



INSTITUTO POLITÉCNICO NACIONAL

Escuela Superior de Ingeniería y Arquitectura
Unidad Ticomán - Ciencias de la Tierra

Operatividad de Programas de Código Abierto en la Ejecución de Inversión Sísmica

T E S I S

Que para optar por el título de
Ingeniero Geofísico

Presenta
Juan de Dios Orellana Orellana

Asesora
Ing. Gabriela De Los Santos Cano

Ciudad de México, Noviembre 2023

Autorización de uso obra

INSTITUTO POLITÉCNICO NACIONAL
Presente

Bajo protesta de decir verdad el que suscribe **Juan de Dios Orellana Orellana** manifiesto de ser autor y titular de los derechos morales y patrimoniales de la obra titulada "**Operatividad de programas de código abierto en la ejecución de inversión sísmica**", en adelante "La Tesis" y de la cual se adjunta copia, por lo que por medio del presente y con fundamento en el artículo 27 fracción II, inciso b) de la Ley Federal del Derecho de Autor, otorgo a el Instituto Politécnico Nacional, en adelante "El IPN", autorización no exclusiva para comunicar y exhibir públicamente total o parcialmente en medio digitales (Publicación en Linea) "La Tesis" por un período indefinido contando a partir de la fecha de la presente autorización, dicho período se renovara automáticamente en caso de no dar aviso expreso a "El IPN" de su terminación.

En virtud de los anterior, "El IPN" deberá reconocer en todo momento mi calidad de autor de "La Tesis".

Adicionalmente, y en mi calidad de autor y titular de los derechos morales y patrimoniales de "La Tesis", manifiesto que la misma es original y que la presente autorización no contraviene ninguna otorgada por el suscrito respecto de "La Tesis" por lo que deslindo toda responsabilidad a "El IPN" en caso de que el contenido de "La Tesis" o la autorización concedida afecte o viole los derechos autorales, industriales, secretos industriales, convenios o contratos de confidencialidad o en general cualquier derecho de propiedad intelectual de terceros y asumo las consecuencias legales y económicas de cualquier demanda o reclamación que puedan derivarse del caso.

Ciudad de Mexico, Noviembre de 2023

Atentamente



Juan de Dios Orellana Orellana



Folio
DIE/263/2023

2023. Año de Francisco Villa
30 Aniversario de la Declaración sobre
la Eliminación de la Violencia contra la Mujer (ONU)
60 Aniversario del CECyT 7 "Cuahtémoc" y del CENAC
90 Aniversario de la Escuela Superior de Ingeniería Textil
40 Aniversario del CIIDIR, Unidad Oaxaca

Asunto
Trámite de titulación

Ciudad de México, 14 de julio de 2023

ING. ROBERTO LOO GUZMÁN
SUBDIRECTOR ACADÉMICO
Presente

Me permito solicitar a usted, gire sus amables instrucciones para que se continúen los trámites de titulación del C. **JUAN DE DIOS ORELLANA ORELLANA**, número de boleta **2019390120**, pasante de la carrera de Ingeniería Geofísica, quien se titulará por la opción de Tesis Individual, con el título:

**"Operatividad De Programas De Código Abierto En La Ejecución
De Inversión Sísmica"**

Se incluye la lista de sinodales para la toma de protesta.

Presidente	Ing. Roberto Loo Guzmán	5250381	
Secretario	Ing. Efrén Murillo Cruz	2966375	
Vocal 1	Ing. Gabriela De Los Santos Cano	3834703	Asesor interno
Vocal 2	M. en C. Julio César Guarneros López	13042526	
Vocal 3	Ing. Alejandro Mendoza Maravillas	1560419	

Sin más por el momento le envío un cordial saludo.

ATENTAMENTE
"La Técnica Al Servicio de la Patria"


M. en C. JULIO CÉSAR GUARNEROS LÓPEZ
Jefe del Depto. de Innovación Educativa

C.c.p. Archivo
JCGL/spd



DEPARTAMENTO DE
INNOVACIÓN
EDUCATIVA



Folio

DGE y T/204/2023.

2023. Año de Francisco Villa
30 Aniversario de la Declaración sobre
la Eliminación de la Violencia contra la Mujer (ONU)
60 Aniversario del CECyT 7 "Cuahtémoc" y del CENAC
90 Aniversario de la Escuela Superior de Ingeniería Textil
40 Aniversario del CIIDIR, Unidad Oaxaca

Asunto

ASESORIA DE TESIS

CDMX, 18 de julio de 2023.

INC. GABRIELA DE LOS SANTOS CANO

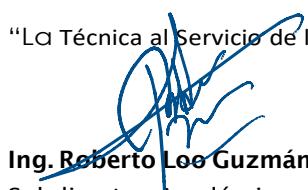
P R E S E N T E

En atención a la solicitud del **C. JUAN DE DIOS ORELLANA ORELLANA**, pasante del Programa Académico de **Ingeniería Geofísica**, con base en el Reglamento de Titulación Profesional del I.P.N., Capítulo II, Artículo 5, Fracción II y el Artículo 7 Inciso I, se le informa que ha sido asignado para asesorar la tesis del pasante mencionado. Por lo que solicito a usted se sirva proponer dentro de un plazo de 30 días calendario, el tema de tesis y contenido a desarrollar por el interesado.

Seguro de contar con su participación, lo saludo cordialmente.

A T E N T A M E N T E

"La Técnica al Servicio de la Patria"


Ing. Roberto Lee Guzmán
Subdirector Académico



c.c.p. M. en C. JULIO CESAR GUARNEROS LÓPEZ- Jefe del Departamento de Innovación Educativa - Interesado

RLG/SBM/mpc



FOLIO

DGE-T/240/2023.

2023. Año de Francisco Villa
30 Aniversario de la Declaración sobre
la Eliminación de la Violencia contra la Mujer (ONU)
60 Aniversario del CECyT 7 "Cuahtémoc" y del CENAC
90 Aniversario de la Escuela Superior de Ingeniería Textil
40 Aniversario del CIIDIR, Unidad Oaxaca

Asunto

Registro de Tema

CDMX, a 19 de septiembre de 2023.

C. JUAN DE DIOS ORELLANA ORELLANA
PASANTE DEL PROGRAMA ACADÉMICO
DE INGENIERÍA GEOFÍSICA
PRESENTE

A continuación, comunico a usted, el tema y contenido que deberá desarrollar en su **Tesis Individual**, indicándole que tiene un año a partir de esta fecha, para elaborarla.

"OPERATIVIDAD DE PROGRAMAS DE CÓDIGO ABIERTO EN LA EJECUCIÓN DE INVERSIÓN SÍSMICA"

INTRODUCCIÓN

- 1.- FUNDAMENTOS TEÓRICOS
- 2.- METODOLOGÍA
- 3.- RESULTADOS
- 4.- CONCLUSIONES
- 5.-APÉNDICE

REFERENCIAS

ATENTAMENTE

"La Técnica al Servicio de la Patria"

Ing. ROBERTO LOO GUZMÁN
Subdirector Académico.

RLG/SBM/mpc



UNIDAD TICOMAN
SUBDIRECCIÓN ACADÉMICA



Folio

DGE Y T/279/2023

2023. Año de Francisco Villa
30 Aniversario de la Declaración sobre
la Eliminación de la Violencia contra la Mujer (ONU)
60 Aniversario del CECyT 7 "Cuauhtémoc" y del CENAC
90 Aniversario de la Escuela Superior de Ingeniería Textil
40 Aniversario del CLIDIR, Unidad Oaxaca

Asunto

Aprobación de Tesis

Ciudad de México, 06 de noviembre de 2023

ING. ROBERTO LOO GUZMAN
SUBDIRECTOR ACADÉMICO
P R E S E N T E

Por este conducto, hacemos constar que el Tema y contenido, por la opción de **TESIS INDIVIDUAL**, con el tema "**OPERATIVIDAD DE PROGRAMAS DE CÓDIGO ABIERTO EN LA EJECUCIÓN DE INVERSIÓN SÍSMICA**", del **C. JUAN DE DIOS ORELLANA ORELLANA** fue revisado y aprobado por los suscritos considerando que está ampliamente desarrollado y puede proceder a su impresión, por lo tanto, la Subdirección a su cargo, puede señalar fecha para realizar el Examen Oral.

TITULARES

ING. ROBERTO LOO GUZMÁN

ING. GABRIELA DE LOS SANTOS CANO

ING. ALEJANDRO MENDOZA MARAVILLAS

ING. EFRÉN MURILLO CRUZ

M. EN C. JULIO CÉSAR GUARNEROS LÓPEZ

*Con gratitud a mis padres
por su incesante apoyo*

Agradecimientos

Al Instituto Politécnico Nacional, mi alma mater, y a mi casa de estudios, la Escuela Superior de Ingeniería y Arquitectura Unidad Ticomán, por brindarme la oportunidad de desarrollarme profesionalmente en sus instalaciones y donde conocí maestros, amigos y compañeros que tuvieron un impacto positivo a mi persona. Gracias a la clase trabajadora de México que, con sus impuestos, sostiene la educación pública de calidad.

A la maestra Gabriela, por sincero interés en mi trabajo, el tiempo dedicado para que este proyecto saliera adelante, gracias a sus comentarios y observaciones.

A mi mamá y papá, por su apoyo constante y aliento a lo largo de mi carrera académica. Su amor y confianza en mí han sido fundamentales en cada paso que he dado en este viaje. Estoy eternamente agradecido por todo lo que han hecho por mí.

A quienes me acompañaron durante todo este transcurso a los cuales les tengo mucho cariño y aprecio: Ramón, Alejandro, Selma♡, Carlos, Demetrio, Raquel, Achoy, Hazel, JuanPa, Yumiko. A mis amigos natales: Andrés, Génesis, Fran, Daniel.

Este trabajo se basa en *software* existente utilizando programas de código abierto la generación, procesamiento y visualización de datos y resultados, aprovecho para agradecer a los desarrolladores de Python y Julia, así como los autores de Devito, JUDI.jl, Matplotlib/PyPlot.jl, SegyIO.jl, PyCall.jl, Equinor-Segyio, y al Laboratorio Sismico para Representación y Modelado - SLIM del Instituto Tecnológico de Georgia, en especial a Mathias Louboutin por atender mis dudas. Y en general a toda la comunidad de código abierto.

Finalmente, mi profundo agradecimiento a Alexandra Elbakyan, la creadora de Sci-Hub, por su valiosa contribución al acceso abierto al conocimiento científico, especialmente para los investigadores y estudiantes del Sur global.

Resumen

El incremento en el poder computacional ha permitido grandes avances en la geofísica, uno de ellos es la implementación de técnicas de inversión de datos sísmicos. El presente trabajo investiga el uso de herramientas computacionales de código abierto en la inversión sísmica, centrándose en las aplicaciones de inversión de forma de onda completa (FWI) y migración de mínimos cuadrados (LSM). Se lleva a cabo una revisión bibliográfica de ambas técnicas y se detallan las ventajas y desventajas en comparación con las técnicas convencionales. Mediante el uso de las herramientas Devito y JUDI se realizan pruebas con el propósito de validar estas técnicas en un orden subsecuente, primero realizando la inversión de forma completa, y al resultado, la migración de mínimos cuadrados.

El estudio se basa en el uso de modelos sintéticos de velocidades: el primero un modelo de capas planas con una caja en el centro, y el segundo, el modelo sintético SEG/EAGE Overthrust. A través de la selección y preparación de los datos sísmicos, se implementan las técnicas de inversión sísmica mencionadas, obteniendo una representación sísmica de cada modelo.

Los resultados obtenidos demuestran la efectividad y aplicabilidad de estas herramientas de código abierto en la inversión sísmica. La inversión de forma de onda completa proporciona una representación detallada y precisa del subsuelo, permitiendo una mejor caracterización de las estructuras geológicas. Por su parte, la migración de mínimos cuadrados ofrece imágenes sísmicas mejoradas y geológicamente coherentes respecto a las técnicas convencionales.

Esta tesis contribuye a demostrar la aplicación de las herramientas de código abierto en inversión sísmica y su utilidad en la enseñanza y aprendizaje de esta.

Abstract

The increase in computational power has allowed significant advances in geophysics, one of which is the implementation of seismic data inversion techniques. This study investigates the use of open-source computational tools in seismic inversion, focusing on the applications of Full Waveform Inversion (FWI) and Least Squares Migration (LSM). A literature review of both techniques is conducted, detailing the advantages and disadvantages compared to conventional techniques. The Devito and JUDI tools are used to implement these techniques sequentially, first performing full waveform inversion and then least squares migration on the result.

The study is based on the use of synthetic velocity models: the first being a model of horizontal layers with an box in the center, and the second being the synthetic SEG/EAGE Overthrust model. Through the selection and preparation of seismic data, the implementation of seismic inversion techniques is carried out to obtain a seismic representation of each study model.

The obtained results demonstrate the effectiveness and applicability of these open-source tools in seismic inversion. Full waveform inversion provides a detailed and accurate representation of the subsurface, allowing for a better characterization of geological structures. On the other hand, least squares migration offers improved seismic images that are geologically coherent compared to conventional techniques.

This study contributes to demonstrating the application of open-source tools in seismic inversion and their utility in teaching and learning in this field.

Índice general

1	Introducción	1
2	Fundamentos teóricos	2
2.1.	Problema directo	2
2.2.	Problema inverso	4
2.3.	Inversión de forma de onda completa	9
2.4.	Migración de mínimos cuadrados	14
2.5.	Herramientas de código abierto	17
3	Metodología	20
3.1.	Datos sísmicos	20
3.2.	Definición del problema sísmico directo	21
3.3.	Definición del problema sísmico inverso	23
3.4.	Definición del problema sísmico inverso lineal	24
3.5.	FWI	25
3.6.	LS-RTM	26
4	Resultados	27
4.1.	Modelo personalizado	27
4.2.	Modelo SEG/EAGE Overthrust	35
5	Conclusiones	43
Anexo A	Apéndice	44
Referencias		45

Introducción

La inversión sísmica es una técnica crucial en la exploración de recursos naturales y la caracterización del subsuelo. En los últimos años, el desarrollo de herramientas computacionales de código abierto ha revolucionado la forma en que se aborda esta disciplina. El presente estudio se enfoca en el análisis y aplicación de dos herramientas de código abierto, Devito y JUDI, en la inversión sísmica, específicamente en las técnicas de inversión de forma de onda completa y migración de mínimos cuadrados.

El objetivo principal de esta investigación es evaluar la efectividad y aplicabilidad de estas herramientas en la resolución de problemas de inversión sísmica, así como evaluar los resultados obtenidos con cada una de ellas. Para ello, se utiliza un enfoque metodológico que involucra la selección y preparación de los datos sísmicos, la implementación de las técnicas mencionadas y el análisis de los resultados obtenidos. Además, se consideran dos modelos sintéticos de velocidades: uno personalizado y el modelo SEG/EAGE Overthrust.

Las herramientas computacionales de código abierto ofrecen ventajas significativas, como la accesibilidad, flexibilidad, reproducibilidad y posibilidad de personalización, lo que las convierte en una opción atractiva para la comunidad científica y los profesionales de la industria. Al comprender y evaluar el desempeño de estas herramientas en la inversión sísmica, se espera contribuir al desarrollo de mejores prácticas y promover su adopción en futuros proyectos y la enseñanza de estos.

Fundamentos teóricos

Podemos definir la inversión como un procedimiento para obtener modelos que describan adecuadamente un conjunto de datos. En el caso de datos geofísicos, las observaciones muestran los efectos de las propiedades de las rocas en los fenómenos físicos, como la propagación de ondas sísmicas. Debido a que la inversión geofísica nos permite extraer información de modelos geológicos a partir de esta serie de datos, este procedimiento es de suma importancia e interés en ciencias de la tierra.

Antes de discutir la inversión sísmica, o problema sísmico inverso, es necesario comprender el problema sísmico directo, ya que el proceso de inversión guarda una estrecha relación con el modelado directo.

2.1. Problema directo

Sheriff (2002) define modelo como “un concepto a partir del cual se pueden deducir efectos para comparar observaciones; en general se utiliza para desarrollar una mejor comprensión de dichas observaciones. El modelo puede ser conceptual, físico o matemático”.

El modelado directo es un proceso que consiste en crear modelos sísmicos sintéticos a partir de información geológica. A menudo para simular el resultado de un estudio sísmico y estimar la respuesta sísmica esperada del modelo de tierra. El problema directo utiliza las propiedades de las rocas del modelo de tierra como entrada, simula la respuesta de la propagación de ondas a medida que las ondas viajan por medio del modelo (Lines & Newrick, 2004). El resultado se compara con secciones sísmicas procesadas para verificar la interpretación que se utilizó para la creación del modelo.

En la sismología de exploración, uno de los métodos más populares utilizados para resolver la ecuación de onda es el método por diferencia finitas, el cual por medio de un *stencil* (patrón de operaciones matemáticas aplicadas a puntos vecinos en una malla) se calculan los valores de velocidad o presión en cada punto de la malla (modelo discretizado) simulando la propagación de ondas a través del subsuelo.

El siguiente ejemplo se desarrolló con ayuda de Devito (Louboutin et al., 2019), que permite resolver la ecuación de onda mediante diferencias finitas utilizando expresiones simbólicas. El modelo consiste en seis capas horizontalmente estratificadas con velocidades de 1.5, 1.9, 2.3, 2.7, 3.1, y 3.5 kilómetros por segundo de la capa más somera a la más profunda, respectivamente. La geometría de adquisición consta de 51 receptores posicionados a lo largo del eje horizontal, desde 0 km hasta 3 km con un distanciamiento de 60 metros entre cada uno, a una profundidad de 20 metros, con un tiempo de grabación es de 3000 milisegundos. Con una única fuente en el centro del modelo, 1.5 km en el eje horizontal, a una profundidad de 50 metros (Figura 2.1).

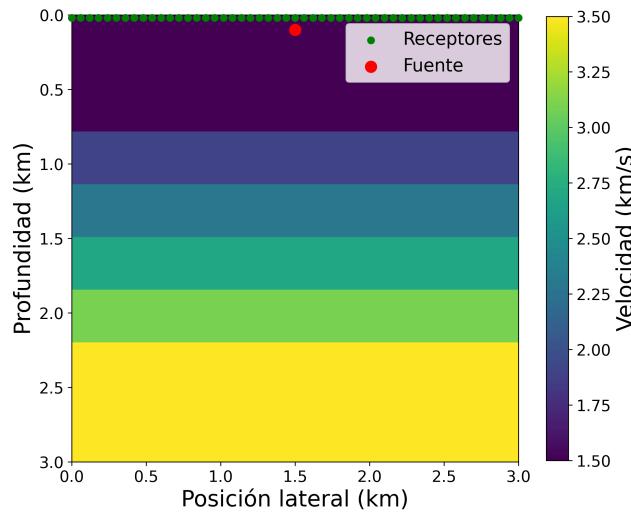


Figura 2.1: Modelo de velocidades

Simulando la propagación de onda a través del modelo, podemos generar una película de instantáneas del campo de onda, obteniendo imágenes de dicho campo en el tiempo discretizado, permitiendo observar el arribo de las ondas directas, las ondas reflejadas e incluso múltiples (Figura 2.2).

El resultado del modelado directo se observa en el registro de tiro común (datos) en la Figura 2.3.

