para los datos sísmicos es una etapa de gran importancia al aplicar el muestro compresivo, dada la naturaleza de los datos de interés se planteó una selección experimental comparando el rendimiento de las transformadas curvelet, Fourier y wavelet, obteniendo como resultado que la transformada curvelet, tiene el mejor desempeño en la representación *sparse* de los datos sísmicos, beneficiando la coherencia hiperbólica y la continuidad de las reflexiones existentes en un *common receiver gather*.

El presente documento está organizado en cuatro secciones que narran una secuencia de pasos necesarios para aplicar el muestreo compresivo en la adquisición de datos sísmicos y sus respectivos resultados. En el capítulo uno se presenta una explicación del muestreo compresivo, el algoritmo de reconstrucción utilizado y la inversión de onda completa. En el capítulo dos se introduce el concepto de adquisición uniforme de datos sísmicos y se plantea la forma como se van realizar las adquisiciones aleatorias. El capítulo tres contiene los resultados de la selección de un dominio para la representación *sparse* de los datos sísmicos, diseño de la matriz de muestreo, selección del parámetro λ y reconstrucción de los datos, además de los resultados obtenidos en la inversión de onda completa usando datos sintéticos reconstruidos. Por último, en el capítulo cuatro se presentan las conclusiones y recomendaciones.

1. MUESTREO COMPRESIVO: TEORÍA

Una señal continua puede representarse y reconstruirse a partir de un conjunto de muestras tomadas uniformemente, como se plantea en el teorema de muestreo de Nyquist, el cual indica que la frecuencia de muestreo debe ser por lo menos el doble de la frecuencia máxima de la señal. El muestreo compresivo es un nuevo modelo en la adquisición de datos, el cual permite representar una señal a partir de un conjunto de muestras menor al que se tendría usando el método tradicional. Esta teoría es relativamente nueva, su auge empezó en el año 2006 con los trabajos de Emmanuel J. Candès[3] y David L. Donoho[4], quienes explicaron el funcionamiento teórico de esta herramienta.

La idea principal del *CS* es reconstruir una señal o imagen de interés, a partir de un conjunto de datos muestreados aleatoriamente, que es menor al que se tendría con los métodos tradicionales. En otros términos, es la reconstrucción de un vector de gran dimensión después de un proceso de reducción dimensional. Esta teoría se basa en dos principios fundamentales: que la señal de interés sea poco densa en un dominio de representación, y que el dominio de representación de la señal de interés sea incoherente con la matriz de muestreo [5].

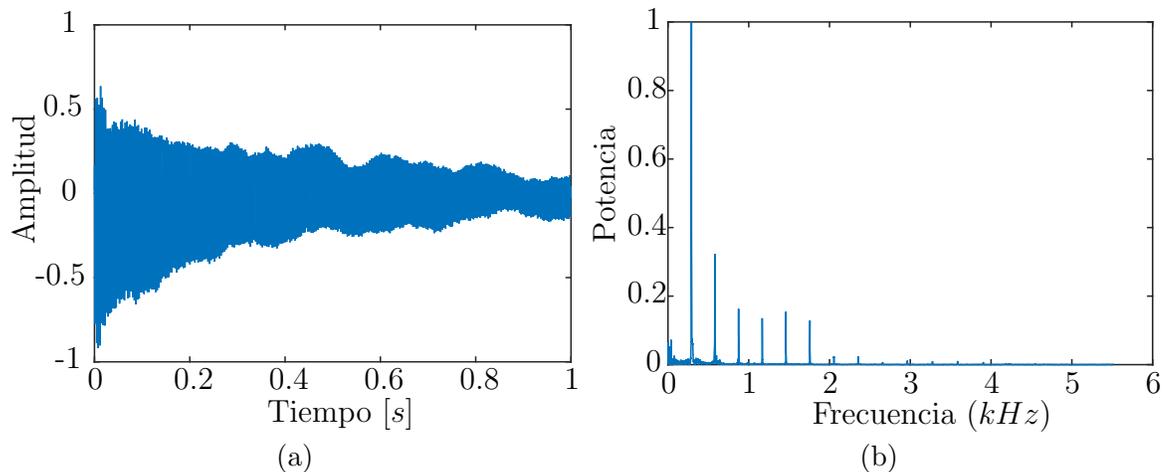
Señal poco densa (*sparse*): es aquella que tiene una baja densidad en los coeficientes de representación, es decir, tiene pocos coeficientes diferentes de cero, en su dominio natural o en algún dominio transformado [5]. En la Figura 1 se muestra como ejemplo una señal de audio que es densa en el dominio del tiempo, pero en el dominio de la frecuencia se representa con pocos coeficientes diferentes de cero, es decir es *sparse*.

Incoherencia (*incoherence*): hace referencia a la baja correlación que debe existir entre cada una de las columnas de la matriz usada para muestrear la señal y la base de representación *sparse*. Con este concepto se extiende la idea de que en el *CS* no se busca observar directamente la señal en su dominio natural, sino muestrear un pequeño número de coeficientes en el dominio de representación *sparse*. [6]

La calidad de la señal reconstruida depende de estos principios de la siguiente forma: Cuanto más *sparse* sea la representación de la señal o más incoherente el muestreo, la señal reconstruida se acerca más a la original.

Una de las etapas más importante dentro del esquema del *CS* es la selección de un dominio adecuado para la representación *sparse* de los datos, debido a que de éste depende en gran medida la calidad de la señal reconstruida. Para seleccionar el dominio más adecuado, se realizó una medición experimental del nivel *sparse* de las señales sísmi-

Figura 1: Señal de audio representada en (a) dominio natural y (b) dominio de la frecuencia.



cas al ser representadas con tres diccionarios diferentes, los cuales son: transformada curvelet, Furier y wavelet.

Los diccionarios a analizar, se escogieron teniendo en cuenta que son transformadas comúnmente utilizadas en el procesamiento de datos sísmicos y en aplicaciones del *CS* en otras áreas.

Transformada curvelet: Es una generalización multidimensional de la transformada wavelet y se utiliza para representar objetos con bordes y datos provenientes de fenómenos de propagación de onda [7]. La descomposición *curvelet* se representa en el dominio espaciotemporal como un conjunto de elipsoides desplazadas a diferentes ángulos y escalas, una descripción más detallada se encuentra en [8].

Transformada de Fourier 2D: Es una herramienta matemática utilizada para transformar señales del dominio temporal o espacial, al dominio de la frecuencia, descompone la señal en una serie de coeficientes que representan las componentes frecuenciales de la señal.

Transformada wavelet: Eepresentación tiempo-frecuencia que descompone la señal en un conjunto de ondículas (*wavelets*), que son generadas a partir de translaciones y cambios de escala de una *wavelet* madre.

Muestreo aleatorio: El muestreo aleatorio de una señal *sparse* se expresa como

$$y = \Phi f, \quad (1.1)$$

donde f es una señal real en tiempo continuo de dimensión N , y es la señal muestreada aleatoriamente de dimensión M y Φ es la matriz de muestreo aleatorio, para $M \ll N$,

sin ninguna columna de ceros. Previamente sabemos que la señal f tiene una representación *sparse* en algún dominio transformado, por lo tanto, existe un diccionario Ψ en el cual $f = \Psi x$, donde x es un vector *sparse*. Las mediciones aleatorias se pueden expresar como $y = \Phi \Psi x$, agrupando Φ y Ψ en un solo termino A , se obtiene

$$y = Ax. \tag{1.2}$$

Reconstrucción: El *CS* es un problema inverso que consiste en recuperar \tilde{f} , a partir de y o recuperar x , a partir de y . En el último caso se tiene un problema lineal de ecuaciones indeterminado, donde solo se sabe que el vector a ser recuperado es *sparse* [4]. La solución se obtiene mediante un algoritmo de optimización convexo, el cual combina la norma l_1 para medir la baja densidad de información y l_2 para medir el error de aproximación. La norma l_1 se define como $\|x\|_1 = \sum_i |x_i|$ y se usa para medir el nivel *sparse* de la señal, esta norma no mide exactamente el nivel *sparse* de la señal, pero es una buena aproximación con un bajo costo computacional [9]. La norma l_2 se define como $\|x\|_2 = \sqrt{\sum_i |x_i|^2}$, se utiliza para medir el ajuste entre la señal reconstruida y la señal observada. El problema de optimización se plantea como

$$\hat{x} = \arg \min_x \|y - Ax\|_2^2 + \lambda \|x\|_1, \tag{1.3}$$

donde λ es un parámetro que regula el nivel *sparse* de la solución, de manera que si λ crece, la solución tiene un mayor nivel *sparse* y si λ decrece, la solución se aproxima a la solución de mínimos cuadrados.

El parámetro de regularización λ se puede calcular en función de los coeficientes de representación *sparse* de la señal. En [10], los autores proponen el siguiente valor teórico de λ :

$$\lambda = 0.1 \|2A^T y\|_\infty, \tag{1.4}$$

la cual representa una fracción del valor máximo que puede tomar λ sin que la solución óptima de (1.3) sea un vector de coeficientes cero

$$\lambda_{max} = \|2A^T y\|_\infty, \tag{1.5}$$

es decir que para $\lambda \geq \lambda_{max}$ la solución óptima del problema (1.3) es cero. La ecuación (1.4) no garantiza un valor óptimo para la reconstrucción por lo tanto se puede complementar con una sintonización empírica, realizando un barrido alrededor de este punto para seleccionar el valor de λ , con el cual se obtiene una reconstrucción de mejor calidad.

1.1. MÉTODOS DE RECONSTRUCCIÓN DE DATOS COMPRIMIDOS

Para resolver el problema dado en (1.3) se han propuesto diferentes métodos como: *Gradient Projection for Sparse Reconstruction (GPSR)* [11], *sparse reconstruction*

by separable approximation (*SpaRSA*) [12], *two-step iterative shrinkage/thresholding* (*TwIST*) [13], entre otros. Utilizados para resolver problemas inversos de gran escala como el del muestreo compresivo.

El *GPSR* [11] se enfoca principalmente en representar (1.3) como un programa cuadrático, y reducir la complejidad de las operaciones hasta alcanzar operaciones básicas entre matrices. Para convertir la ecuación (1.3) en un programa cuadrático, los autores proponen: dividir x en parte positiva u y parte negativa v

$$\begin{aligned} x &= u - v, \\ \text{para } u &\geq 0, \quad v \geq 0, \end{aligned} \tag{1.6}$$

cumpliendo la relación de $u_i = (x_i)_+$ y $v_i = (-x_i)_+$ para $i = 1, 2, \dots, n$. Sustituyendo la nueva representación de x se tiene que $\|x\|_1 = \mathbf{1}_n^T u + \mathbf{1}_n^T v$, donde $\mathbf{1}_n = [1, 1, \dots, 1]^T$ es un vector de n unos, por lo tanto podemos describir (1.3) como un *bound-constrained quadratic program* (*BCQP*)

$$\begin{aligned} \min_{u,v} \|y - A\Psi(u - v)\|_2^2 + \lambda \mathbf{1}_n^T u + \lambda \mathbf{1}_n^T v, \\ \text{para } u &\geq 0, \quad v \geq 0. \end{aligned} \tag{1.7}$$

El problema (1.7) se puede escribir con una notación estándar de *BCQP*

$$\begin{aligned} \min_z c^T z + \frac{1}{2} z^T B z \equiv F(z), \\ \text{para } z &\geq 0, \end{aligned} \tag{1.8}$$

donde

$$z = \begin{bmatrix} u \\ v \end{bmatrix}, \quad b = A^T y, \quad c = \lambda \mathbf{1}_{2n} + \begin{bmatrix} -b \\ b \end{bmatrix} \quad B = \begin{bmatrix} A^T A & -A^T A \\ -A^T A & A^T A \end{bmatrix}. \tag{1.9}$$

Para resolver el problema (1.8) se utiliza la proyección de gradiente *GP* (por sus siglas en inglés, *Gradient Projection*) con el enfoque de Barzilai y Borwein (BB) [14], el cual utiliza una Hessiana calculada de forma simple y aproximada. Inicialmente se escoge un parámetro $\alpha^{(k)}$ y se establece

$$w^{(k)} = (z^{(k)} - \alpha^{(k)} \nabla F(z^{(k)}))_+. \tag{1.10}$$

Luego se escoge un segundo escalar $\lambda^{(k)} \in [0, 1]$ y se establece

$$z^{(k+1)} = z^{(k)} + \lambda^{(k)} (w^{(k)} - z^{(k)}). \tag{1.11}$$

Con el enfoque de BB se calcula el avance en cada paso, con la expresión $\delta^{(k)} = -H_k^{-1}\nabla F(Z^{(k)})$, donde H_k es una aproximación a la Hessiana de F en $z^{(k)}$. Barzilai y Borwein [14] proponen una forma rápida y aproximada para calcular la Hessiana como múltiplo de la identidad $H_k = \eta^{(k)}I$, donde $\eta^{(k)}$ se escoge para que esta aproximación tenga un comportamiento similar al de la Hessiana real del último paso, como lo expresa

$$\nabla F(z^{(k)}) - \nabla F(z^{(k-1)}) \approx \eta^{(k)} [z^{(k)} - z^{(k-1)}], \quad (1.12)$$

escogiendo $\eta^{(k)}$ para satisfacer esta condición en el sentido de mínimos cuadrados, el avance se expresa como

$$z^{(k+1)} = z^{(k)} - (\eta^{(k)})^{-1} \nabla F(z^{(k)}). \quad (1.13)$$

Se escoge λ_k en (1.11) como el minimizador exacto en el intervalo $[0, 1]$ y se escoge $\eta^{(k)}$ en cada paso, como se describe anteriormente, exceptuando para $\alpha^{(k)} = (\eta^{(k)})^{-1}$ que está restringida en el intervalo $[\alpha_{min}, \alpha_{max}]$. El valor de $\alpha^{(k)}$ se calcula utilizando

$$\nabla F(z^{(k)}) - \nabla F(z^{(k-1)}) = B [z^{(k)} - z^{(k-1)}]. \quad (1.14)$$

El algoritmo del *GPSR* con el enfoque de BB, fue propuesto e implementado por Figueiredo, Nowak y Wright en [11], y se muestra a continuación.

Paso 0 (inicialización): Dado $z^{(0)}$, se escogen los parámetros $\alpha_{min}, \alpha_{max}, \alpha^{(0)} \in [\alpha_{min}, \alpha_{max}]$ y se fija $k = 0$.

Paso 1: Cálculo del paso:

$$\delta^{(k)} = (z^{(k)} - \alpha^{(k)} \nabla F(z^{(k)})) - z^{(k)}. \quad (1.15)$$

Paso 2 (búsqueda lineal): Hallar el escalar $\lambda^{(k)}$ que minimice $F(z^{(k)} + \lambda^{(k)}\delta^{(k)})$ en el intervalo $\lambda^{(k)} \in [0, 1]$ y establecer $z^{(k+1)} = z^{(k)} + \lambda^{(k)}\delta^{(k)}$.

Paso 3 (actualizando α): Calcular

$$\begin{aligned} \gamma^{(k)} &= (\delta^{(k)})^T B \delta^{(k)}; \\ \text{si } \gamma^{(k)} &= 0, \text{ entonces } \alpha^{(k+1)} = \alpha_{max}, \text{ en otros casos} \\ \alpha^{(k+1)} &= \text{mid} \left\{ \alpha_{min}, \frac{\|\delta^{(k)}\|_2^2}{\gamma^{(k)}}, \alpha_{max} \right\}. \end{aligned} \quad (1.16)$$

Paso 4: Determinar si la solución aproximada de $z^{(k+1)}$ cumple con la prueba de convergencia, de lo contrario establecer $k \leftarrow k + 1$ y regresar al **paso 1**.

El algoritmo completo está escrito en Matlab y se encuentra disponible para descargar gratuita en <http://www.lx.it.pt/mtf/GPSR/>.

1.2. INVERSIÓN DE ONDA COMPLETA

Una de las etapas más importantes en el procesamiento de datos sísmicos es la estimación del modelo de velocidades asociado a la zona de interés, ya que se requiere de un buen modelo de velocidades para obtener la posición correcta de las estructuras geológicas en la imagen migrada. La inversión de onda completa (*FWI*, por sus siglas en inglés) es un método propuesto por Tarantola en 1984,[15] para obtener modelos de velocidades mediante un problema de optimización que busca minimizar (en el sentido de los mínimos cuadrados) la función de ajuste J , la cual mide el error entre los datos observados d y los datos modelados $F(m)$, que se van obteniendo con el modelo de velocidad estimado m .

$$J(m) = \frac{1}{2} \|d - F(m)\|^2, \quad (1.17)$$

donde $F(\cdot)$ es el operador de modelado directo.

Los datos modelados se pueden generar en el dominio del tiempo mediante la solución numérica de la ecuación de onda completa, esta solución numérica debe representar lo mejor posible la propagación de la onda sísmica en el subsuelo terrestre, pero debido a limitantes computacionales se deben realizar ciertas aproximaciones que permitan su implementación. El subsuelo terrestre se modela comúnmente como un medio acústico isótropo [16], que se representa con la ecuación de onda acústica con densidad constante:

$$\frac{\partial^2 p(x, t)}{\partial t^2} = v^2(x) \nabla^2 p(x, t) + s(x, t), \quad (1.18)$$

donde $p(x, t)$ es el campo de presión, $v(x)$ representa el modelo de velocidad acústica, ∇^2 es el operador laplaciano, $s(x, t)$ es la fuente, x la posición y t el tiempo.

La función de costo (1.17) se minimiza por medio de un proceso iterativo donde se actualiza el modelo de velocidades m siguiendo el método de Newton [16], a través de la ecuación de actualización:

$$m_{k+1} = m_k - H_k^{-1} g_k, \quad (1.19)$$

donde k indica el paso de iteración, H_k es la matriz hessiana que contiene la segunda derivada parcial de $J(m)$ evaluadas en m_k y g_k es el gradiente de la función de ajuste ($\partial J / \partial m$ evaluada en m_k). El gradiente se puede calcular de forma eficiente utilizando el método del estado adjunto [17], para el caso isótropo con densidad constante g está dado por

$$g = \frac{\partial J}{\partial m} = \frac{2}{v^3} \int_0^T \beta \frac{\partial^2 p(x, t)}{\partial t^2} dt, \quad (1.20)$$

donde T es el tiempo de adquisición y β es el campo que representa el residual retro propagado, y se halla mediante una retro propagación de $F(m) - d$ en la ecuación (1.18).

Debido al costo computacional y tamaño de la matriz Hessiana, se hace inviable su cálculo. Una alternativa es utilizar un método iterativo generalizado como el propuesto

en[18].

$$m_{k+1} = m_k + \alpha_k h_k, \tag{1.21}$$

donde α_k es un escalar que indica la longitud del paso de búsqueda y h_k es la dirección de búsqueda [16], calculadas con métodos de gradiente como L-BFGS [19].

2. ADQUISICIÓN DE DATOS SÍSMICOS

Los datos sísmicos son el resultado de un proceso geofísico conocido como exploración sísmica, el cual utiliza un conjunto de fuentes y receptores para aplicar una perturbación al subsuelo y captar la señal de respuesta. Convencionalmente las fuentes y los receptores se ubican de forma regular, manteniendo una separación constante entre cada uno de ellos a lo largo de la línea de adquisición [20]. La ubicación de las fuentes es un tema de especial interés ya que no siempre es posible ubicarlas en la posición donde deberían ir, causando irregularidades. Además el costo de una exploración sísmica se da en función del número de fuentes, por lo tanto, si reducimos su cantidad necesaria, se reduce el costo y tiempo de la adquisición.

Utilizando el muestreo compresivo como estrategia para la adquisición de datos sísmicos, podemos plantear una distribución aleatoria de fuentes, reduciendo el número de fuentes necesarias y facilitando su ubicación con respecto a una adquisición convencional. Durante el desarrollo de este proyecto de investigación se abarcó el problema de aplicar el muestreo aleatorio en la distribución de las fuentes y posteriormente reconstruir los datos observados de modo que se obtenga un conjunto de datos sísmicos similar ó igual al de una adquisición convencional.

2.1. DEFINICIONES GEOFÍSICAS

Receptor (*receiver*): Dispositivo que detecta la energía sísmica en forma de movimiento terrestre u onda de presión en los fluidos y la transforma en corriente eléctrica [21].

Fuente (*source*): Dispositivo que perturba el subsuelo terrestre inyectando la energía necesaria para la adquisición de datos sísmicos, pueden ser cargas explosivas, cañones de aire o vibradores [21].

Traza sísmica: Es la información sísmica capturada por un canal o receptor que representa la respuesta del subsuelo ante una perturbación causada por una fuente [21].

Gather: Agrupación o recolección de trazas sísmicas que comparten un parámetro de adquisición, como posición de la fuente o el punto medio común [21].

Offset: En la adquisición de datos sísmicos superficiales corresponde a la distancia entre la fuente y los geófonos [21].

Midpoint: Punto medio entre una fuente y un receptor en la superficie terrestre [21].

Fold: Es una medida de la redundancia o trazas acumuladas que aportan información a un *midpoint* [21].

Header: Información complementaria de los datos sísmicos que indican la localización, parámetros de adquisición, procesamiento y demás información relacionada con los registros sísmicos [21], el amarre de esta información se hace con cada traza sin importar si es redundante dentro de un mismo *gather*.

Geófono: Dispositivo utilizado en la adquisición sísmica terrestre, que detecta el movimiento del terreno, producido por una onda sísmica y la transforma en impulsos eléctricos [21].

Hidrófono: Dispositivo diseñado para detectar la energía sísmica, que se presenta como cambios de presión de la columna de agua, durante una adquisición sísmica marina [21].

2.2. EXPLORACIÓN SÍSMICA

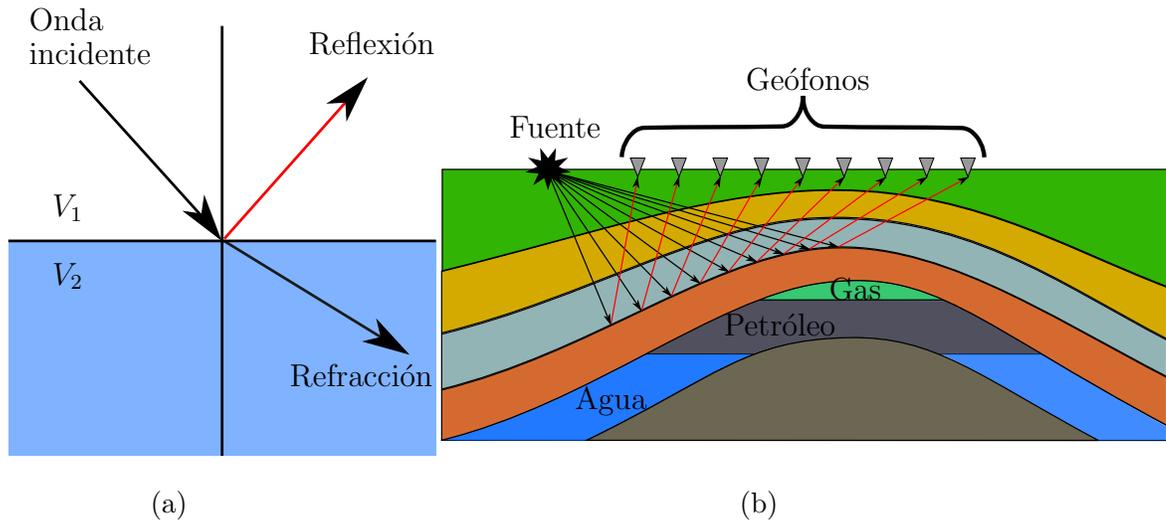
La sísmica de reflexión es el método geofísico más usado a nivel mundial para la exploración de hidrocarburos, se basa en la propagación de ondas elásticas a través del subsuelo las cuales al pasar por una interfaz entre dos capas de roca con diferentes propiedades físicas y litológicas (reflector), se reflejan y refractan como lo indica la ley de Snell. En la Figura 2 (a) se ilustra este comportamiento. La energía reflejada contiene la información de interés, y al regresar a la superficie es captada por un conjunto de sensores llamados geófonos, como se muestra en la Figura 2 (b).

