



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS DE LA TIERRA
INSTITUTO DE GEOFÍSICA

Tomografía de Forma de Onda Completa de la Subducción de la Placa
de Cocos bajo el Centro de México

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS DE LA TIERRA

PRESENTA:

Armando Espíndola Carmona

Director de Tesis:

Dr. José Carlos Ortiz Alemán

Instituto de Geofísica

Ciudad Universitaria Cd. México, Agosto 2017



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Resumen

La Tomografía Sísmica de Forma de Onda Completa actualmente se ha convertido en una técnica muy utilizada para conocer la estructura de la Tierra; esto ha sido posible gracias al gran avance del cómputo de alto rendimiento y la precisión de los métodos numéricos. En este trabajo se aplicó la Tomografía Sísmica de Forma de Onda Completa utilizando una formulación adjunta para obtener un modelo tridimensional de la subducción de la placa de Cocos bajo la placa de Norteamérica. El modelo inicial se construyó mediante la recopilación de los estudios sísmológicos y geológicos acerca de la zona de subducción. La solución del problema directo se obtuvo utilizando el método de diferencias finitas escalonadas propuesto por Madariaga [1976] con una precisión en la aproximación espacial de octavo orden y de segundo orden en tiempo. El mejoramiento del modelo de velocidades se realizó mediante el método del gradiente conjugado, minimizando el desajuste en los tiempos de arribo \mathcal{T} así como la norma L_2 entre los datos observados y sintéticos. Las diferencias entre los sismogramas sintéticos calculados utilizando el modelo mejorado y los sismogramas observados muestran una reducción en el error para frecuencias menores a 0.25 [Hz]. Con los resultados obtenidos en este trabajo fue posible obtener un modelo tridimensional con resolución mejorada, donde es posible apreciar la geometría aproximada del slab de subducción.

Índice general

Índice	i
Introducción	1
1. Inversión de Forma de Onda Completa: Teoría	3
1.1. Método Adjunto	4
1.1.1. Formulación General	5
1.1.2. Aplicación en la ecuación de la elastodinámica	8
1.1.3. Kernels de sensibilidad	11
1.1.4. Fuente adjunta y funcionales de desajuste	14
1.1.4.1. Función de Green Adjunta	14
1.1.4.2. Norma \mathbf{L}_2	16
1.1.4.3. Correlación cruzada - Tiempos de arribo $\Delta\mathcal{T}$	17
1.2. Optimización: Gradiente conjugado	19
1.2.1. Gradiente Conjugado	20
1.2.2. Selección del modelo de prueba ν_t	21
1.2.2.1. Selección del paso de prueba : Polinomio cuadrático	21
2. Zona de Subducción Cocos-Rivera: Estudios Previos	23
2.1. Configuración tectónica	23
2.1.1. Placa de Cocos	24
2.1.1.1. Configuración del slab de subducción	25
2.1.2. Placa de Rivera	27

2.1.2.1. Configuración del slab de subducción	27
3. Metodología	29
3.1. Problema directo	29
3.2. Base de Datos	32
3.3. Modelo inicial \mathbf{M}_0	34
3.4. Análisis de sensibilidad	38
3.5. Procesamiento de datos	40
3.5.1. Construcción fuentes adjuntas	41
3.5.2. Post-procesamiento kernels de sensibilidad	42
4. Resultados	45
5. Discusión y Conclusiones	55
Referencias	59

Introducción

La Faja Volcánica Trans Mexicana (FVTM) es una de las provincias geológicas más estudiadas del territorio mexicano debido a que es la expresión superficial del vulcanismo asociada a la actual subducción de las placas Rivera y Cocos bajo la placa Norteamericana, además de ser parte de un límite convergente muy particular en el que el arco volcánico no es paralelo a la trinchera [Ferrari, 2004; Ferrari et al., 2012]. Se han realizado una gran cantidad de estudios tratando de caracterizar el subsuelo bajo esta región. Algunos de los estudios que se han realizado para obtener un modelo de la subducción de la placa de Cocos son los siguientes: localización de hipocentros [Pardo y Suárez, 1995], tomografía sísmica clásica [Husker y Davis, 2009], perfiles gravimétricos [Campos-Enriquez y Sánchez-Zamora, 2000; Urrutia-Fucugauchi y Flores-Ruiz, 1996], estudios electromagnéticos, específicamente magnetotelúricos [Jödicke et al., 2006], utilizando el método funciones receptor [Pérez-Campos et al., 2008; Melgar y Pérez-Campos, 2011], entre otros. Sin embargo, a pesar de todos los estudios previos realizados en esta zona para obtener un modelo que defina con mayor detalle la estructura y geometría de las placas involucradas en este proceso, sólo se han obtenido modelos con resolución limitada especialmente en la región donde cambia el ángulo de subducción. De esta forma, existe un área de oportunidad para la utilización de nuevos métodos de reconstrucción de imágenes para obtener modelos detallados de la geometría del slab.