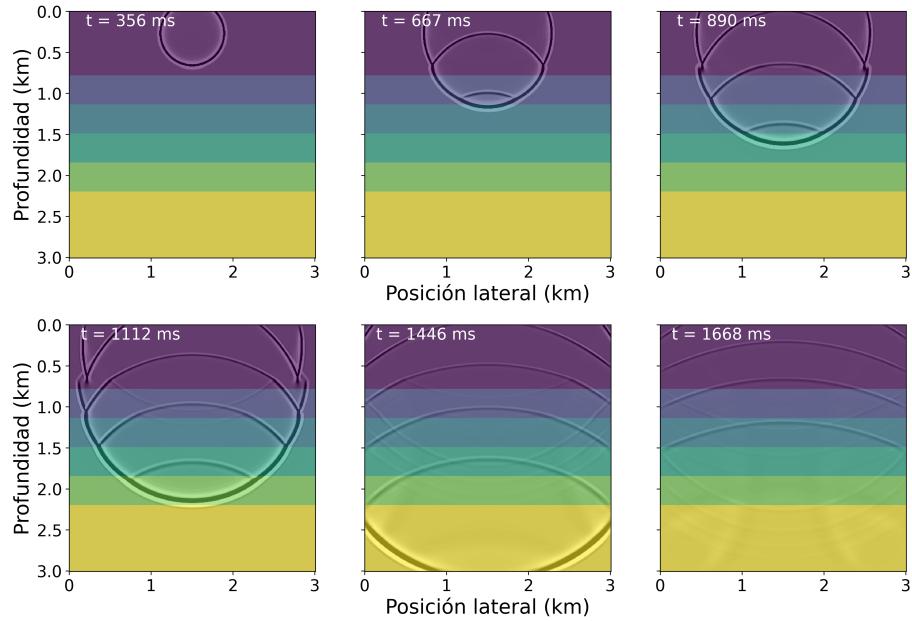


Figura 2.2: Simulación del campo de ondas

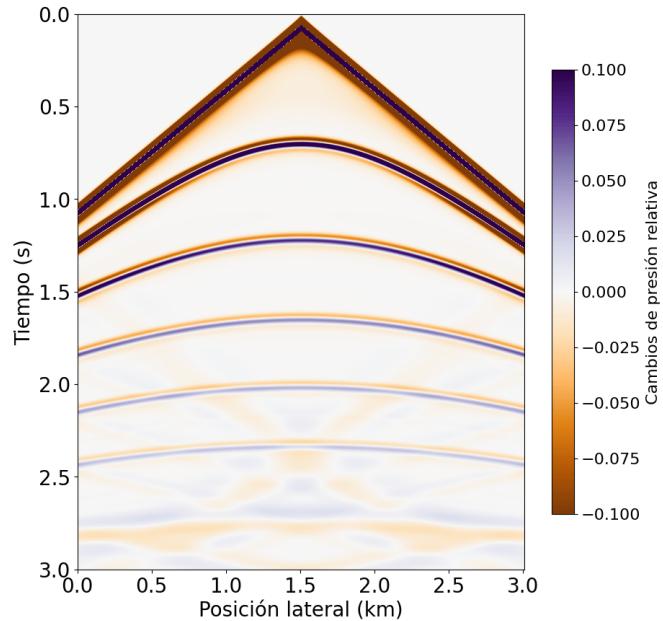


Figura 2.3: Registro de tiro común

2.2. Problema inverso

El proceso de inversión se relaciona estrechamente con el modelado directo, la relación de ambos procesos puede visualizarse en la Figura 2.4.

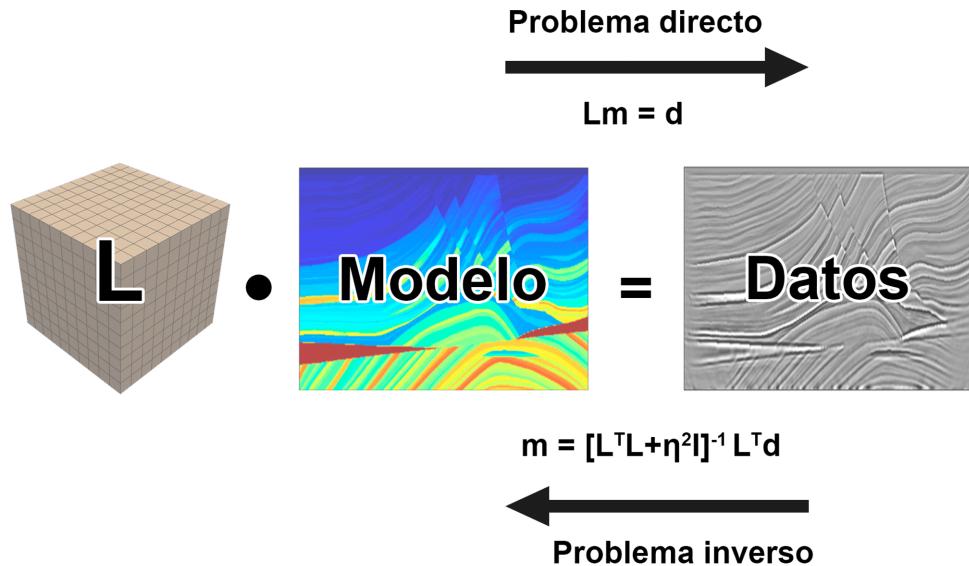


Figura 2.4: Modelado directo e inverso

El problema inverso discreto en geofísica se puede definir de manera general como la inversión del vector de modelo \mathbf{m} de los datos registrados en campo \mathbf{d} , donde \mathbf{d} está relacionado con \mathbf{m} por

$$\mathbf{L}\mathbf{m} = \mathbf{d} \quad (2.1)$$

donde \mathbf{L} representa el operador de modelado directo no lineal que depende implícitamente de \mathbf{m} y, que predice los datos \mathbf{d} del modelo \mathbf{m} . Si el operador de modelado directo es no lineal, por lo general se linealiza mediante un modelo de fondo \mathbf{m}_0 .

De acuerdo con Hadamard (1902) los modelos matemáticos correctamente planteados, como la versión linealizada de la Ecuación 2.1, de sistemas físicos deben tener las siguientes características:

1. **Existe una solución.** Para los problemas geofísicos, no existe una solución porque los datos tienen mucho ruido, y conducen a un sistema de ecuaciones inconsistente y sobredeterminado. Esto se resuelve encontrando una solución que minimice, por ejemplo, una suma de los residuos al cuadrado de los datos y un parámetro de ajuste (mínimos cuadrados amortiguados).
2. **Solución única.** En los experimentos geofísicos, se cuentan con fuentes y receptores ubicados en una pequeña parte del objeto de estudio, esto a menudo resulta en un espacio nulo no vacío para \mathbf{L} y, por ende, en valores propios cero para $\mathbf{L}^T \mathbf{L}$. Para mitigar este problema se utiliza la regularización, técnicas para abordar problemas de inversión mal condicionados, *ill-posed*, (cuya solución no es única) o con insuficiente información. Estas técnicas

agregan información adicional o imponen restricciones en el proceso de inversión, con el objetivo de obtener soluciones estables y consistentes, como la regularización de Tikhonov.

3. **La solución siempre debe ser dependiente de los datos, de lo contrario la solución es inestable.** Una solución inestable es aquella en la cual pequeños cambios en el nivel de ruido de los datos resultan en cambios discontinuos en el modelo.

Para la inversión geofísica, autores como Fletcher (1987) y Meyer et al. (2004) proponen replantear el problema de manera que se busque el modelo óptimo que minimice el desajuste de datos y las funciones de penalización del modelo, conocidas como funciones objetivo. Por lo general, la función objetivo es una función no lineal de los parámetros del modelo, el procedimiento de solución está en la clase de métodos no lineales. Schuster (2017) muestra seis pasos para representar la implementación de la mayoría de los algoritmos de inversión geofísica, mostrados a continuación.

1. **Discretización del modelo \mathbf{m} .** El modelo, ya sea de velocidad o lentitud, de resistividad eléctrica, de gravedad o de la propiedad a medir, de tierra se discretiza en una malla. Por ejemplo, la Figura 2.5 muestra la discretización de un modelo de velocidades. Cada punto de la malla es una celda asociada a un valor desconocido de la propiedad. Por conveniencia se asume que el valor de la propiedad es constante en la celda j -ésima, lo que resulta en un vector de modelo \mathbf{m} de tamaño $N \times 1$ de N valores desconocidos.

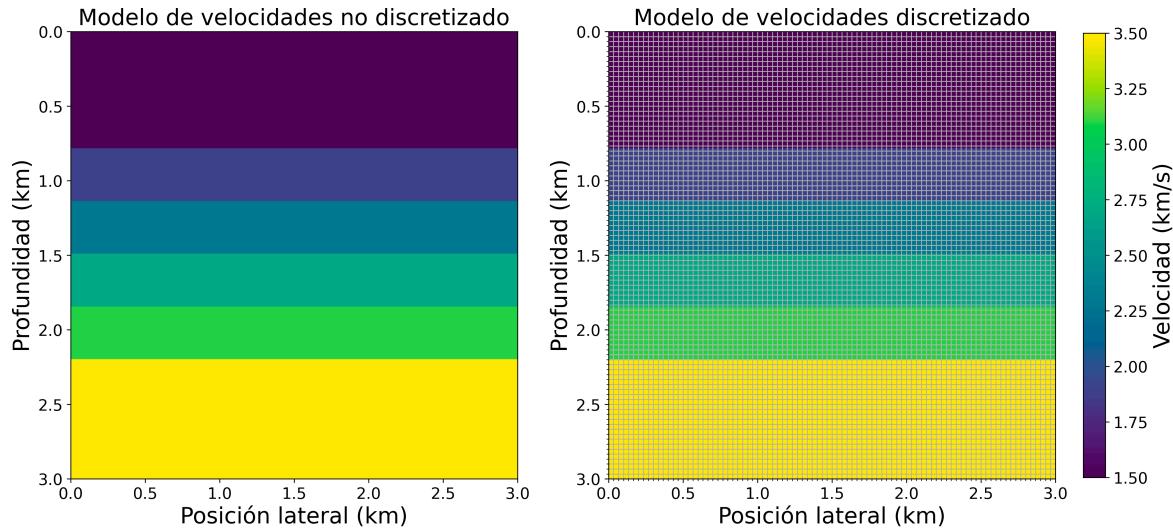


Figura 2.5: Discretización de un modelo de velocidades

2. **Discretización de los datos \mathbf{d} .** Para cada conjunto de M se observan i -ésima mediciones, formando un vector de datos \mathbf{d} de tamaño $M \times 1$.

3. **Operador de modelado discreto L.** El problema inverso se resuelve asumiendo un modelo inicial y un operador de modelado \mathbf{L} para generar los datos estimados. Los datos a lo largo de la i -ésima medición es la suma de todas las mediciones individuales d_i en cada celda:

$$\sum_j l_{ij} m_j = d_i \quad \text{o} \quad \mathbf{Lm} = \mathbf{d} \quad (2.2)$$

\mathbf{L} es una matriz de tamaño $M \times N$

4. **Linealización.** La Ecuación 2.2 es una ecuación no lineal, l_{ij} depende del modelo utilizado. Por ende, \mathbf{L} depende implícitamente de \mathbf{m} , al invertir el sistema de ecuaciones no se llegará a una respuesta plausible, al menos que el modelo inicial se asemeje al modelo real. Esto se resuelve comenzando con un modelo \mathbf{m}_0 que se acerque al modelo real, y linealizar la relación entre los datos y el modelo. Así, los datos se pueden invertir para obtener un modelo más preciso \mathbf{m}_1 . Utilizando \mathbf{m}_1 como nuevo modelo inicial, este procedimiento se repite hasta que que los datos y el modelo converjan. Para linealizar, se comienza definiendo el j -ésimo parámetro de modelo como m_j , y expandiendo la i -ésima medición de datos d_i a una serie de Taylor de primer orden, sobre un modelo inicial cercano al modelo real:

$$d_i(\mathbf{m}) \approx d_i(\mathbf{m}_0) + \sum_j \left[\frac{\partial d_i(\mathbf{m})}{\partial m_j} \right] \Big|_{\mathbf{m}_0} \delta m_j \quad (2.3)$$

en donde $\mathbf{m} = \mathbf{m}_0 + \delta\mathbf{m}$. Reagrupando los términos, obtenemos la ecuación linealizada

$$\delta d_i(\mathbf{m}) = \sum_j \left[\frac{\partial d_i(\mathbf{m})}{\partial m_j} \right] \Big|_{\mathbf{m}_0} \delta m_j \quad (2.4)$$

en la forma matricial

$$\delta \mathbf{d}(\mathbf{m}) = \mathbf{L} \delta \mathbf{m} \quad (2.5)$$

Aquí, el residual $\delta d_i = [d_i(\mathbf{m}) - d_i(\mathbf{m}_0)]$ es el i -ésimo componente de la diferencia entre los datos observados $d_i(\mathbf{m})$ y los datos calculados $d_i(\mathbf{m}_0)$. La perturbación en el modelo $\delta\mathbf{m} = \mathbf{m} - \mathbf{m}_0$ es la diferencia entre el modelo actual \mathbf{m} y el modelo inicial \mathbf{m}_0 . La matriz \mathbf{L} es la matriz Jacobiana y sus elementos son la primera derivada parcial $l_{ij} = \partial d_i(\mathbf{m}_0)/\partial m_j$, también conocida como la derivada Fréchet, determina la sensibilidad de los i -ésimo datos a las perturbaciones del modelo en la j -ésima celda.

5. **Solución regularizada.** Los datos grabados contienen ruido, lo que deriva en un sistema de ecuaciones inconsistente y sobredeterminado. Además, la solución puede ser inestable. Para remediar parcialmente estos problemas, buscamos un modelo que minimice de la mejor forma la función objetivo ϵ ,

que es la suma del término de penalidad y la norma p del residual elevado a la p -ésima potencia:

$$\epsilon = \frac{1}{p} \left\| \overbrace{\mathbf{L}\delta\mathbf{m} - \delta\mathbf{d}}^{\text{vector residual}} \right\|_p^p + \eta^2 g(\mathbf{m}), \quad (2.6)$$

donde η^2 es un pequeño escalar positivo y $g(\mathbf{m})$ una función de penalidad que se vuelve más pequeña a medida que el modelo estimado se acerca al modelo real, p es un entero positivo y el vector residual $\mathbf{L}\delta\mathbf{m} - \delta\mathbf{d}$ es la diferencia entre los datos calculados y observados. El término de penalidad a veces se expresa como $g(\mathbf{m}) = \frac{1}{p} \|\mathbf{m} - \mathbf{m}'\|_p^p$ para un modelo \mathbf{m}' . A menudo, $p \rightarrow 2$ y la función objetivo se determina para ser la suma del cuadrado del residual y el término de penalidad $\|\mathbf{L}\delta\mathbf{m} - \delta\mathbf{d}\|^2 + \eta^2 \|\delta\mathbf{m}\|^2$. En este caso, la solución regularizada se conoce como mínimos cuadrados amortiguados

$$\delta\mathbf{m} = [\mathbf{L}^\top \mathbf{L} + \eta^2 \mathbf{I}]^{-1} \mathbf{L}^\top \delta\mathbf{d} \quad (2.7)$$

donde \mathbf{I} es la matriz identidad. Este proceso de solución regularizada se conoce como el método de Leverberg-Marquardt.

6. **Solución regularizada iterativa.** Los datos tienen una relación no lineal con el modelo, por tanto, la solución de la Ecuación 2.7 se encuentra mediante un esquema iterativo

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha [\mathbf{L}^\top \mathbf{L} + \eta^2 \mathbf{I}]^{-1} \mathbf{L}^\top \delta\mathbf{d}^{(k)} \quad (2.8)$$

donde α es el tamaño del paso que se utiliza para actualizar el modelo en cada iteración para acercarse a la solución óptima. $\mathbf{m}^{(k)}$ representa el modelo en la k -ésima iteración. La matriz $[\mathbf{L}^\top \mathbf{L}]$, conocida como matriz Hessiana \mathbf{H} por lo general, es muy costosa de calcular, almacenar, e invertir, por ejemplo, el modelo de velocidad asociado a una adquisición sísmica puede requerir una malla de $1000 \times 1000 \times 1000 = 10^9$ puntos de valores desconocidos de velocidad (\mathbf{L}), entonces la matriz Hessiana ($\mathbf{H} = \mathbf{L}^\top \mathbf{L}$), requerirá el cálculo y almacenamiento de una matriz de 10^{18} elementos. A menudo es aproximada por sus componentes diagonales $[\mathbf{L}^\top \mathbf{L}]_{ij} \approx [\mathbf{L}^\top \mathbf{L}]_{ii} \delta_{ij}$ para dar como resultado, la solución precondicionada del descenso más rápido (*steepest descent*)

$$m_i^{(k+1)} = m_i^{(k)} - \frac{\alpha [\mathbf{L}^\top \delta\mathbf{d}^{(k)}]_i}{[\mathbf{L}^\top \mathbf{L}]_{ii} + \eta^2} \quad (2.9)$$

Cuando hablamos de inversión sísmica, la podemos definir como el procedimiento para reconstruir las propiedades del subsuelo a partir de datos sísmicos, por ejemplo, reflexión, refracción, tomografía, entre otros. (Aki et al., 1977; Bishop et al., 1985; Tarantola, 2005). La Figura 2.6 muestra un diagrama de

Modelado e inversion de datos sísmicos

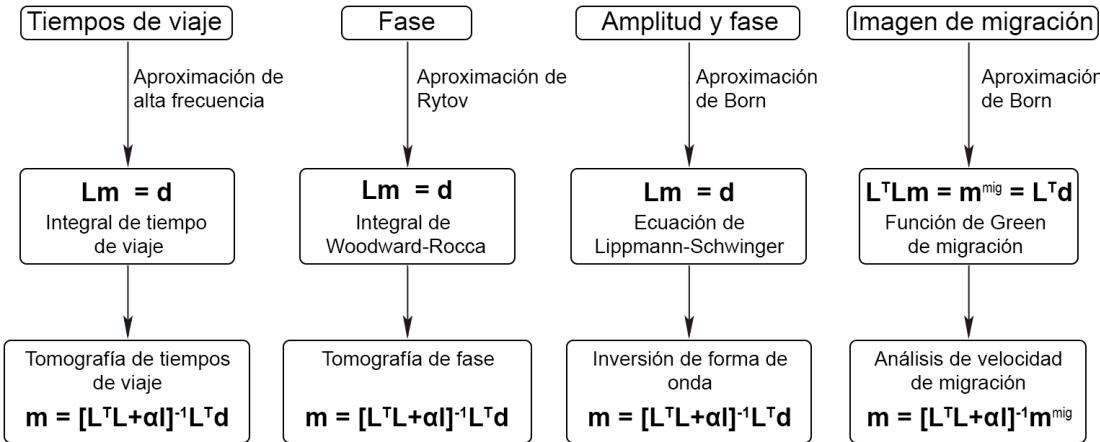


Figura 2.6: Diagrama de inversión para cuatro tipos de datos sísmicos por medio de mínimos cuadrados no lineales. Modificado de Schuster (2017)

organización para la inversión de cuatro tipos de datos sísmicos utilizando el método de mínimos cuadrados no lineales. Al ser una inversión no lineal, la función objetivo se encuentra plagada de muchos mínimos locales, por ende, una solución iterativa tiende a estancarse en un mínimo local y nunca alcanza el mínimo global, o el modelo real. En el presente trabajo, discuto la inversión de forma de onda completa (*full waveform inversion - FWI*) y la migración de tiempo reverso por mínimos cuadrados (*least-squares reverse time migration - LSRTM*).

2.3. Inversión de forma de onda completa

Conocido como FWI, es un método avanzado de representación sísmica de alta resolución para la construcción de modelos de velocidades que ofrece calidad previamente inalcanzable. En contraste con los métodos tradicionales, que se basan en la información de los tiempos de llegada de las ondas sísmicas, FWI está basado en utilizar todo el contenido de las trazas sísmicas para extraer parámetros físicos del medio muestreado por las ondas sísmicas.