Las ondas elásticas se generan por medio de fuentes impulsivas como explosiones o camiones vibradores en el caso terrestre y cañones de aire en el caso marino. Como resultado del experimento sísmico se obtienen una serie de trazas que corresponden a las ondas captadas por cada geófono, las trazas se pueden ordenar de diferentes formas para facilitar su análisis, la más común y como se obtienen originalmente es agrupadas por fuente común (*CSG* por sus siglas en inglés, *common shot gather*). Un ejemplo sintético se presenta en la Figura 3, donde las hipérbolas representan los diferentes reflectores y el ápice de la hipérbola indica la posición de la fuente.

2.3. ADQUISICIÓN SÍSMICA 2D CONVENCIONAL

Las técnicas de adquisición sísmica 2D convencionales se caracterizan por tener una distribución uniforme de sensores y fuentes a lo largo de una línea de exploración, como se muestra en el ejemplo de la Figura 4 (a). Para determinar los intervalos de separación entre fuentes (S_i) y receptoras (R_i), se utiliza información previa del terreno como ángulos de buzamiento, velocidades promedio de propagación de la onda hasta el

Figura 2: (a) representación gráfica de la ley de snell, (b) experimento sísmico terrestre.



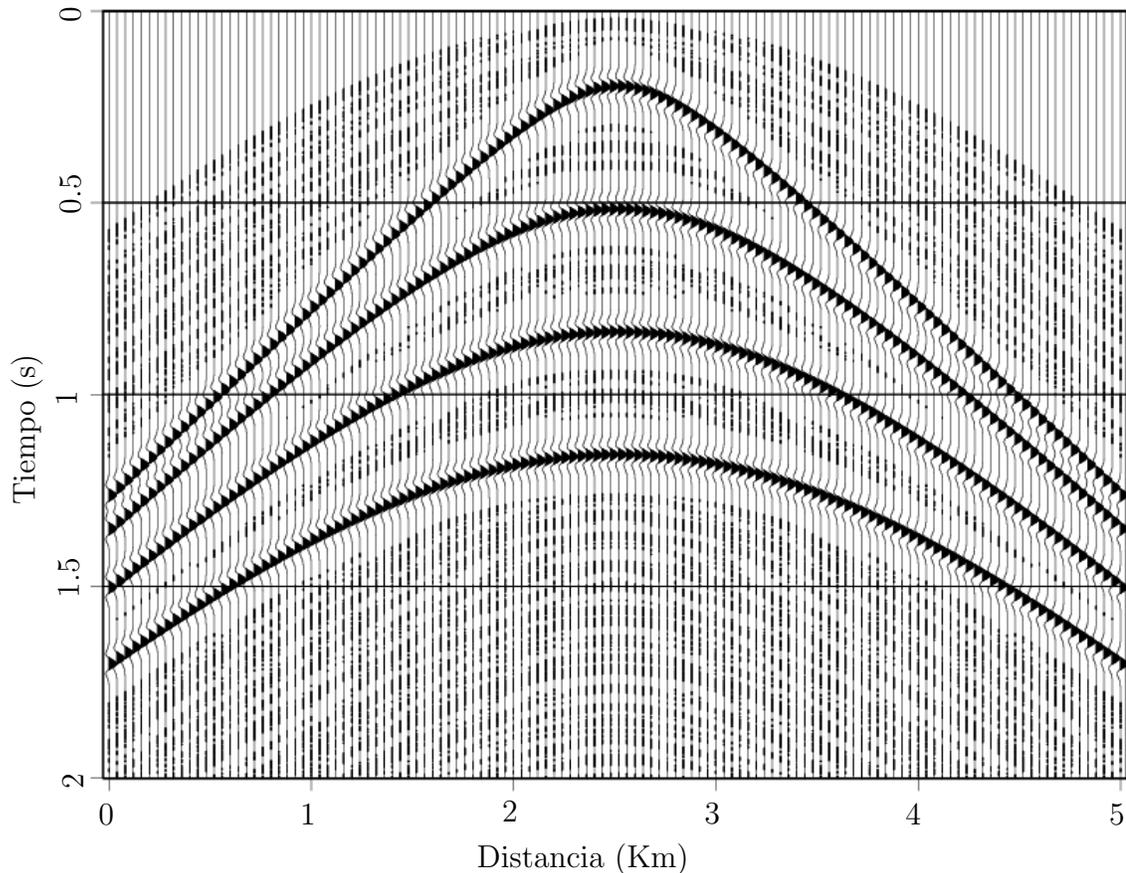
objetivo, frecuencia máxima que se puede recuperar y la resolución espacial deseada. La teoría detallada de cómo se seleccionan todos los parámetros de adquisición se encuentra en [20][23].

Para evitar problemas de *aliasing* espacial se tiene en cuenta la teoría de Nyquist, garantizando que se muestreen por lo menos dos trazas por longitud de onda aparente que regresa a la superficie. La longitud de onda aparente corresponde a los eventos más pequeños que se desean iluminar dentro de la zona de interés [24].

La regularidad en la ubicación de fuentes y receptores se debe cumplir con especial cuidado, debido a que las irregularidades en la posición de las fuentes o sensores dificulta el procesamiento de los datos [24], este efecto se puede ver fácilmente en el comportamiento al agrupar los datos en *cdp bin gathers* donde aparecen huecos en la información.

2.4. ADQUISICIÓN SÍSMICA 2D USANDO MUESTREO COMPRESIVO

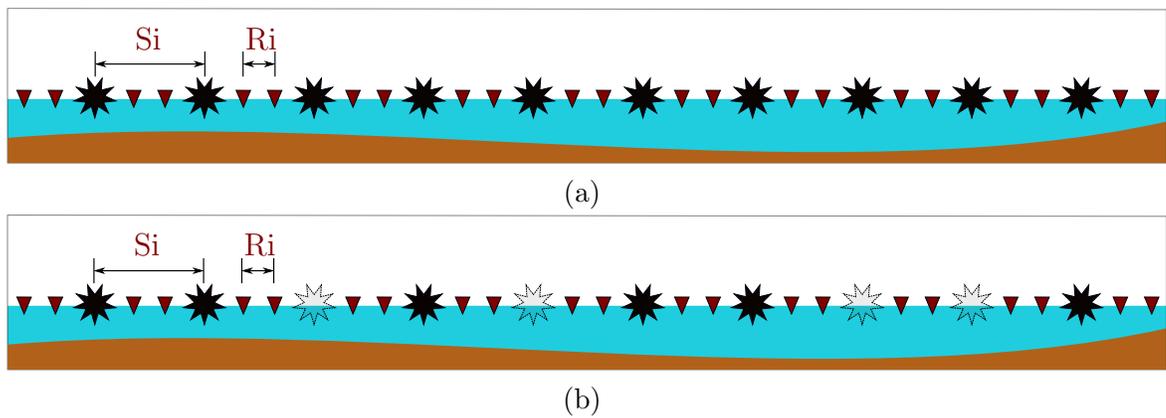
El muestreo compresivo como técnica para la adquisición de datos sísmicos permite implementar una distribución aleatoria de fuentes a lo largo de una línea de adquisición, permitiendo omitir un porcentaje de fuentes con respecto a la adquisición convencional, pero sin perder datos ya que las fuentes eliminadas se reconstruyen en una etapa posterior, por lo tanto no se afecta la regularidad de los datos y se pueden aplicar todas las técnicas de procesamiento convencional. La reconstrucción de los datos se realiza ordenando las trazas por receptor común (*CRG* por sus siglas en inglés, *common re-*

Figura 3: *Common shot gather* sintético. Adaptado de[22]

ceiver gather) en los cuales las fuentes eliminadas se reflejan como trazas faltantes o huecos de información en el *gather*.

El muestreo aleatorio o distribución aleatoria de fuentes, se realiza de manera regular, eliminando fuentes previamente establecidas en un esquema de adquisición sísmica convencional, como se puede observar en el ejemplo de las Figuras 4 (a) y (b). Para establecer cuáles fuentes se van a eliminar se propone realizar un diseño de la adquisición aleatoria teniendo en cuenta ciertas reglas que favorezcan la posterior reconstrucción de los datos. Como primera consideración no se deben eliminar fuentes en los bordes de la línea de adquisición ya que los *CRG* correspondientes a los receptores de esas zonas agrupan pocas trazas y no cuentan con la información suficiente para una buena reconstrucción, también se restringe la cantidad de fuentes contiguas que pueden ser eliminadas para evitar huecos de información muy grandes, adicionalmente, se generan las posiciones de las fuentes a eliminar siguiendo una distribución de probabilidad uniforme para garantizar que los huecos en la información no se concentren en ninguna zona de la línea de adquisición.

Figura 4: Líneas de adquisición sísmica con una distribución de fuentes: (a) uniforme y (b) aleatoria. Las estrellas sólidas representan las fuentes que se colocan en campo y las huecas representan las fuentes que se eliminan aleatoriamente. Los triángulos representan los receptores.



3. PRUEBAS Y RESULTADOS

3.1. SELECCIÓN DEL DOMINIO DE REPRESENTACIÓN *SPARSE*

En el capítulo 1 se presento una definición de señal *sparse* pero esta definición debe ampliarse, debido que al representar una señal en diferentes dominios no se obtienen pocos coeficientes diferentes de cero, sin embargo, la energía sí se concentra en pocos coeficientes, lo cual permite descartar los coeficientes de menor amplitud sin perder un porcentaje elevado de la información. La cantidad de coeficientes que se pueden descartar sin destruir la señal indican qué tan *sparse* es la representación en cada dominio transformado.

Con base en el planteamiento anterior se formuló un experimento que consiste en representar un dato sísmico con las transformadas curvelet, Fourier y wavelet, conservar un factor ρ de los coeficientes más significativos y aplicar la transformada inversa de cada diccionario sobre estos coeficientes para regresar al dominio natural y realizar las mediciones correspondientes.

Se compararon tres dominios de representación, los cuales son: Transformada curvelet con 5 escalas y 64 ángulos, transformada wavelet de 2 escalas con una ondícula *symlet* 12 y transformada de Fourier *2D*. Los dominios para analizar se escogieron teniendo en cuenta que son transformadas comúnmente utilizadas en aplicaciones del CS en otras áreas, los parámetros específicos de cada transformada se establecieron por medio de experimento complementarios iguales al planteado en el párrafo anterior. La transformada curvelet se encuentra disponible en el software CurveLab [25] y las demás transformadas están disponibles en el software Matlab.

3.2. DESCRIPCIÓN Y ACONDICIONAMIENTO DE LOS DATOS DE PRUEBA

Se utilizó un conjunto de datos sísmicos 2D marino de 314 fuentes con una geometría de adquisición uniforme tradicional, agrupados en 1816 *gathers* receptor común de 2400 muestras y periodo de muestreo 2[ms], y se muestran en la Figura 5. Para los análisis de *sparsity* y selección del parámetro λ se extrajeron aleatoriamente 18 *gathers* del conjunto total de datos, como muestra representativa para realizar todas las pruebas de selección de parámetros ya que trabajar con todo el conjunto de datos se vuelve inviable dentro de una metodología de reconstrucción por su costo computacional.

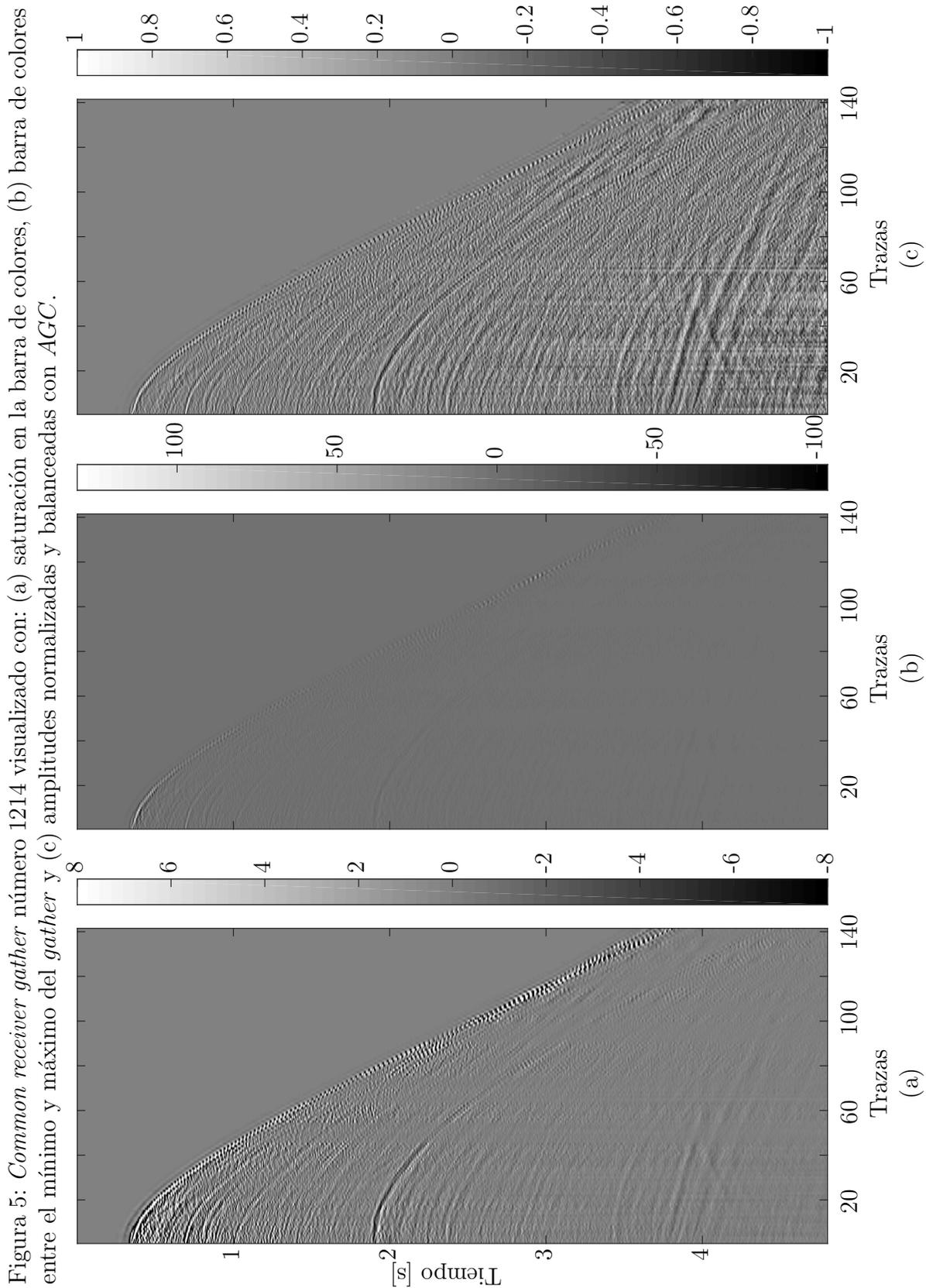


Figura 5: *Common receiver gather* número 1214 visualizado con: (a) saturación en la barra de colores, (b) barra de colores entre el mínimo y máximo del *gather* y (c) amplitudes normalizadas y balanceadas con AGC.

La amplitud de los datos sísmicos reales decrece a medida que incrementa el tiempo y *offset*, llegando a cambiar hasta 100 veces el rango de amplitudes de las trazas con el menor y mayor *offset* dentro de un mismo *gather*, estas pérdidas de energía se deben a fenómenos de propagación como la divergencia esférica y atenuación. Este fuerte cambio en las amplitudes no se nota en la mayoría de imágenes presentadas en este informe como es el caso de la Figura 5(a), debido a que se aplica una saturación en la barra de colores que permite obtener una buena visualización de los eventos sísmicos contenidos en el *gather*. De no aplicarse este cambio en la barra de colores, se obtendría una imagen como la Figura 5(b), en la cual solo se identifica la primera reflexión y las múltiples, perdiendo de vista la mayoría de las hipérbolas de reflexión.

Los cambios en las amplitudes también afectan la representación *sparse* de los datos sísmicos, al transformar un *gather* a los diferentes dominios, los coeficientes que representan los eventos de las zonas de menor amplitud son considerablemente menores que los coeficientes que representan las zonas de mayor amplitud, sin implicar que los coeficientes más grandes sean los más significativos o que contengan la mayor cantidad de información del *gather*, lo cual no concuerda con la definición de señal *sparse* dada al inicio de esta sección. Por ésto se debe aplicar un balance de amplitudes para evitar una falsa representación *sparse*, como se ve en las Figuras 6,7 y 8, donde para un valor de ρ pequeño se empiezan a eliminar los eventos de las zonas de mayor *offset* y tiempo, conservando solo los eventos de mayor amplitud como las primeras reflexiones y las múltiples, este comportamiento de las amplitudes debe ser corregido, pues al no tener una correcta representación *sparse* de la señal no se cumple con uno de los principios fundamentales para aplicar el *CS*.

Para balancear las amplitudes se puede aplicar una corrección como se realiza en el procesamiento de datos sísmicos siguiendo modelos de atenuación, pero las correcciones con modelos simplificados no balancean por completo las amplitudes y los modelos de atenuación más complejos requieren de modelos de velocidad que no se tienen en el momento de la adquisición. Por estas razones se propone aplicar una ganancia controlada automáticamente, *AGC* por sus siglas en inglés, para balancear las amplitudes a lo largo de cada traza y una normalización por traza que equilibre las amplitudes en los diferentes *offset* del *gather*, en el anexo 4 se explica su funcionamiento. En la Figura 5(c), se muestra un *gather* con las amplitudes normalizadas y completamente balanceadas, utilizando el método propuesto anteriormente.

Se debe tener cuidado al aplicar el *AGC* para favorecer la representación *sparse* y reconstrucción de datos sísmicos, debido a que este balance de amplitudes no corresponde a ningún comportamiento físico de la propagación de onda, siendo únicamente una herramienta de procesamiento que realiza eventos, por lo tanto, una vez terminada la reconstrucción se debe aplicar un operador inverso para regresar la amplitud de los datos a su estado inicial. El operador correspondiente a las trazas eliminadas se obtiene aplicando una interpolación entre los operadores de las trazas adyacentes.

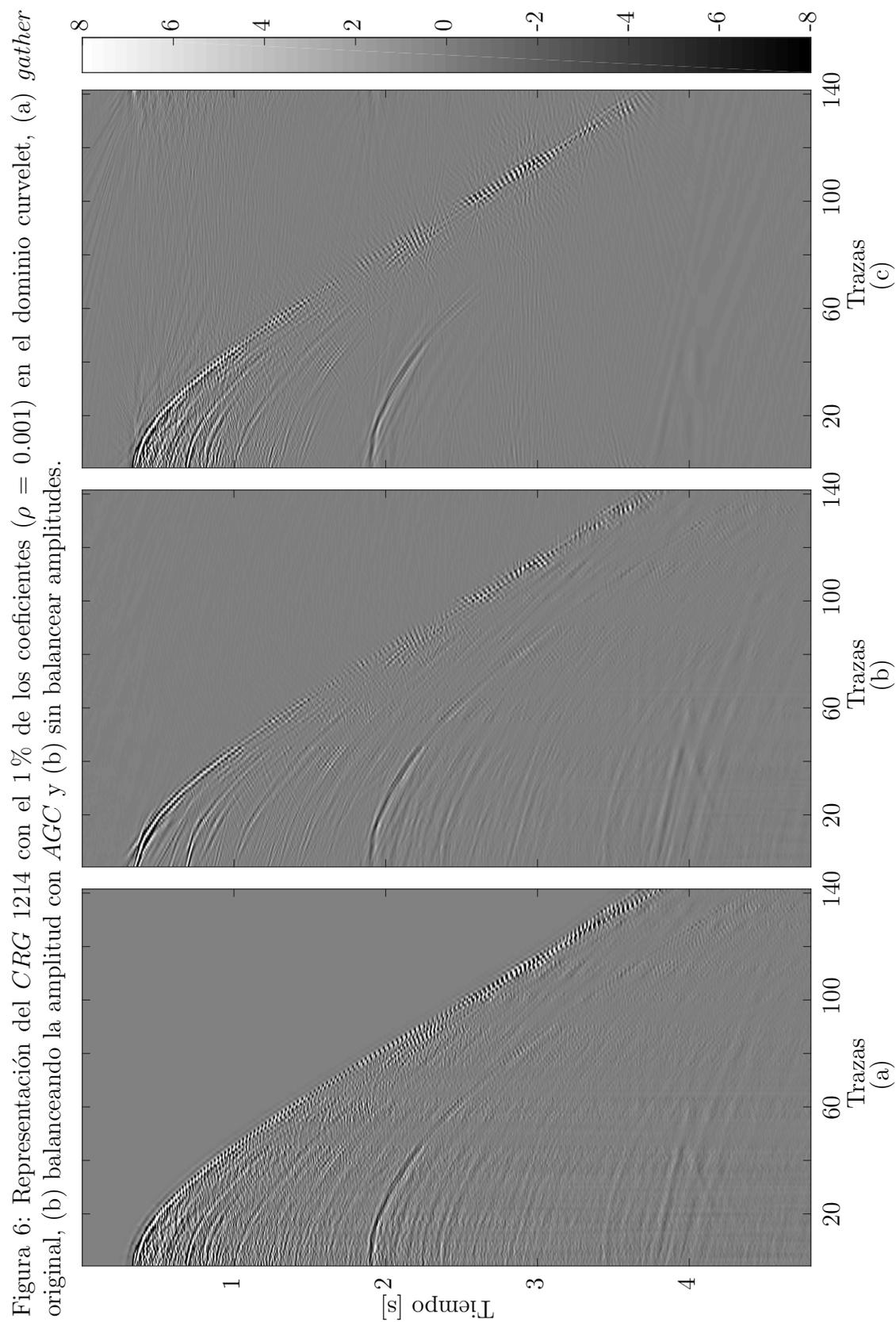
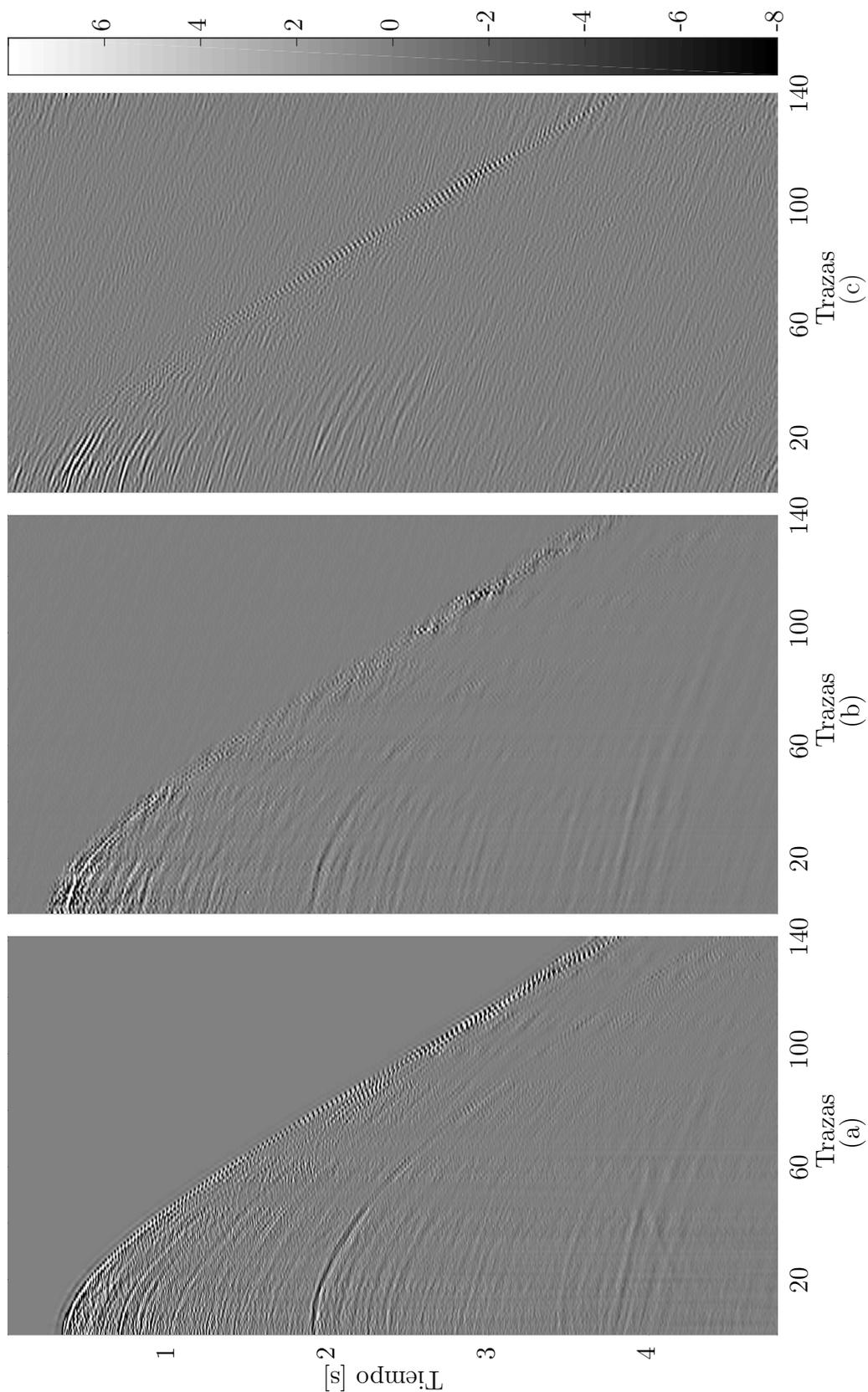


Figura 7: Representación del *CRG* número 1214 con el 1% de los coeficientes ($\rho = 0.001$) en el dominio de Fourier, (a) *gather* original, (b) balanceando la amplitud con *AGC* y (c) sin balancear amplitudes.