Actualmente, debido al creciente desarrollo de la sismología teórica, de los métodos numéricos y principalmente del cómputo de alto rendimiento, se han conseguido avances sustanciales en la simulación de la propagación de ondas en medios geológicos que van desde escalas de decenas de kilómetros hasta modelos globales de la Tierra [Komatitsch y Tromp, 2002; Peter et al., 2011]. Estos avances permiten una alta precisión y eficacia en el cálculo de sismogramas sintéticos, con el empleo de técnicas robustas como la Inversión de la Forma de Onda Completa o FWI (*Full Wave*

form Inversion), por sus siglas en inglés.

El método de Inversión de la Forma de Onda Completa (FWI) es una variante de la tomografía sísmica clásica (teoría de rayos); la FWI se caracteriza por la solución numérica de las ecuaciones de la elastodinámica y el uso de toda la información contenida en los datos generados por el modelo numérico y los datos observados [Fichtner et al., 2008]. También se caracteriza por el mejoramiento en las imágenes tomográficas en cada iteración, tomando en cuenta la relación no lineal que existe entre los parámetros y los datos sintéticos [Virieux y Operto, 2009].

El método de FWI se ha utilizado tanto en escalas locales, regionales y globales, como también en sísmica de exploración [Luo y Schuster, 1991; Fichtner et al., 2008; Tape et al., 2010; Zhu, 2013], dando buenos resultados en cuanto a resolución y precisión de los modelos obtenidos. Debido al éxito de la técnica en diferentes escenarios, geológicamente hablando, surge la motivación de utilizar esta metodología para intentar mejorar los modelos que se tienen actualmente para la zona de subducción.

El presente estudio tiene como objetivo aplicar el método de la FWI a los datos sismológicos de la red *MesoAmerican Subduction Experiment*, MASE, para estimar la geometría del slab de subducción bajo el sector centro-oriental de la FVTM e intentar caracterizar la región de transición de la inclinación del slab.

El capítulo 1 de este trabajo detalla los conceptos teóricos en que se fundamenta la técnica FWI, el desarrollo matemático de la formulación adjunta, la construcción de la fuente adjunta definida por la función de desajuste elegida entre los datos, el cálculo de los kernels de sensibilidad y por último el método de optimización utilizado en la inversión. En el capítulo 2 se realiza un resumen de los antecedentes sismológicos y tectónicos del área de estudio. Posteriormente, en el capítulo 3 se describe el procesamiento aplicado a los datos sintéticos y observados para su uso óptimo en el método de inversión; aquí también se detallan los parámetros utilizados en cada iteración y el post-procesamiento de los kernels de sensibilidad y el gradiente de la función de desajuste en cada paso intermedio del proceso de optimización. Finalmente, la última parte de este trabajo se dedica a mostrar los resultados obtenidos mediante la FWI, así como a su discusión y comparación con los estudios previos en la zona de estudio.

Capítulo 1

Inversión de Forma de Onda Completa: Teoría

La Inversión de Forma de Onda Completa, FWI, basada en la solución numérica de las ecuaciones de la elastodinámica, se remonta a los 80's Tarantola [1986] en la problemática de obtener modelos 1-D y 2-D en sísmica de exploración. La transición actual que permite pasar de modelos 2-D a 3-D se debe principalmente al avance en conjunto de la capacidad computacional con la que hoy se cuenta y al éxito en la simulación de la propagación de ondas en medios geológicos complejos Komatitsch y Tromp [2002]; Peter et al. [2011]. Las simulaciones numéricas de la ecuación de onda, por ejemplo, por medio del método de diferencias finitas, nos permiten obtener el campo de ondas completo generado por un medio geológico complejo. Toda esta información contenida en los sismogramas sintéticos puede ser utilizada por la técnica FWI, que intenta minimizar a la funcional $\sum_{r=1}^N \int [u(\mathbf{x}^r, t) - u^0(\mathbf{x}^r, t)]^2 dt$, que representa la norma L_2 acumulativa, y que mide el error entre los datos sintéticos u y los datos observados u^0 , para un número N de receptores, ubicados en \mathbf{x}^r .