El objetivo de la inversión de onda completa es minimizar la discrepancia entre los datos sísmicos observados y los datos sísmicos calculados, mejorando de forma iterativa el modelo de velocidad del subsuelo (Hill & Rüger, 2019). De manera ideal, el resultado de FWI será un modelo de velocidad en donde las trazas sísmicas modeladas son iguales a las trazas sísmicas observadas, tanto en tiempos de reflexión como en las amplitudes. Esto requiere de un alto poder computacional y un correcto modelo inicial, ya que es muy sensible a las condiciones iniciales,

además, requiere de datos de alta resolución y cobertura (ancho de banda y *offsets* largos) para obtener resultados precisos.

El pionero en los fundamentos matemáticos para la inversión de forma de onda completa es Albert Tarantola (Tarantola, 1984), que reformula el problema de imagen de migración de Claerbout (1976) como un problema de optimización local, el objetivo es minimizar, mediante mínimos cuadrados, la discrepancia entre los datos observados y los datos modelados. En su trabajo, Tarantola muestra que el gradiente de la función de ajuste a lo largo del cual se busca el modelo de perturbaciones, puede construirse mediante la correlación cruzada entre el campo de onda incidente emitido desde la fuente y los campos de onda residuales retropropagados (desde el subsuelo hasta la superficie). El modelo de perturbaciones obtenido después de la primera iteración de la optimización local, se asemeja a una imagen migrada obtenida mediante migración de tiempo reverso (RTM). Una diferencia es que el campo de onda sísmico registrado en el receptor se retropropaga en la migración de tiempo reverso, mientras que en la inversión de forma de onda de Tarantola (1984) se retropropaga la discrepancia de los datos. Cuando se agregan las perturbaciones de velocidad al modelo de velocidad inicial, se obtiene un modelo de velocidad actualizado, que se utiliza como modelo inicial para la siguiente iteración de minimización de la función de discrepancia.

La información completa contenida en los sismogramas puede ser considerada en la optimización, llegando a la inversión de forma de onda completa, donde se modela la ecuación de onda completa en cada iteración de la optimización, en el modelo final de la iteración previa. En la optimización de FWI, se involucran todos los tipos de ondas, incluyendo las de refracción, las reflexiones hipercríticas y las ondas multiesparcidas, como los múltiplos (Virieux & Operto, 2009).

Hay muchas razones por las cuales FWI no ha sido reconocido como una técnica eficiente en representación sísmica, una de ellas debido a que las aplicaciones en sus inicios, se limitaron a datos de reflexión sísmica. En adquisiciones de *offsets* cortos, los campos de ondas sísmicos no son sensibles a longitudes de onda intermedias, por ende, la optimización no puede reconstruir de manera adecuada el modelo de velocidades real de manera iterativa. Solo cuando un modelo inicial suficientemente preciso es proporcionado, se converge al modelo de velocidades a través de dichas actualizaciones en cada iteración.

Unicamente con el beneficio de *offsets* largos y de transmisión de datos para la reconstrucción de longitudes de onda intermedias y grandes de la estructura, FWI ha alcanzado su madurez (Mora, 1987, 1988; Pratt, 1999). FWI intenta caracterizar un espectro de número de onda amplio y continuo en cada punto del modelo, unificando las tareas de construcción del macro modelo y migración en un solo procedimiento (Virieux & Operto, 2009). Sin embargo, la aplicación de FWI a datos de *offsets* largos sigue siendo un desafío debido a las crecientes no linealidades introducidas por los campos de onda propagados a lo largo de varias decenas de

longitudes de onda y diferentes ángulos de incidencia (Sirgue, 2006).

Podemos ver la inversión de forma de onda completa como un flujo de trabajo que consiste en los pasos siguientes (la Figura 2.7 muestra a grandes rasgos este flujo):

1. Generar los datos (disparos) sintéticos mediante el modelado directo sobre el modelo de velocidades inicial. Hablando en términos generales, con un modelo de velocidades perfecto, los datos sintéticos generados son efectivamente iguales a los datos observados. Esto no sucede en la realidad.
2. Calcular la diferencia aritmética entre las trazas sísmicas de los datos observados (registro de campo) y los datos calculados. El desajuste entre los datos solo debe incluir los efectos causados por el error de velocidad en el modelo actual.
3. Realizar un proceso similar a la migración de tiempo reverso, llamado retropropagación, obteniendo así, una medida del error en el modelo de velocidad actual, que muestra donde el modelo es demasiado rápido o lento (gradiente).
4. Determinar, luego de una búsqueda demandante, el valor óptimo del escalar a aplicar al campo de gradiente para convertir el campo de gradiente a cambios de velocidad medidas en unidades físicas y actualizar el modelo. Este escalar es necesario pues el error del campo de gradiente es relativo no absoluto.
5. Añadir todas estas actualizaciones al modelo de velocidades inicial y generar un nuevo modelo de velocidades.

Ejemplo: Modelo BP 2004

Peng et al. (2018) muestran la capacidad de la inversión de forma de onda completa con la información correcta, haciendo uso del modelo BP 2004 (Billette & Brandsberg-Dahl, 2005). La Figura 2.8a muestra un fragmento del modelo de velocidad, en el que existe una geometría muy compleja del cuerpo salino. El modelo inicial proporcionado para el proceso de FWI es una versión muy suavizada sin la presencia del cuerpo salino (Fig. 2.8b). El resultado de la inversión de forma de onda completa se muestra en la Figura 2.8c, en el que se observa un modelo final muy cercano al modelo real, con la excepción de los límites definidos de las estructuras. Para lograr esto, los autores generaron los datos sísmicos para el proceso de inversión con *offsets* ultralargos, 30 km, y con presencia de señal de baja frecuencia, hasta 0.5 Hz. En la realidad, estos parámetros son casi imposibles de satisfacer, especialmente las bajas frecuencias. La frecuencia más baja utilizable es de 4 Hz para adquisiciones marinas, llegando, en extremos, hasta los 2 Hz (Dellinger, 2016; Mandroux et al., 2013). Las Figuras 2.9b y 2.9c muestran los

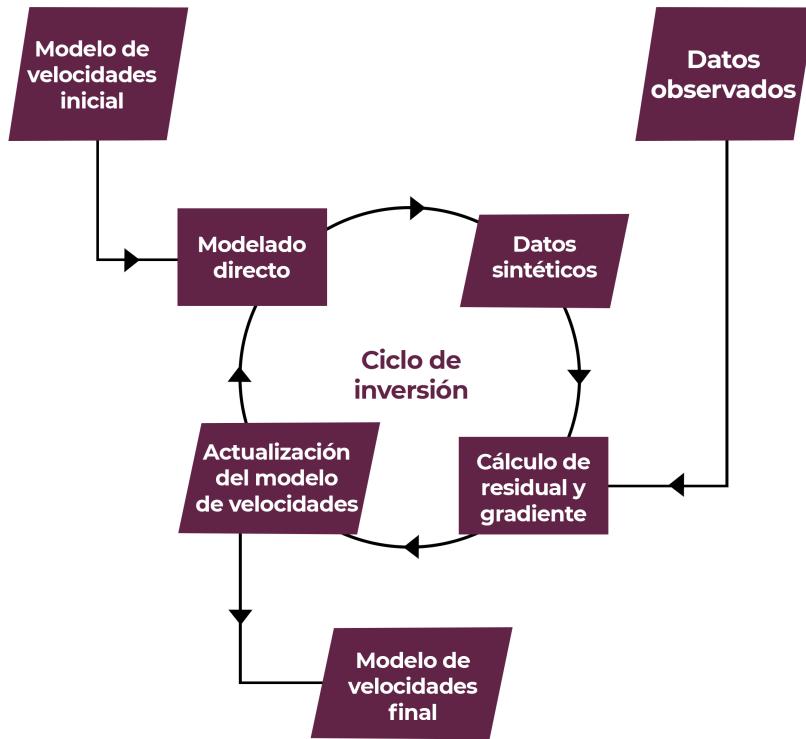


Figura 2.7: Diagrama de flujo de FWI

resultados de inversión de datos sísmicos con una frecuencia mínima utilizable de 2 y 4 Hz respectivamente. Incluso con *offsets* muy lejanos de 30 km, la inversión fracasa en recuperar el modelo de velocidad real.

Los autores señalan que, hasta cierto punto, la falta de frecuencias bajas podría superarse si se consiguiera un mejor modelo inicial con los métodos tradicionales. La Figura 2.10a muestra un nuevo modelo inicial con una interpretación errónea del cuerpo salino, en la parte superior. En contraste con la inversión del primer modelo inicial, la inversión de forma de onda completa corrige la interpretación errónea con datos de 2 Hz y *offsets* ultralargos de 30 km (Fig. 2.10b). La Fig. 2.10c muestra el resultado de FWI con datos de 4 Hz y *offsets* de 9 km, donde la inversión presenta fallas en recuperar el modelo real en la parte más profunda, debido a la falta de penetración de las ondas continuamente refractadas.

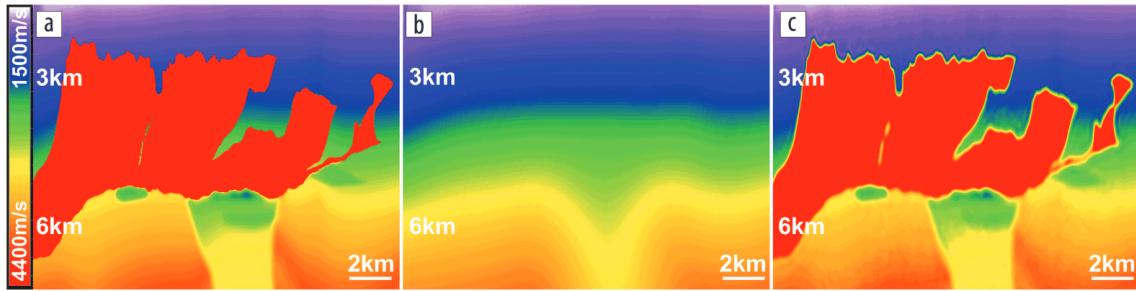


Figura 2.8: a) Parte del modelo de velocidades BP 2004. b) Modelo (inicial) fuertemente suavizado. c) Modelo resultado de FWI con frecuencia a partir de 0.5 Hz. Tomada de Peng et al. (2018)

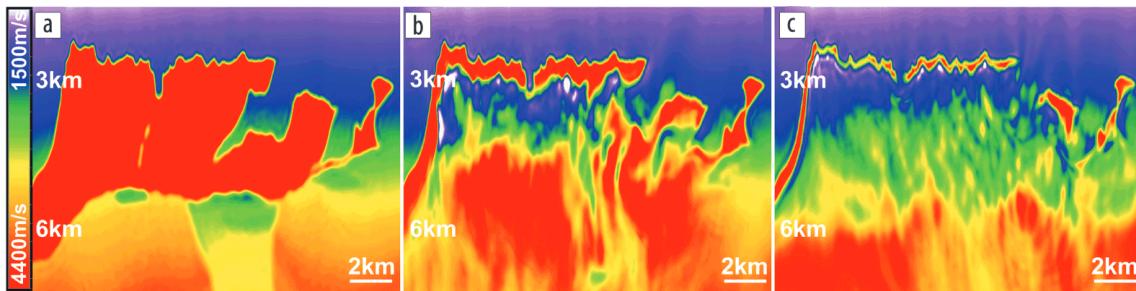


Figura 2.9: Modelo resultado de FWI con frecuencia a partir de a) 0.5 Hz. b) 2 Hz. c) 4 Hz. Tomada de Peng et al. (2018)

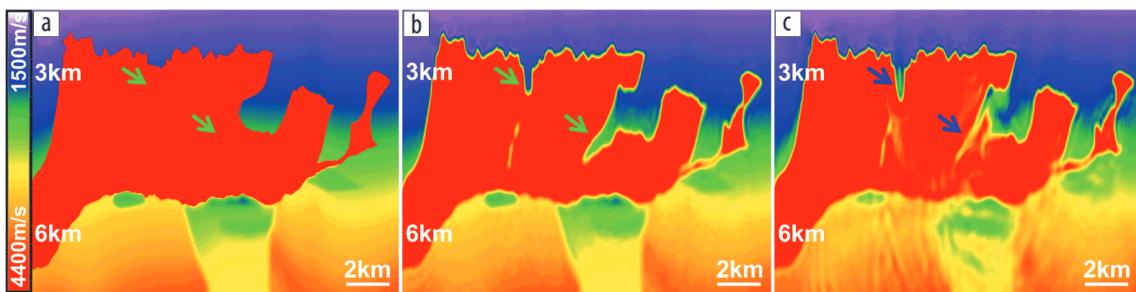


Figura 2.10: a) Modelo de velocidades BP 2004 con errónea interpretación de la sal (flechas verdes). b) Resultado de FWI con datos de 2 Hz y *offset* de 30km. c) Resultado de FWI con datos de 4 Hz y *offset* de 9km. Tomada de Peng et al. (2018)

2.4. Migración de mínimos cuadrados

Antes de discutir la migración de mínimos cuadrados, es conveniente hablar sobre la migración de tiempo reverso.

La migración es la técnica de procesamiento que posiciona las reflexiones sísmicas en su posición correcta en el subsuelo (Bording & Lines, 1997). La migración de tiempo reverso, tiene sus orígenes en la 52^a Reunión Anual de la Sociedad de Geofísicos de Exploración (*Society of Exploration Geophysicists*) presentada por McMechan (1982) como migración de tiempo reverso, donde se hace uso de diferencias finitas para resolver la ecuación de onda con la reversión en el tiempo de las trazas sísmicas utilizadas como fuentes variables en el límite de la superficie. Esto permite que el campo de onda pueda ser retropropagado al punto en profundidad donde las reflexiones se originaron. Las ideas de retropropagación fueron explicadas por McMechan (1983) usando el ejemplo del punto refractor. Bording y Lines (1997) señala que “la belleza de la migración de tiempo reverso es que es tan general como la solución de ecuación de onda”. Ya que la ecuación de onda es indiferente a si el tiempo avanza o retrocede, las ondas pueden propagarse hacia atrás en el tiempo (de manera computacional). Glanz (1994) mostró que esta técnica tiene aplicaciones fuera de la exploración geofísica, el Laboratorio de Ondas y Acústica de la Universidad de París construyó un aparato que digitaliza una onda sonora entrante, la invierte y la devuelve a su fuente por su trayectoria original, como si el tiempo retrocediera. En resumen, la migración de tiempo reverso es una técnica de migración antes de apilar que utiliza la ecuación de onda completa, lo que permite considerar el viaje de las ondas en cualquier dirección, su implementación consta de tres pasos: sobre el modelo de velocidades dado, llevar a cabo las propagaciones de ondas directas por cada fuente, resultando en una serie de sismogramas por cada receptor para cada una de las fuentes. Luego realizar la retropropagación para cada una de las fuentes, en las que los receptores actúan como fuentes, y utilizan como función fuente a los sismogramas registrados en las propagaciones del primer paso. A medida se realizan las retropropagaciones se consideran los campos de ondas obtenidos en ambas simulaciones mediante una condición de imagen. Sumando los resultados obtenidos para las fuentes, se obtiene la imagen de migración RTM para el modelo de velocidades dado.

Gracias al incremento notorio en el poder de cómputo, el procesamiento sísmico lentamente está cambiando a flujos de trabajo que hacen uso de la inversión, como ya vimos para la estimación de velocidad, FWI, sino también para la estimación del modelo de reflectividad (migración), este es el caso de la migración de mínimos cuadrados.

Las imágenes de migración convencionales (Kirchhoff, RTM, etc.) pueden presentar:

1. Artefactos de migración causados por submuestreo o geometría irregular de adquisición.
2. Resolución deficiente producto de una apertura de grabación limitada.
3. Artefactos causados por la inestabilidad de la ondícula fuente.
4. Degradación de la amplitud debida a la dispersión geométrica, atenuación y variaciones en la iluminación del subsuelo.

Huang et al. (2014) señalan que estos problemas pueden ser mitigados, en parte, por la migración de mínimos cuadrados (LSM), conocida como inversión sísmica linealizada, que tiene por objetivo, invertir linealmente, datos sísmicos para la distribución de reflectividad del subsuelo. Dong et al. (2012) demostraron que la migración de mínimos cuadrados genera una imagen de reflectividad correspondiente a la geología, con amplitudes físicas, superando el problema de la degradación de amplitudes.

La migración de mínimos cuadrados es un algoritmo de representación sísmica equivalente a la inversión linealizada de la forma de onda, que trata de encontrar la mejor imagen de reflectividad \mathbf{m} dado un modelo fijo de velocidad de fondo suavizado (Huang et al., 2014). Esto se logra minimizando la siguiente función objetivo:

$$\epsilon = \|\mathbf{d} - \mathbf{L}\mathbf{m}\|^2 + \eta^2 g(\mathbf{m}) \quad (2.10)$$

donde \mathbf{d} son los registros en campo de reflexión, \mathbf{L} representa el operador de modelado directo lineal, que depende del modelo de velocidad de fondo suavizado, y $\eta^2 g(\mathbf{m})$ es una regularización que acerca a la solución a adoptar características específicas de \mathbf{m} .

Huang et al. (2014) muestran, mediante ejemplos, los beneficios de la migración de mínimos cuadrados en comparación con los métodos convencionales. Haciendo uso del modelado de Kirchhoff, generan conjuntos de trazas sísmicas sintéticas para un punto dispersor enterrado a una profundidad de 500 m, para luego migrarlos mediante Kirchhoff y LSM. La imagen generada por la migración de Kirchhoff (Fig. 2.11a) cuenta con muchos artefactos causados por el submuestreo y una apertura limitada de datos, algunas veces denominado huella de adquisición. Por otro lado, la imagen de LSM (Fig. 2.11b) minimiza estos artefactos de submuestreo, y logra una resolución espacial de acuerdo con el ancho de banda completo del espectro de la fuente, si el modelo de velocidad es suficientemente preciso.

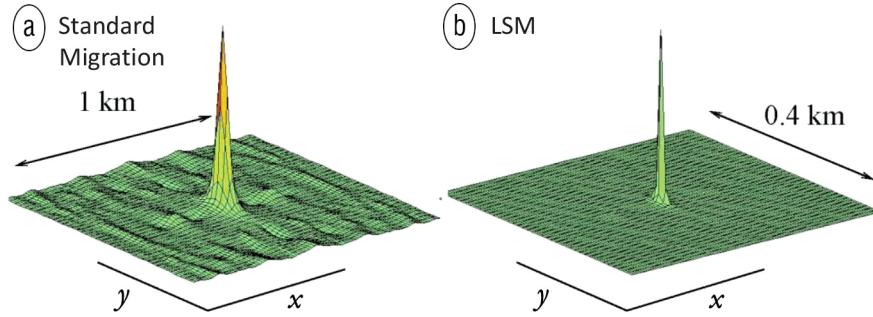


Figura 2.11: Funciones de dispersión de la a) migración de Kirchhoff b) LSM para un punto dispersor enterrado a 500 m. Los datos constan de 176 conjuntos de trazas, cada uno con 176 trazas registradas en superficie, y el eje vertical representa la amplitud de migración. Tomada de Huang et al. (2014)

La migración de tiempo reverso por mínimos cuadrados (LS-RTM) es una técnica de inversión sísmica, que tiene como primera solución la imagen obtenida por la migración de tiempo reverso. Este modelo de reflectividad es usado para generar los datos sintéticos (modelado directo), este paso es una de-migración, con el modelado de Born. El residual es calculado sustrayendo los datos sintéticos de los datos de campo. Luego, el gradiente es calculado usando RTM convencional con el residual. El gradiente se usa para ajustar el modelo de reflectividad. Este procedimiento, el cual se puede observar en la Figura 2.12, se puede repetir hasta alcanzar un nivel deseado de residual o un número máximo de iteraciones. Por cada iteración de LS-RTM, se requiere una RTM y un modelado directo.