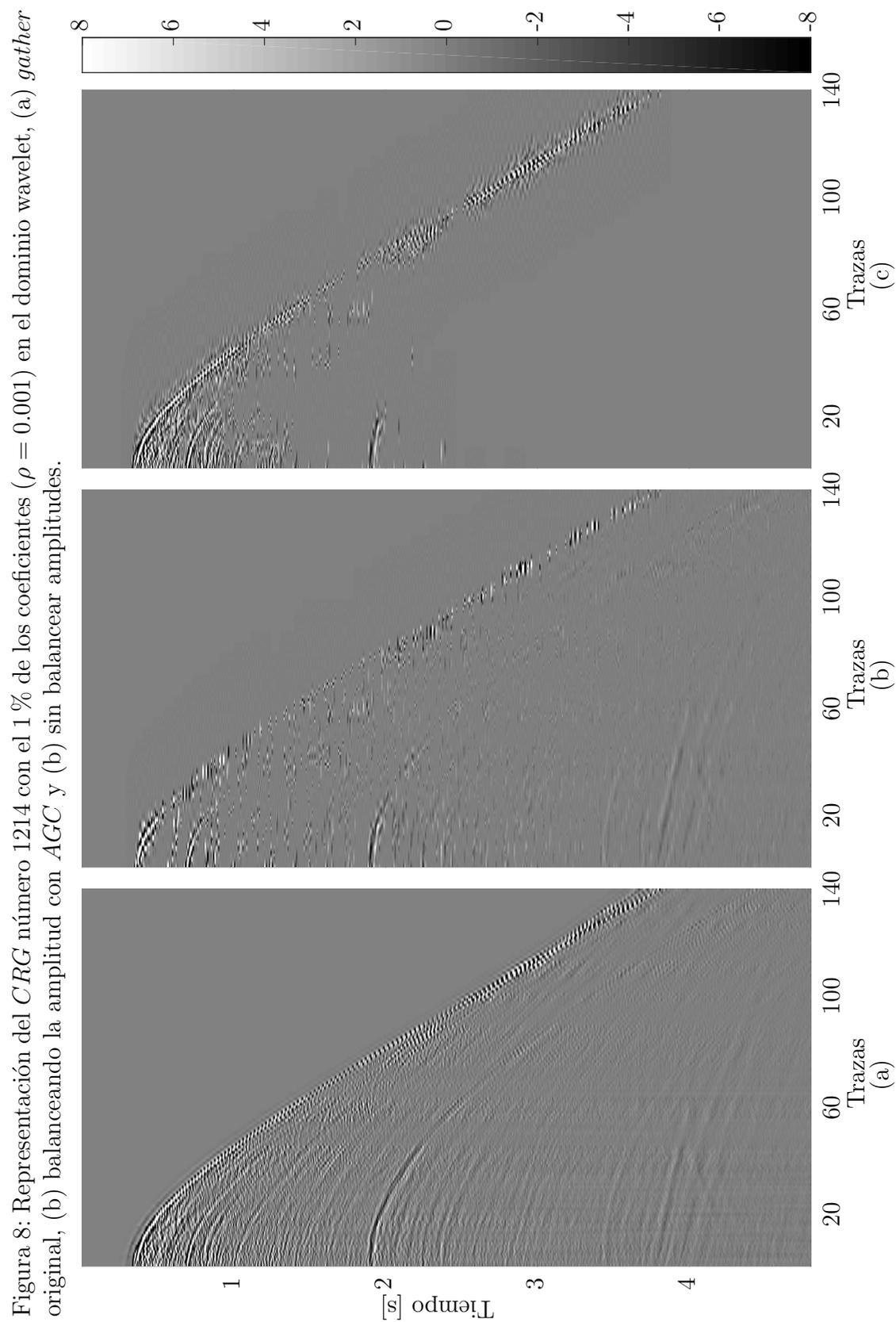


Figura 8: Representación del *CRG* número 1214 con el 1% de los coeficientes ($\rho = 0.001$) en el dominio wavelet, (a) *gather* original, (b) balanceado la amplitud con *AGC* y (c) sin balancear amplitudes.

3.3. MEDICIÓN DEL ERROR DE APROXIMACIÓN

La relación señal a ruido (SNR por sus siglas en inglés) es una métrica utilizada para cuantificar que tan buena es la reconstrucción de una señal tras realizarse algún tipo de procesamiento, considerando que la señal inicial es libre de ruido y que el ruido inducido solo se debe al procesamiento. De esta forma se puede utilizar el SNR para cuantificar que tan *sparse* es la representación de una señal (*sparcity*) [2], en función del factor de (*sparcity*) $\rho = k/p$, donde k es el número de coeficientes diferentes de cero que se conservan y p el número de coeficientes total de la representación. Se define como

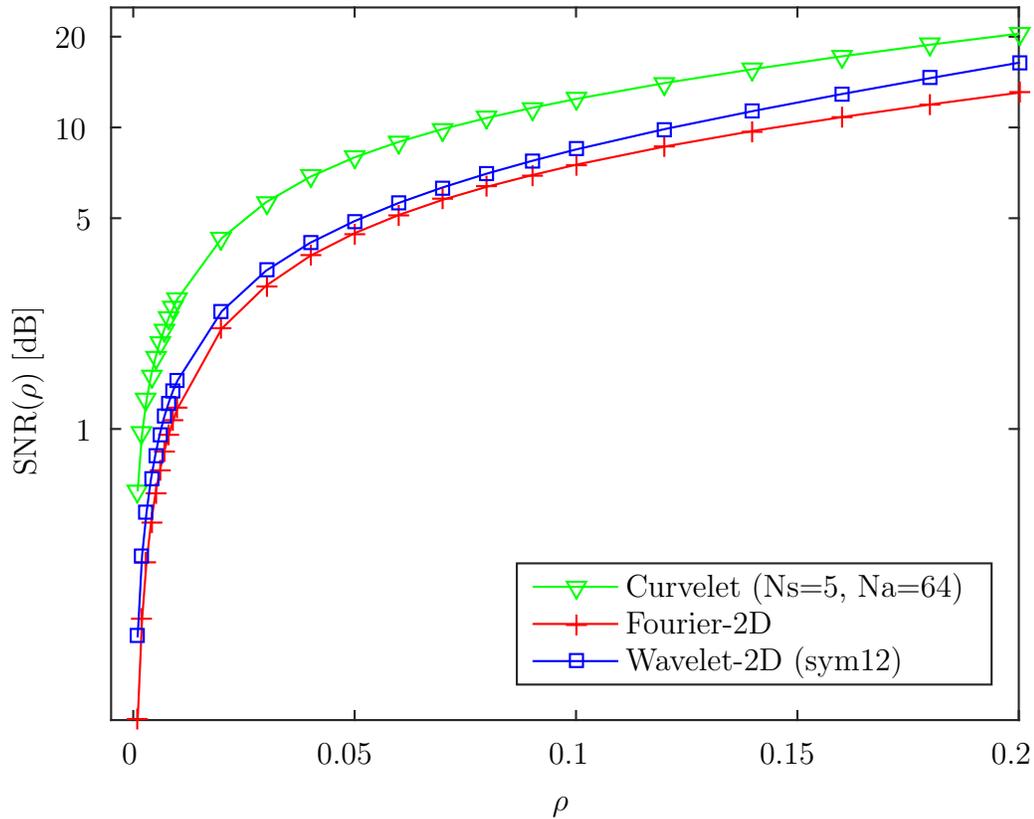
$$SNR(\rho) = -20 \log \frac{\|f - f_\rho\|}{\|f\|}, \quad (3.1)$$

donde f es la señal sísmica inicial y f_ρ es la señal a la cual se le forzó una *sparcity* ρ .

Para establecer cual diccionario de representación es el más adecuado se realizó una serie de mediciones del nivel *sparse* de los *CRG* del subconjunto de datos de prueba, con el experimento planteado al inicio de esta sección, variando ρ de 0.002 a 0.2, los resultados se muestran en la Figura 9. Cuanto menor sea el valor de ρ , más coeficientes son ignorados forzando que la representación sea más *sparse*, en consecuencia, el valor de SNR tiende a ser menor, debido a que se está eliminando la información que no se alcanza a representar con los coeficientes de mayor magnitud.

Analizando los resultados en la Figura 9 se observa que la curva de la transformada *curvelet* alcanza un mayor SNR en todos los valores de ρ establecidos, en comparación a las demás transformadas, lo cual significa que es el diccionario con la mejor representación *sparse* de los datos sísmicos, concentrando la mayor cantidad de información en pocos coeficientes. Comparando visualmente los *gathers* obtenidos para el mismo ρ , se aprecia que la transformada *curvelet* conserva mejor la continuidad de los eventos, sin destruirlos como hace la transformada wavelet y Fourier. En las Figura 6(b),7(b) y 8(b) se muestra un ejemplo, para un ρ en el codo de las curvas de SNR, en el que se observa fácilmente como la transformada wavelet elimina las hipérbolas del *gather* dejando solamente pequeñas manchas lineales que siguen la forma de la primera reflexión (fondo marino) y la múltiple. La transformada de Fourier conserva la tendencia de los eventos, pero pierde por completo la resolución, difuminando los eventos de tal forma que parecen mezclarse y replicarse en toda la imagen.

La transformada *curvelet* a diferencia de las demás no causa cambios tan fuertes en el *gather* conservando la forma y continuidad de los eventos mejor que las demás. Sin embargo, sí se pierde información, pero no se concentra en un rango de frecuencias o en una zona específica como las demás transformadas, eliminando los datos que no coinciden con un comportamiento hiperbólico o que tiene poca coherencia con las elipses

Figura 9: Nivel *sparse* de los datos sísmicos agrupados por receptor común.

de representación que utiliza la transformada, lo cual se puede interpretar como una supresión de ruido donde se eliminan los datos que no aportan en la construcción de las hipérbolas de los eventos sísmicos. En la Figura 6(b) se observa fácilmente como sobresalen más algunos eventos en comparación al *gather* original en la Figura 6(a). Como efecto adverso de esta transformada aparecen una serie de líneas con pendiente opuesta a las hipérbolas de reflexión, que son creadas por la transformada curvelet al encontrar coherencia entre puntos de diferentes eventos, las zonas más susceptibles a la aparición de estos artefactos son donde las hipérbolas tienen mayor pendiente y aparecen problemas de *aliasing* espacial en la visualización del evento.

Se seleccionó la transformada curvelet como base de representación *sparse* para realizar los experimentos de reconstrucción de datos sísmicos, presentando el mejor rendimiento en comparación a las de más transformadas analizadas.

3.4. MATRIZ DE MUESTREO

La distribución de las fuentes que se eliminan al pasar de una adquisición de datos sísmicos convencional a una aleatoria se representa por medio de una matriz de muestreo aleatorio A , la cual indica con unos y ceros que fuentes se conservan y cuales se eliminan, esta matriz se utiliza para la simulación de las adquisiciones aleatorias y para la posterior reconstrucción de los datos.

El diseño de la matriz de muestreo se hace teniendo en cuenta las consideraciones expuestas en la sección 2.4, para favorecer la calidad de las trazas reconstruidas. Se inició por establecer que los CRG a ser reconstruidos deben tener un mínimo de 30 trazas para que contengan la información necesaria en la reconstrucción, este valor se sugiere para el dato sísmico trabajado y en otros casos puede llegar a variar. Con ayuda de la información de la posición externa de las fuentes y receptores se obtiene el rango de las posibles fuentes a eliminar. Se fija en dos el número máximo de fuentes adyacentes que se pueden eliminar cuando se aplica el muestreo compresivo para reducir hasta un 30% el número de fuentes de la adquisición, para evitar que se creen huecos de mas de dos trazas en los CRG que son difíciles de reconstruir y alteran la calidad del *gather*.

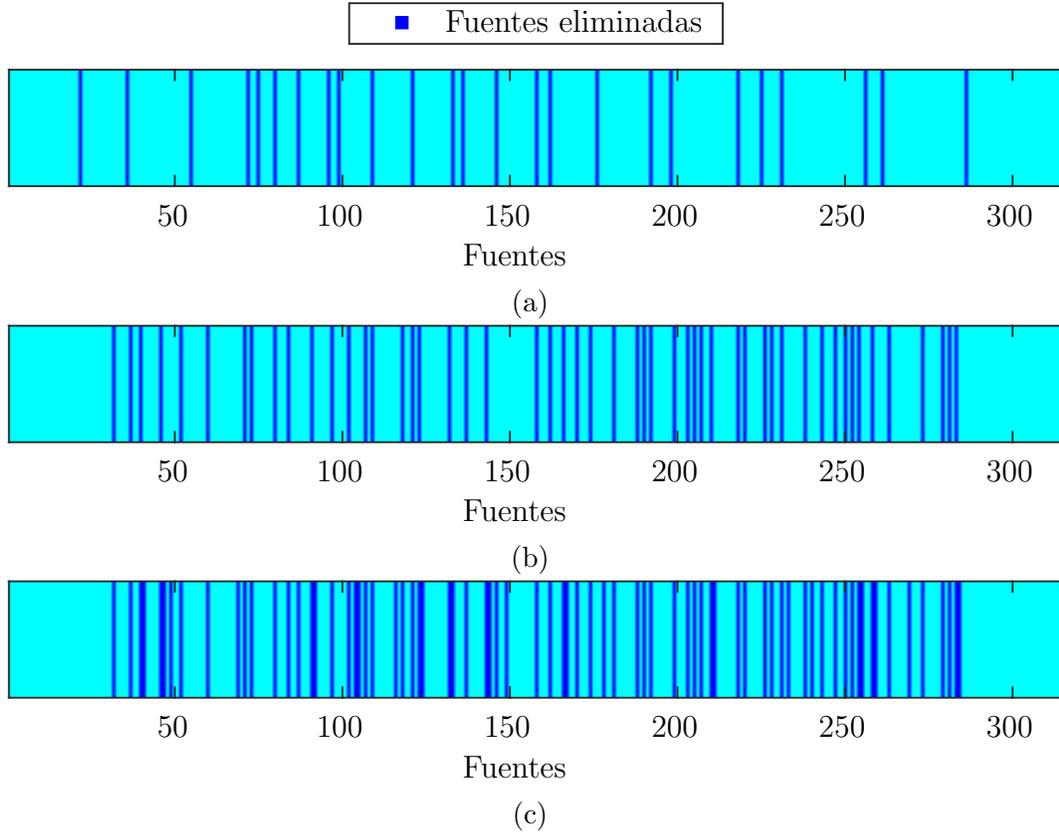
La posición de las fuentes a eliminar se obtiene por medio de un generador de números aleatorios con distribución uniforme, el cual esta condicionando al rango de las posibles fuentes a eliminar y al máximo de fuentes adyacentes que se admite eliminar sin causar huecos demasiado grandes en la información. Como resultado se obtienen las matrices de muestreo aleatorio de la Figura 10 para los factores de muestreo aleatorio de $sr = 0.1$, $sr = 0.2$ y $sr = 0.3$, donde

$$sr = \frac{\# \text{ de fuentes aleatoria}}{\# \text{ de fuentes uniformes}} \quad (3.2)$$

Debido a que la reconstrucción de los datos se realiza agrupando las trazas por receptor común, se requiere transformar la matriz de muestreo aleatorio en un conjunto de submatrices A_k que representen el patrón de adquisición en los CRG , esto se debe a que la adquisición y el muestreo aleatorio se realiza en el dominio de la fuente común, y la reconstrucción en el dominio del receptor común, por lo tanto se requiere tantas submatrices de muestreo A_k , como *gather*s se vayan a reconstruir, teniendo en cuenta que cada matriz A_k se deriva de la matriz de muestreo aleatorio A diseñada previamente.

Para realizar esta transformación se hace un amarre entre la matriz A y la posición de la fuente con que se generó cada traza del CRG , al agrupar todas las matrices A_k en un cubo y visualizar la cara superior se resume el comportamiento de todas las submatrices de muestreo para el dato sísmico, como se muestra en la Figura 11. Con esta información se puede analizar que las submatrices A_k están fuertemente correlacionadas y que el porcentaje de fuentes que se eliminan en una adquisición aleatoria no coincide con el porcentaje de trazas faltantes en cada CRG , por el contrario, este valor se vuelve

Figura 10: Matriz de muestreo aleatorio para el dato real, eliminado el (a)10%, (b)20% y (c)30% de las fuentes. Las barras representan las fuentes que se conservan y los huecos las fuentes que se omiten en la adquisición aleatoria.

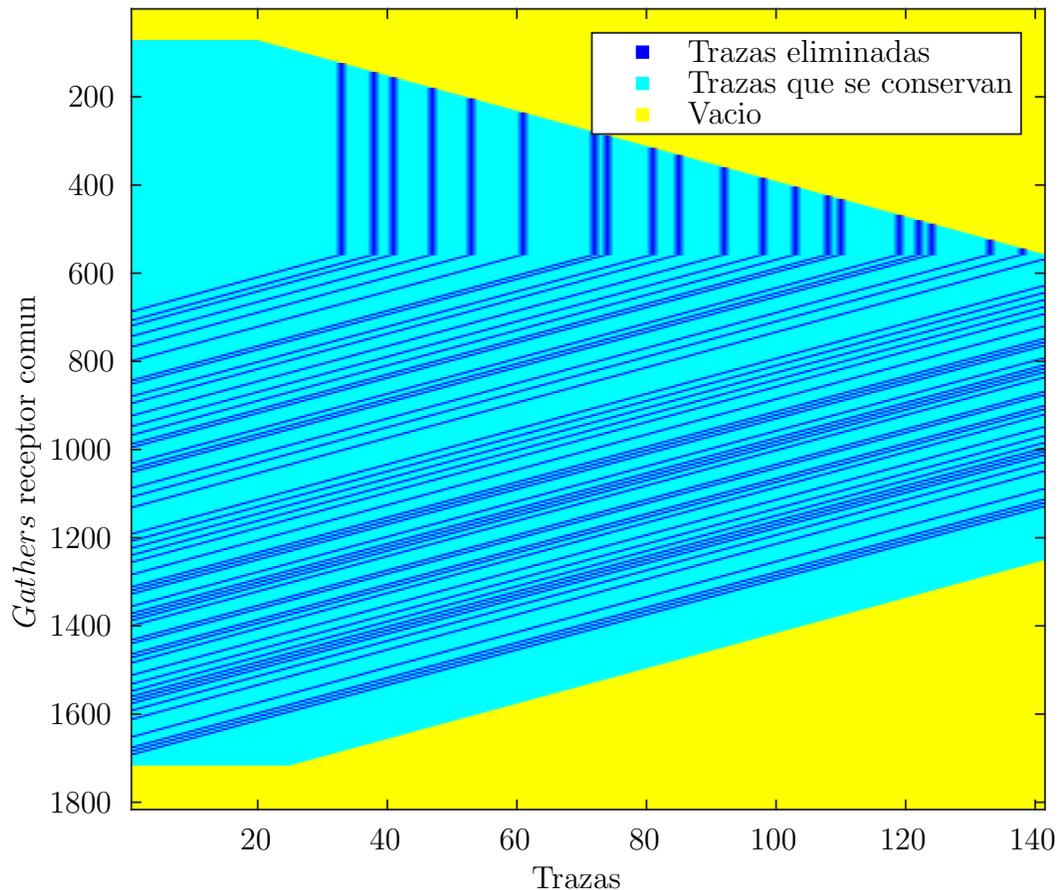


variable para cada *gather*, como se muestra en la Figura 12, siguiendo como tendencia que los *gathers* con menor número de trazas son los que tienen el menor porcentaje de huecos y los *gather* con mayor número de trazas presentan el mayor porcentaje de trazas faltantes.

3.5. SELECCIÓN DEL PARÁMETRO DE RECONSTRUCCIÓN λ .

El parámetro λ cumple una función de balance en el problema de optimización de la ecuación (1.3), controlando que tan *sparse* es la solución y cuanto se ajusta a los datos observados. Como este balance se realiza escalando la norma l_1 de x , el cual es un vector de coeficientes curvelet que representan el dato sísmico reconstruido, el valor que debe tomar λ para obtener la mejor solución va a depender de la amplitud de los datos sísmicos, el porcentaje de trazas que se hayan eliminado, la cantidad de eventos y calidad del dato sísmico.

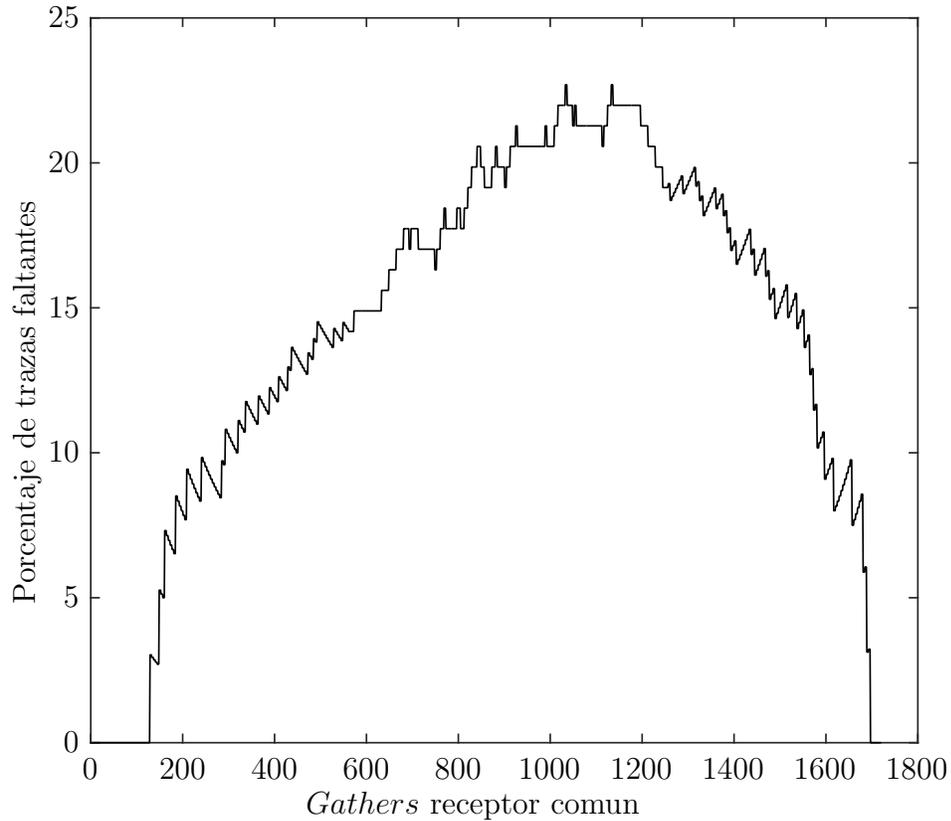
Figura 11: Vista superior del cubo de submatrices A_k , derivadas de la matriz de muestreo aleatorio de la Figura 10 (b).