Para el proceso de minimización de la funcional se emplean métodos de optimización iterativos tanto locales como globales. Algoritmos que han probado ser eficientes para este tipo de problema, en específico para la FWI, son: el método denominado *steepest decent* [Tarantola, 1986] y variantes del método del gradiente conjugado [Tape et al., 2010].

Uno de los pilares del método de FWI es la formulación adjunta, que fue desarrollada originalmente en la teoría de control óptimo Lions [1968] y utilizada posteriormente en el campo de las geociencias Chavent et al. [1975]. La formulación adjunta de la FWI permite calcular el gradiente de cualquier funcional diferenciable por medio de la solución del problema directo y su problema adjunto. En otras palabras son solamente necesarias dos simulaciones para obtener el gradiente de la funcional, lo cual es mucho más eficiente que calcular el gradiente por medio de diferencias finitas. Esta última forma de calcular el gradiente resulta impráctica debido a que el espacio de parámetros del modelo es muy grande, y el costo computacional es muy alto. Las aplicaciones del método adjunto no sólo incluyen la sismología sino también meteorología y modelado de aguas subterráneas [Plessix, 2006].

Este capítulo tiene como objetivo introducir los conceptos teóricos relacionados con la formulación adjunta de la FWI, la construcción de la fuente adjunta definida por la función de desajuste elegida entre los datos, el cálculo de los kernels de sensibilidad y, por último, el método de optimización utilizado en la inversión. Para este propósito se utiliza la notación propuesta por Fichtner (2010).

1.1. Método Adjunto

El método adjunto es una herramienta matemática que nos permite calcular el gradiente de cualquier funcional de manera eficiente y práctica, siendo independiente del problema físico que se este resolviendo. En este apartado se aborda de manera general la formulación matemática del método, para posteriormente obtener las ecuaciones generales adjuntas que hagan posible pasar de manera simple al problema sismológico.

En teoría de control se considera una variable física, \mathbf{u} , que es la salida de un sistema dinámico, cuyo comportamiento depende de los parámetros del modelo, \mathbf{m} , mismo que está relacionado con la variable física \mathbf{u} por medio de un sistema físico denotado por un operador \mathbf{L} ,

$$\mathbf{L}(\mathbf{u}(\mathbf{m}), \mathbf{m}) = \mathbf{f}, \quad (1.1)$$

donde \mathbf{f} representa a las fuerzas externas. El objetivo es buscar los parámetros del modelo \mathbf{m} en

función de la variable física \mathbf{u} , tal que el sistema opere de manera óptima. En términos matemáticos el concepto óptimo significa minimizar o maximizar, según sea el caso, la funcional de un problema específico $\chi(\mathbf{m}) = \chi[\mathbf{u}(\mathbf{m})]$. Una perturbación en la funcional χ , se produce en respuesta a una perturbación a los parámetros del modelo \mathbf{m} , denotada como $\mathbf{m} + \varepsilon\delta\mathbf{m}$. Esta perturbación puede ser aproximada por medio de la derivada de χ con respecto a \mathbf{m} en la dirección de la perturbación $\delta\mathbf{m}$, la cual queda definida de la siguiente manera [Arfken y Weber, 1999]:

$$\nabla_m \chi(\mathbf{m})\delta\mathbf{m} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\chi(\mathbf{m} + \varepsilon\delta\mathbf{m}) - \chi(\mathbf{m})]. \quad (1.2)$$

Debido a que en muchas aplicaciones, incluyendo a la tomografía sísmica, el número de parámetros del modelo es muy grande y la solución del problema directo Ec.(1.1) es computacionalmente demandante, es prácticamente imposible aproximar el gradiente de la funcional en todas las direcciones de $\delta\mathbf{m}$ por medio de diferencias finitas. Esta es la razón principal por la que los métodos adjuntos resultan demasiado útiles.

1.1.1. Formulación General

Considerando la variable física \mathbf{u} , que depende del vector de posición $\mathbf{x} \in G \subset \mathbb{R}^3$, el tiempo $t \in T = [t_0, t_1]$ y los parámetros del modelo $\mathbf{m} \in \mathfrak{M}$:

$$\mathbf{u} = \mathbf{u}(\mathbf{m}; \mathbf{x}, t). \quad (1.3)$$

El espacio de modelos \mathfrak{M} engloba todos los posibles parámetros del modelo y el punto y coma en la Ec.(1.3) indica que la variable física \mathbf{u} depende del tiempo y el espacio. En sismología \mathbf{u} representa el campo de desplazamientos, que se encuentran relacionado con los parámetros por medio de la ecuación de onda $\mathbf{L}(\mathbf{u}(\mathbf{m}), \mathbf{m}) = \mathbf{f}$, que se puede encontrar sujeta a fuerzas externas \mathbf{f} . Los parámetros que se modelan en sismología son la densidad ρ y el tensor elástico \mathbf{C} .