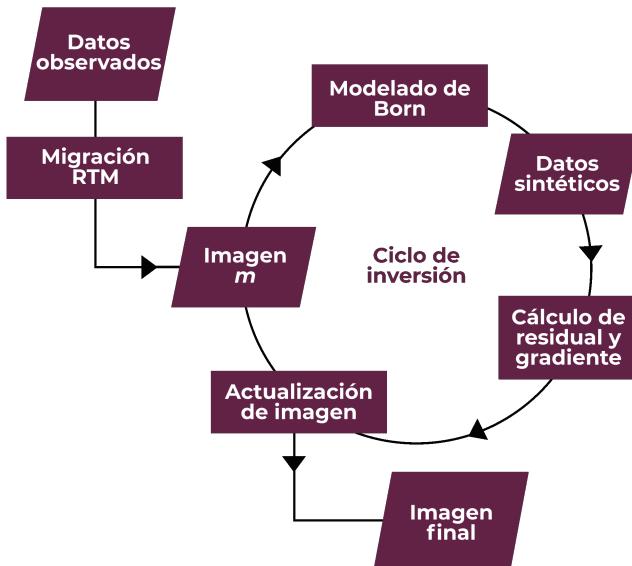


Figura 2.12: Flujo de trabajo de LS-RTM

Ejemplo: Datos de campo - Cuenca Santos, Brasil

Uno de los casos de estudios disponibles en el sitio web de la empresa noruega PGS es el de la Cuenca Santos ubicada en Brasil. Los yacimientos de esta cuenca se localizan en acumulaciones de carbonatos presalinos bajo secuencias evaporíticas (Arasanipalai et al., 2020). En la Figura 2.13 observamos una comparación de migración de tiempo reverso (RTM) y migración de mínimos cuadrados (LSM). Se aprecia un incremento en la resolución de los reflectores detallando varias facies sísmicas en el yacimiento presalino, incluyendo una mejoría en la definición del sistema de fallas.

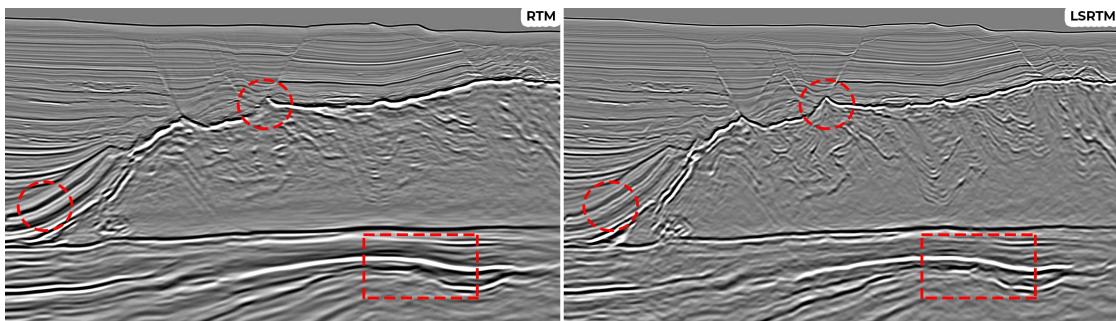


Figura 2.13: Comparación de migración a) RTM y b) LSM para datos de la cuenca Santos, Brasil. LSM mejora la resolución de los reflectores del yacimiento presalino. Modificada de PGS (2020).

2.5. Herramientas de código abierto

Hasta el momento, se mencionó el incremento en el poder computacional que ha permitido llevar a cabo flujos de trabajo de inversión de datos sísmicos. En el presente trabajo, la totalidad de implementaciones y resultados se obtuvieron haciendo uso de una serie de herramientas computacionales, que pertenecen al grupo de código abierto y libres. Además, adoptando el principio de reproducibilidad en la investigación, los algoritmos, datos y resultados, están disponibles para cualquier usuario que desee hacer uso de ellos.

Existe una herramienta fundamental, que permite la reproducibilidad de resultados, Jupyter. Creada por Fernando Pérez en 2001 bajo el nombre de *IPython* (Interactive Python), evolucionó hasta lo que hoy conocemos como Jupyter. Llegando a ser nombrada como uno de los diez códigos computacionales que cambiaron la ciencia (Perkel, 2021), por medio de sus cuadernos de trabajo (*notebooks*) permite la creación y uso de documentos interactivos que combinan código, texto y gráficas, resultando de mucha utilidad en la investigación geofísica, donde el análisis de datos y la visualización son componentes claves. Los *notebooks*

permiten explorar y manipular datos en tiempo real, permitiendo al usuario con acceso a estos cuadernos y los datos, seguir la secuencia de comandos y alcanzar los mismos resultados que la investigación previa.

La Sociedad de Geofísicos de Exploración (SEG) tiene una colección de tutoriales de geofísica en formatos de cuadernos de trabajo a través de su portal de GitHub. Además, de su uso en la academia, los cuadernos de trabajo de Jupyter, están adquiriendo relevancia en la industria. Empresas como Equinor, petrolera estatal noruega, y Chevron, petrolera estadounidense, cuentan con repositorios empleando esta herramienta. Washbourne et al. (2021), presentaron su flujo de optimización Chevron para la representación e inversión sísmica (*Chevron Optimization Framework for Imaging and Inversion*), esquema de trabajo de código abierto para el modelado e inversión sísmica capaz de utilizarse en computación en la nube. Estos cuadernos muestran el procedimiento de uso para realizar inversión de forma de onda completa y migración de tiempo reverso.

Este trabajo se realizó en Julia (Bezanson et al., 2017), lenguaje de programación de alto nivel y de código abierto que ha adquirido popularidad en el campo de la computación científica y numérica. Fue desarrollado específicamente para abordar los desafíos de rendimiento y productividad en este campo. Una de las características destacadas de Julia es su capacidad para combinar la facilidad de uso y la expresividad de lenguajes como Python con el rendimiento cercano al de lenguajes de bajo nivel como C y Fortran. Esto se logra mediante la compilación just-in-time (JIT), que permite que el código Julia se ejecute de manera eficiente. Julia cuenta con una sintaxis limpia y legible, similar a otros lenguajes científicos como MATLAB y Python. Además, ofrece una amplia gama de bibliotecas y paquetes especializados para diversas áreas de la computación científica, como el procesamiento de datos, la visualización, el aprendizaje automático y la optimización. Una de las fortalezas de Julia es su capacidad para manejar cómodamente arreglos multidimensionales y operaciones numéricas avanzadas. Esto lo convierte en una herramienta poderosa para el análisis de datos, la simulación numérica y la resolución de problemas matemáticos complejos. También, Julia se destaca por su enfoque en la paralelización y el cálculo distribuido, lo que lo convierte en una opción atractiva para trabajar en sistemas de alto rendimiento y en entornos de computación distribuida.

Una de las bibliotecas de Julia, y la principal aquí utilizada es Julia Devito Inversion, JUDI, un marco de trabajo para el modelado e inversión sísmica a gran escala, diseñado para permitir la rápida integración de algoritmos a códigos ejecutables y eficientes, con la capacidad de adaptación a problemas tridimensionales de tamaño industrial. Las ecuaciones de onda en JUDI, se resuelven con ayuda de Devito, una biblioteca para cálculos de diferencias finitas automatizados (Louboutin et al., 2019).

JUDI fue implementado para la investigación y desarrollo de flujos de trabajo de

Enfoque tradicional

$$m \frac{\partial^2 u}{\partial t^2} + \eta \frac{\partial u}{\partial t} - \Delta u = 0$$

```
void kernel(...) {
    ...
    <impenetrable code with aggressive
    performance optimizations written
    by rockstars, gurus, ninjas,
    unicorns and celestial beings>
    ...
}
```

Definición simbólica

$$m \frac{\partial^2 u}{\partial t^2} + \eta \frac{\partial u}{\partial t} - \Delta u = 0$$

Devito

```
eqn = m * u.dt2 + eta * u.dt - u.laplace
```

```
void kernel(...) { ... }
```

Figura 2.14: Definición simbólica de operadores. Modificada de Luporini et al. (2020).

inversión sísmica, permitiendo la rápida incorporación de expresiones matemáticas a secuencias de comandos. Construido con la idea de abstracción, presenta múltiples capas para cada aspecto de complejidad:

1. Operadores lineales sin matriz.
2. Alto nivel de flexibilidad en la paralelización, siendo adaptable a diferentes ambientes computacionales, como clusters o computación en la nube.
3. Definiciones simbólicas de operadores directos y adjuntos, y ecuaciones de onda (Fig. 2.14).

3

Metodología

3.1. Datos sísmicos

Se utilizaron dos modelos sintéticos para evaluar la eficacia y desempeño de las herramientas en los problemas de inversión de forma de onda completa y migración de tiempo reverso por mínimos cuadrados.

Modelo sintético personalizado

El primero, un modelo personalizado bidimensional, que consta de tres capas planas con un cuadrado, en el centro. El modelo tiene dimensiones de 1km por 1km, discretizado en una malla de celdas de 10m por 10m. Las velocidades de onda P tienen valores de 1.5, 2, 3 km/s de la capa más somera a la más profunda, mientras que el cuadrado del centro tiene una velocidad de onda P de 2.5 km/s.

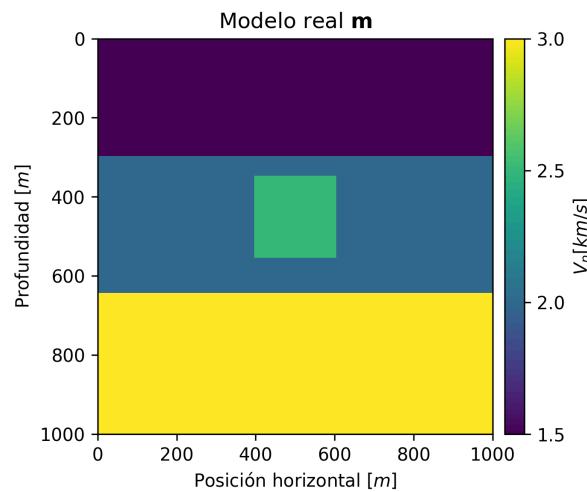


Figura 3.1: Modelo de velocidades sintético personalizado.

Modelo sintético SEG/EAGE Overthrust

El segundo modelo utilizado en este trabajo es el modelo sintético SEG/EAGE Overthrust (Aminzadeh et al., 1994). Tiene dimensiones de 10km por 3km, en este trabajo, para optimizar el tiempo de procesamiento se discretizó en una malla de celdas de 25m por 25m. Fue desarrollado como parte del proyecto conjunto de la SEG y la Asociación Europea de Geocientíficos e Ingenieros (EAGE). Se basa en un modelo sintético de un yacimiento petrolero con una estructura geológica compleja, caracterizada por una zona de cabalgamiento. En su mayoría, las capas tienen velocidades constantes que van desde los 2.5 km/s hasta los 5.5 km/s, con una columna de agua en los primeros 500 metros de profundidad.

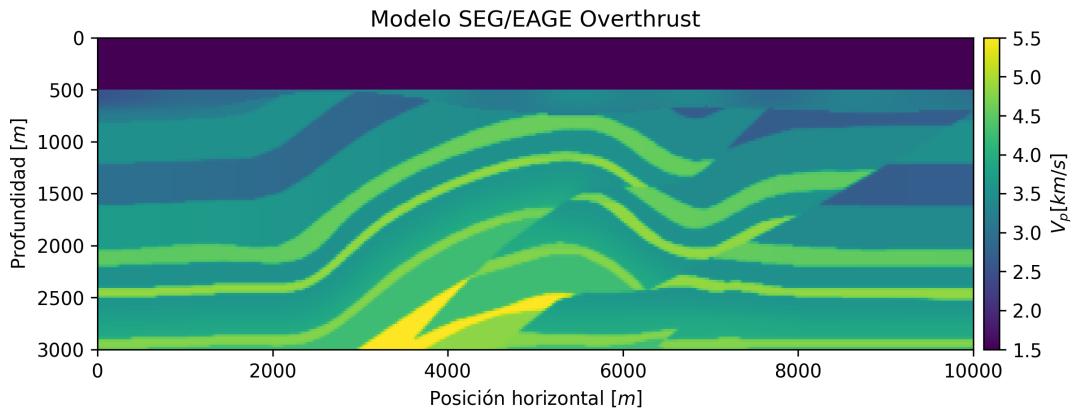


Figura 3.2: Modelo de velocidades sintético SEG/EAGE Overthrust y = 5,650 m.

3.2. Definición del problema sísmico directo

A diferencia de la sismología global, la sismología de exploración se basa en la perturbación mediante una fuente de energía controlada, que generan ondas sonoras y/o elásticas que viajan a través del subsuelo. Las ondas se dispersan y reflejan en la estructura geológica, lo que provoca que parte del campo de onda vuelvan a la superficie, donde son registradas por un conjunto de receptores. Después de dispararse una fuente, cada receptor registra cambios relativos en presión como función del tiempo. Estos datos se grafican como una imagen de una matriz bidimensional, en la cual cada columna corresponde a un receptor, en función del tiempo, y la intensidad del color indica los cambios de presión (Fig. 2.3). En una adquisición sísmica cada fuente se dispara a media que se desplaza por el área de estudio, los datos observados se registran en cada posición de la fuente y se denominan registros de tiro común (*shot record*). De forma matemática, cada registro de tiro común se indica por vectores $\mathbf{d}_i^{\text{obs}} \in \mathbb{R}^{n_d}$, donde $i = 1, \dots, n_s$ es el

índice de la fuente que va de 1 al número total de fuentes n_s , o como un único vector concatenado \mathbf{d}_{obs} . La dimensión de los datos n_d esta dado por el número de receptores n_r por en número de muestras de tiempo n_t . El objetivo de la inversión sísmica es recuperar una parametrización física del subsuelo a partir de los datos sísmicos registrados. Como ya se mencionó, esto se consigue minimizando la discrepancia (desajuste) entre los datos observados y los datos calculados. El problema directo está definido como un registro de tiro común $\mathbf{d}_i^{\text{calc}}$ calculado mediante un operador de modelado directo \mathcal{F} (Witte, 2020):

$$\mathbf{d}_i^{\text{calc}} = \mathcal{F}(\mathbf{m}, \mathbf{q}_i) \quad (3.1)$$

El vector $\mathbf{m} \in \mathbb{R}^n$ representa la parametrización vectorizada del modelo del subsuelo, velocidad de onda P. El escalar n es el número total de puntos en la malla, en dos o tres dimensiones. El vector $\mathbf{q}_i \in \mathbb{R}^{n_t}$ representa el patrón de tiempo de la fuente sísmica, la cual se asume como una función delta en el espacio con un patrón de radiación constante, lo que la convierte en una función únicamente del tiempo, con n_t siendo el número de muestras de tiempo. El operador de modelado directo \mathcal{F} se define como la solución de una ecuación de onda discretizada para un modelo y fuente dada:

$$\mathcal{F}(\mathbf{m}, \mathbf{q}_i) = \mathbf{P}_r \mathbf{F}(\mathbf{m})^{-1} \mathbf{P}_s^\top \mathbf{q}_i \quad (3.2)$$

La matriz $\mathbf{P}_s \in \mathbb{R}^{n_s \times N}$ es el operador de proyección de fuente, cuya transposición, denotada por \top , emite el patrón de la fuente dependiente del tiempo en la ubicación actual de la fuente. El operador de proyección de la fuente mapea el vector dependiente del tiempo $\mathbf{q}_i \in \mathbb{R}^{n_t}$ a un vector con las dimensiones del campo de ondas $\mathbf{u} \in \mathbb{R}^N$ dependiente del tiempo y del espacio, donde $N = n * n_t$ es el número puntos de la malla por el número de muestras de tiempo. $\mathbf{P}_r \in \mathbb{R}^{n_d \times N}$ es el operador de proyección de receptor, que muestrea el campo de onda dependiente del tiempo en la ubicación de los receptores. El operador $\mathbf{F}(\mathbf{m}) \in \mathbb{R}^{N \times N}$ representa la ecuación de onda discretizada dependiente del tiempo:

$$\underbrace{\left(\mathbf{m} \frac{\partial^2}{\partial t^2} - \nabla^2 \right)}_{\mathbf{F}(\mathbf{m})} \mathbf{u} = \mathbf{P}_s^\top \mathbf{q}_i \quad (3.3)$$

Los operadores $\frac{\partial^2}{\partial t^2} \in \mathbb{R}^{N \times N}$ y $\nabla^2 \in \mathbb{R}^{N \times N}$ representan la segunda derivada respecto al tiempo y el Laplaciano, ambos discretizados mediante diferencias finitas. Aquí, $\mathbf{F}(\mathbf{m})$ está escrito como un operador lineal actuando sobre el campo de onda \mathbf{u} , que se puede obtener aplicando el inverso del operador al lado derecho de la ecuación, $\mathbf{u}_i = \mathbf{F}(\mathbf{m})^{-1} \mathbf{P}_s^\top \mathbf{q}_i$. Los datos sísmicos se obtienen aplicando el operador de restricción del receptor al campo de ondas, es decir $\mathbf{d}_i^{\text{calc}} = \mathbf{P}_r \mathbf{u}_i$.

3.3. Definición del problema sísmico inverso

Retomando el problema sísmico inverso, tiene como objetivo la recuperación de los parámetros del modelo \mathbf{m} del conjunto de datos sísmicos observados $\mathbf{d}_i^{\text{obs}}$. El problema se puede formular como un problema de optimización no lineal, en el cual el objetivo es minimizar el desajuste entre los datos observados y los datos calculados (Tarantola, 1984; Virieux & Operto, 2009):

$$\min_{\mathbf{m}} \sum_{i=1}^{n_s} \phi \left[\mathbf{d}_i^{\text{obs}} - \mathcal{F}(\mathbf{m}, \mathbf{q}_i) \right] \quad (3.4)$$

donde $\phi : \mathbb{R}^{n_d} \rightarrow \mathbb{R}$ es una función de desajuste suavizada y por lo general se elige la función de desajuste de mínimos cuadrados $\phi = \frac{1}{2} \|\cdot\|^2$. Esta versión del problema inverso se conoce como formulación de estado reducido o de estado adjunto. En este estado reducido, el problema sísmico inverso no lineal (Ec. 3.4) se llama inversión de forma de onda completa (FWI), ya que el objetivo es la predicción de todas las formas de ondas en los datos sísmicos observados (reflexiones, difracciones, refracciones, de retorno, etc.) invirtiendo los parámetros del modelo \mathbf{m} .

FWI se resuelve comúnmente utilizando algoritmos de optimización basados en gradientes, como gradiente descendente, métodos de Gauss-Newton o los métodos Quasi-Newton. El gradiente de la Ecuación 3.4 respecto a los parámetros del modelo \mathbf{m} se calcula usando el método del estado adjunto, o estado reducido. Para el caso en que ϕ sea el desajuste de mínimos cuadrados, el gradiente (estado adjunto) se define como (Witte, 2020):

$$\mathbf{g} = \sum_{i=1}^{n_s} \left(\frac{\partial \mathcal{F}(\mathbf{m}, \mathbf{q}_i)}{\partial \mathbf{m}} \right)^{\top} \left(\mathbf{d}_i^{\text{obs}} - \mathcal{F}(\mathbf{m}, \mathbf{q}_i) \right) \quad (3.5)$$

donde $\mathbf{g} \in \mathbb{R}^n$ es el gradiente de la función objetivo de FWI y $\frac{\partial \mathcal{F}}{\partial \mathbf{m}} \in \mathbb{R}^{n_d \times n}$ es la derivada parcial del operador de modelado directo respecto a \mathbf{m} , que se conoce como el Jacobiano y corresponde físicamente a la dispersión linealizada de Born. Calcular el gradiente de la función objetivo de FWI implica resolver una ecuación de onda directa y adjunta para cada índice i de la suma de las ubicaciones de las fuentes, n_s siendo del orden de 10^2 a 10^5 . Además, la relación entre \mathbf{m} y los datos calculados es no lineal, y la función objetivo es no convexa, por lo que se deben tomar pasos adicionales (regularización, restricciones, etc.). Una variación con alta importancia práctica es el problema sísmico inverso linealizado.