Como la magnitud de los datos sísmicos se balancea y normaliza previamente a la reconstrucción, no se espera ningún efecto si se tienen diferencias de amplitud en las trazas o disparos del dato inicial. La cantidad de eventos, calidad del dato y porcentaje de trazas eliminadas, son factores que no se pueden generalizar, por tanto se debe seleccionar un nuevo parámetro λ para cada dato sísmico aleatorio, en la sección anterior se explicó que el porcentaje de trazas eliminadas en los *CRG* de una misma adquisición aleatoria no se mantiene constante, por lo tanto podrían existir diferentes valores de λ adecuados para una misma adquisición, sin embargo no es viable determinar un λ para cada *CRG*, optando por utilizar un subconjunto de datos de prueba para establecer un λ promedio que se adapte a la reconstrucción de todos los *CRG* de una determinada adquisición aleatoria sin comprometer en gran medida la calidad de los mismos, es decir que se busca obtener un valor intermedio que se adapte a la reconstrucción de los *gathers* con diferente porcentaje de trazas eliminadas.

Como metodología para la selección de λ se propuso iniciar por calcular el valor teórico para todos los *gathers* obtenidos tras aplicar las matrices de muestreo aleatorio diseñadas en la sección 3.4, y promediar estos valores para conocer el rango en el que

Figura 12: Porcentaje de trazas eliminadas en cada *CRG*, al aplicar las sub matrices de muestreo de la Figura 11.



se encuentra el valor más adecuado. En la Tabla 1 se presentan los valores teóricos de λ obtenidos para cada adquisición aleatoria. Seguido a esto, se realiza una sintonización empírica entre los valores cercanos al teórico para obtener un valor experimental que será utilizado para reconstruir el dato.

Tabla 1: Valor promedio de λ teórico para los diferentes porcentajes de muestreo aleatorio. Utilizando la transformada curvelet como base de representación *sparse*.

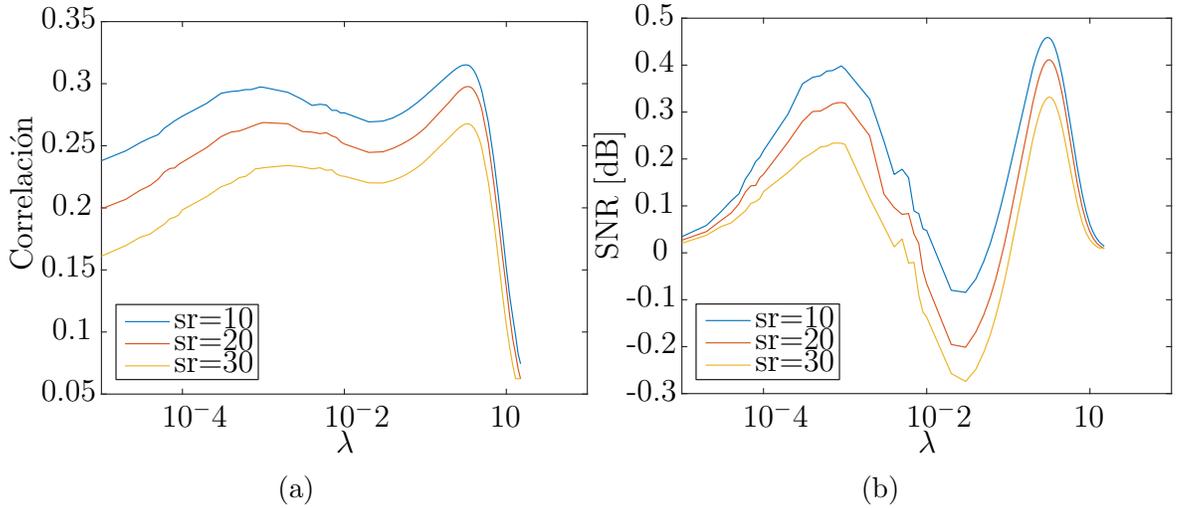
λ teórico		
SR=10 %	SR=20 %	SR=30 %
0.4511	0.4091	0.3835

La sintonización empírica consiste en reconstruir el subconjunto de datos de prueba variando los valores de λ desde 1×10^5 hasta 1.5. Para seleccionar el valor más adecuado se realiza una comparación visual entre los *CRG* reconstruidos y se calculan los valores de correlación y SNR de las trazas reconstruidas con respecto a las originales,

Tabla 2: Valores de λ experimental, con los cuales se obtiene el mejor promedio de correlación y SNR.

	λ experimental		
	SR=10 %	SR=20 %	SR=30 %
Correlación	0.32	0.33	0.34
SNR	0.3	0.32	0.32

Figura 13: Curvas de (a) correlación y (b) SNR promedio de las trazas sísmicas de prueba reconstruidas con diferentes valores de λ .



en un escenario no académico donde solo se tiene los datos aleatorios se pueden eliminar algunas trazas extra y utilizarlas como medida de control, para complementar la comparación visual con un factor numérico. El promedio de los resultados obtenidos para la correlación y SNR se muestra en las Figuras 13(a) y 13(b) respectivamente y en la Tabla 2 se presenta un resumen de λ donde se obtuvieron los valores máximos.

Al contrastar visualmente los resultados de las diferentes reconstrucciones obtenidas se evidencia que para los valores de λ presentados en la Tabla 2 se tiene una reconstrucción que favorece la continuidad hiperbólica de los eventos, en especial para las bajas frecuencias y tiende a omitir los lóbulos de la traza para los cuales no encuentra ninguna coherencia con las trazas vecinas, lo cual es un efecto de supresión de ruido inherente a la promoción del nivel *sparse*. Para valores menores de λ se observa que las trazas reconstruidas no se ajustan muy bien a la traza original y para valores mayores se destruye la mayor parte de la información, lo cual concuerda con la teoría. En la Figura 14 se presenta un *CRG* reconstruido con tres valores diferentes de λ para comparar como varía el resultado obtenido en función de este parámetro.

Tomando en cuenta los resultados teóricos, experimentales y la comparación visual, se seleccionaron los valores de λ correspondientes a los máximos de correlación como los más indicados para la reconstrucción de los datos, puesto que el SNR no resultó ser una métrica muy diciente en cuestión de calidad del *gather* sísmico reconstruido, además que sus valores son muy bajos siendo poco confiables, y la correlación aporta directamente información sobre la fase y amplitud de los eventos sísmicos considerándolo el parámetro más importante a tener en cuenta en la reconstrucción. Estos valores de λ serán los utilizados para realizar la reconstrucción del dato sísmico completo con cada uno de los factores de muestreo aleatorio.

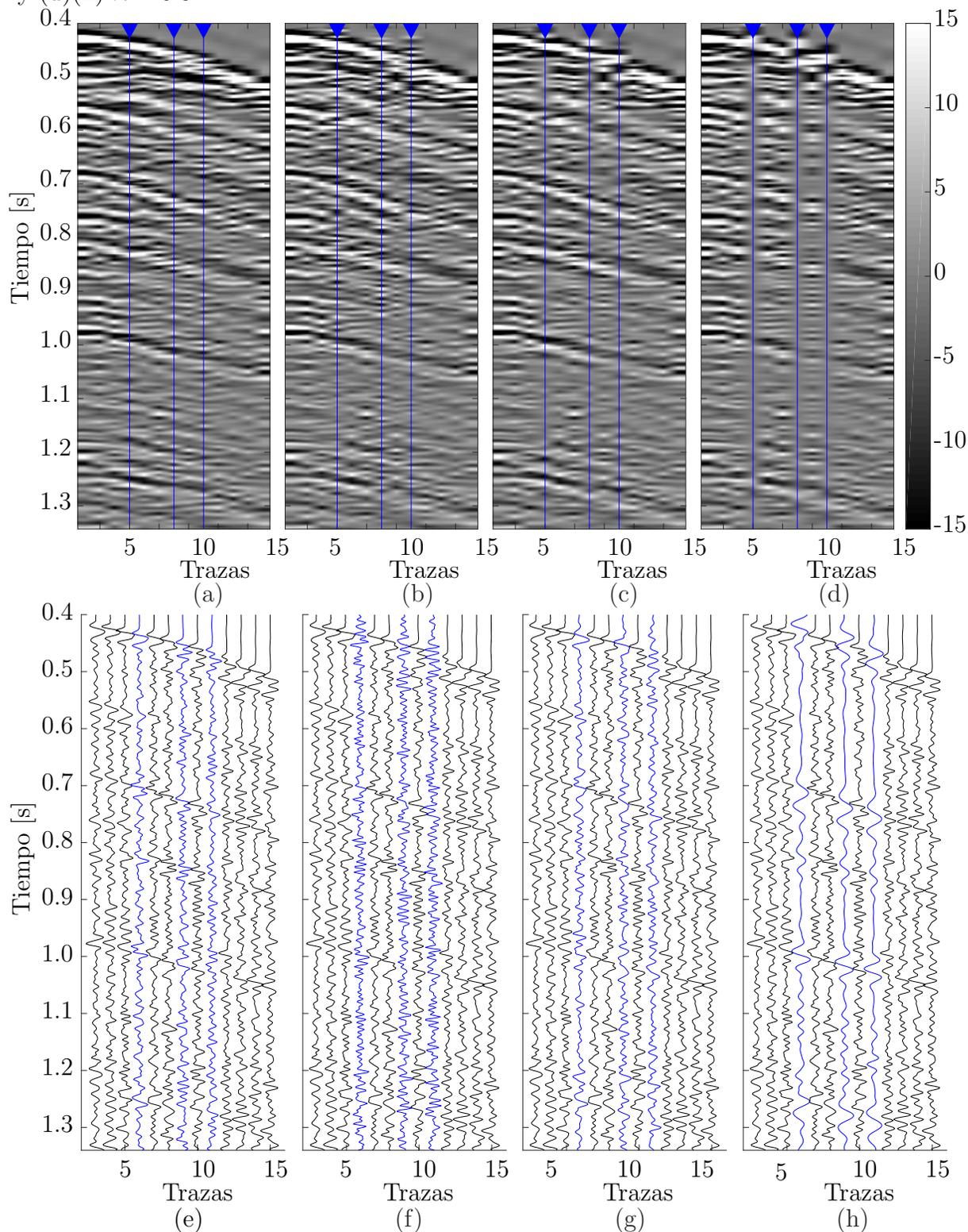
3.6. VALIDACIÓN DEL DISEÑO DE ADQUISICIÓN USANDO *CS*

La adquisición aleatoria de datos sísmicos tiene implícita una etapa de reconstrucción de datos, donde se completan o llenan los huecos en la información causados por la aleatoriedad en la ubicación de las fuentes. En esta etapa, se plantearon una serie de pruebas para reconstruir el dato sísmico completo, tras simular la adquisición aleatoria con las matrices de muestreo diseñadas en la sección 3.4. Tras reconstruir el dato sísmico completo, se reordenaron las trazas por fuente común, para visualizar los disparos que fueron reconstruidos y realizar las respectivas comparaciones con los disparos de la adquisición uniforme.

Se reconstruyó el dato sísmico real en dos escenarios diferentes, el primero es considerando todo el ancho de banda y el segundo con un ancho de banda limitado a $15Hz$. Esta reconstrucción se propone para datos que van a ser utilizados en *FWI* donde solo se utiliza la información de baja frecuencia. Como prueba complementaria para evaluar las prestaciones de la adquisición aleatoria se planteó realizar una *FWI* con datos sintéticos adquiridos de forma aleatoria.

El proceso de reconstrucción consta principalmente de dos etapas, en la primera etapa se acondiciona el dato muestreado aleatoriamente, balanceando las amplitudes de las trazas observadas y se guarda el operador de balance de cada traza para aplicarlo de forma inversa al terminar la reconstrucción, para obtener el operador de balance de las trazas que van a ser reconstruidas se aplica una interpolación lineal entre sus trazas adyacentes, obteniendo el operador de forma aproximada. En la segunda etapa se realiza la reconstrucción con el algoritmo GPRS, el cual tiene como entrada: un *CRG* muestreado aleatoriamente, λ , la matriz de muestreo y el diccionario de representación *sparse*, al terminar este proceso se retira el balance de amplitudes para regresar el dato sísmico a su rango dinámico inicial.

Figura 14: Sección del número *CRG* 1214 reconstruido con diferentes valores λ para comparar visualmente los efectos de variar este parámetro, las línea azules representan las trazas que fueron reconstruidas. (a)(e) *CRG* original, (b)(f) $\lambda = 0.003$, (c)(g) $\lambda = 0.3$ y (d)(h) $\lambda = 0.8$



3.7. RECONSTRUCCIÓN DE DATOS SÍSMICOS REALES

Se realizó la reconstrucción del dato sísmico descrito en la sección 3.2, tras simular las adquisiciones aleatorias con las tres matrices de muestreo diseñadas en la sección 3.4. La reconstrucción se realiza con el dato agrupado en *CRG* y al terminar el proceso, se reordenan las trazas por *CSG* para visualizar los disparos completos que fueron reconstruidos. En las Figuras 15, 16 y 17 se muestra un *CSG* que fue eliminado en las tres matrices de muestreo aleatorio para comparar el resultado obtenido en sus respectivas reconstrucciones, cuando se elimina un porcentaje de fuentes diferente.

En las imágenes se observa que las reconstrucciones fueron exitosas obteniendo *gather*s similares al original, en cuanto a la tendencia y forma de los reflectores sísmicos, los cuales son reconstruidos en su mayoría. Al analizar detalladamente las imágenes se observan diferencias, la principal es que se eliminan eventos de alta frecuencia lo cual se ve reflejado en la imagen de la resta punto a punto donde aparecen líneas muy finas, además se eliminan valores que no tenían coherencia con las hipérbolas de los eventos sísmicos, lo cual se puede interpretar como un *denoise* que ayuda a resaltar los eventos sísmicos en ciertas zonas del *gather*, también se da el caso en el que el lóbulo de una traza no coincide con el de las trazas adyacentes que están representando un evento sísmico, y dicho lóbulo es modificado en la reconstrucción para que tenga coherencia con las trazas adyacentes en la representación del evento. Estos casos deben ser analizados con especial cuidado cuando se da una valoración cuantitativa de la calidad de los datos reconstruidos ya que estas diferencias pueden ser beneficiosas o adversas y las métricas como el SNR no las pueden diferenciar, por lo tanto se debe complementar con una valoración cualitativa de un experto que de una visión más general al respecto de la calidad de dicho *gather* reconstruido. Esta aclaración aplica para el caso académico donde se tiene el dato original como referencia.

Visualmente es difícil diferenciar entre los *CSG* reconstruidos a partir de las tres adquisiciones aleatorias, pero alternando las imágenes se logran observar ligeros cambios en la continuidad de los reflectores indicando que la calidad del *gather* reconstruido es inversamente proporcional a la cantidad de fuentes eliminadas, efecto que se observa especialmente en las trazas de los bordes.

Adicionalmente se calculó el espectro en frecuencia de los *gather*s sísmico reconstruidos, para compararlo con el espectro del *gather* original y conocer en qué medida se está conservando el ancho de banda del dato, en la Figura 19 se presentan los resultados evidenciando que para las bajas frecuencias se recuperó en buena medida el ancho de banda de los datos, y para frecuencias más altas, por encima de los 35Hz , se empieza a perder amplitud. Este resultado corrobora las conclusiones cualitativas expresadas anteriormente al respecto de la calidad de los datos sísmicos reconstruidos.

Durante el proceso de reconstrucción se observó un efecto adverso en algunas zonas

de los *CRG*, donde la reconstrucción encuentra coherencia entre lóbulos de las trazas para ángulos en los que no se espera tener eventos sísmicos, las zonas donde más se evidencian es alrededor de los primeros arribos, como se observa en la Figura 18, lo cual es un efecto adverso de la reconstrucción que no se observa en los *CSG*, pero si están afectando la calidad del dato reconstruido. Estos efectos pueden ser en parte debido a problemas de *aliasing* en los *CRG* y se podría solucionar restringiendo los ángulos de la transformada curvelet de manera que solo represente el *gather* sísmico con hipérbolas en ángulos coherentes con los eventos sísmicos.

Como figura de mérito complementaria para la reconstrucción de los datos sísmicos se realizó un apilado bruto de los datos reconstruidos, para comparar como se afecta este producto final cuando se trabaja con las diferentes adquisiciones aleatorias planteadas y con sus respectivos datos reconstruidos, el apilado se realizó con los mismos parámetros para todos los datos y los resultados se muestran en las Figuras 20 y 21. Debido a que el apilado de los datos es un proceso que compacta horizontalmente la energía contenida en los *gather* punto medio común, se conserva la tendencia de la mayoría de las fuentes y no se aprecia fácilmente la diferencia entre los apilados, pero al comparar las Figuras 20(c) y 21(c), con la resta se observan cambios significativos, los cuales indican que se pierde menos información cuando se hace el apilado con los datos reconstruidos en comparación al apilado de los datos aleatorios.

Para facilitar la comparación visual se señaló con elipses las zonas donde se evidencian los mayores cambios. En la elipse de tiempo $4s$ se observa como el apilado reconstruido no resalta ningún evento, mientras que en el apilado aleatorio si se observan eventos residuales, lo cual significa que se perdió parte de la información que representaba dichos eventos. En las elipses de tiempo $0.8s$ y $1.8s$ se observa que en ambos apilados se perdió información, pero en el del dato reconstruidos el grosor de los reflectores residuales es significativamente menor en comparación a los del apilado del dato aleatorio, dando a conocer que en el apilado reconstruido se está perdiendo información de alta frecuencia y en el apilado de los datos aleatorios se pierde toda la información, puesto que no se cuenta con un porcentaje de las fuentes.

Figura 15: *Common source gather* número 159 reconstruido tras eliminar un 10% ($sr=0.1$) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

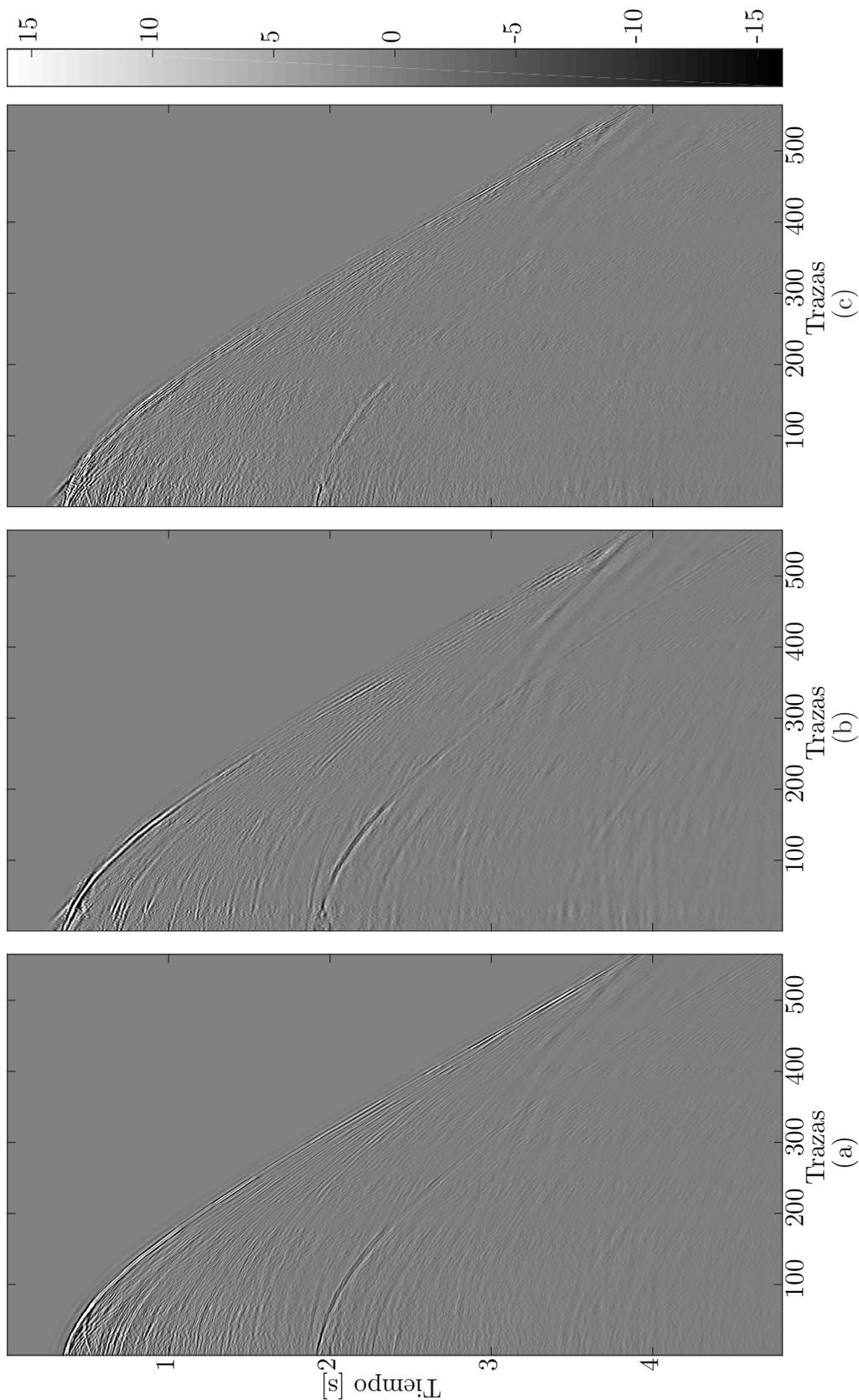


Figura 16: *Common source gather* número 159 reconstruido tras eliminar un 20% ($sr=0.2$) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

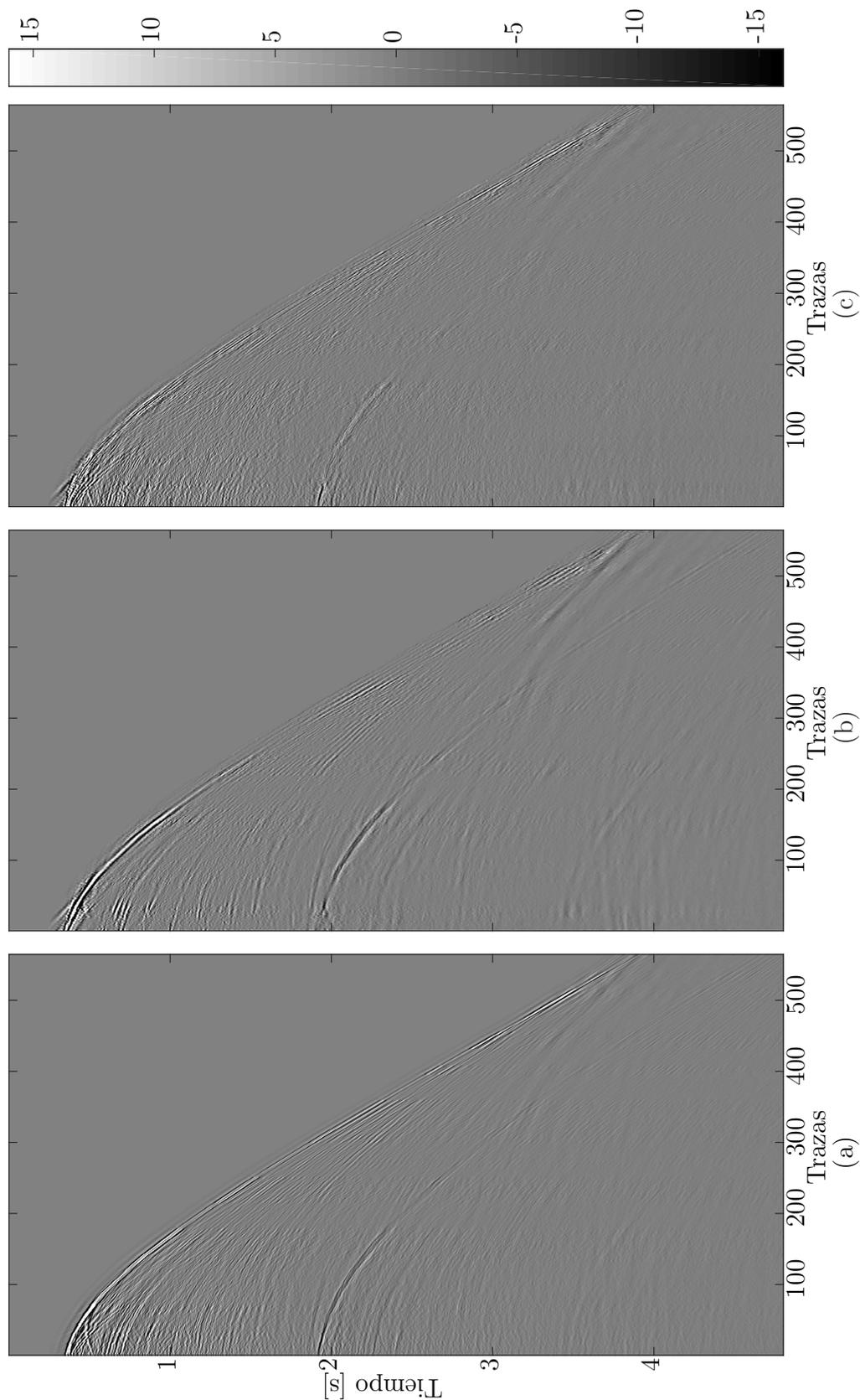


Figura 17: *Common source gather* número 159 reconstruido tras eliminar un 30% ($sr=0.3$) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