Para el método adjunto no es de utilidad el uso de los desplazamientos \mathbf{u} directamente, sino que es necesario el uso de una funcional objetivo $\chi(\mathbf{x})$, tal que pueda ser usada para cuantificar la similitud de las variables físicas observadas $\mathbf{u}^0(\mathbf{x}, t)$ y los datos sintéticos $\mathbf{u}(\mathbf{x}, t)$ obtenidos por la solución numérica del modelo físico en cuestión. Una de las funcionales que se utiliza en sismología

es la norma L_2 , que cuantifica la diferencia entre los datos observados y sintéticos, en la posición de los receptores $\mathbf{x} = \mathbf{x}^r$:

$$\chi(\mathbf{m}) = \frac{1}{2} \int_T \int_G [\mathbf{u}(\mathbf{m}; \mathbf{x}, t) - \mathbf{u}^0(\mathbf{x}, t)]^2 dt d^3\mathbf{x}. \quad (1.4)$$

La Ec.(1.4) es un caso especial de una funcional objetivo, que es muy utilizada en varias áreas de la geofísica. Sin pérdida de generalidad, para el método adjunto se define la funcional χ de la siguiente forma:

$$\chi(\mathbf{m}) = \int_T \int_G \chi_1[\mathbf{u}(\mathbf{m}; \mathbf{x}, t)] dt d^3\mathbf{x} = \langle \chi_1(\mathbf{m}) \rangle, \quad (1.5)$$

en donde χ_1 es cualquier funcional para mantener la generalidad del desarrollo; también se introdujo $\langle \cdot \rangle$ para tener una notación corta de la integral de $T \times G$. Para poder obtener el modelo óptimo que se busca, es necesario calcular la derivada de la funcional $\chi(\mathbf{m})$ con respecto de los parámetros \mathbf{m} en la dirección de la perturbación $\delta\mathbf{m}$, siguiendo la regla de la cadena:

$$\nabla_m \chi \delta\mathbf{m} = \nabla_u \chi \delta\mathbf{u} = \langle \nabla_u \chi_1 \cdot \delta\mathbf{u} \rangle. \quad (1.6)$$

donde

$$\delta\mathbf{u} := \nabla_m \mathbf{u} \delta\mathbf{m}. \quad (1.7)$$

La Ec.(1.7) define la derivada de \mathbf{u} con respecto a \mathbf{m} en la dirección de $\delta\mathbf{m}$. Una de las dificultades del uso de la Ec.(1.6) directamente para el cálculo del gradiente es que aparece el término $\delta\mathbf{u}$, que es difícil de calcular numéricamente, debido a que si se utiliza una aproximación de primer orden de la Ec. (1.6), es necesario calcular $\nabla_m \mathbf{u}$ para cada dirección posible de $\delta\mathbf{m}$. Sin embargo, actualmente es casi imposible calcular la derivada de $\chi(\mathbf{m})$ con respecto a los parámetros \mathbf{m} para modelos relativamente grandes; de esta forma, no es posible utilizar la Ec.(1.6) tal cual. Es necesario buscar la manera de eliminar $\delta\mathbf{u}$ de la Ec.(1.6). Con este fin derivamos la Ec.(1.1) con respecto los parámetros \mathbf{m} , dando como resultado

$$\nabla_m \mathbf{L} \delta\mathbf{m} + \nabla_u \mathbf{L} \delta\mathbf{u} = 0. \quad (1.8)$$

El termino de fuentes externas \mathbf{f} desaparece en la Ec.(1.8) debido a que no depende de los parámetros del modelo \mathbf{m} . Posteriormente se multiplica por la izquierda la Eq. (1.8) por el operador adjunto de \mathbf{u} , \mathbf{u}^\dagger , y se aplica el operador $\langle \cdot \rangle$, obteniendo:

$$\langle \mathbf{u}^\dagger \cdot \nabla_m \mathbf{L} \delta \mathbf{m} \rangle + \langle \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \delta \mathbf{u} \rangle = 0. \quad (1.9)$$