3.4. Definición del problema sísmico inverso lineal

En el problema sísmico inverso lineal, la suposición fundamental es que un modelo \mathbf{m} por partes (Fig. 3.3a), puede separarse en un componente suavizado $\mathbf{m}_0 \in \mathbb{R}^n$, que contiene los números de onda cortos (Fig. 3.3b) y una componente de números de onda largos $\delta\mathbf{m} \in \mathbb{R}^n$, que contiene los grandes contrastes de los parámetros (Fig. 3.3c). En consecuencia, el operador de modelado directo lineal $\mathcal{F}(\mathbf{m}, \mathbf{q}_i)$ se aproxima mediante la suma de \mathcal{F} evaluado en \mathbf{m}_0 más el efecto del Jacobiano sobre $\delta\mathbf{m}$:

$$\mathcal{F}(\mathbf{m}, \mathbf{q}_i) \approx \mathcal{F}(\mathbf{m}_0, \mathbf{q}_i) + \frac{\partial \mathcal{F}(\mathbf{m}_0, \mathbf{q}_i)}{\partial \mathbf{m}} \delta\mathbf{m} + \mathcal{O}(\delta\mathbf{m}^\top \delta\mathbf{m}) \quad (3.6)$$

La intuición física detrás de esta separación de escalas es que los datos sísmicos pueden ser (aproximadamente) divididos en contribuciones de ondas transmitidas que se propagan a través de \mathbf{m}_0 (ondas directas y de retorno), mientras que las reflexiones y ondas difractadas son causadas por las perturbaciones del modelo en $\delta\mathbf{m}$ (Claerbout, 2004). Los términos de orden superior de la Ecuación 3.6 describen datos que resultan de la dispersión múltiple.

El Jacobiano $\frac{\partial \mathcal{F}(\mathbf{m}, \mathbf{q}_i)}{\partial \mathbf{m}}$, el mismo operador lineal de la Ecuación 3.5, se obtiene derivando la Ecuación 3.3 respecto a \mathbf{m} , que da como resultado:

$$\frac{\partial \mathcal{F}(\mathbf{m}, \mathbf{q}_i)}{\partial \mathbf{m}} = -\mathbf{P}_r \mathbf{F}(\mathbf{m})^{-1} \text{diag}\left(\frac{\partial \mathbf{F}(\mathbf{m})}{\partial \mathbf{m}} \mathbf{F}(\mathbf{m})^{-1} \mathbf{P}_s^\top \mathbf{q}_i\right) \quad (3.7)$$

El Jacobiano \mathbf{J} mapea las perturbaciones en el modelo a las perturbaciones, en los datos sísmicos, y a diferencia de \mathcal{F} , es un operador lineal (Witte, 2020).

Las perturbaciones del modelo $\delta\mathbf{m}$ con respecto con respecto a un modelo de fondo (suavizado) \mathbf{m}_0 se llama imagen sísmica. Su valor práctico radica en que proporciona información estructural del subsuelo, que están asociadas con cambios rápidos en la velocidad de onda y/o densidad. Por lo tanto, si se dispone de una buena aproximación de \mathbf{m}_0 (por ejemplo, obtenido mediante FWI), la imagen sísmica se puede obtener mediante la inversión linealizada de la forma de onda resolviendo el siguiente problema de mínimos cuadrados lineal:

$$\min_{\delta\mathbf{m}} \sum_{i=1}^{n_s} \frac{1}{2} \|\delta\mathbf{d}_i^{\text{obs}} - \mathbf{J}(\mathbf{m}_0, \mathbf{q}_i) \delta\mathbf{m}\|^2 \quad (3.8)$$

Donde el vector $\delta\mathbf{d}_i^{\text{obs}} = \mathbf{d}_i^{\text{obs}} - \mathcal{F}(\mathbf{m}_0, \mathbf{q}_i)$ son los datos observados linealizados, datos sísmicos donde, idealmente, solo se tienen las ondas reflejadas. Esta inversión sísmica linealizada se conoce como migración de tiempo reverso de mínimos cuadrados (LS-RTM) ya que el residual debe ser retropropagado en el tiempo para calcular el gradiente de la función objetivo.

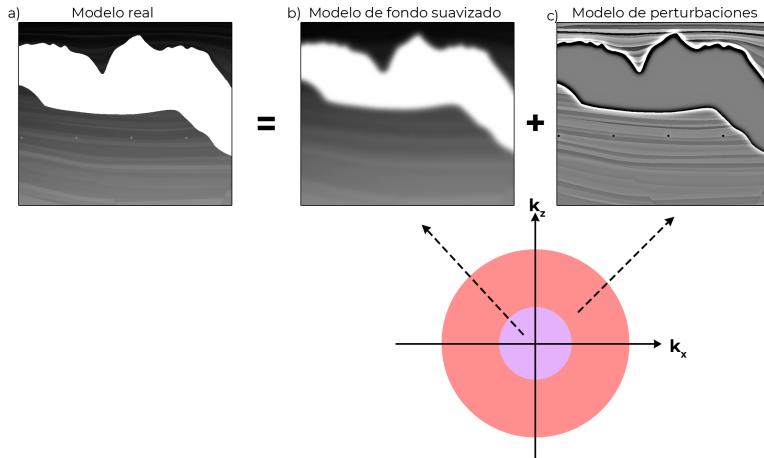


Figura 3.3: Modelo de tierra real compuesto por la contribución de modelos de velocidad de fondo suavizado (números de ondas corto) y reflectividad o modelo de perturbaciones (números de ondas largos). Modificada de Long (2020)

3.5. FWI

La función objetivo de la inversión de forma de onda completa es:

$$\min_{\mathbf{m}} \phi \left[\mathbf{d}_{\text{obs}} - \mathbf{P}_r \mathbf{F}(\mathbf{m})^{-1} \mathbf{P}_s^\top \mathbf{q} \right] \quad (3.9)$$

el gradiente de esta ecuación (3.9) respecto a los parámetros del modelo \mathbf{m} se calcula usando el método del estado adjunto, aplicando el adjunto del Jacobiano (operador de migración) al residual entre los datos observados y calculados. Con JUDI, es posible implementar esta función objetivo en una función de Julia llamada `fwi_objective`, que toma el modelo actual, la fuente y los datos observados como argumentos de entrada, genera los datos calculados para el modelo actual, posterior, calcula el desajuste entre los datos, así como el gradiente.

En el siguiente paso, corresponde implementar el algoritmo de optimización, el gradiente descendiente estocástico (SGD), la idea principal es ajustar gradualmente los parámetros en la dirección del gradiente negativo de la función objetivo, lo que eventualmente lleva a encontrar un mínimo local o global. El SGD es una variante del gradiente descendiente, que utiliza una muestra aleatoria de datos (en este caso disparos) en cada iteración, en lugar de hacer uso de todo el conjunto de datos. El SGD es de mucha utilidad cuando se trabaja con conjuntos de datos grandes o cuando el modelo tiene una gran cantidad de parámetros, ya que reduce la carga computacional al calcular el gradiente solo para una muestra a la vez, en lugar de para todo el conjunto de datos.

El primer paso en el bucle de principal (Alg. 1) es seleccionar una serie aleatoria de fuentes y disparos, para los cuales se calcula el gradiente y el valor de la función

objetivo. Luego, se pasa el subconjunto de fuentes y datos a la función objetivo para evaluar la serie actual de disparos. Se realiza una búsqueda para determinar el tamaño del paso adecuado para actualizar el modelo. El último paso corresponde a la aplicación de la actualización al modelo.

Algoritmo 1 FWI por gradiente descendiente estocástico

```

for  $j = 1$  to  $n$  do
    Selección aleatoria de una serie de disparos  $s(j) \in [1 \dots n_s]$ 
     $f_j, g_j = \text{fwi\_objetive}(\text{model}, q_i, d_i^{\text{obs}})$ 
     $\text{step} = \text{linesearch}(\text{model}, g)$ 
     $\text{model} = \text{proyección}(\text{model} + \text{step})$ 
end for

```

3.6. LS-RTM

Para el caso de migración de tiempo reverso de mínimos cuadrados, se consideró una función objetivo con precondicionadores \mathbf{M}_l^{-1} y \mathbf{M}_r^{-1} que corresponden al espacio del modelo y de los datos respectivamente (Witte, 2020):

$$\min_{\widehat{\delta\mathbf{m}}} \frac{1}{2} \|\mathbf{M}_l^{-1} \mathbf{J} \mathbf{M}_r^{-1} \widehat{\delta\mathbf{m}} - \mathbf{M}_l^{-1} \delta\mathbf{d}\|^2 \quad (3.10)$$

donde $\delta\mathbf{m} = \mathbf{M}_r^{-1} \widehat{\delta\mathbf{m}}$ es la imagen que se desea recuperar, \mathbf{J} representa el Jacobiano, $\delta\mathbf{d}$ son los datos observados linealizados, disparos donde todos los eventos excepto las reflexiones se han removido. El algoritmo de optimización escogido para minimizar la función, es el gradiente descendiente estocástico (Alg. 2). En cada iteración se realiza una selección aleatoria de registros de tiro común. Se calcula el residual y gradiente para la selección actual de las fuentes. La iteración finaliza con el cálculo del tamaño del paso y la actualización del modelo.

Algoritmo 2 LS-RTM por Gradiente descendiente estocástico

```

for  $j = 1$  to  $n$  do
    Selección aleatoria de una serie de disparos  $s(j) \in [1 \dots n_s]$ 
     $\mathbf{r}_j = \mathbf{M}_l^{-1} \mathbf{J}_{s(j)} \mathbf{M}_r^{-1} \mathbf{x}_j - \mathbf{M}_l^{-1} \delta\mathbf{d}_{s(j)}$ 
     $\mathbf{g}_j = \mathbf{M}_r^{-T} \mathbf{J}_{s(j)}^{-T} \mathbf{M}_l^{-T} \mathbf{r}_j$ 
     $t_j = \frac{\|\mathbf{r}_j\|^2}{\|\mathbf{g}_j\|^2}$ 
     $\mathbf{x}_{j+1} = \mathbf{x}_j - t_j \mathbf{g}_j$ 
end for

```

4

Resultados

4.1. Modelo personalizado

Parámetros

Parámetros del modelo	
Disparos	21
Receptores	101
Frecuencia	4 Hz, 15 Hz
Longitud de grabación	1000 ms
Intervalo de muestreo	2 ms
Dimensiones del modelo	1km × 1km
Iteraciones FWI	30
Iteraciones LS-RTM	30

Tabla 4.1: Parámetros del modelo personalizado

Geometría de adquisición

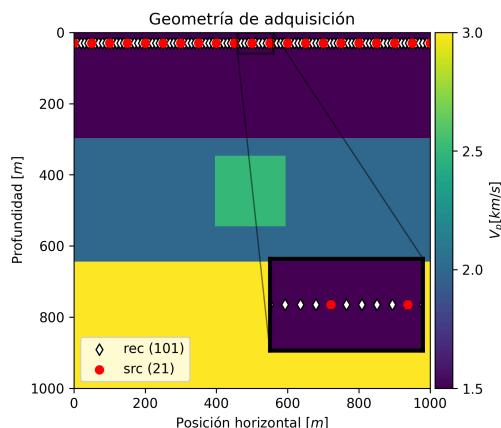


Figura 4.1: Geometría de adquisición del modelo personalizado.

Datos observados (4 Hz)

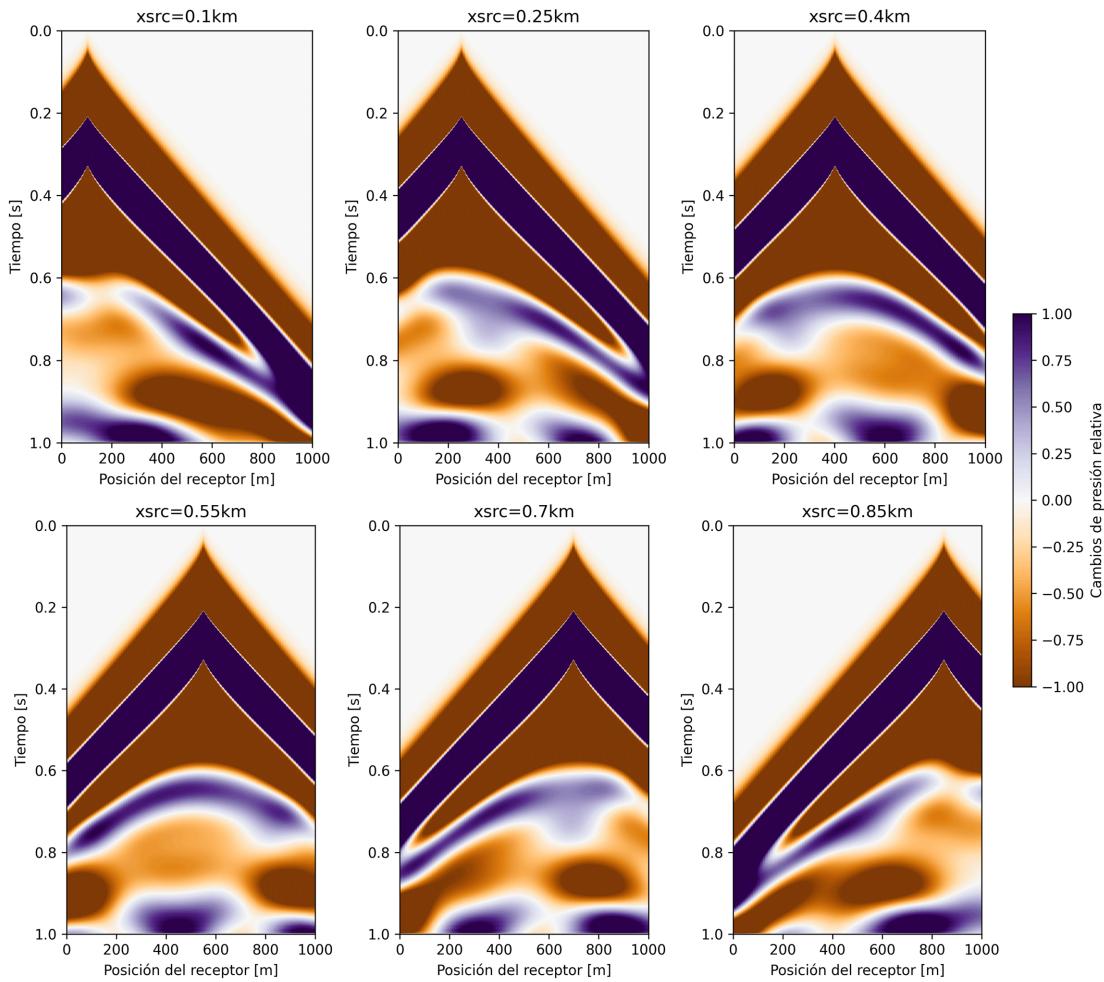


Figura 4.2: Registros de disparos para distintas fuentes. Datos de baja frecuencia - 4 Hz.

Datos observados (15 Hz)

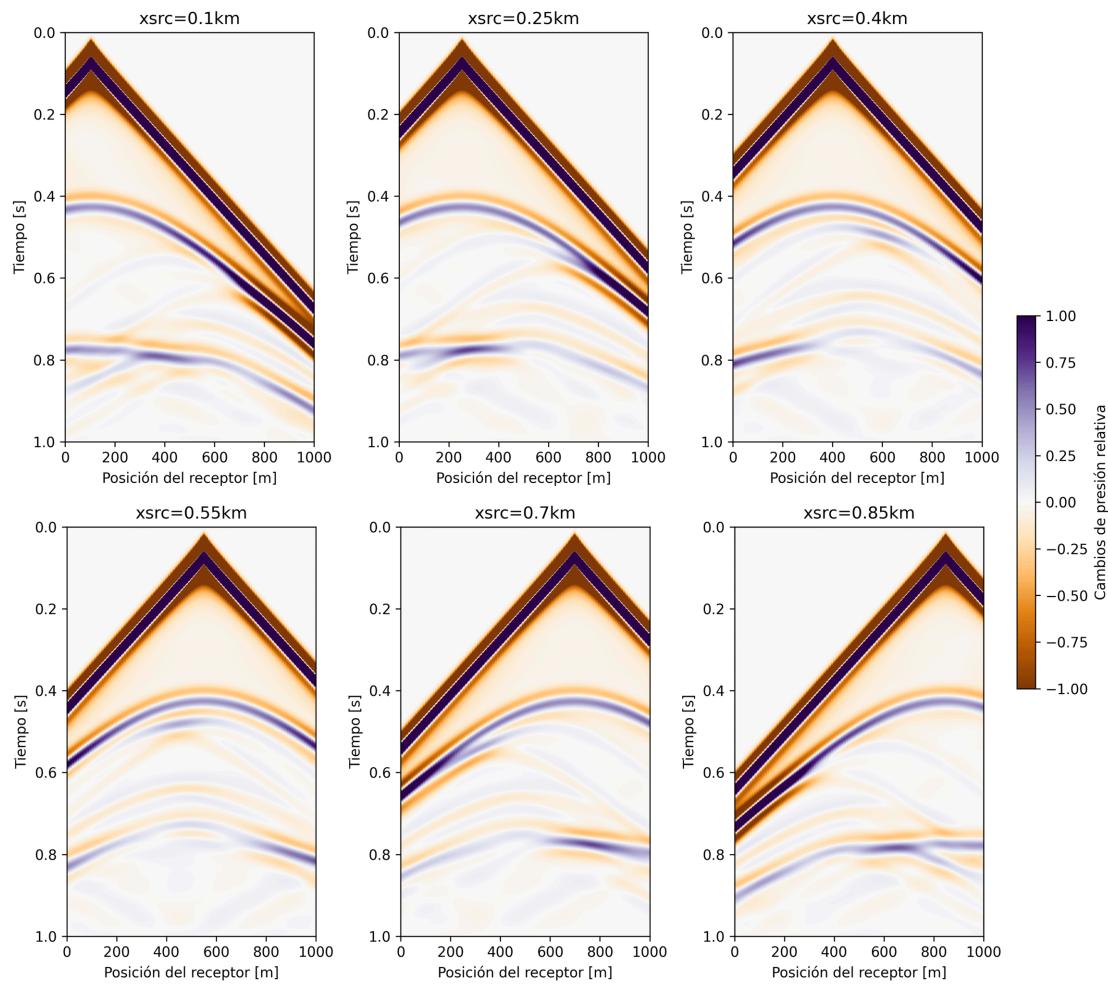


Figura 4.3: Registro de disparos para distintas fuentes. Datos de alta frecuencia - 15 Hz.