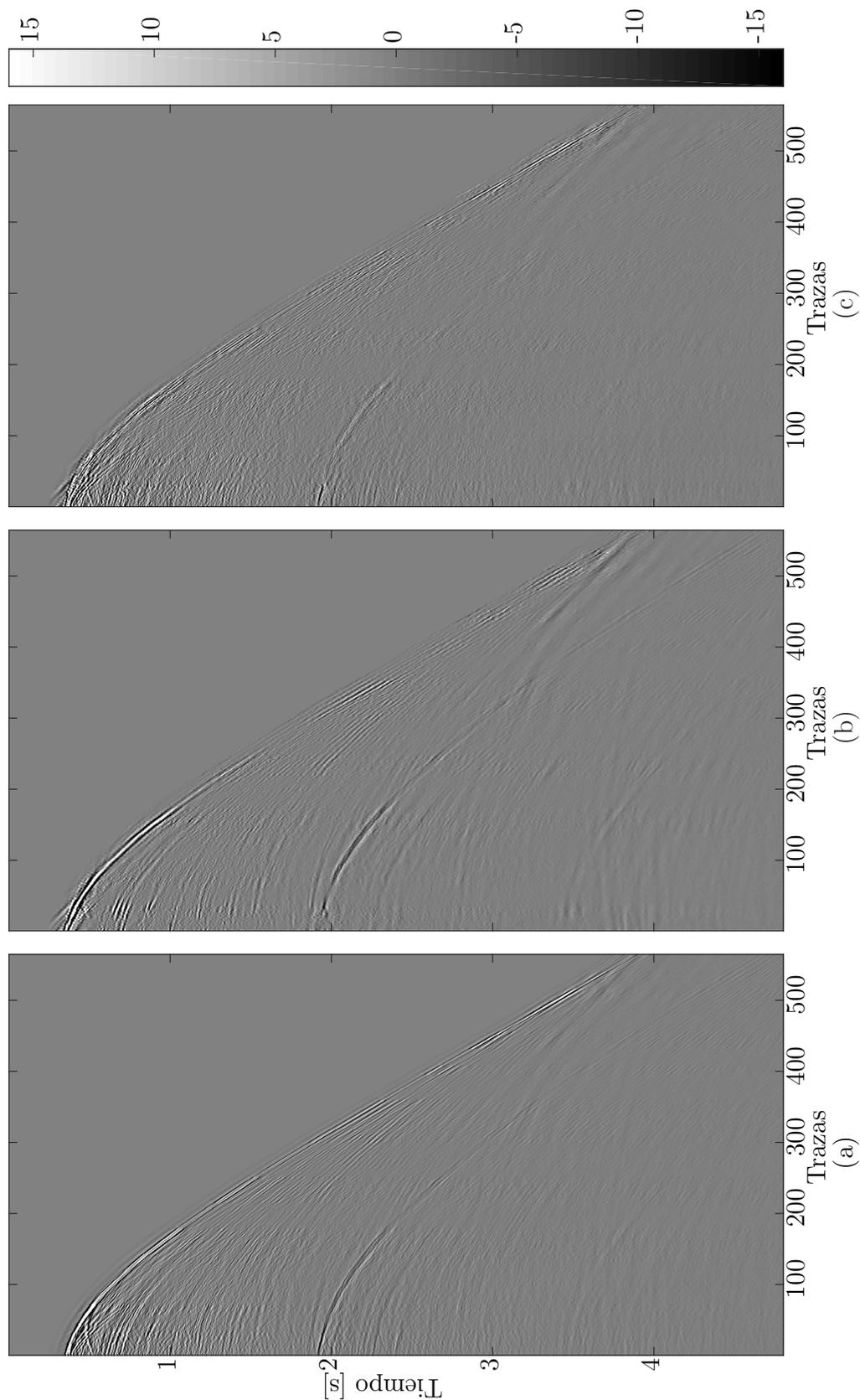


Figura 18: *Common receiver gathers* reconstruidos donde se evidencia la aparición de eventos anómalos en ángulos opuestos a los eventos sísmicos. (a) *Gather* original, *gather* reconstruido tras eliminar (b) 9%, (c) 21% y (d) 29% de las trazas.

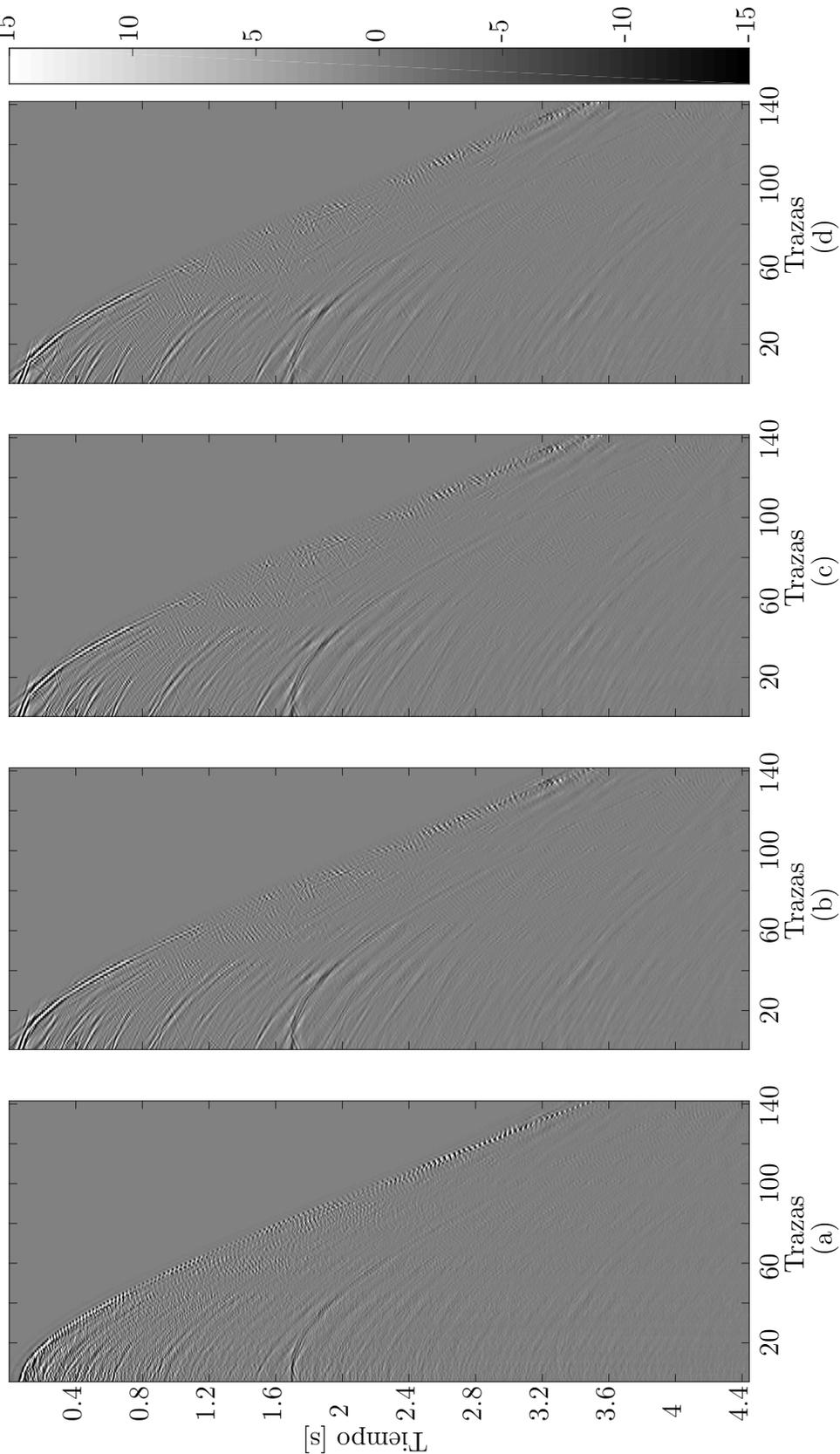


Figura 19: Espectro en frecuencia del *CSG* número 159 original y los *CSG* reconstruidos, evidenciando que las reconstrucciones recuperan las bajas frecuencias, sin embargo, las altas se están perdiendo.

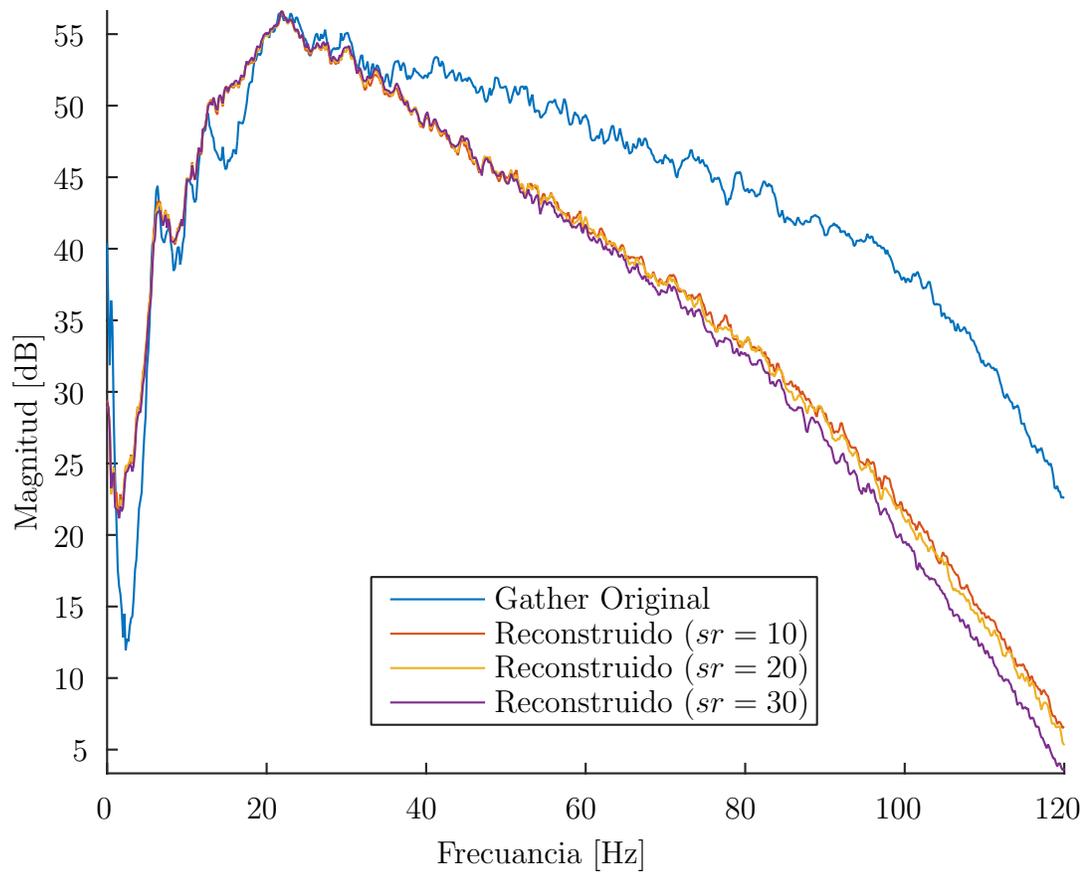


Figura 20: Sección sísmica apilada utilizando los datos (a) originales y (b)reconstruidos a partir de la adquisición aleatoria con $sr = 0.3$. (c) Diferencia entre la imagen (a) y (b), resaltando las zonas donde se tienen más cambios, con respecto a la Figura 21.

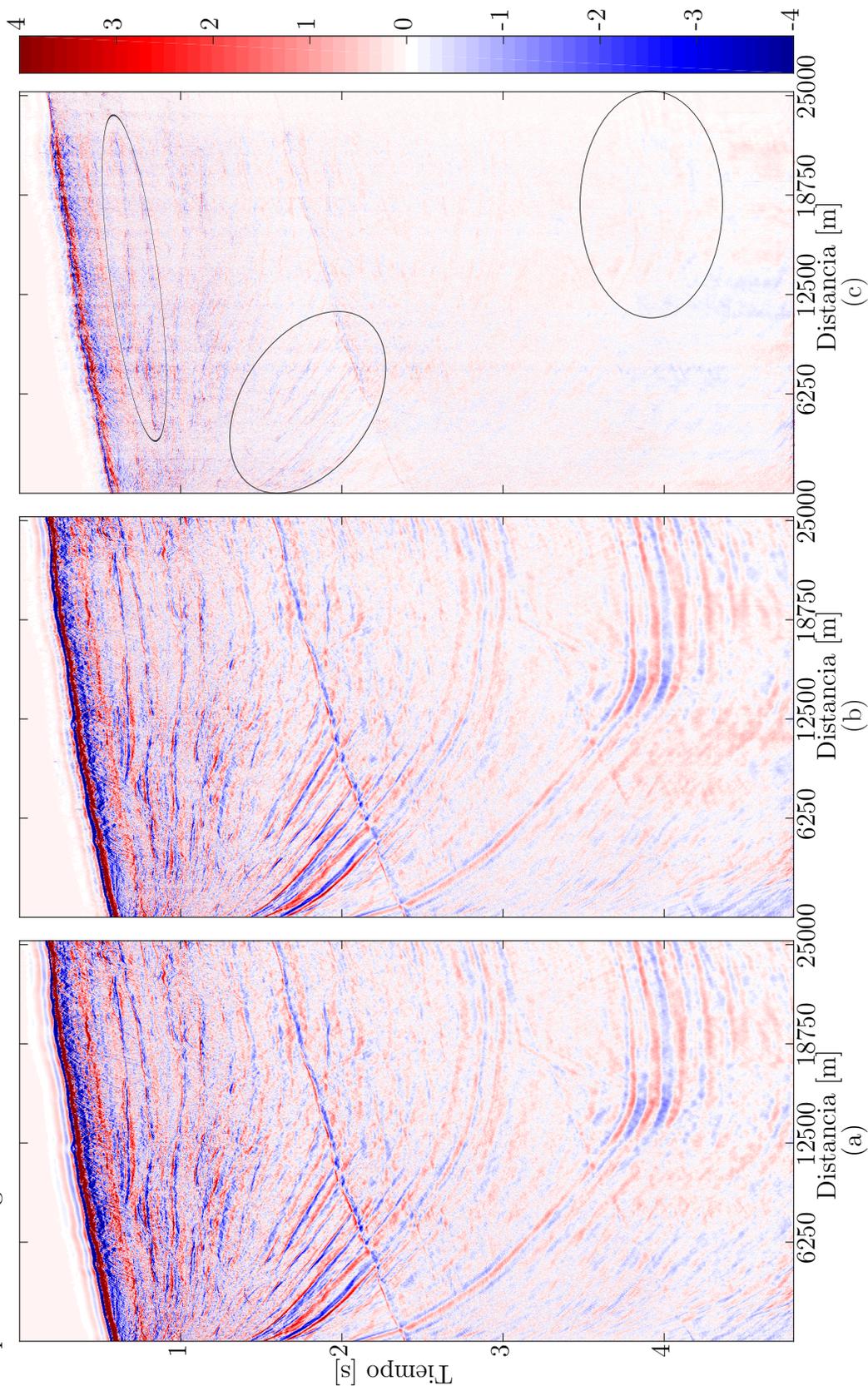
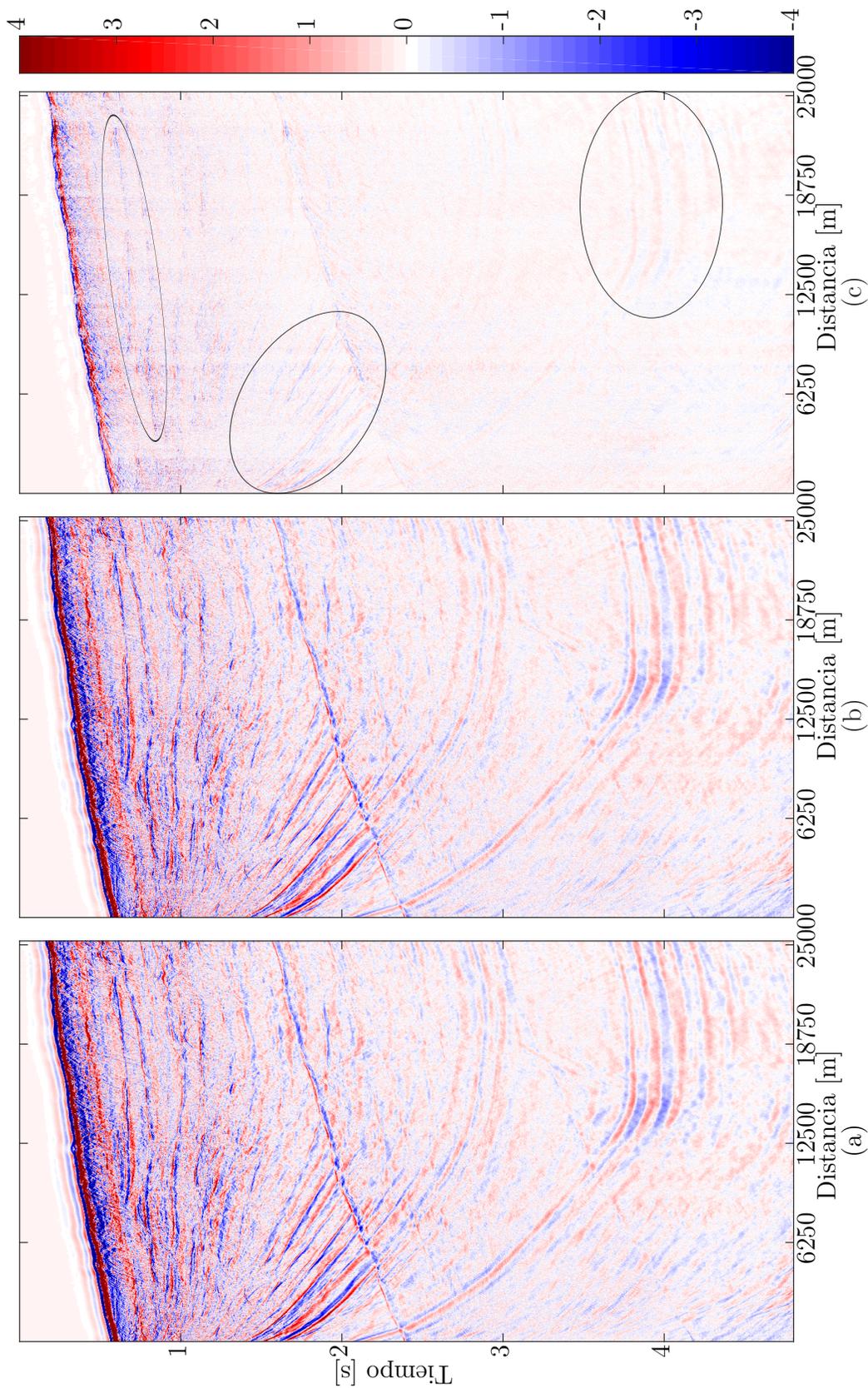


Figura 21: Sección sísmica apilada utilizando los datos (a) originales y (b) adquiridos aleatoriamente con $sr = 0.3$. (c) Diferencia entre la imagen (a) y (b), resaltando con elipses las zonas donde se tienen más cambios, con respecto a la Figura 20.



3.8. RECONSTRUCCIÓN DE DATOS SÍSMICOS REALES CON ANCHO DE BANDA LIMITADO

Se planteó un experimento en el cual los datos observados tienen un ancho de banda limitado a las bajas frecuencias. Para analizar cómo se desempeña el esquema de reconstrucción planteado bajo dichas condiciones, el experimento está orientado a reconstruir datos que van a ser utilizados en *FWI*, proceso que solo requiere de bandas muy específicas de frecuencia [16] que generalmente está por debajo de los $15Hz$ [26].

El dato sísmico inicial se obtuvo aplicando un filtro Ormsby con puntos de corte $2 - 4 - 15 - 25$ al dato real y se simuló la adquisición aleatoria con las matrices de muestreo diseñadas en la sección 3.4. Para establecer el parámetro de reconstrucción λ se empleó el procedimiento planteado en la sección 3.5. Los resultados se presentan en las Tablas 3 y 4 y en las Figuras 22 y 23. Para los datos de baja frecuencia se observa que la reconstrucción tiene un mejor desempeño, obteniendo valores de correlación y SNR más altos que en el caso anterior y un rango de posibles λ más grande, comparando visualmente los resultados se concluye que las trazas reconstruidas se ajustan en forma y amplitud a las trazas originales en un rango de $0.003 < \lambda < 0.5$, lo que significa que se tiene una tolerancia para la selección de este parámetro sin afectar en gran medida la calidad de la reconstrucción, se seleccionaron los valores correspondientes al máximo de correlación y SNR que en este caso son los mismos, para los tres porcentajes de fuentes eliminadas y son cercanos a los valores del caso anterior.

Tabla 3: Valor promedio de λ teórico para los diferentes porcentajes de muestreo aleatorio, del dato sísmico con ancho de banda limitado.

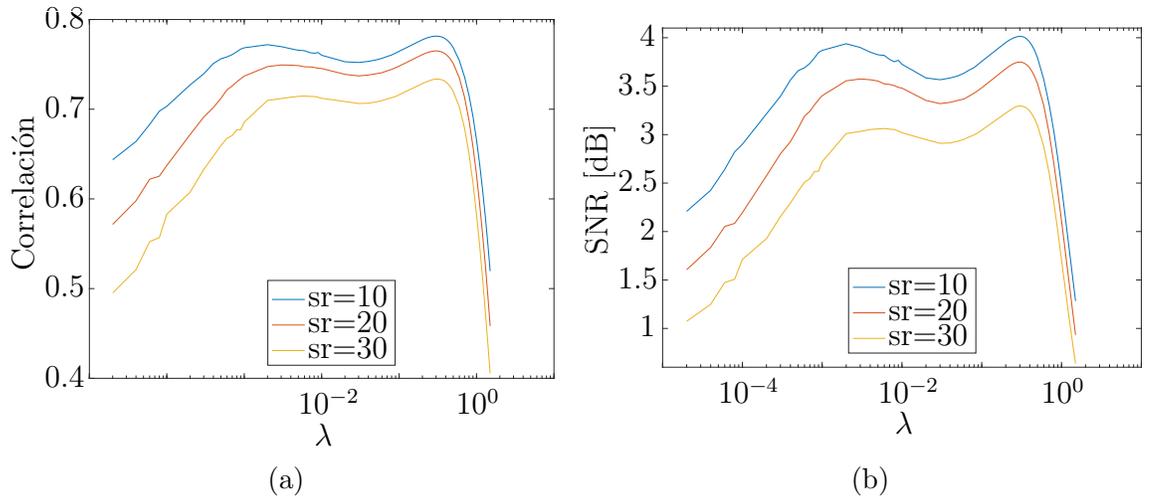
	λ teórico		
	SR=10 %	SR=20 %	SR=30 %
	1.0248	0.9349	0.8754

Los resultados de la reconstrucción completa se presentan en las Figuras 24, 25 y 26, en las imágenes se observa que el proceso se llevó a cabo exitosamente, reconstruyendo *gathers* sísmicos que visualmente son difíciles de diferenciar con el original, lo cual indica que se recuperó satisfactoriamente la información perdida al eliminar fuentes, sin embargo en las imágenes de diferencias se alcanza a observar eventos de baja amplitud debido a que la reconstrucción no alcanza a recuperar exactamente las mismas amplitudes del dato original. Este efecto se observa fácilmente en el espectro en frecuencia de los *CSG* reconstruidos en la Figura 27 donde se observa que el espectro de los datos reconstruidos mantiene la misma forma del original, sin embargo cambia la amplitud

Tabla 4: Valores de λ experimental, con los cuales se obtiene el mejor promedio de correlación y SNR , en la reconstrucción del dato sísmico con ancho de banda limitada.