Sumando la Ec.(1.6) y la Ec.(1.9)

$$\nabla_m \chi \delta \mathbf{m} = \langle \mathbf{u}^\dagger \cdot \nabla_m \mathbf{L} \delta \mathbf{m} \rangle + \langle \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \delta \mathbf{u} \rangle + \langle \nabla_u \chi_1 \cdot \delta \mathbf{u} \rangle. \quad (1.10)$$

Se puede reescribir la Ec.(1.10) usando los operadores adjuntos $\nabla_u \chi_1^\dagger$ y $\nabla_u \mathbf{L}^\dagger$, los cuales satisfacen las siguientes relaciones, respectivamente:

$$\langle \nabla_u \chi_1 \cdot \delta \mathbf{u} \rangle = \langle \delta \mathbf{u} \cdot \nabla_u \chi_1^\dagger \rangle. \quad (1.11)$$

y

$$\langle \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \delta \mathbf{u} \rangle = \langle \delta \mathbf{u} \cdot \nabla_u \mathbf{L}^\dagger \mathbf{u}^\dagger \rangle, \quad (1.12)$$

para cualquier $\delta \mathbf{u}$ y \mathbf{u}^\dagger . Por lo tanto, se obtiene

$$\nabla_m \chi \delta \mathbf{m} = \langle \delta \mathbf{u} \cdot (\nabla_u \chi_1^\dagger + \nabla_u \mathbf{L}^\dagger \mathbf{u}^\dagger) \rangle + \langle \mathbf{u}^\dagger \cdot \nabla_m \mathbf{L} \delta \mathbf{m} \rangle. \quad (1.13)$$

Ahora, es posible eliminar el termino $\delta \mathbf{u}$ de la Ec.(1.13), si se puede determinar el campo \mathbf{u}^\dagger que satisface la Ec.(1.14), a la cual se le conoce como la *ecuación adjunta* de (1.1).

$$\nabla_u \mathbf{L}^\dagger \mathbf{u}^\dagger = -\nabla_u \chi_1^\dagger. \quad (1.14)$$

en donde \mathbf{u}^\dagger y $-\nabla_u \chi_1^\dagger$ son el campo adjunto y la fuente adjunta, respectivamente. Cuando se resuelve la *ecuación adjunta*, se obtiene el campo adjunto \mathbf{u}^\dagger y entonces la derivada de la funcional objetivo, Ec.(1.15), se reduce a

$$\nabla_m \chi \delta \mathbf{m} = \langle \mathbf{u}^\dagger \cdot \nabla_m \mathbf{L} \delta \mathbf{m} \rangle. \quad (1.15)$$

Por construcción, la Ec.(1.15) puede ser calculada para cualquier dirección de $\delta\mathbf{m}$ sin conocer explícitamente $\delta\mathbf{u}$. Sin embargo, este beneficio conlleva el costo de calcular el campo adjunto \mathbf{u}^\dagger de la Ec.(1.14).

1.1.2. Aplicación en la ecuación de la elastodinámica

En la sección anterior se desarrollaron los fundamentos teóricos del método adjunto en un modo general, que puede ser aplicado para cualquier funcional χ y para cualquier modelo físico que se quiera trabajar. En el contexto de la geofísica, específicamente en sismología, se define el operador del sistema físico \mathbf{L} como la ecuación de onda y en donde \mathbf{m} representa las propiedades físicas de la Tierra.

La ecuación de onda se define de la siguiente manera:

$$\rho(\mathbf{x})\ddot{\mathbf{u}} - \nabla \cdot \sigma(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t), \quad \mathbf{x} \in G \subset \mathbb{R}^3, \quad t \in [t_0, \infty] \subset \mathbb{R}. \quad (1.16)$$

La Ec.(1.16) relaciona el campo de desplazamientos \mathbf{u} con los parámetros, que en este caso son la densidad ρ y el tensor de esfuerzos σ y una densidad de fuerza externa definida como \mathbf{f} . Asumiendo un medio visco-elástico lineal, el tensor de esfuerzos σ se relaciona con el gradiente del campo de desplazamientos $\nabla\mathbf{u}$ por la relación constitutiva [Aki y Richards, 2002]

$$\sigma(\mathbf{x}, t) = \int_{\tau=t_0}^{\infty} \dot{\mathbf{C}}(\mathbf{x}, t - \tau) : \nabla\mathbf{u}(\mathbf{x}, \tau) d\tau. \quad (1.17)$$