FWI

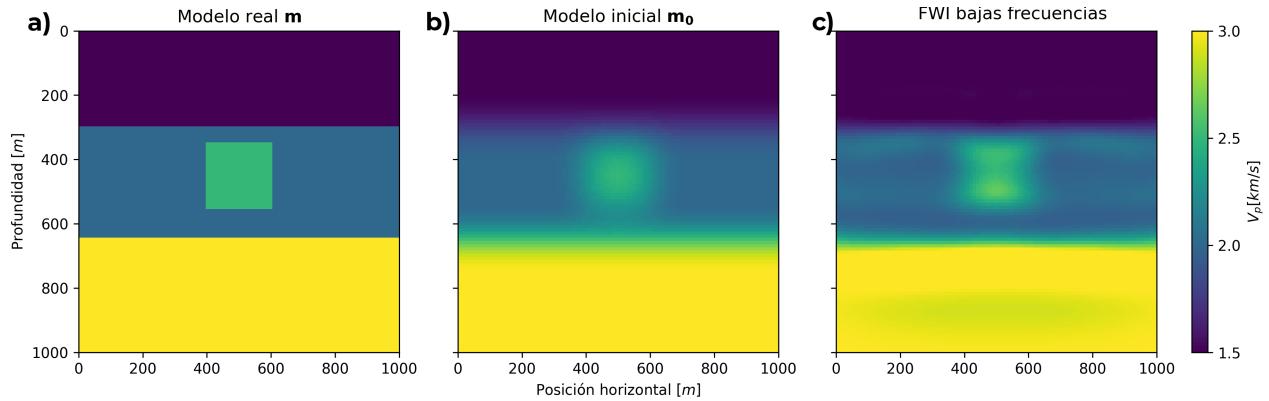


Figura 4.4: Resultados de FWI con datos de baja frecuencia - 4 HZ. a) Modelo real b) Modelo inicial (modelo real suavizado) y c) FWI con bajas frecuencias.

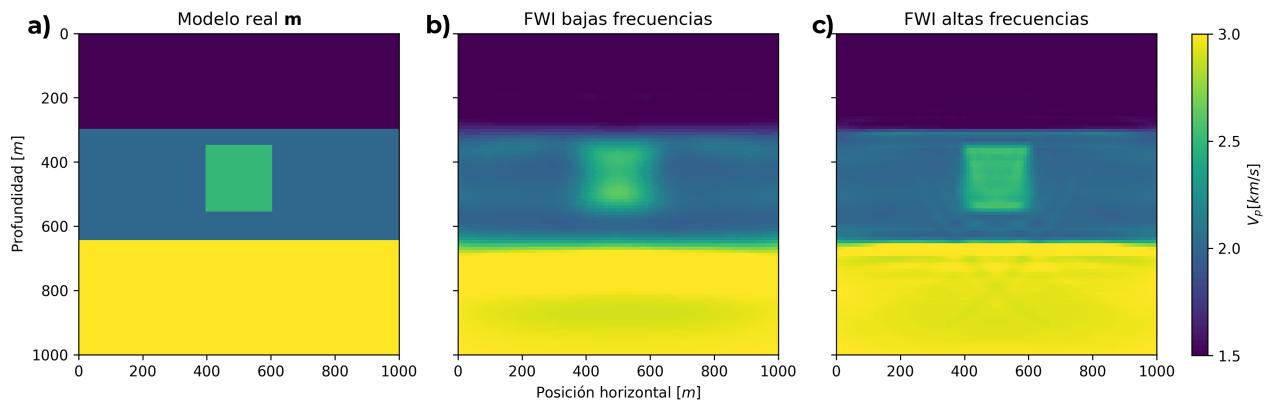


Figura 4.5: Resultados de FWI con datos de alta frecuencia con modelo inicial de FWI de baja frecuencia - 4 y 15 HZ. a) Modelo real b) Modelo inicial - resultado de FWI de bajas frecuencias y c) FWI final (4 y 15 Hz).

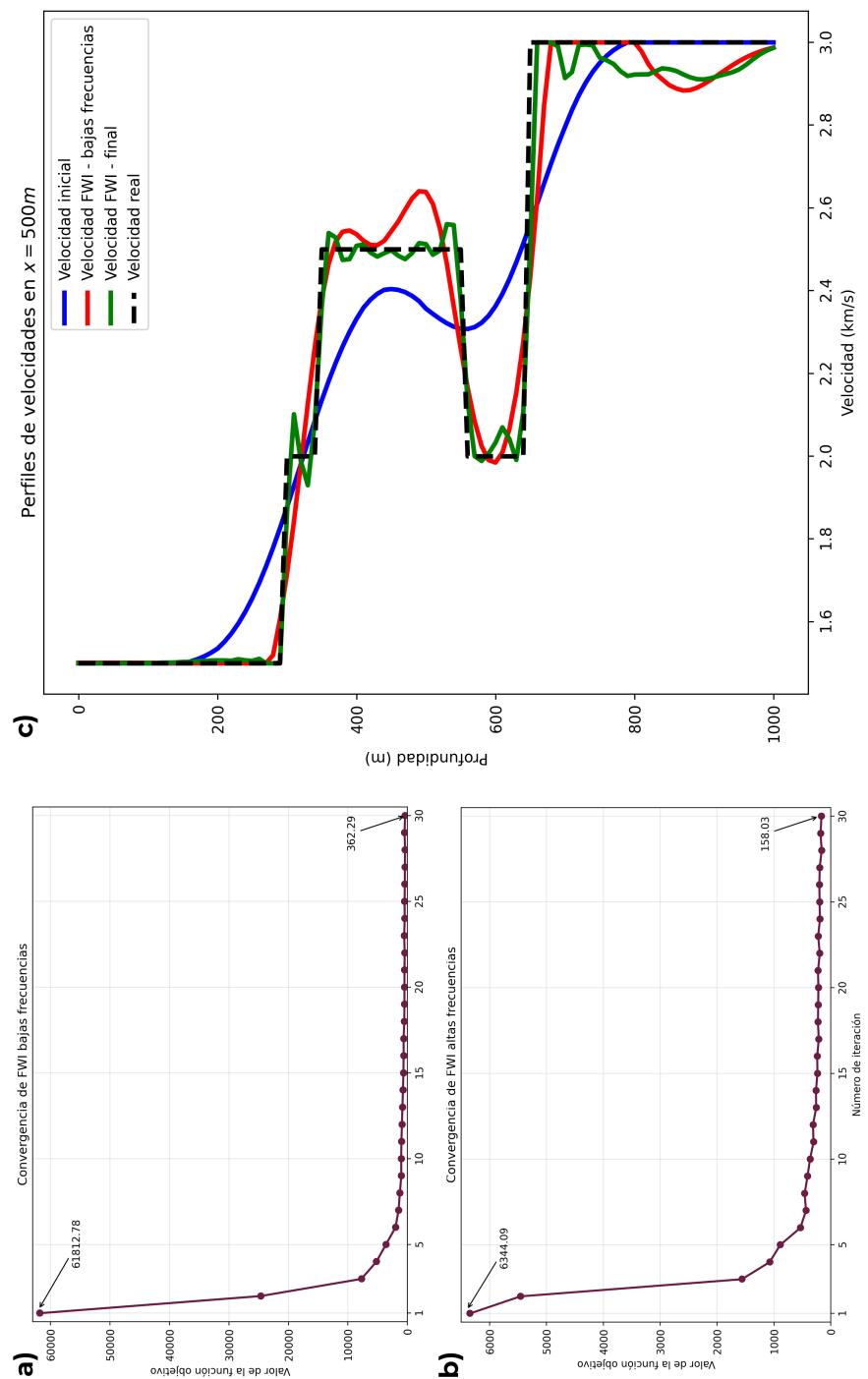


Figura 4.6: a) Función objetivo a lo largo de las iteraciones de FWI para bajas frecuencias. b) Función objetivo a lo largo de las iteraciones de FWI para altas frecuencias. c) Comparación de velocidades: Inicial, FWI bajas frecuencias, FWI altas frecuencias y real en $x = 500$ m. Se puede observar una mejora en la velocidad del resultado final de inversión de forma de onda completa.

LS-RTM

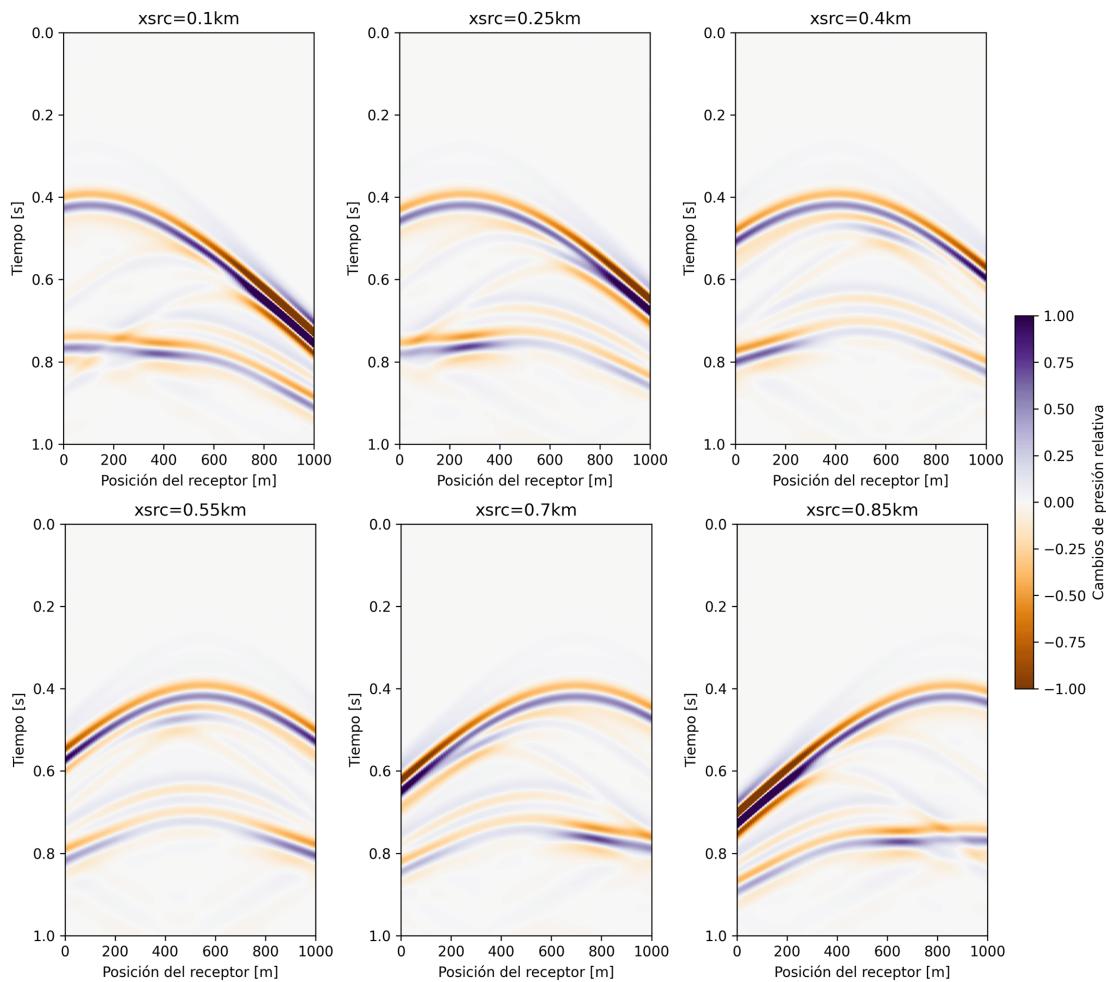


Figura 4.7: Registros de disparos para distintas fuentes. Datos de alta frecuencia linealizados (Born) - 15 Hz.

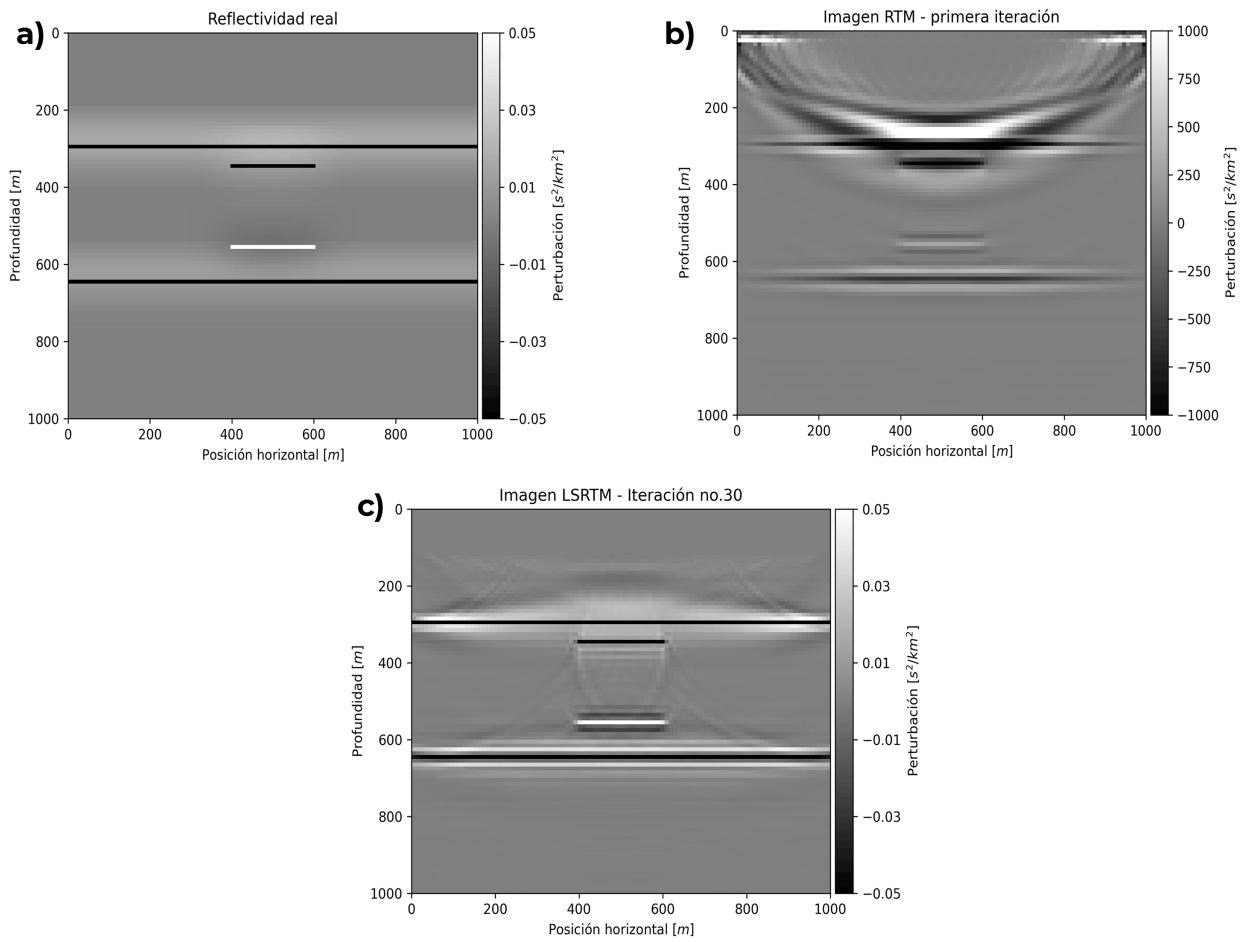


Figura 4.8: Resultados de LS-RTM con datos de alta frecuencia - 4 y 15 HZ. a) Modelo real b) RTM y c) LS-RTM final.

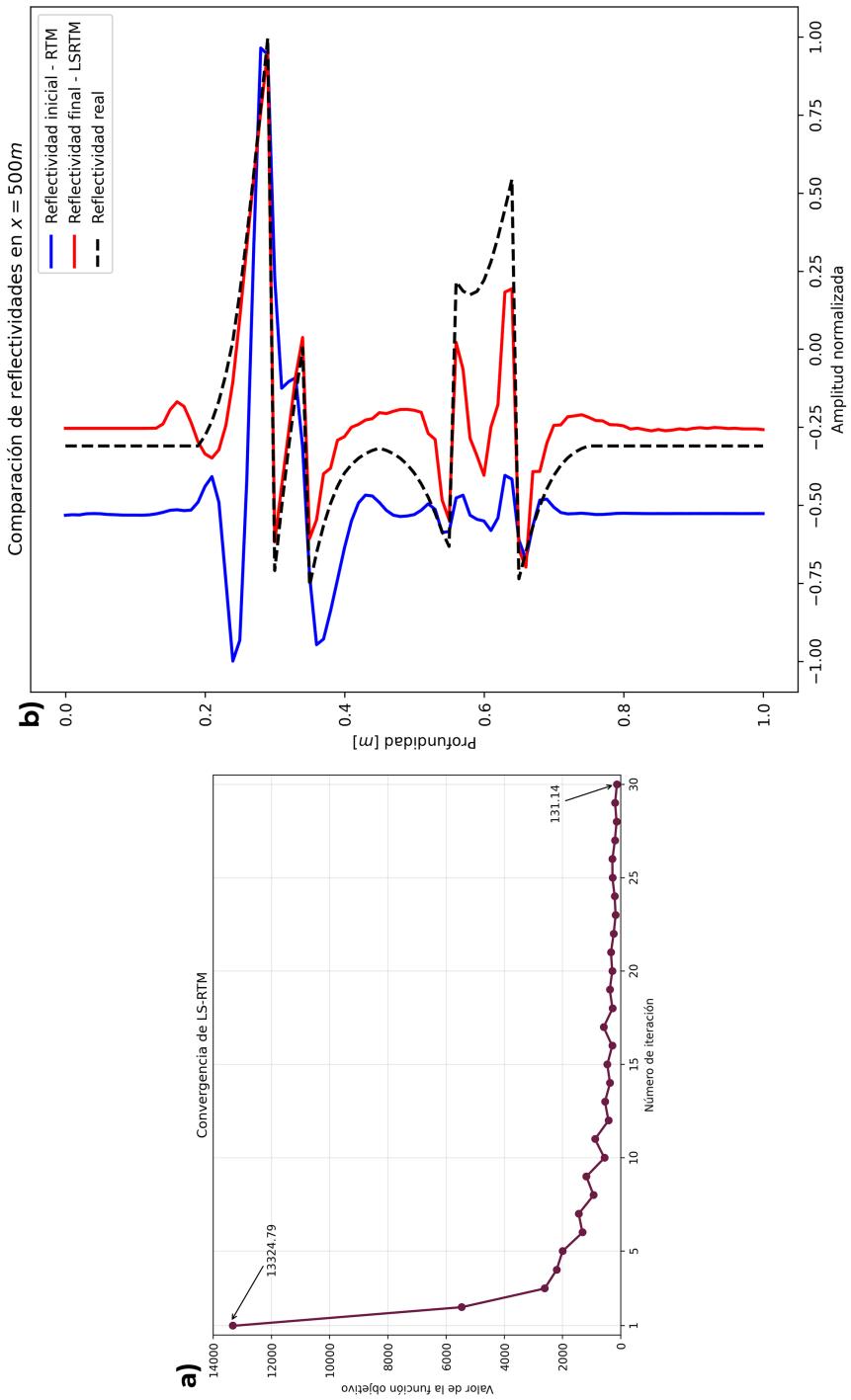


Figura 4.9: a) Función objetivo a lo largo de las iteraciones de LSRTM. b) Comparación de reflectividades normalizadas: RTM, LSRTM y real en $x = 500$ m. Se aprecia que la reflectividad obtenida mediante RTM no cuenta con las amplitudes correspondientes a la geología.