	λ experimental		
	SR=10 %	SR=20 %	SR=30 %
Correlación	0.3	0.3	0.3
SNR	0.3	0.3	0.3

Figura 22: Curvas de (a) correlación y (b) SNR promedio de las trazas sísmicas con ancho de banda limitada al ser reconstruidas con diferentes valores de λ .



de algunos puntos. Se esperaba obtener una mayor diferencia entre los CSG reconstruidos tras eliminar diferentes porcentajes de fuentes, pero la naturaleza de cómo se realiza la reconstrucción no permite apreciar fácilmente dicho efecto, puesto que un CSG reconstruido es la unión de muchas trazas que fueron reconstruidas en diferentes CRG con diferentes porcentajes de trazas faltantes, lo cual indica que se tienen trazas con diferentes calidades de reconstrucción, que al ser analizadas como un conjunto se pierden de vista.

Comparando las secciones sísmicas apiladas se observan fácilmente múltiples diferencias. En la Figura 28 se presenta la sección apilada con los datos reconstruidos, mostrando que en comparación al original se pierde poca información y las diferencias se concentran en el fondo marino, a diferencia del dato aleatorio apilado en la Figura 29 donde es significativa la pérdida de información.

Figura 23: Sección de un *CRG* con ancho de banda limitada reconstruido con diferentes valores de λ para comparar visualmente los efectos de variar este parámetro, las líneas azules representan las trazas que fueron reconstruidas. (a)(e) *CRG* original, (b)(f) $\lambda = 0.0003$, (c)(g) $\lambda = 0.3$ y (d)(h) $\lambda = 0.8$

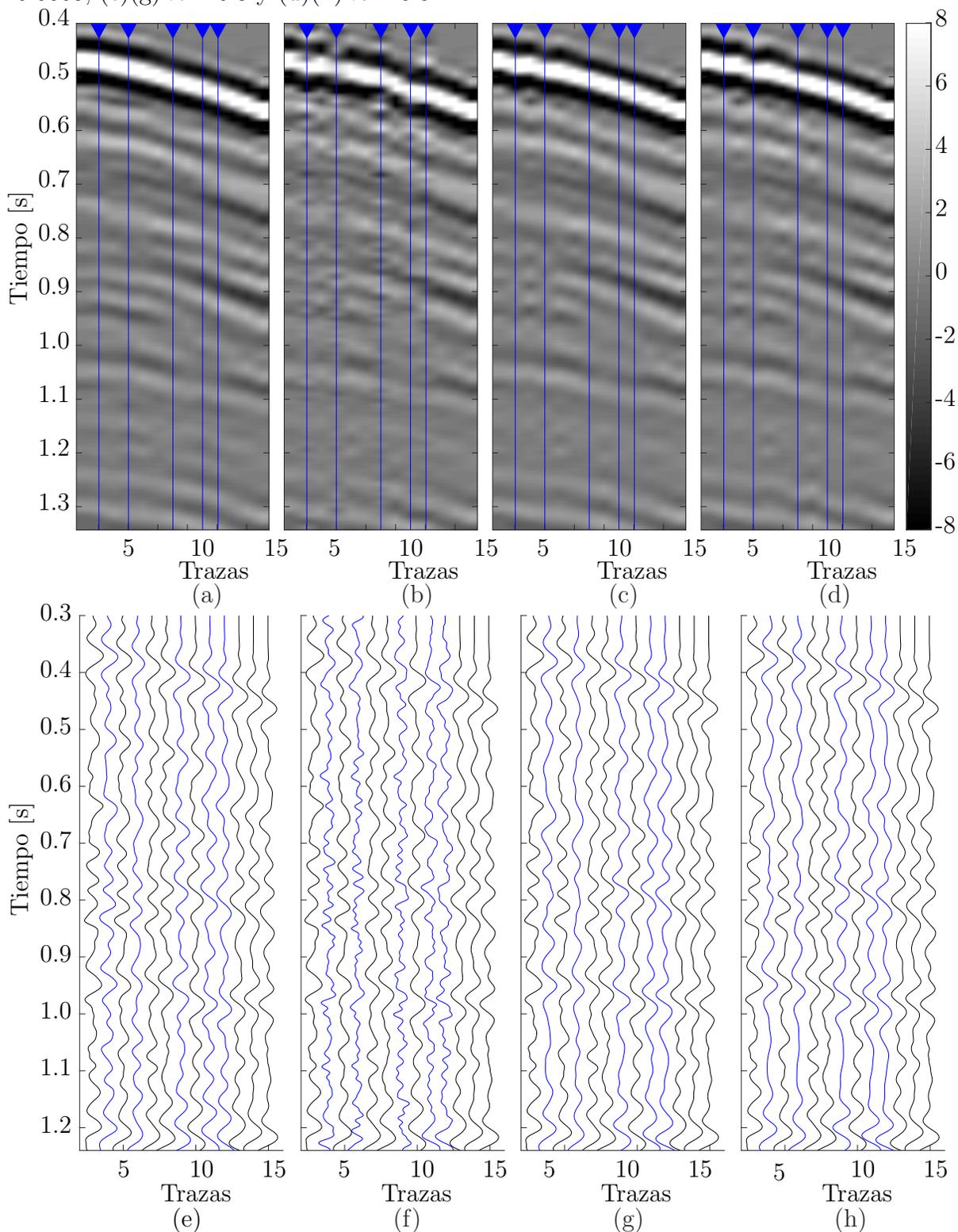


Figura 24: *Common source gather* número 159 con ancho de banda limitado que fue reconstruido tras eliminar un 10% (sr=0.1) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

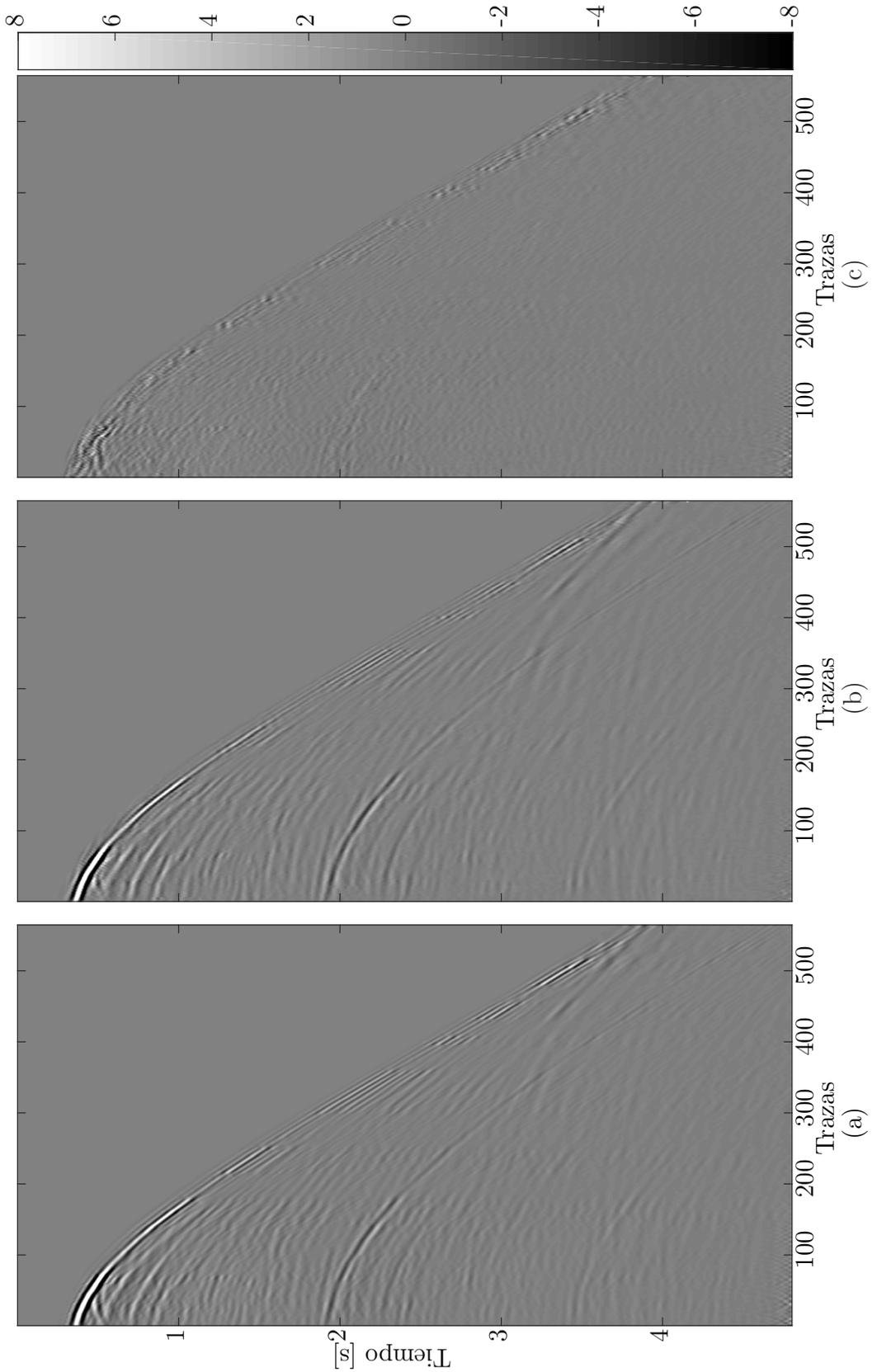


Figura 25: *Common source gather* número 159 con ancho de banda limitado que fue reconstruido tras eliminar un 20% ($sr=0.2$) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

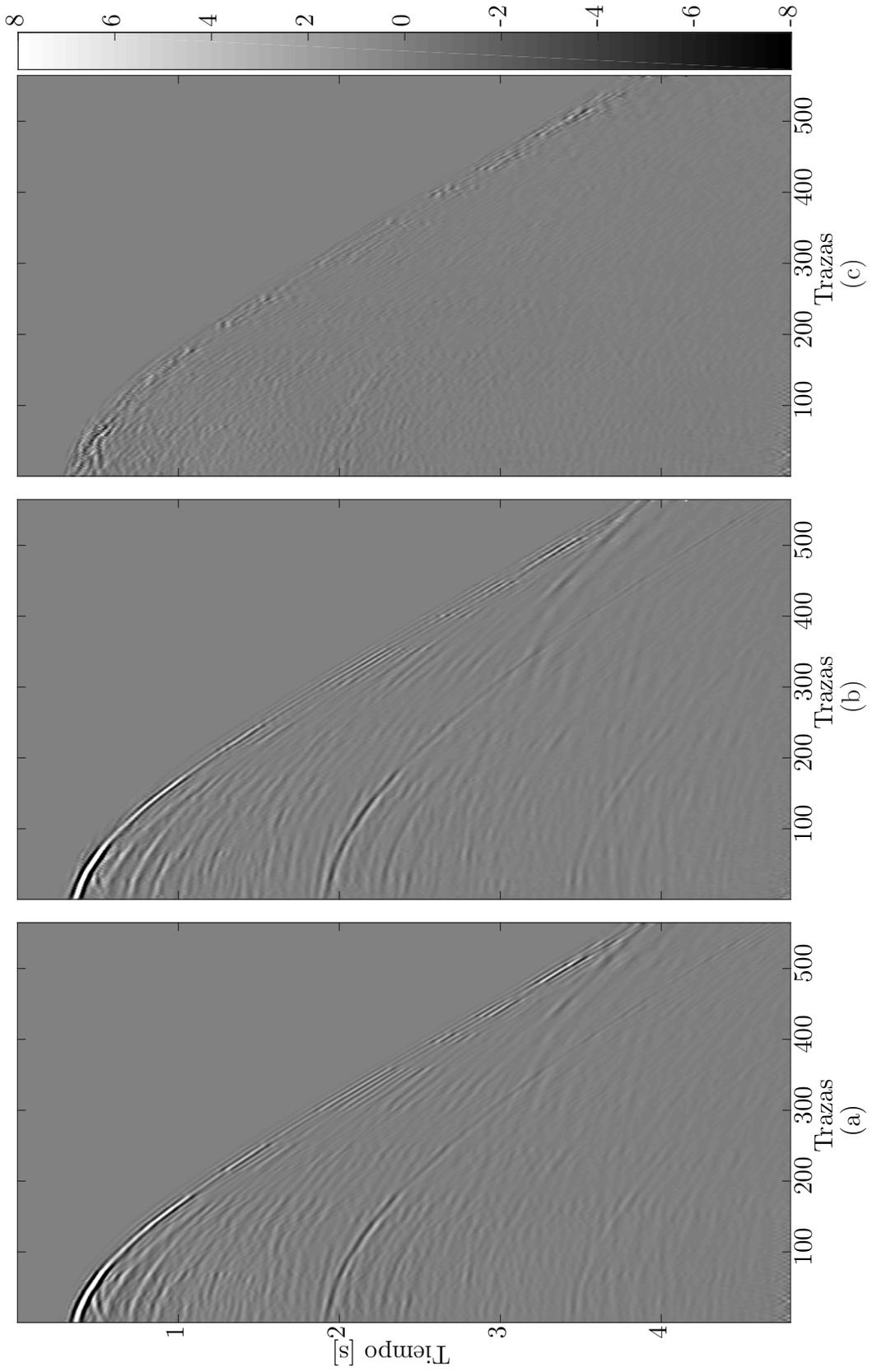


Figura 26: *Common source gather* número 159 con ancho de banda limitado que fue reconstruido tras eliminar un 30% (sr=0.3) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

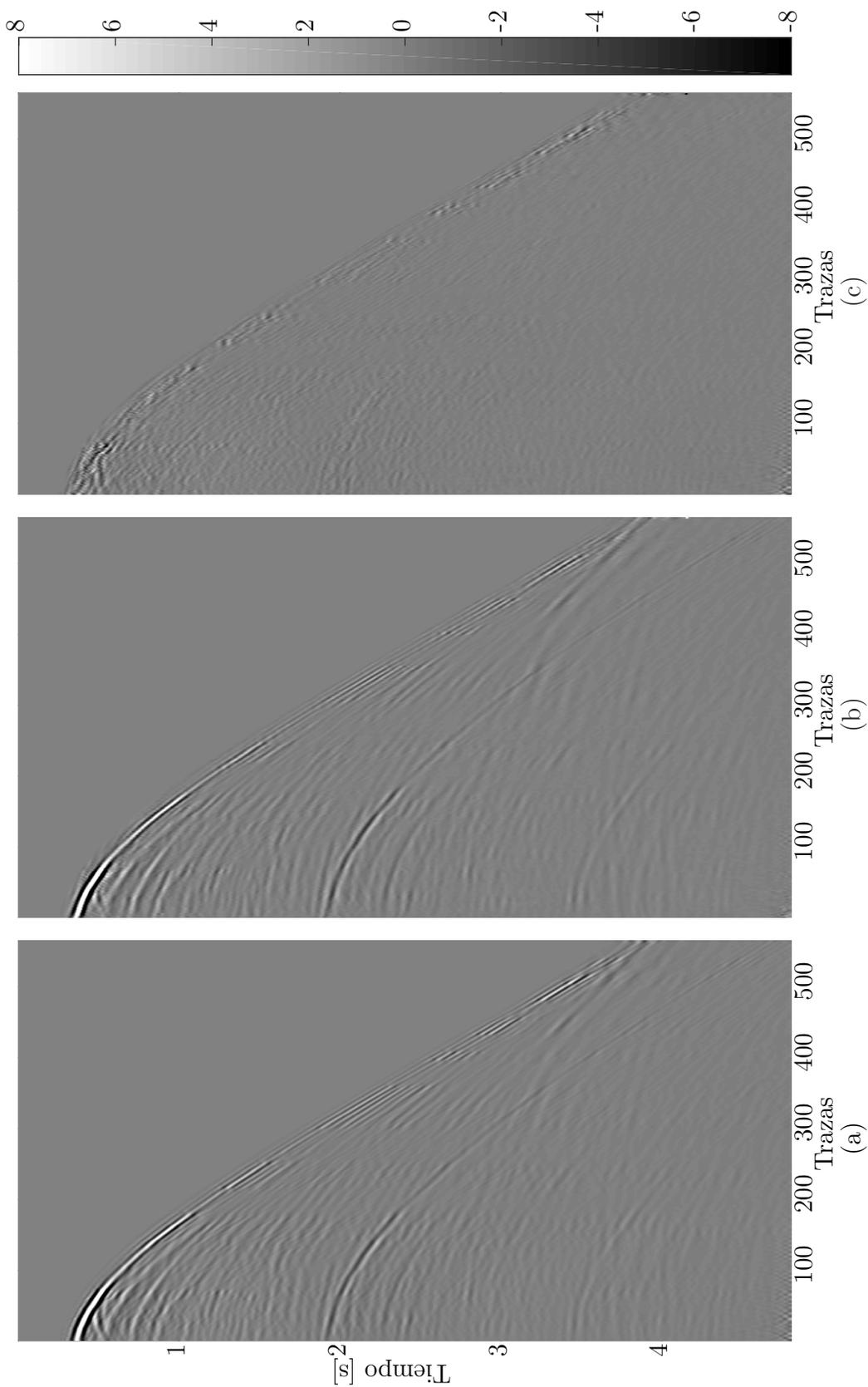


Figura 27: Espectro en frecuencia del *CSG* original número 159 y los *CSG* reconstruidos con ancho de banda limitado, evidenciando que se recuperaron las bajas frecuencias y que tienen el mismo comportamiento del ancho de banda del dato original. Los lóbulos por encima de $25Hz$ no se tienen en cuenta puesto que están por encima de la banda de rechazo del filtro que se le aplicó al dato original.

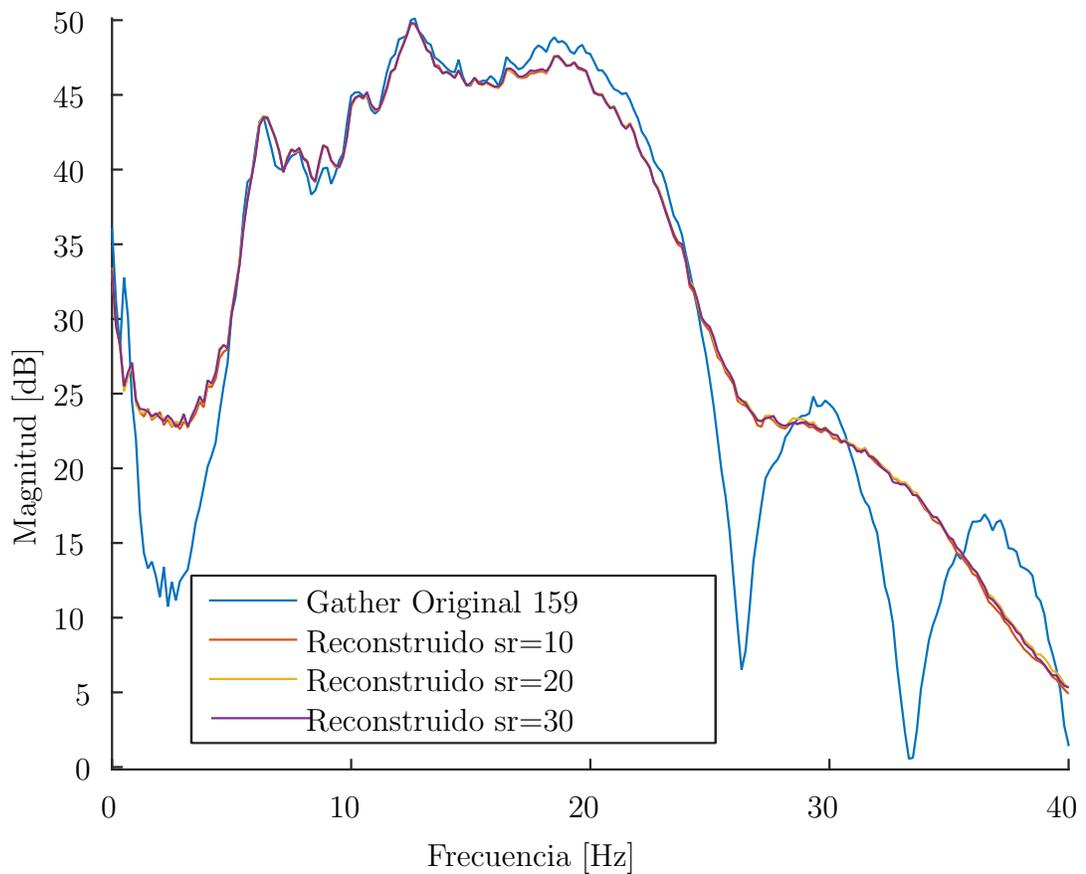


Figura 28: Sección sísmica apilada utilizando los datos (a) originales con ancho de banda limitado y (b) reconstruidos a partir de la adquisición aleatoria con $sr = 0.3$. (c) Diferencia entre la imagen (a) y (b), resaltando con elipses las zonas donde se tienen más cambios, con respecto a la Figura 29.