Sustituyendo la Ec.(1.17) en la Ec.(1.16) es posible obtener la ecuación de onda en terminos del operador \mathbf{L} :

$$\mathbf{L}(\mathbf{u}, \rho, \mathbf{C}) = \mathbf{f}, \quad (1.18)$$

$$\mathbf{L}(\mathbf{u}, \rho, \mathbf{C}) = \rho(\mathbf{x})\ddot{\mathbf{u}} - \nabla \cdot \int_{\tau=t_0}^{\infty} \dot{\mathbf{C}}(\mathbf{x}, t - \tau) : \nabla\mathbf{u}(\mathbf{x}, \tau) d\tau. \quad (1.19)$$

El operador de ecuación de onda \mathbf{L} viene definido con condiciones iniciales y condiciones de frontera

$$\mathbf{u}|_{t \leq t_0} = \dot{\mathbf{u}}|_{t \leq t_0} = \mathbf{0}, \quad \mathbf{n} \cdot \boldsymbol{\sigma}|_{\mathbf{x} \in \partial G} = \mathbf{0}. \quad (1.20)$$

De acuerdo a la Sección 1.1.1, para poder calcular el gradiente de la funcional objetivo χ , con respecto los parámetros ρ y \mathbf{C} es necesario encontrar el operador adjunto de la Ec.(1.9). Recordando la Ec.(1.12), se puede obtener la siguiente relación:

$$\langle \mathbf{u} \cdot \nabla_u \mathbf{L}^\dagger \mathbf{u}^\dagger \rangle = \langle \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \mathbf{u} \rangle. \quad (1.21)$$

La versión explícita de la parte derecha de la Ec.(1.21) se muestra en la Ec.(1.22), donde se omiten la dependencia espacial de las variables por cuestiones prácticas y de claridad. El objetivo principal es separar el campo de desplazamientos \mathbf{u} de la Ec.(1.22) tal que no sea diferenciado el campo \mathbf{u} .

$$\begin{aligned} \langle \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \mathbf{u} \rangle &= \int_T \int_G \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \mathbf{u} \, d^3 \mathbf{x} \, dt = \int_T \int_G \rho \mathbf{u}^\dagger \cdot \ddot{\mathbf{u}} \, d^3 \mathbf{x} \, dt \\ &\quad - \int_T \int_G \mathbf{u}^\dagger \cdot \left[\nabla \cdot \int_{\tau=t_0}^t \dot{\mathbf{C}}(t-\tau) : \nabla \mathbf{u}(\tau) \, d\tau \right] \, dt \, d^3 \mathbf{x} \end{aligned} \quad (1.22)$$

Iniciando con el primer término de la derecha de la Ec.(1.22), e integrando por partes, se obtiene

$$\begin{aligned} \langle \mathbf{u}^\dagger \cdot \rho \ddot{\mathbf{u}} \rangle &= \int_T \int_G \rho \mathbf{u}^\dagger \cdot \ddot{\mathbf{u}} \, d^3 \mathbf{x} \, dt = \int_T \int_G \rho \mathbf{u} \cdot \ddot{\mathbf{u}}^\dagger \, d^3 \mathbf{x} \, dt \\ &\quad + \int_G \rho \dot{\mathbf{u}} \cdot \mathbf{u}^\dagger \, d^3 \mathbf{x} \Big|_{t=t_1} - \int_G \rho \dot{\mathbf{u}} \cdot \dot{\mathbf{u}}^\dagger \, d^3 \mathbf{x} \Big|_{t=t_1}, \end{aligned} \quad (1.23)$$

donde ya se evaluaron las condiciones iniciales, Ec.(1.20), y definiendo las *condiciones terminales* del campo adjunto

$$\mathbf{u}^\dagger|_{t \geq t_1} = \dot{\mathbf{u}}^\dagger|_{t \geq t_1} = \mathbf{0}, \quad (1.24)$$

con lo que se obtiene el primer termino del operador adjunto:

$$\langle \mathbf{u}^\dagger \cdot \rho \ddot{\mathbf{u}} \rangle = \langle \mathbf{u} \cdot \rho \ddot{\mathbf{u}}^\dagger \rangle. \quad (1.25)$$