4.2. Modelo SEG/EAGE Overthrust

Parámetros

Parámetros del modelo	
Disparos	16
Receptores	201
Frecuencia	4 Hz, 8 Hz
Longitud de grabación	2000 ms
Intervalo de muestreo	4 ms
Dimensiones del modelo	10km × 3km
Iteraciones FWI	30
Iteraciones LS-RTM	30

Tabla 4.2: Parámetros del modelo SEG/EAGE Overthrust

Geometría de adquisición

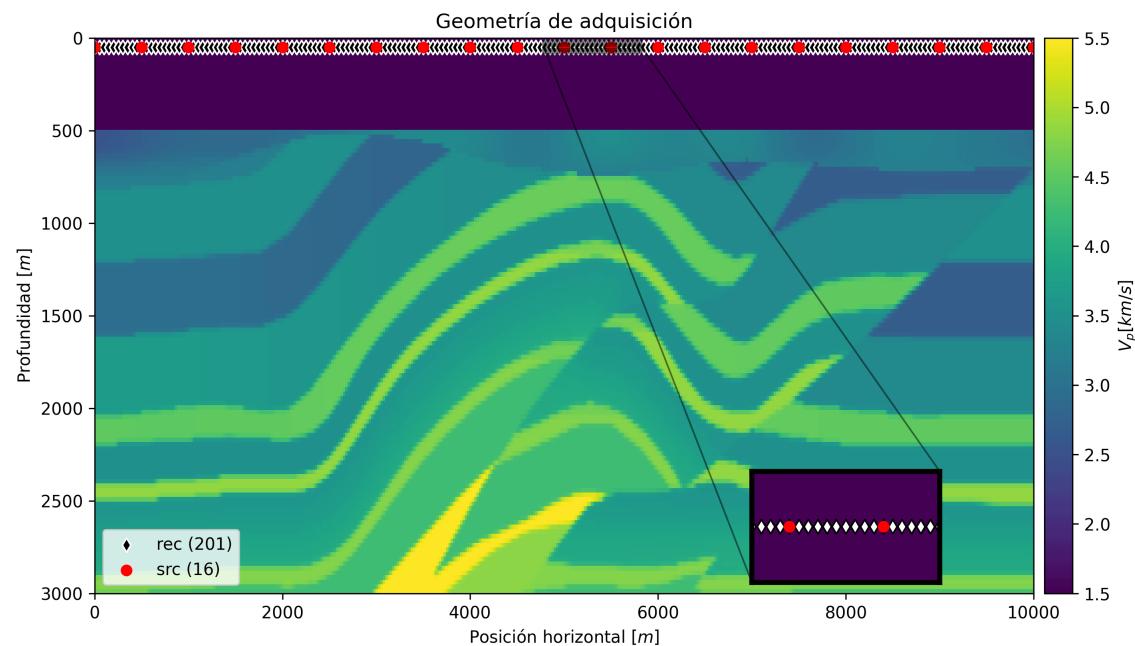


Figura 4.10: Geometría de adquisición del modelo SEG/EAGE Overthrust.

Datos observados (4 Hz)

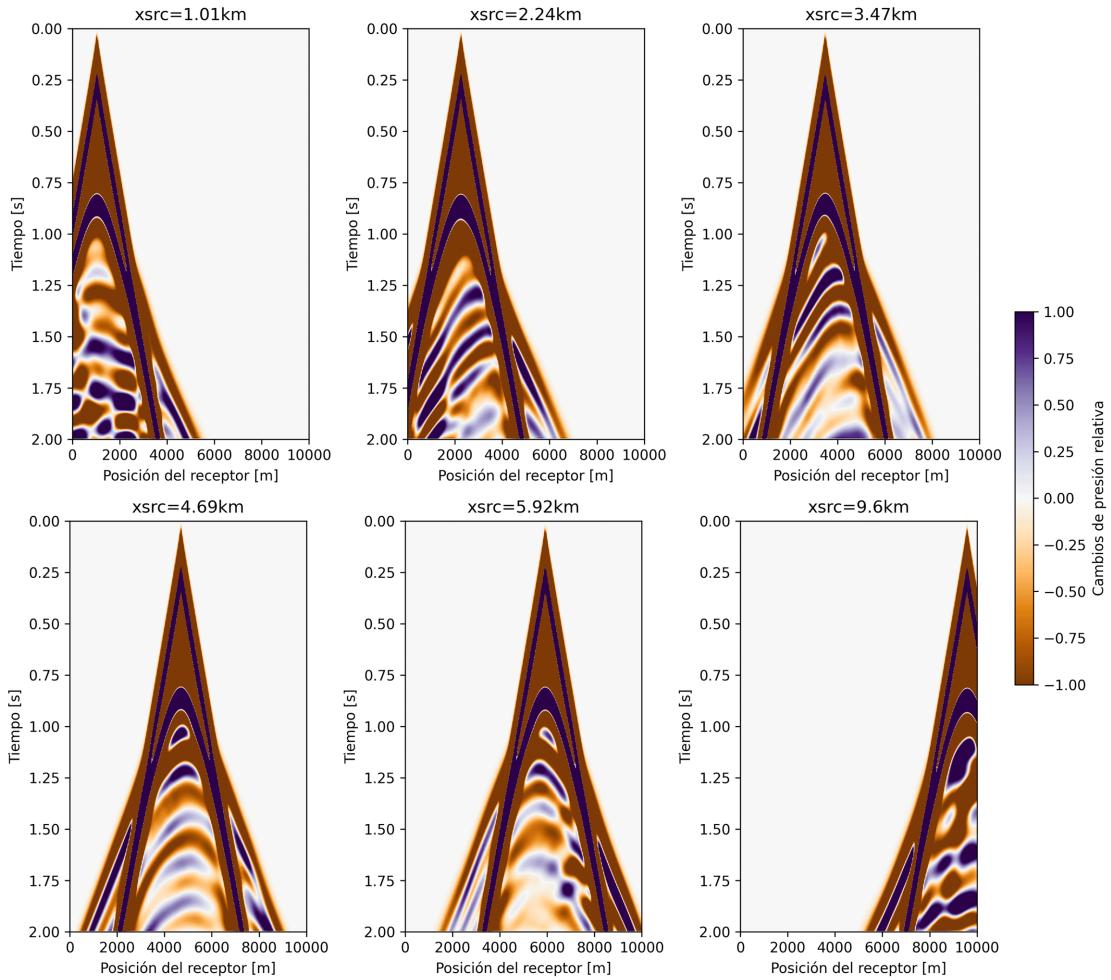


Figura 4.11: Registros de disparos para distintas fuentes. Datos de baja frecuencia - 4 Hz.

Datos observados (8 Hz)

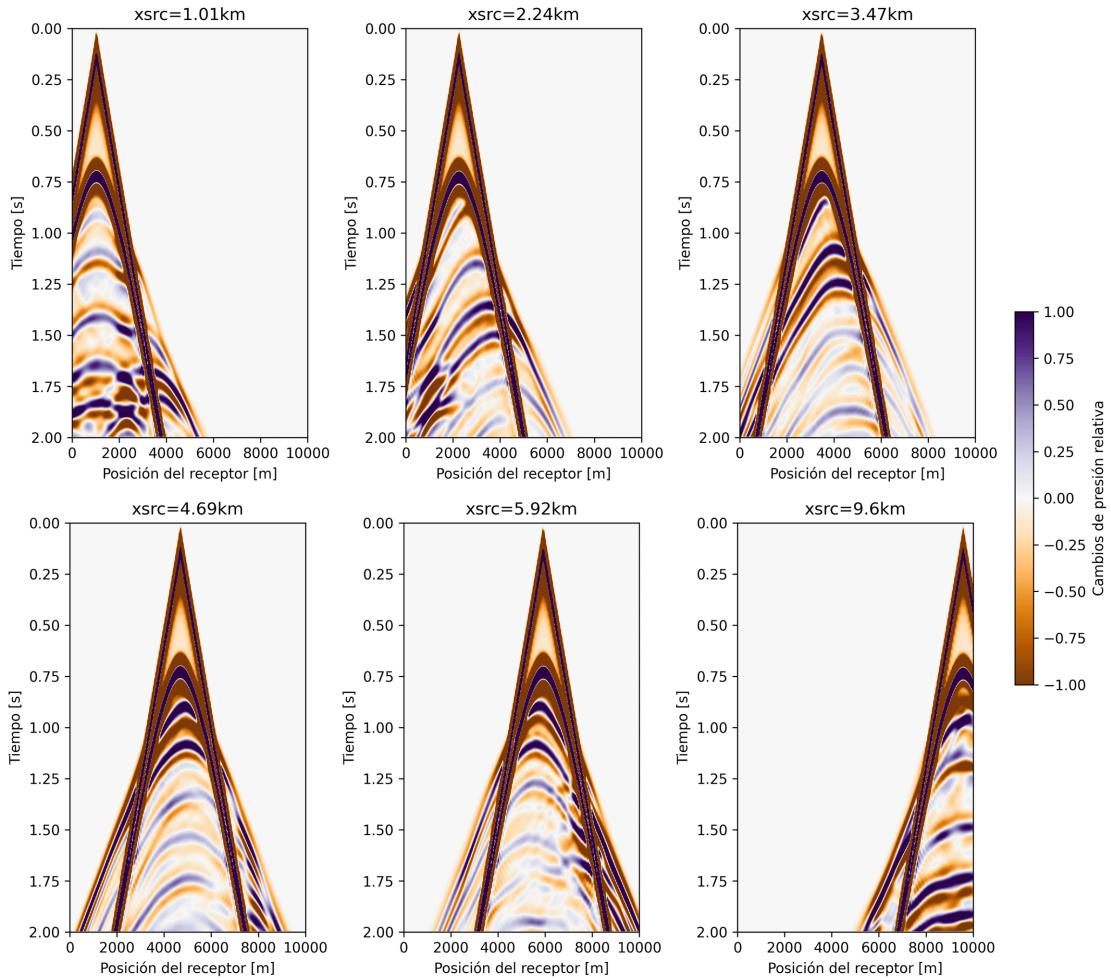


Figura 4.12: Registros de disparos para distintas fuentes. Datos de alta frecuencia - 8 Hz.

FWI

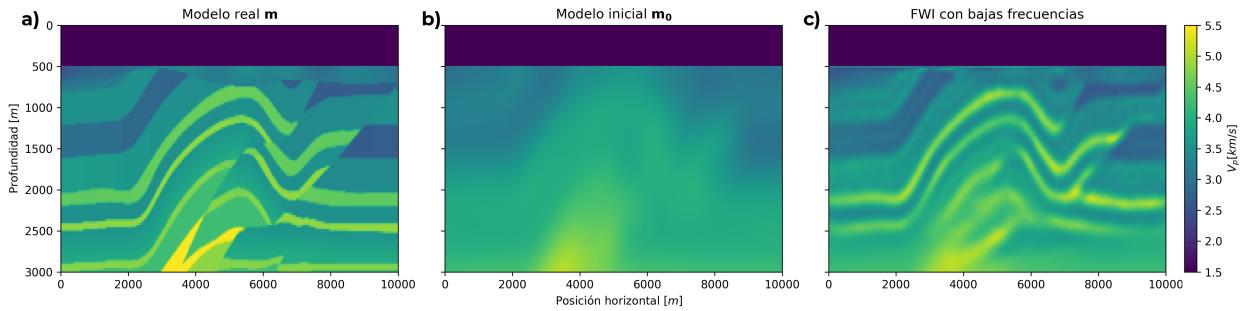


Figura 4.13: Resultados de FWI con datos de baja frecuencia - 4 HZ. a) Modelo real
b) Modelo inicial (modelo real suavizado) y c) FWI.

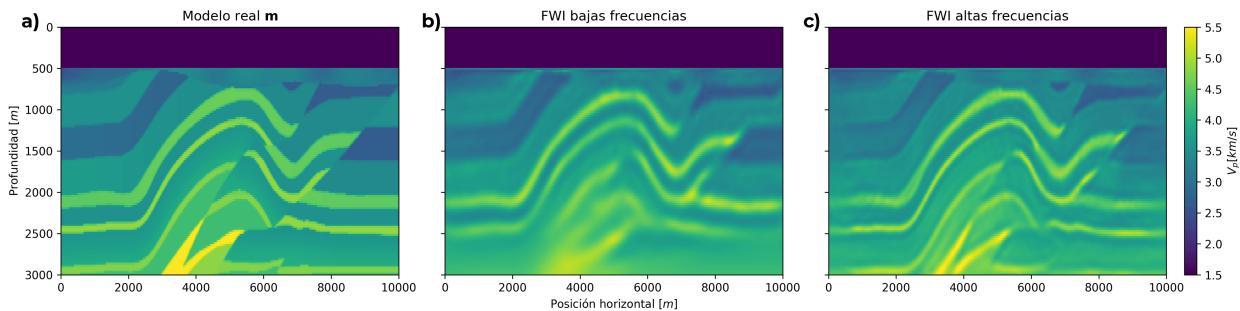


Figura 4.14: Resultados de FWI con datos de alta frecuencia con modelo inicial de FWI de baja frecuencia - 4 y 8 HZ. a) Modelo real b) Modelo inicial - resultado de FWI de bajas frecuencias y c) FWI final (4 y 8 Hz).

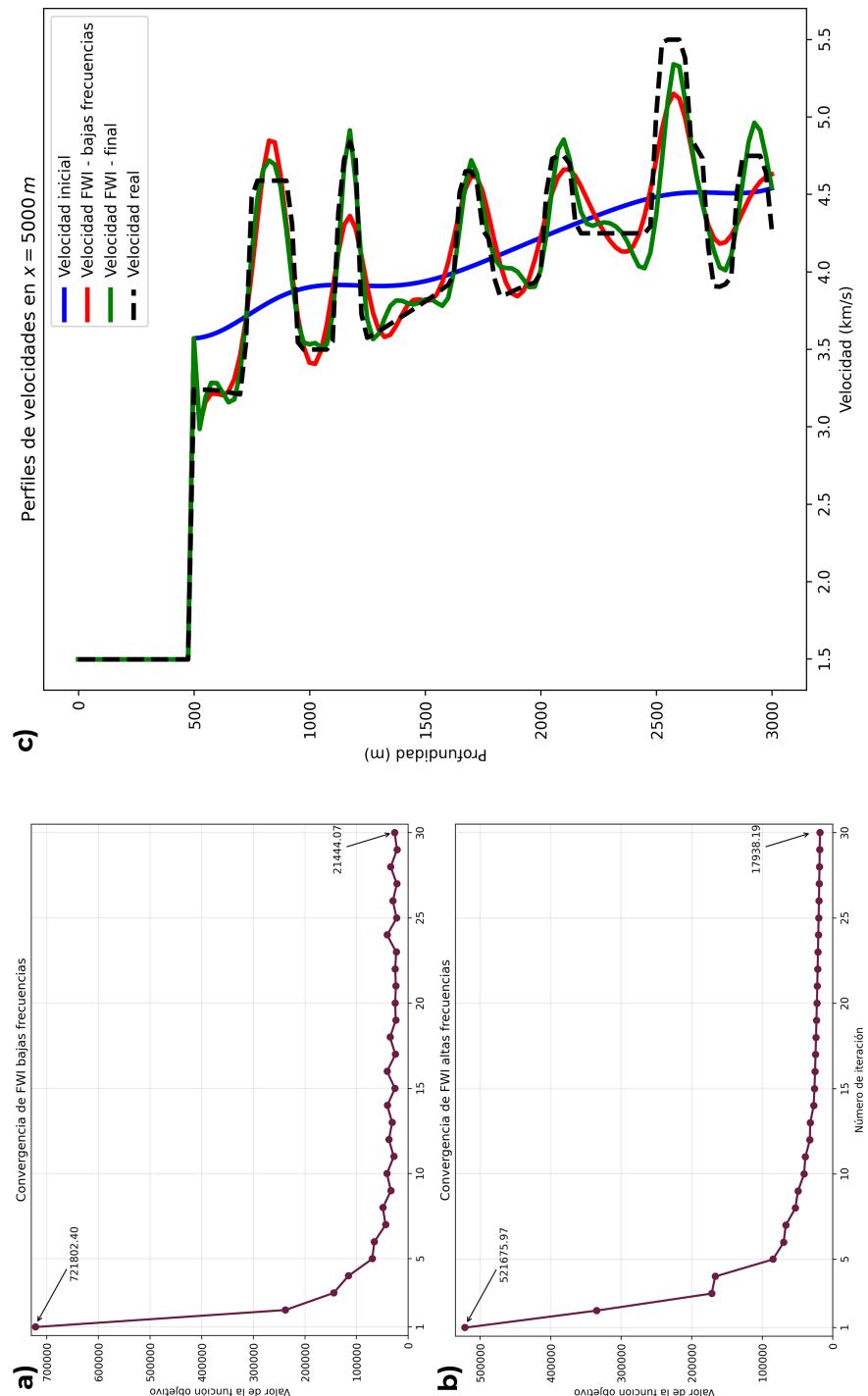


Figura 4.15: a) Función objetivo a lo largo de las iteraciones de FWI para bajas frecuencias. b) Función objetivo a lo largo de las iteraciones de FWI para altas frecuencias. c) Comparación de velocidades: Inicial, FWI bajas frecuencias, FWI altas frecuencias y real en $x = 5000\text{ m}$. Se aprecia una mejora en la velocidad del resultado final de inversión de forma de onda completa.

LS-RTM

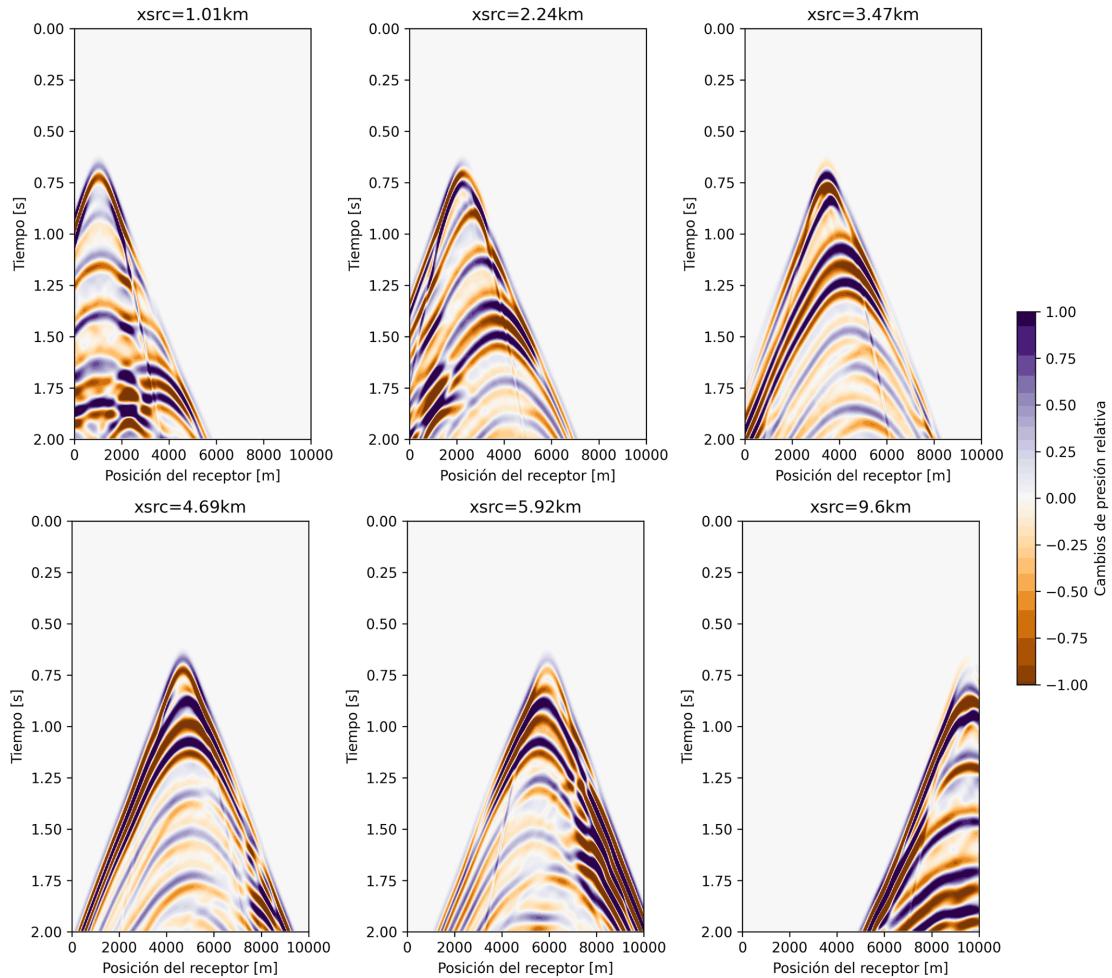


Figura 4.16: Registros de disparos para distintas fuentes. Datos de alta frecuencia linealizados (Born) - 8 Hz.