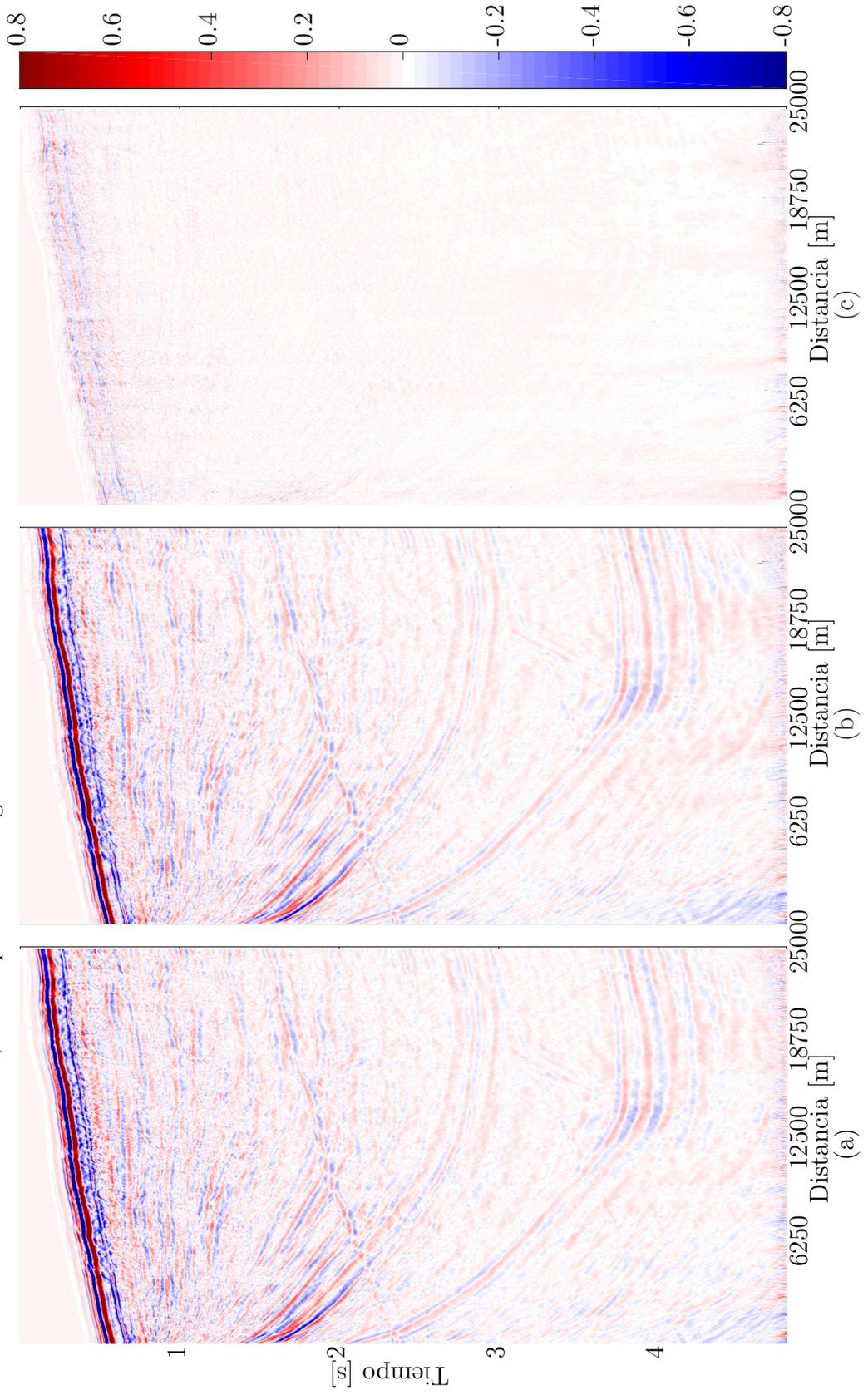
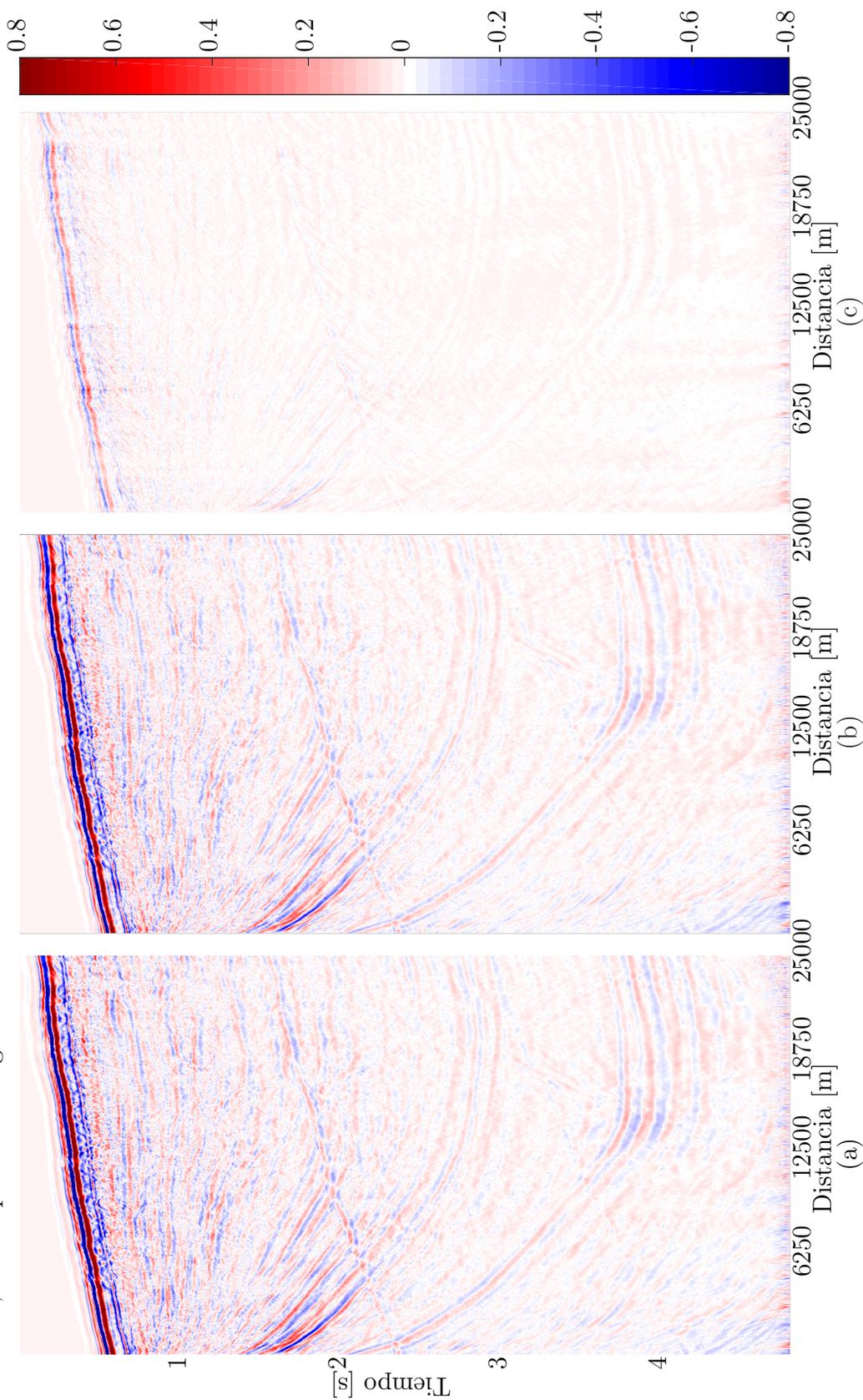


Figura 29: Sección sísmica apilada utilizando los datos (a) originales con ancho de banda limitado y (b) adquiridos aleatoriamente con $sr = 0.3$. (c) Diferencia entre la imagen (a) y (b), resaltando con elipses las zonas donde se tienen más cambios, con respecto a la Figura 28.



3.9. RECONSTRUCCIÓN DE DATOS SÍSMICOS SINTÉTICOS PARA *FWI*

Se realizó un experimento sintético para analizar los efectos de utilizar datos adquiridos de forma aleatoria en etapas complejas del procesamiento sísmico como es el caso de *FWI*, el experimento consiste en diseñar tres adquisiciones aleatorias y reconstruir sus datos con los siguiendo las indicaciones expuestas en este libro, para finalmente comparar los modelos de velocidad obtenidos con *FWI* cuando los datos de entrada son: los reconstruidos, aleatorios y convencionales.

El diseño de la adquisición sintética se realizó emulando una adquisición marina sobre el modelo de velocidades Marmousi ii[27] con una ondícula Ricker de $6Hz$ como fuente, distribuyendo 130 fuentes con una separación de 50 metros en la adquisición de control que es uniforme y 400 sensores activos separados 12.5 metros. Basándose en la adquisición de control se diseñaron tres adquisiciones aleatorias, reduciendo un 10%, 20% y 30% el número de fuentes. El parámetro de reconstrucción λ se seleccionó siguiendo el procedimiento descrito en la sección 3.5, encontrando que los valores más adecuados para las reconstrucciones son 0.1, 0.2 y 0.3 respectivamente, cabe resaltar que estos valores pueden variar en un rango cercano, puesto que en la comparación visual para los λ hasta una década por debajo del seleccionado no se aprecia una diferencia significativa y se utilizó la medida de correlación entre las trazas originales y reconstruidas, sin embargo estas medidas no se tendrían en un escenario real.

En la Figura 30 se compara un *CSG* reconstruido en el que se puede observar que la reconstrucción fue exitosa, recuperando la mayoría de los eventos sísmicos y conservando las frecuencias y amplitud de las trazas sísmicas, como se observa en la Figura 31, en la Figura 30(c) se observan algunos eventos para los cuales no se ajustó por completo la amplitud o tenían un comportamiento cercano al lineal y fueron eliminados, tendencia que se mantienen en todos los *gather* de las diferentes adquisiciones aleatorias diseñadas.

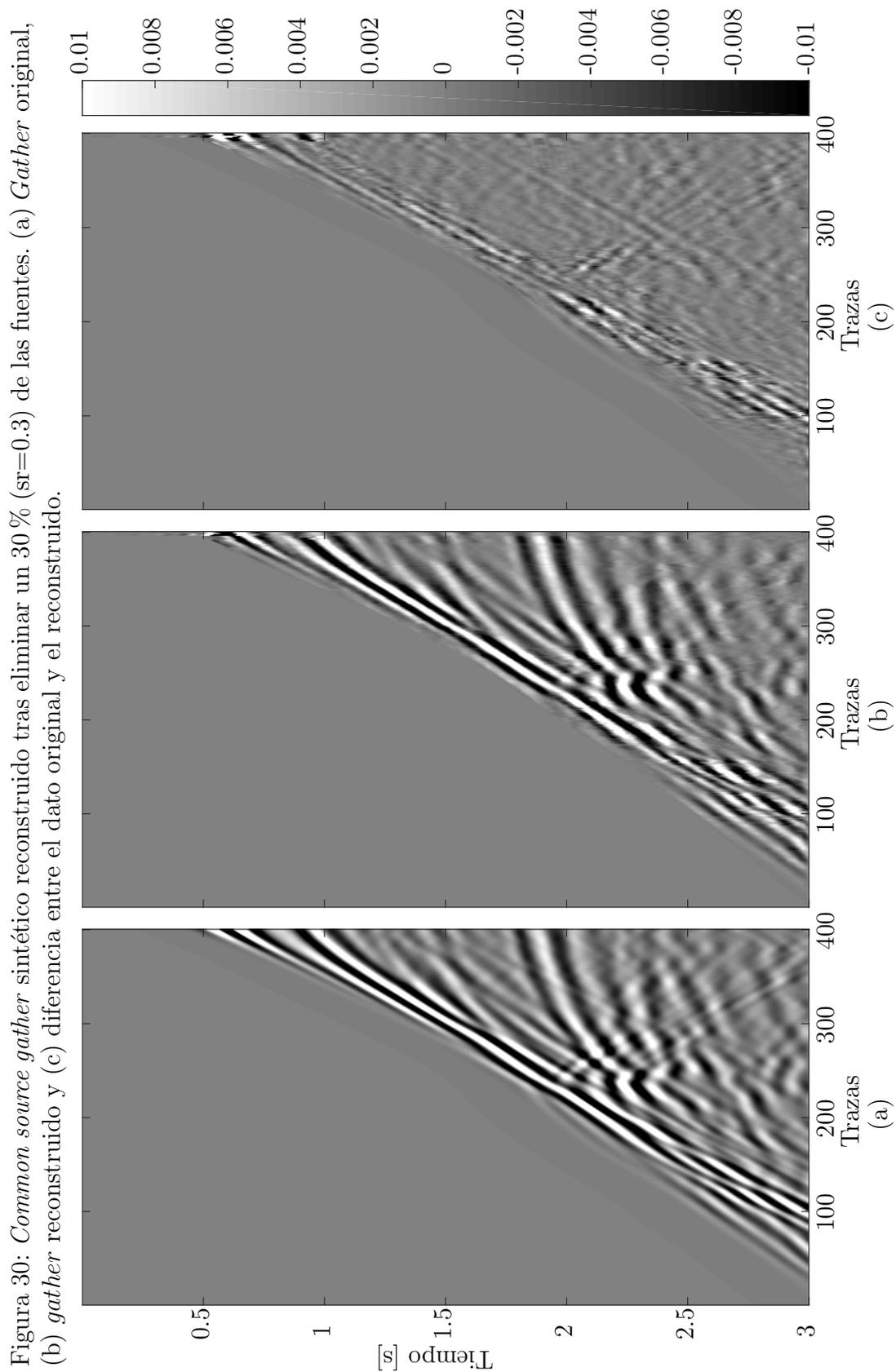
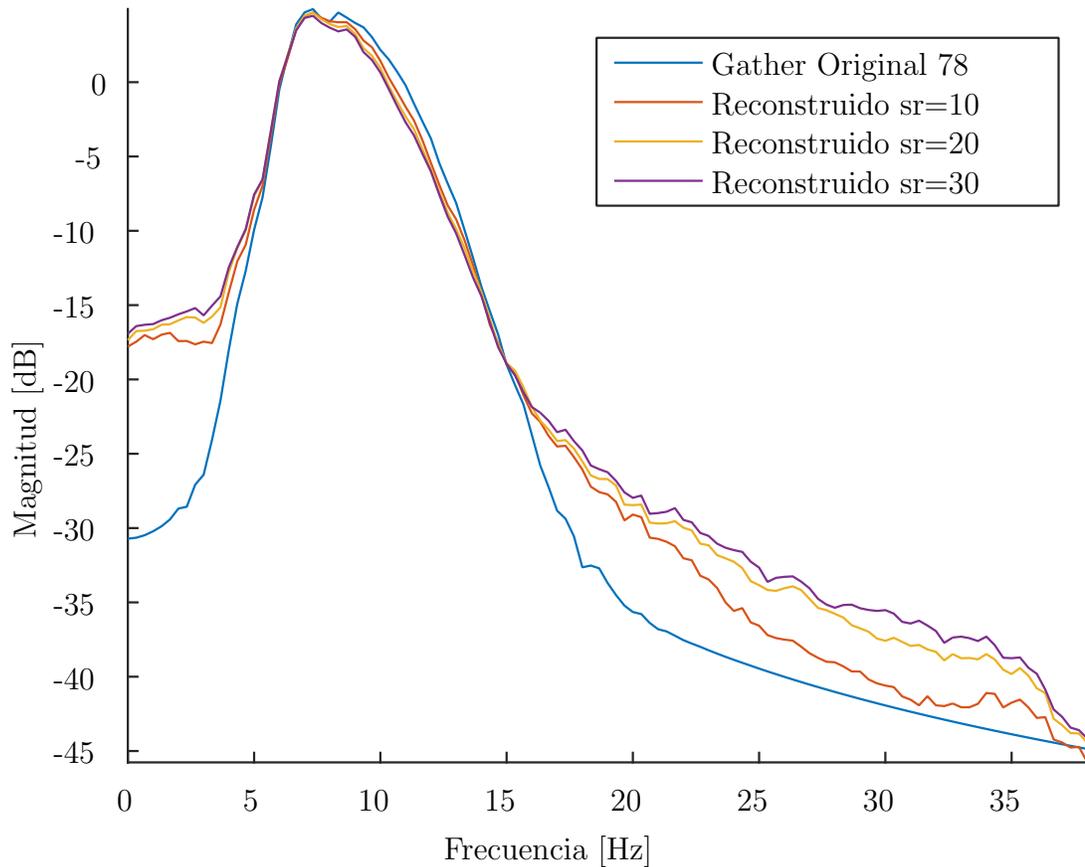


Figura 30: *Common source gather* sintético reconstruido tras eliminar un 30 % ($sr=0.3$) de las fuentes. (a) *Gather* original, (b) *gather* reconstruido y (c) diferencia entre el dato original y el reconstruido.

Figura 31: Espectro en frecuencia de *CSG* sintético y sus respectivos *CSG* reconstruidos a partir de las diferentes adquisiciones aleatorias.



Para realizar el proceso de *FWI* se utilizó un código suministrado por el grupo de investigación CPS el cual realiza la inversión en tiempo sobre un clúster GPU [26], se usaron diez iteraciones de *FWI* con alfa óptimo encontrado usando el método L-BFGS [16], utilizando el mismo modelo de velocidades inicial de la Figura 32(b). Para todas las inversiones, los resultados para la adquisición con $sr = 0.3$ se muestran en las Figuras 32, 33 y 34. Comparando visualmente los resultados se observa que entre los modelos obtenidos con los datos de control y el obtenido con los datos reconstruidos no existe una gran diferencia en cuanto a la forma de las estructuras, es decir que con los dos conjuntos de datos la *FWI* llegó a resultados similares. En la Figura 33(c) se muestran las diferencias entre los dos modelos, evidenciando que los valores de velocidad están alrededor de cero a excepción de la zona superior derecha donde se tiene la diferencia más considerable.

En el caso de la *FWI* con los datos aleatorios sin reconstruir el resultado no es tan bueno, como se observa en la Figura 34, en la cual se ve que la *FWI* alcanza a definir la forma global de las estructuras, igual que con el dato de control pero aparecen una serie de artefactos verticales que asemejan estrías en las estructuras y

en la superficie coinciden con los puntos de las fuentes eliminadas. Estos artefactos se deben a problemas de iluminación en la *FWI*, causados por las fuentes faltantes que no aportan información al gradiente, estos artefactos se evidencia fácilmente en la imagen de diferencias donde aparecen una serie de manchas oscuras.

En la tabla 5 se resumen los resultados de comparar numéricamente los modelos obtenidos con respecto al modelo de velocidad original, encontrando que en todos los casos la norma l_2 de la diferencia entre el modelo original y el obtenido con los datos reconstruidos es menor al valor de la norma l_2 de la diferencia entre el modelo original y el obtenido con los datos aleatorios sin reconstruir. Se comprueba con este experimento que al utilizar datos reconstruidos para el procesamiento sísmico se mejora la calidad de los resultados con respecto a procesar el dato aleatorio, es decir con huecos en la información, sin importar que la información reconstruida no sea exactamente igual a la original.

Tabla 5: Norma l_2 de la diferencia entre el modelo de velocidad real y los modelos obtenidos con *FWI*. Los valores están normalizados con respecto a la norma l_2 entre el modelo original y el obtenido con los datos de control (muestreo uniforme), y su valor es: 33382.

sr	Norma l_2	
	Modelo con los <i>gathers</i> reconstruido	Modelo sin reconstrucción de los <i>gathers</i>
0.1	1.0028	1.0056
0.2	1.0136	1.0138
0.3	1.0195	1.0247

Figura 32: Modelos de velocidad (a) original, (b) inicial y (c) obtenido al aplicar *FWI* sobre los datos de control (adquisición uniforme).

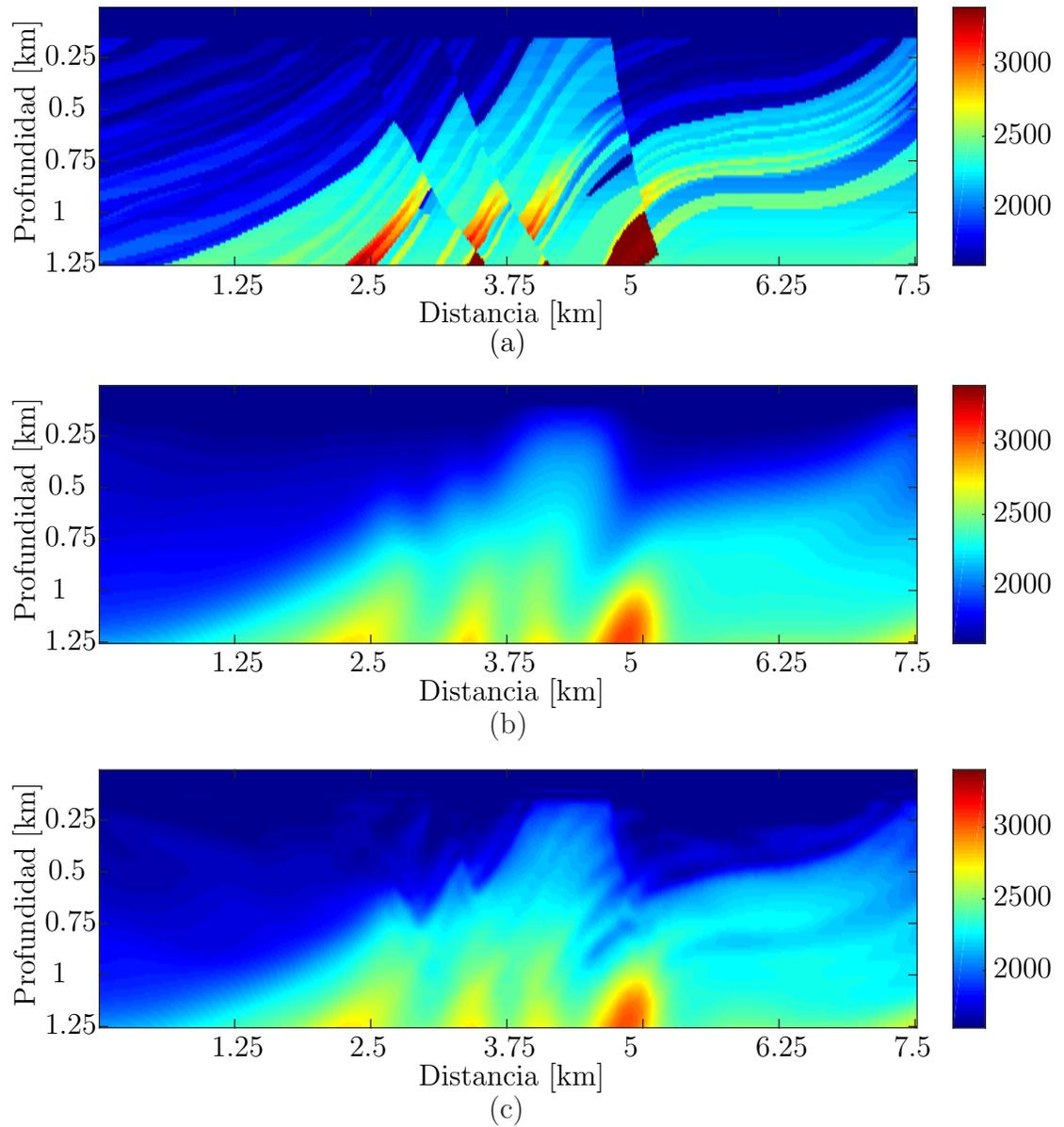


Figura 33: Modelo de velocidad (a) de control, (b) obtenido con los datos reconstruidos de la adquisición aleatoria con $sr = 0.3$ y (c) diferencia entre el modelo a y b.

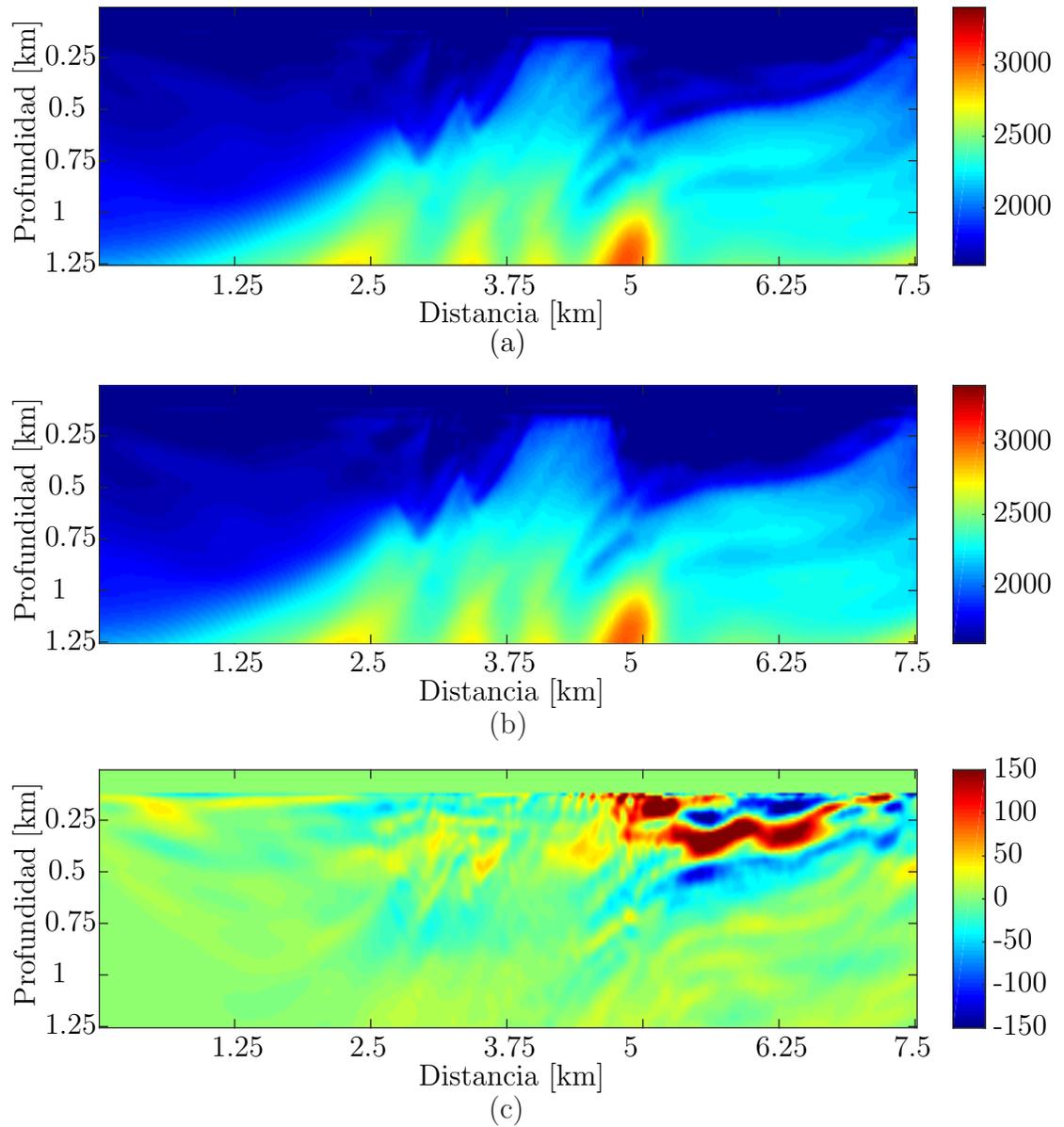
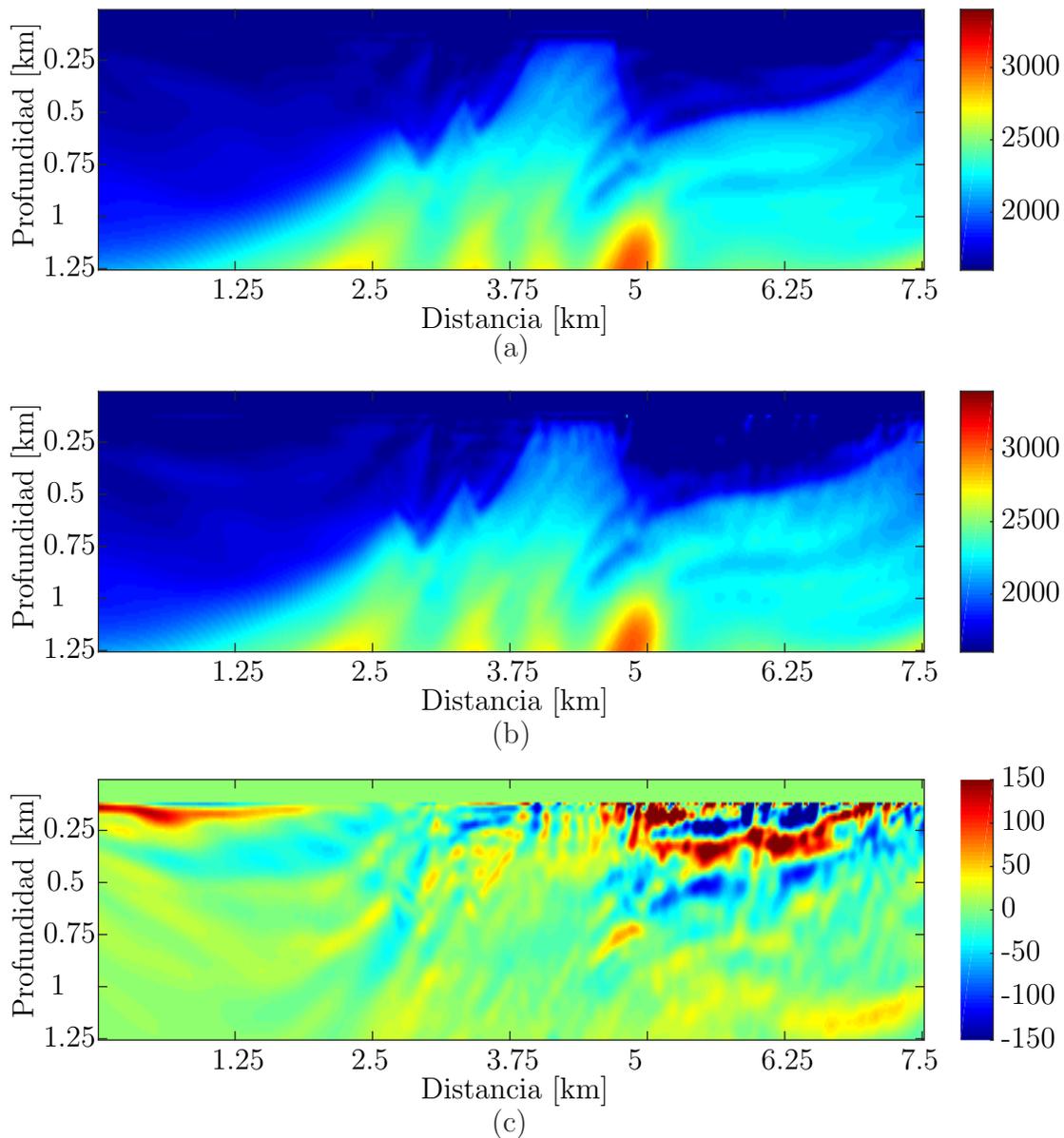