Siguiendo el procedimiento matemático del término derecho en la Ec.(1.22), desarrollado por Fichtner (2010), se obtiene

$$\begin{aligned} \Upsilon = \langle \mathbf{u}^\dagger \cdot (\nabla \cdot \sigma) \rangle &= \int_{\partial G} \int_{t=t_0}^{t_1} \mathbf{u}^\dagger(t) \cdot \left[\int_{\tau=t_0}^t \dot{\mathbf{C}}(t-\tau) : \nabla \mathbf{u}(\tau) d\tau \right] \cdot \mathbf{n} dt d^2\mathbf{x} \\ &\quad - \int_{\partial G} \int_{\tau=t_0}^{t_1} \mathbf{u}(\tau) \cdot \left[\int_{t=\tau}^{t_1} \dot{\mathbf{C}}(t-\tau) : \nabla \mathbf{u}^\dagger(t) dt \right] \cdot \mathbf{n} d\tau d^2\mathbf{x} \\ &\quad + \int_G \int_{\tau=t_0}^{t_1} \mathbf{u}(\tau) \cdot \left[\nabla \cdot \int_{t=\tau}^{t_1} \dot{\mathbf{C}}(t-\tau) : \nabla \mathbf{u}^\dagger(t) dt \right] d\tau d^3\mathbf{x}. \end{aligned} \quad (1.26)$$

El primer termino de la Ec.(1.26) es igual a cero, debido a que el tensor de esfuerzos normal a la superficie $\sigma \cdot \mathbf{n} = \mathbf{0}$ en ∂G , como consecuencia de la condición de superficie libre. Sabiendo que el campo de desplazamientos adjunto \mathbf{u}^\dagger solo tiene condiciones terminales, es posible imponer condiciones de frontera al campo adjunto, tales que el segundo termino de la derecha de la Ec.(1.26) sea igual a cero:

$$\mathbf{n} \cdot \sigma^\dagger \Big|_{\mathbf{x} \in \partial G} = \mathbf{0}, \quad (1.27)$$

con tensor adjunto σ^\dagger definido por

$$\sigma^\dagger(\mathbf{x}, t) = \int_{\tau=t}^{t_1} \dot{\mathbf{C}}(\mathbf{x}, \tau - t) : \nabla \mathbf{u}^\dagger(\mathbf{x}, \tau) d\tau. \quad (1.28)$$

El termino restante ya se encuentra en la forma que se requiere, por lo que el siguiente paso es sustituir la Ec.(1.26) y la Ec.(1.25) en la Ec.(1.22), obteniendo

$$\langle \mathbf{u} \cdot \nabla_u \mathbf{L}^\dagger \mathbf{u}^\dagger \rangle = \langle \mathbf{u} \cdot \rho \ddot{\mathbf{u}}^\dagger \rangle - \Upsilon = \langle \mathbf{u} \cdot (\rho \ddot{\mathbf{u}}^\dagger - \nabla \cdot \sigma^\dagger) \rangle. \quad (1.29)$$

De la Ec.(1.29) se puede observar que el operador adjunto $\nabla_u \mathbf{L}^\dagger$ se encuentra dado por

$$\nabla_u \mathbf{L}^\dagger \mathbf{u}^\dagger = \rho(\mathbf{x}) \ddot{\mathbf{u}}^\dagger - \nabla \cdot \sigma^\dagger(\mathbf{x}, t). \quad (1.30)$$

Por lo tanto, para poder calcular la derivada de la funcional objetivo, es necesario solucionar la ecuación adjunta

$$\rho(\mathbf{x}) \ddot{\mathbf{u}}^\dagger - \nabla \cdot \sigma^\dagger(\mathbf{x}, t) = -\nabla_u \chi_1^\dagger, \quad (1.31)$$

sujeta a la condiciones terminales y de frontera

$$\mathbf{u}^\dagger|_{t \geq t_1} = \dot{\mathbf{u}}^\dagger|_{t \geq t_1} = \mathbf{0} \quad \mathbf{n} \cdot \boldsymbol{\sigma}^\dagger|_{\mathbf{x} \in \partial G} = \mathbf{0}. \quad (1.32)$$

Si se considera un medio no disipativo, el operador de la ecuación de onda \mathbf{L} es auto-adjunto, i.e. que $\mathbf{L} = \mathbf{L}^\dagger$. Considerando lo anterior, el operador adjunto \mathbf{L}^\dagger sigue representando a la ecuación de onda, por lo cual puede resolverse empleando los mismos métodos numéricos usados en la ecuación de onda dada por el operador \mathbf{L} .