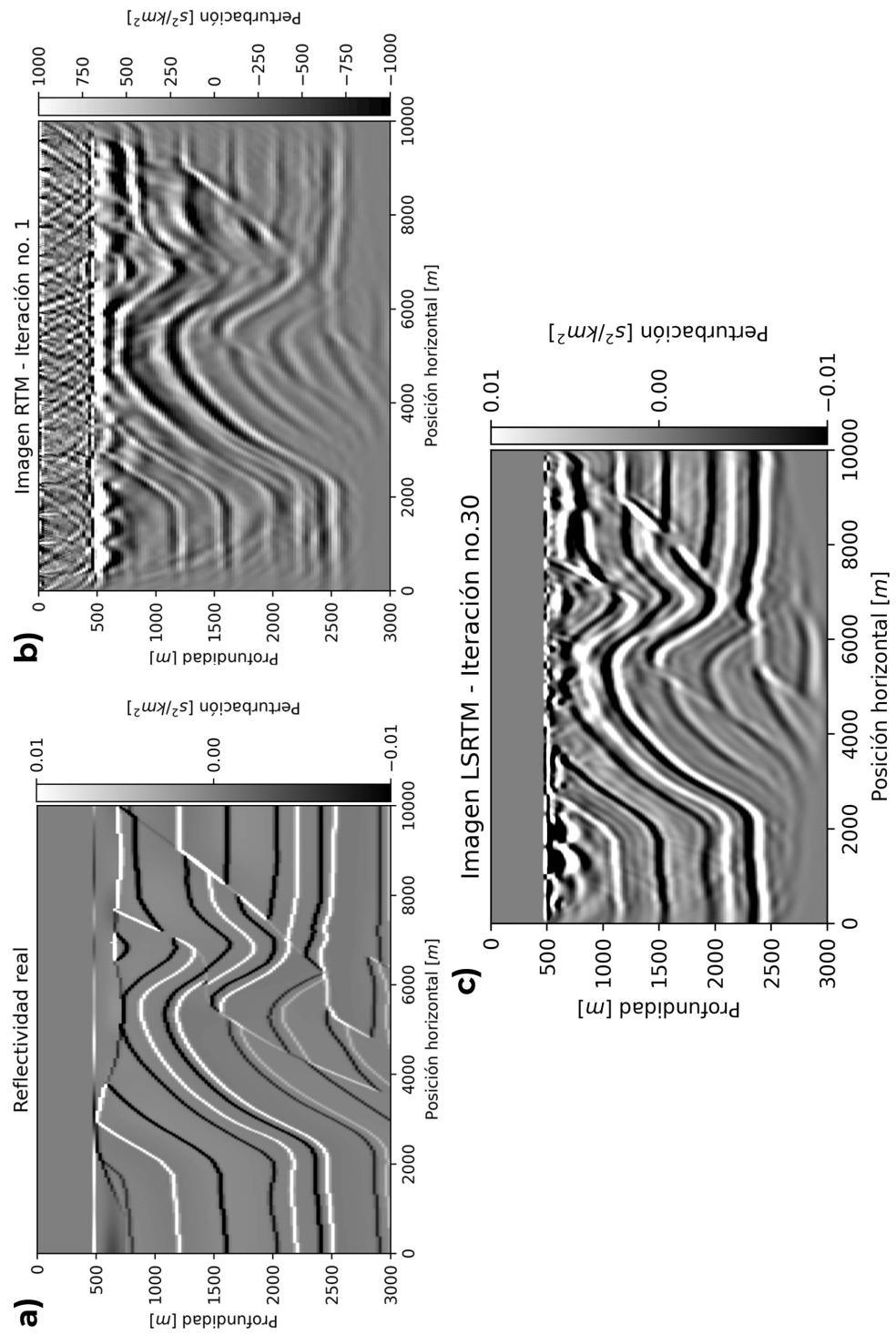


Figura 4.17: Resultados de LS-RTM con datos de alta frecuencia - 4 y 15 HZ. a) Modelo real b) RTM y c) LS-RTM final.

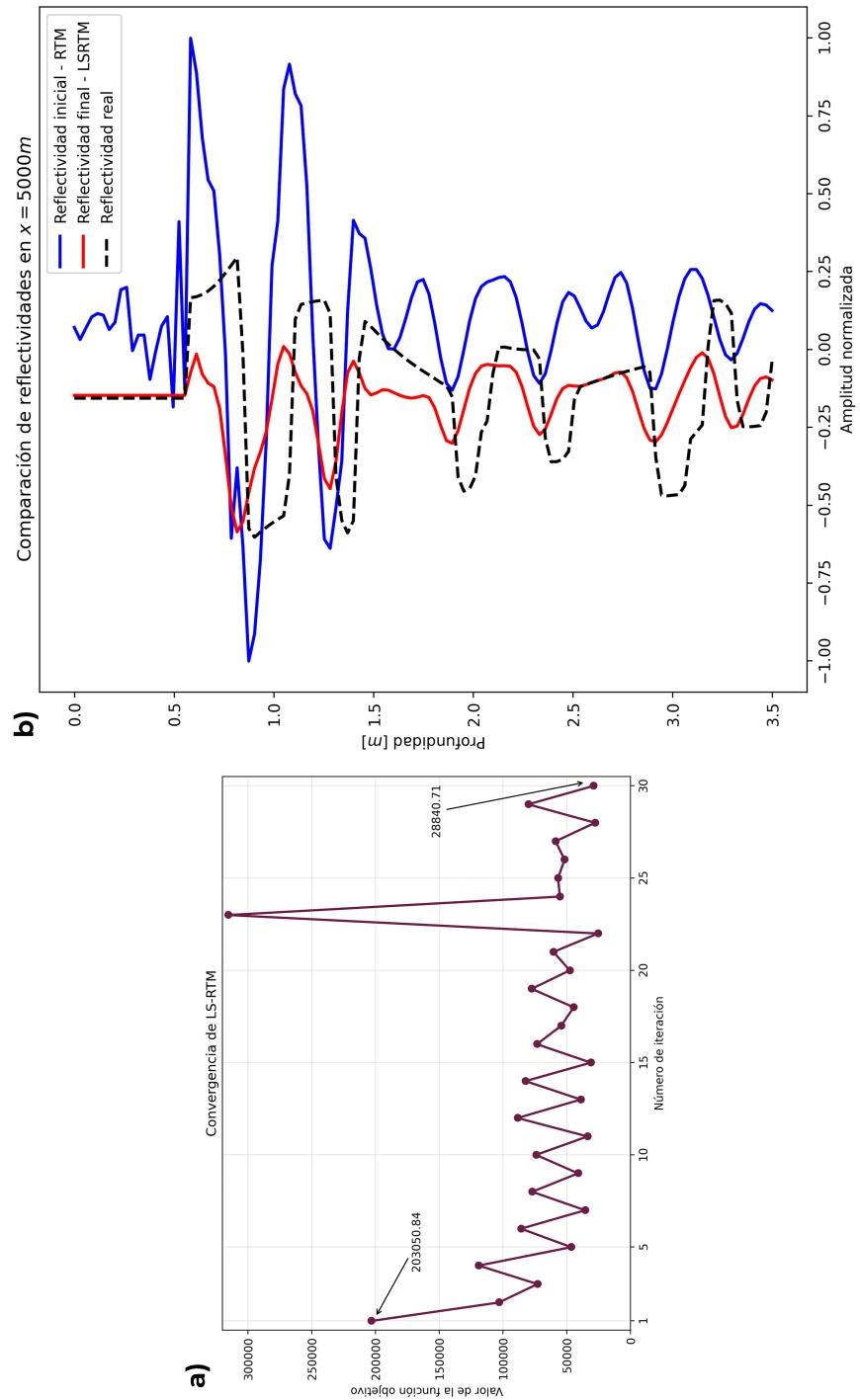


Figura 4.18: a) Función objetivo a lo largo de las iteraciones de LSRTM. b) Comparación de reflectividades normalizadas: RTM, LSRTM y real en $x = 500\text{ m}$. La imagen sísmica obtenida mediante LSRTM, a diferencia de la RTM, no cuenta con artefactos de migración en la columna de agua (primeros 500 m) además de obtener una reflectividad mas cercana a la real.

Conclusiones

Se investigó y analizó la aplicación de las herramientas computacionales de código abierto Devito y JUDI en la inversión sísmica, específicamente en la inversión de forma de onda completa y migración de mínimos cuadrados.

- Se encontró que la inversión de forma de onda completa proporciona una representación detallada y precisa del subsuelo cuando se utilizan datos de bajas frecuencia, seguido de datos de altas frecuencias, permitiendo una mejor caracterización de las estructuras geológicas complejas.
- Al comparar los resultados de la inversión de forma de onda completa en los modelos, se observa una diferencia significativa. En el primero, con un rango de *offsets* limitado (1 km), no se detectan ondas de retorno en la zona más profunda, lo que indica una falta de precisión en esa área. En cambio, en el segundo, con un rango de *offsets* más amplio (10 km), se obtiene mayor información, incluyendo la zona mas profunda con las ondas de retorno. Estos resultados resaltan la importancia de considerar los parámetros de adquisición de sísmica, como la distribución de *offsets*, para obtener datos adecuados para el proceso de inversión.
- Por otro lado, la migración de mínimos cuadrados permitió mejorar las imágenes sísmicas, comparadas con un método convencional como RTM, utilizando los modelos de velocidades obtenidos con FWI. Se eliminaron los artefactos de migración y se obtuvieron amplitudes geologicamente coherentes.
- Los resultados obtenidos demuestran la efectividad y aplicabilidad de las herramientas de código abierto en el estudio y resolución de problemas de inversión sísmica. Además de destacar la importancia de estas herramientas, ya que ofrecen accesibilidad, flexibilidad, reproducibilidad y posibilidad de personalizar, lo que permite a investigadores, alumnos y profesionales adaptar las herramientas a sus necesidades y explorar diferentes enfoques.

En resumen, este estudio demuestra que las herramientas de código abierto son eficaces y aplicables en la inversión sísmica. Al utilizar una combinación de inversión de forma de onda completa y migración de mínimos cuadrados se obtienen resultados precisos. Estas conclusiones respaldan el uso de herramientas de código abierto en la comunidad estudiantil, científica e industria sísmica, promoviendo su adopción y contribuyendo al avance de esta investigación.

A

Apéndice

Recursos computacionales

Cómputo

- SO: Ubuntu (WSL)
- Kernel: x86_64 Linux 5.15.90.1
- CPU: AMD Ryzen™ 5 7600X
- GPU: NVIDIA RTX 3060 Ti 8 GB VRAM
- RAM: 32 GB DDR5 6000 MHz

Lenguajes de programación

- Python 3.9.2
- Julia 1.9.0
- CUDA 12.1

Entorno de trabajo

- Visual Studio Code 1.79.2
- Jupyter 2023.5.1101742258

- Pylance 2023.6.40
- Julia Extension 1.47.2

Bibliotecas

- Devito 4.8.1
- NumPy 1.24.3
- SymPy 1.11.1
- Matplotlib 3.7.1
- JUDI 3.2.3
- PyCall 1.95.1
- PyPlot 2.11.1
- IJulia 1.24.0
- SegyIO 0.8.3
- SlimOptim 0.2.2
- JOLI 0.7
- ImageFiltering 0.7.4
- SetIntersectionProjection 0.2.4

Comparación de tiempos de ejecución de inversión sísmica

Tiempos de ejecución [min]			
Modelo 1			
	CPU	CPU + GPU	Factor de reducción de tiempo
FWI bajas frecuencias	8.17	1.98	4.12
FWI altas frecuencias	8.33	2.04	4.08
LS-RTM	44.32	10.49	4.23

Modelo 2			
	CPU	CPU + GPU	Factor de reducción de tiempo
FWI bajas frecuencias	13.14	3.57	3.68
FWI altas frecuencias	14.13	4.09	3.45
LS-RTM	75.17	19.62	3.83

Tabla A.1: Tiempos de ejecución para los procesos de inversión sísmica

Referencias

- Aki, K., Christoffersson, A., & Husebye, E. S. (1977). Determination of the three-dimensional seismic structure of the lithosphere. *Journal of Geophysical Research*, 82(2), 277-296. <https://doi.org/10.1029/jb082i002p00277>
- Aminzadeh, F., Burkhard, N., Nicoletis, L., Rocca, F., & Wyatt, K. (1994). SEG/EAEG 3-D modeling project: 2nd update. *The Leading Edge*, 13(9), 949-952. <https://doi.org/10.1190/1.1437054>
- Arasanipalai, S., Mavilio, A. V., & Lebit, H. (2020). Improved presalt imaging using innovative model-driven imaging technology. *First Break*, 38(3), 83-89. <https://doi.org/10.3997/1365-2397.fb2020023>
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, 59(1), 65-98. <https://doi.org/10.1137/141000671>
- Billette, F., & Brandsberg-Dahl, S. (2005). The 2004 BP Velocity Benchmark. *67th EAGE Conference & Exhibition*. <https://doi.org/10.3997/2214-4609-pdb.1.b035>
- Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A., & Wyld, H. W. (1985). Tomographic determination of velocity and depth in laterally varying media. *GEOPHYSICS*, 50(6), 903-923. <https://doi.org/10.1190/1.1441970>
- Bording, R. P., & Lines, L. R. (1997). *Seismic Modeling and Imaging with the Complete Wave Equation*. Society of Exploration Geophysicists. <https://doi.org/10.1190/1.9781560801870>
- Claerbout, J. (1976). *Fundamentals of geophysical data processing*. McGraw-Hill.
- Claerbout, J. (2004). *Earth Sounding Analysis: Processing Versus Inversion*. Blackwell Scientific Publications.
- Dellinger, J. (2016). Challenges to extending the usable seismic bandwidth at the seafloor in the deepwater Gulf of Mexico. *SEG Technical Program Expanded Abstracts 2016*. <https://doi.org/10.1190/segam2016-13762860.1>
- Dong, S., Cai, J., Guo, M., Suh, S., Zhang, Z., Wang, B., & Li, Z. (2012). Least-squares reverse time migration: towards true amplitude imaging and improving the resolution. *SEG Technical Program Expanded Abstracts 2012*. <https://doi.org/10.1190/segam2012-1488.1>
- Fletcher, R. (1987). *Practical Methods Of Optimization*. John Wiley & Sons.

- Glantz, J. (1994). Time-Reversed Sound Waves Resonate Among Physicists. *Science*, 265(5171), 474-475. <https://doi.org/10.1126/science.265.5171.474>
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13(4), 49-52.
- Hill, S. J., & Rüger, A. (2019). *Illustrated Seismic Processing Volume 1: Imaging*. Society of Exploration Geophysicists. <https://doi.org/10.1190/1.9781560803621>
- Huang, Y., Dutta, G., Dai, W., Wang, X., Schuster, G. T., & Yu, J. (2014). Making the most out of least-squares migration. *The Leading Edge*, 33(9), 954-960. <https://doi.org/10.1190/tle33090954.1>
- Lines, L. R., & Newrick, R. T. (2004). *Fundamentals of Geophysical Interpretation*. Society of Exploration Geophysicists. <https://doi.org/10.1190/1.9781560801726>
- Long, A. (2020). Back-Propagating into the Future. Least-Squares Migration (LSM), Full Waveform Inversion (FWI) and Neural Networking Training: Imaging and Characterization by Inversion. *PGS Technical Library*. <https://www.pgs.com/media-and-events/news/industry-insights--back-propagating-into-the-future/>
- Louboutin, M., Lange, M., Luporini, F., Kukreja, N., Witte, P. A., Herrmann, F. J., Velesko, P., & Gorman, G. J. (2019). Devito (v3.1.0): an embedded domain-specific language for finite differences and geophysical exploration. *Geoscientific Model Development*, 12(3), 1165-1187. <https://doi.org/10.5194/gmd-12-1165-2019>
- Luporini, F., Nelson, R., Louboutin, M., Bisbas, G., Caunt, E., & Gorman, G. (2020). Devito: a DSL and compiler for automated generation of production-grade wave propagators. *Domain-Specific Languages in High-Performance Computing*.
- Mandroux, F., Ting, C.-O., Montgomery, B., & Lenart, A. (2013). Staggered marine acquisition design for complex imaging. *SEG Technical Program Expanded Abstracts 2013*. <https://doi.org/10.1190/segam2013-0354.1>
- McMechan, G. A. (1982). Research workshop 4 - migration: Fundamental issues and future developments. *SEG Technical Program Expanded Abstracts 1982*. <https://doi.org/10.1190/1.1827145>
- McMechan, G. A. (1983). Migration by extrapolation of time-dependent boundary values. *Geophysical Prospecting*, 31(3), 413-420. <https://doi.org/10.1111/j.1365-2478.1983.tb01060.x>
- Meyer, M., Hermand, J., Le Gac, J., & Asch, M. (2004). Penalization method for WAPE adjoint-based inversion of an acoustic field. *Proceedings of the 7th European Conference on Underwater Acoustics, ECUA 2004*, 243-48.
- Mora, P. (1987). Nonlinear two-dimensional elastic inversion of multioffset seismic data. *GEOPHYSICS*, 52(9), 1211-1228. <https://doi.org/10.1190/1.1442384>
- Mora, P. (1988). Elastic wave-field inversion of reflection and transmission data. *GEOPHYSICS*, 53(6), 750-759. <https://doi.org/10.1190/1.1442510>
- Peng, C., Wang, M., Chazalnoel, N., & Gomes, A. (2018). Subsalt imaging improvement possibilities through a combination of FWI and reflection FWI. *The Leading Edge*, 37(1), 52-57. <https://doi.org/10.1190/tle37010052.1>

- Perkel, J. M. (2021). Ten computer codes that transformed science. *Nature*, 589(7842), 344-348. <https://doi.org/10.1038/d41586-021-00075-2>
- PGS. (2020). Seamless and Accurate Regional Presalt Imaging. <https://www.pgs.com/company/resources/case-studies/santos-vision/>
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *GEOPHYSICS*, 64(3), 888-901. <https://doi.org/10.1190/1.1444597>
- Schuster, G. T. (2017). *Seismic Inversion*. Society of Exploration Geophysicists. <https://doi.org/10.1190/1.9781560803423>
- Sheriff, R. E. (2002). *Encyclopedic Dictionary of Applied Geophysics*. Society of Exploration Geophysicists. <https://doi.org/10.1190/1.9781560802969>
- Sirgue, L. (2006). The Importance of Low Frequency and Large Offset in Waveform Inversion. *68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006*. <https://doi.org/10.3997/2214-4609.201402146>
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *GEOPHYSICS*, 49(8), 1259-1266. <https://doi.org/10.1190/1.1441754>
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial; Applied Mathematics. <https://doi.org/10.1137/1.9780898717921>
- Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *GEOPHYSICS*, 74(6), WCC1-WCC26. <https://doi.org/10.1190/1.3238367>
- Washbourne, J., Kaplan, S., Merino, M., Albertin, U., Sekar, A., Manuel, C., Mishra, S., Chenette, M., & Loddoch, A. (2021). Chevron optimization framework for imaging and inversion (COFII) – An open source and cloud friendly Julia language framework for seismic modeling and inversion. *First International Meeting for Applied Geoscience & Energy Expanded Abstracts*. <https://doi.org/10.1190/segam2021-3594362.1>
- Witte, P. A. (2020). *Software and algorithms for large-scale seismic inverse problems* [Tesis doctoral, Georgia Institute of Technology].