Figura 34: Modelo de velocidad (a) de control, (b) obtenido con los datos aleatorios de $sr = 0.3$ sin reconstruir y (c) diferencia entre el modelo a y b. En la figura (b) y (c) Se observan una serie de anomalías verticales en el modelo, que son producidas por la pérdida de iluminación producto de las fuentes faltantes.



4. CONCLUSIONES

Se aplicó el muestreo compresivo de forma exitosa para la reconstrucción de datos sísmicos, planteando una estrategia para las etapas de diseño de la adquisición aleatoria, preconditionamiento de los datos y selección de los parámetros requeridos para la reconstrucción, además se realizó la selección de un dominio de representación *sparse* para datos sísmicos 2D.

La transformada curvelet es el dominio de representación con el cual se obtuvo el mejor desempeño numérico y visual en la representación *sparse* de los datos sísmicos, favoreciendo la continuidad de los eventos con coherencia hiperbólica, efecto que no se vio reflejados con las transformadas de Fouier y wavelet. Las fuertes variaciones en la amplitud de las trazas sísmicas a medida que aumenta el tiempo y el *offset* es un factor que afecta la reconstrucción y representación *sparse* de los datos, quitándole relevancia a los eventos de baja amplitud sin considerar que es debida a la atenuación y divergencia esférica de la propagación de onda, es decir que sí son eventos relevantes y se deben conservar, volviéndose necesario balancear las amplitudes previamente a la reconstrucción. Como herramienta de balance se propuso utilizar un *AGC* para acondicionar el dato previamente a la reconstrucción, y retirar dicha ganancia al finalizar el proceso para que las amplitudes del dato observado y reconstruidos no se vean afectadas.

El muestreo aleatorio se puede aplicar de tres formas diferentes en la adquisición de datos sísmicos 2D: distribución de las fuentes, distribución de los receptores y periodo de muestreo. El caso que se consideró más relevante es cambiar la distribución de las fuentes para reducir el número de disparos y flexibilizar la ubicación los mismos, la posterior reconstrucción de los datos se realiza agrupando las trazas en *CRG* donde cada fuente eliminada se ve como una única traza faltante. La matriz de muestreo aleatoria de las fuentes se debe diseñar teniendo en cuenta ciertas consideraciones que favorezcan una posterior reconstrucción de los datos, las que encontramos más relevantes son: no eliminar fuentes en los extremos de la línea de adquisición puesto que se tendrían huecos en los *CRG* que agrupan menos trazas, no eliminar más de dos fuentes contiguas para que no se presenten huecos de más de dos trazas en los *CRG* y seleccionar las fuentes a eliminar con un generador de números aleatorios con densidad de distribución uniforme para que los huecos no se concentren en ninguna zona.

Se diseñaron tres adquisiciones aleatorias para analizar el comportamiento de los *CSG* reconstruidos cuando se eliminan diferentes porcentajes de fuentes, concluyendo que las reconstrucciones de los datos aleatorios fueron exitosas obteniendo *CSG* similares al original en los tres casos. Al comparar los mismos *CSG* reconstruidos en las tres adquisiciones se observa que son difíciles de diferenciar entre sí, efecto que no se

esperaba puesto que fueron reconstruidos a partir de datos con diferentes números de fuentes, efecto que se atribuye al cambio en la agrupación de las trazas para la reconstrucción, causando que los *CRG* tengan un porcentaje variable de trazas eliminadas y al volver a ordenar las trazas en *CSG* su calidad no se puede atribuir a un único porcentaje de muestreo. La reducción del ancho de banda de las trazas sísmicas reconstruidas es un comportamiento general en todos los *gathers*, reduciendo la magnitud de las frecuencias altas y reconstruyen mejor los eventos de baja frecuencia que son los de mayor energía y un comportamiento más marcado en el *gather*.

Se obtuvo un resultado favorable al comparar el apilado de los datos reconstruidos con respecto al apilado del dato aleatorio y el convencional, observando que no se altera la forma de los reflectores y solo se pierde una pequeña parte de la amplitud de los reflectores más delgados (atribuidos a las altas frecuencias), mientras que en la sección apilada con los datos aleatorios se observa una mayor diferencia en la amplitud de los reflectores de baja y alta frecuencia.

La *FWI* es una herramienta de procesamiento susceptible ante la pérdida de información, viéndose afectada en gran medida cuando se pierden disparos en el proceso de adquisición, y es en este escenario donde el muestreo compresivo puede jugar un papel muy importante recuperando la información perdida o flexibilizando la geometría de adquisición para evitar obstáculos externos que pueden impedir la ubicación de las fuentes.

REFERENCIAS

- [1] Zwartjes P. and Gisolf A., “Fourier reconstruction with sparse inversion,” *Geophysical Prospecting*, vol. 55, no. 2, pp. 199–221, 2007.
- [2] Herrmann Felix J., “Randomized sampling and sparsity: getting more information from fewer samples,” *Geophysics*, vol. 75, pp. WB173–WB187, 12 2010.
- [3] Candès E, Romberg J, and Tao T, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *Information Theory, IEEE Transactions on*, vol. 52, pp. 489–509, Feb 2006.
- [4] Donoho D.L., “Compressed sensing,” *Information Theory, IEEE Transactions on*, vol. 52, pp. 1289–1306, April 2006.
- [5] Candès E.J and Wakin M.B, “An introduction to compressive sampling,” *Signal Processing Magazine, IEEE*, vol. 25, pp. 21–30, March 2008.
- [6] Candès E. and Romberg J., “Sparsity and incoherence in compressive sampling,” *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007. cited By 986.
- [7] Candès Emmanuel J. and Demanet Laurent, “The curvelet representation of wave propagators is optimally sparse,” *Communications on Pure and Applied Mathematics*, vol. 58, no. 11, pp. 1472–1528, 2005.
- [8] Candès Emmanuel, Demanet Laurent, Donoho David, and Ying Lexing, “Fast discrete curvelet transforms,” *Multiscale Modeling & Simulation*, vol. 5, no. 3, pp. 861–899, 2006.
- [9] Paredes, J.L. Arce, G.R., “Compressive sensing signal reconstruction by weighted median regression estimates,” *Signal Processing, IEEE Transactions on*, vol. 59, pp. 2585–2601, June 2011.
- [10] Kim S. J., Koh K., Lustig M., Boyd S., and Gorinevsky D., “An interior-point method for large-scale ell_1 -regularized least squares,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, pp. 606–617, Dec 2007.
- [11] Figueiredo M.A.T., Nowak R.D., and Wright S.J., “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, pp. 586–597, Dec 2007.

-
- [12] Wright S. J., Nowak R. D., and Figueiredo M. A. T., “Sparse reconstruction by separable approximation,” *IEEE Transactions on Signal Processing*, vol. 57, pp. 2479–2493, July 2009.
- [13] Bioucas-Dias J. M. and Figueiredo M. A. T., “A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration,” *IEEE Transactions on Image Processing*, vol. 16, pp. 2992–3004, Dec 2007.
- [14] Barzilai Jonathan and Borwein Jonathan M., “Two-point step size gradient methods,” *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.
- [15] Tarantola Albert, “Inversion of seismic reflection data in the acoustic approximation,” *GEOPHYSICS*, vol. 49, no. 8, pp. 1259–1266, 1984.
- [16] dos Santos Adriano W. G. and Pestana Reynam C., “Time-domain multiscale full-waveform inversion using the rapid expansion method and efficient step-length estimation,” *GEOPHYSICS*, vol. 80, no. 4, pp. R203–R216, 2015.
- [17] Plessix R.-E., “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications,” *Geophysical Journal International*, vol. 167, no. 2, pp. 495–503, 2006.
- [18] Ma Yong and Hale Dave, “Quasi-newton full-waveform inversion with a projected hessian matrix,” *GEOPHYSICS*, vol. 77, no. 5, pp. R207–R216, 2012.
- [19] Liu Dong C. and Nocedal Jorge, “On the limited memory bfgs method for large scale optimization,” *Mathematical Programming*, vol. 45, pp. 503–528, Aug 1989.
- [20] Herrera Yajaira, Cooper Norman, and Mustagh Resources Ltd, *Manual para la adquisición y procesamiento de sísmica terrestre y sus aplicaciones en Colombia*. Universidad Nacional de Colombia, 2010.
- [21] Oilfield Glossary. <http://www.glossary.oilfield.slb.com/>, 2016. Revisado: Julio del 2016.
- [22] Ahay. http://www.ahay.org/RSF/book/cup/rna2d/paper_html/node5.html, 2016. Revisado: Enero del 2016.
- [23] Cooper Norm, “A world of reality—designing land 3d programs for signal, noise, and prestack migration,” *The Leading Edge*, vol. 23, no. 10, pp. 1007–1014, 2004.
- [24] Yilmaz Oz, *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists., 2001.
- [25] Candès Emmanuel, Demanet Laurent, Donoho David, and Ying Lexing, “Curvelet 2.1.2.” <http://www.curvelet.org/>, 2008. Revisado: Julio del 2016.
-

- [26] Noriega R. F., Ramirez A. B., Abreo S. A., and Arce G. R., “Implementation strategies of the seismic full waveform inversion,” en *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1567–1571, March 2017.
- [27] Martin Gary S., Wiley Robert, and Marfurt Kurt J., “Marmousi2: An elastic upgrade for marmousi,” *The Leading Edge*, vol. 25, no. 2, pp. 156–166, 2006.
- [28] Candès Emmanuel, Demanet Laurent, Donoho David, and Ying Lexing, “Curvelab 2.1.2,” 2008.
- [29] Maslen Guy, “Agc: The equaliser.” <http://seismicreflections.globeclaritas.com/2013/04/agc-equaliser.html>, 2016. Revisado: Julio del 2016.
- [30] Paredes J. L. and Arce G. R., “Compressive sensing signal reconstruction by weighted median regression estimates,” *IEEE Transactions on Signal Processing*, vol. 59, pp. 2585–2601, June 2011.
- [31] SubSurfWiki. <http://subsurfwiki.org/wiki/Gather>, 2016. Revisado: Julio del 2016.

BIBLIOGRAFÍA

- Ahay. http://www.ahay.org/RSF/book/cup/rna2d/paper_html/node5.html, 2016.
Revisado: Enero del 2016
- Barzilai Jonathan and Borwein Jonathan M., “Two-point step size gradient methods,” *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988
- Bioucas-Dias J. M. and Figueiredo M. A. T., “A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration,” *IEEE Transactions on Image Processing*, vol. 16, pp. 2992–3004, Dec 2007
- Candès E, Romberg J, and Tao T, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *Information Theory, IEEE Transactions on*, vol. 52, pp. 489–509, Feb 2006
- Candès E. and Romberg J., “Sparsity and incoherence in compressive sampling,” *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007. cited By 986
- Candès E.J and Wakin M.B, “An introduction to compressive sampling,” *Signal Processing Magazine, IEEE*, vol. 25, pp. 21–30, March 2008
- Candès Emmanuel J. and Demanet Laurent, “The curvelet representation of wave propagators is optimally sparse,” *Communications on Pure and Applied Mathematics*, vol. 58, no. 11, pp. 1472–1528, 2005
- Candès Emmanuel, Demanet Laurent, Donoho David, and Ying Lexing, “Curveletab 2.1.2,” 2008
- Candès Emmanuel, Demanet Laurent, Donoho David, and Ying Lexing, “Fast discrete curvelet transforms,” *Multiscale Modeling & Simulation*, vol. 5, no. 3, pp. 861–899, 2006
- Cooper Norm, “A world of reality—designing land 3d programs for signal, noise, and prestack migration,” *The Leading Edge*, vol. 23, no. 10, pp. 1007–1014, 2004
- Donoho D.L., “Compressed sensing,” *Information Theory, IEEE Transactions on*, vol. 52, pp. 1289–1306, April 2006
- dos Santos Adriano W. G. and Pestana Reynam C., “Time-domain multiscale full-

waveform inversion using the rapid expansion method and efficient step-length estimation,” *GEOPHYSICS*, vol. 80, no. 4, pp. R203–R216, 2015

Figueiredo M.A.T., Nowak R.D., and Wright S.J., “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, pp. 586–597, Dec 2007

Herrera Yajaira, Cooper Norman, and Mustagh Resources Ltd, *Manual para la adquisición y procesamiento de sísmica terrestre y sus aplicaciones en Colombia*. Universidad Nacional de Colombia, 2010

Herrmann Felix J., “Randomized sampling and sparsity: getting more information from fewer samples,” *Geophysics*, vol. 75, pp. WB173–WB187, 12 2010

Kim S. J., Koh K., Lustig M., Boyd S., and Gorinevsky D., “An interior-point method for large-scale ell_1 -regularized least squares,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, pp. 606–617, Dec 2007

Liu Dong C. and Nocedal Jorge, “On the limited memory bfgs method for large scale optimization,” *Mathematical Programming*, vol. 45, pp. 503–528, Aug 1989

Ma Yong and Hale Dave, “Quasi-newton full-waveform inversion with a projected hessian matrix,” *GEOPHYSICS*, vol. 77, no. 5, pp. R207–R216, 2012

Martin Gary S., Wiley Robert, and Marfurt Kurt J., “Marmousi2: An elastic upgrade for marmousi,” *The Leading Edge*, vol. 25, no. 2, pp. 156–166, 2006

Maslen Guy, “Agc: The equaliser.” <http://seismicreflections.globeclaritas.com/2013/04/agc-equaliser.html>, 2016. Revisado: Julio del 2016

Noriega R. F., Ramirez A. B., Abreo S. A., and Arce G. R., “Implementation strategies of the seismic full waveform inversion,” en *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1567–1571, March 2017

Oilfield Glossary. <http://www.glossary.oilfield.slb.com/>, 2016. Revisado: Julio del 2016

Paredes J. L. and Arce G. R., “Compressive sensing signal reconstruction by weighted median regression estimates,” *IEEE Transactions on Signal Processing*, vol. 59, pp. 2585–2601, June 2011

Plessix R.-E., “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications,” *Geophysical Journal International*, vol. 167,

no. 2, pp. 495–503, 2006

SubSurfWiki. <http://subsurfwiki.org/wiki/Gather>, 2016. Revisado: Julio del 2016

Tarantola Albert, “Inversion of seismic reflection data in the acoustic approximation,” *GEOPHYSICS*, vol. 49, no. 8, pp. 1259–1266, 1984

Wright S. J., Nowak R. D., and Figueiredo M. A. T., “Sparse reconstruction by separable approximation,” *IEEE Transactions on Signal Processing*, vol. 57, pp. 2479–2493, July 2009

Yilmaz Oz, *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Society of Exploration Geophysicists., 2001

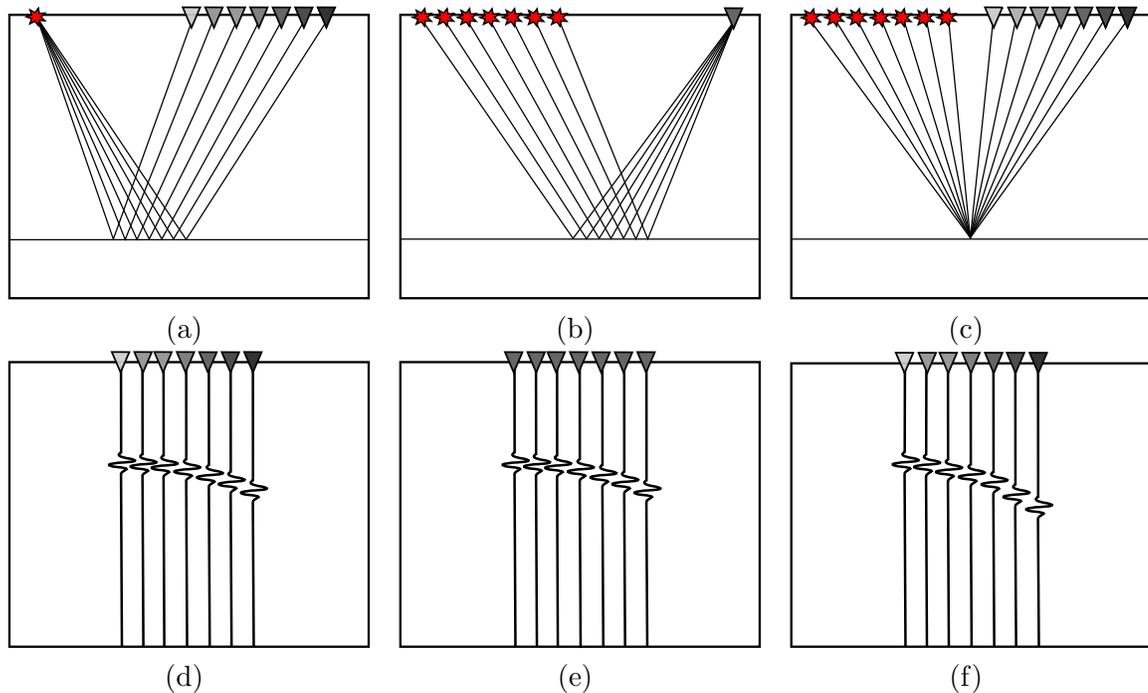
Zwartjes P. and Gisolf A., “Fourier reconstruction with sparse inversion,” *Geophysical Prospecting*, vol. 55, no. 2, pp. 199–221, 2007

ANEXOS

A. AGRUPACIÓN DE DATOS SÍSMICOS

Los datos sísmicos se representan por medio de *gathers*, que son la agrupación u ordenamiento de trazas sísmicas con características geométricas en común. Entre las características más utilizadas para ordenar los datos sísmicos se encuentran: fuente común, receptor común y punto medio común, en la Figura 1 se muestra un ejemplo de cómo es el comportamiento geométrico de cada una de estas características.

Figura 1: Características geométricas para la agrupación de datos sísmicos: (a) Fuente común, (b) receptor común y (c) punto medio común. Comportamiento de los respectivos *gathers* sísmicos: (d) *Gather* fuente común, (e) *gather* receptor común y (e) *gather* punto medio común. Adaptada de [31]



Fuente común: corresponden a agrupar todas las trazas sísmicas obtenidas ante la activación de una única fuente, las trazas se ordenan de menor a mayor distancia entre la fuente y el receptor (*offset*), como se muestra en la Figura 1 (a) y (d). Es la agrupación más común de las trazas ya que corresponden a la forma como se adquieren los datos en campo, obteniendo tantos *gathers* como fuentes se hallan utilizado y se les conoce como *common shot gather (CSG)* o agrupación fuente común.

Receptor común: consiste en agrupar todas las trazas sísmicas captadas por un mismo receptor ante la activación individual de las diferentes fuentes utilizadas en la adquisición, como se muestra en la Figura 1 (b) y (e), se les conoce como *common receiver gather (CRG)*. El orden de las trazas en el *gather* se da en función del *offset* entre la fuente y receptor. Al ordenar los datos de esta forma se obtiene tantos *gathers* como receptores o puntos de adquisición se hallan utilizado, este término es diferente al número de canales.

Punto medio común: utilizando la posición de las parejas fuente-receptor se puede calcular un punto medio que corresponde a la mitad del desplazamiento horizontal del rayo que representa la propagación de la energía y en un medio isótropo de capas paralelas corresponde a la posición horizontal de los puntos de reflexión del rayo, como se ve en la Figura 1 (c), por lo tanto, se asume que la traza sísmica contiene información acerca de las estructuras bajo el punto medio común. Al agrupar las trazas por el punto medio común en el cual aportan información se obtiene un *common mid point gather (CMP)*, como en la Figura 1 (f).

B. *AUTOMATIC GAIN CONTROL (AGC)*

Es una de las ganancias más utilizadas para facilitar la visualización de los *gather* sísmicos que tienen valores muy altos en la amplitud y se dificulta la visualización de los eventos más tenues, el *AGC* balancea las amplitudes a lo largo de una traza utilizando una ventana de tiempo deslizante [24]. Se calcula el valor promedio de la amplitud dentro de la ventana de tiempo, luego se mueve la ventana una muestra, en la dirección del tiempo y se repite el proceso hasta recorrer toda la traza, al finalizar se obtiene una función de ganancia $g(t)$ que se puede multiplicar por un factor de escala para obtener la traza directamente normalizada al aplicar la función de $g(t)$. En la Figura 2 (a) y (b) se muestra un ejemplo del funcionamiento del *AGC* en una traza del conjunto de datos utilizados en esta investigación y su respectiva función de ganancia $g(t)$.

El tamaño típico de la ventana es de 250 – 500 *ms* de largo [29], una ventana más corta tiene un efecto más pronunciado, pero pueden llegar a dañar la forma de onda en la traza. *AGC* eleva el nivel de ruido en la columna de agua, aire o espacio libre antes de registrar la primera reflexión, haciendo conveniente aplicar un top mute después del *AGC*.

Figura 2: (a) Comparación de una traza sísmica a la cual se aplicó un balance de amplitudes utilizando *AGC*, en negro se presenta la traza original con la amplitud normalizada y en azul la misma traza después de aplicar la *AGC*. (b) Función de ganancia $g(t)$ extraída de la traza original en la figura (a).