1.1.3. Kernels de sensibilidad

La expresión más general de la derivada de la funcional objetivo $\chi(\mathbf{m})$ en la dirección de $\delta\mathbf{m}$, definida en las sección 1.1.1, está dada por la Ec.(1.15) y sustituyendo el operador de onda definido en la Ec.(1.19), obtenemos una expresión explicita de la derivada

$$\nabla_m \chi \delta\mathbf{m} = \int_T \int_G \delta\rho \mathbf{u}^\dagger \cdot \ddot{\mathbf{u}} d^3\mathbf{x} dt - \int_T \int_G \mathbf{u}^\dagger \cdot \left[\nabla \cdot \int_{\tau=t_0}^t \delta\dot{\mathbf{C}}(t-\tau) : \nabla\mathbf{u}(\tau) d\tau \right] dt d^3\mathbf{x}, \quad (1.33)$$

con $\delta\mathbf{m} = (\delta\rho, \delta\mathbf{C})$ e integrando por partes, se obtiene una expresión simétrica, la cual facilita su manipulación para el próximo desarrollo

$$\nabla_m \chi \delta\mathbf{m} = - \int_T \int_G \delta\rho \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}} d^3\mathbf{x} dt + \int_T \int_G \left[\int_{\tau=t_0}^t \nabla\mathbf{u}^\dagger : \delta\dot{\mathbf{C}}(t-\tau) : \nabla\mathbf{u}(\tau) d\tau \right] dt d^3\mathbf{x}. \quad (1.34)$$

Los kernels de sensibilidad asociados con la Ec.(1.34) son

$$K_\rho = - \int_T \mathbf{u}^\dagger \cdot \dot{\mathbf{u}} dt \quad (1.35)$$

y

$$\mathbf{K}_\mathbf{C} = - \int_T \nabla\mathbf{u}^\dagger(t) \otimes \nabla\mathbf{u}(t+\tau) dt, \quad (1.36)$$

donde \otimes denota el producto tensorial. Los dos kernels no son cero en la zona de primera influencia, donde el campo de desplazamientos adjunto y regular interactúan entre un tiempo t_0 y t_1 . La zona

de primera influencia, como se muestra en la Fig.(1.1) es la región en donde la perturbación del modelo, $\delta\mathbf{m}$, produce un campo de desplazamientos \mathbf{u} de primer orden que afecta las medidas del receptor. Una perturbación fuera de esta zona influencia no tiene un efecto de primer orden en el receptor.

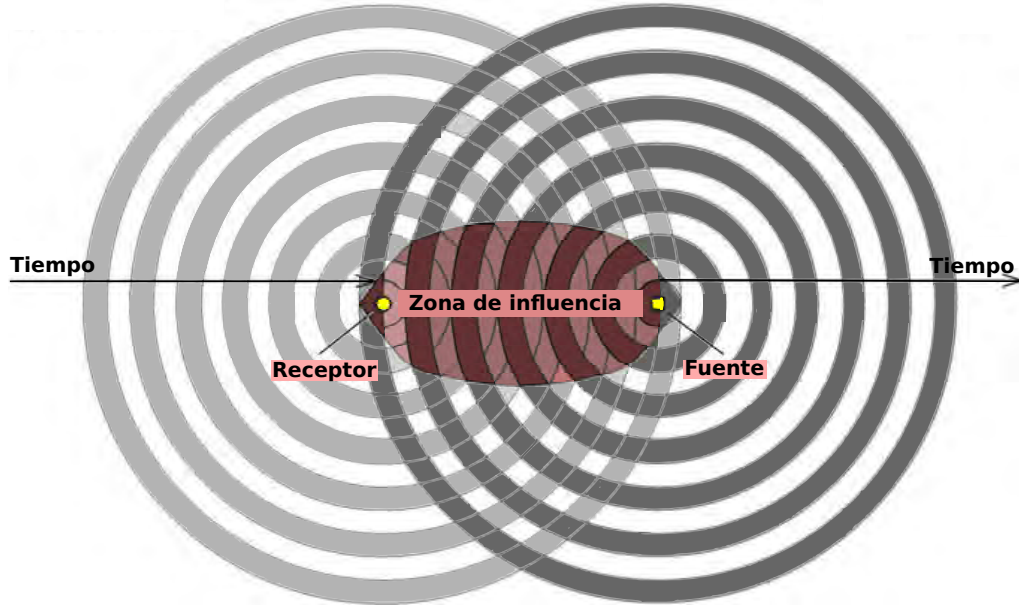


Figura 1.1: Interacción del campo adjunto con el campo regular de desplazamientos, modificada de Fichtner [2010].

Las ecuaciones de los kernels de sensibilidad definidas anteriormente son considerando un medio anelástico y anisótropo; para el caso isotrópico y elástico los esfuerzos no dependen del tiempo, por lo que es posible modificar el tensor de constantes elásticas y sus perturbaciones como:

$$\mathbf{C}(\mathbf{x}, t) = \mathbf{C}(\mathbf{x})H(t), \quad \delta\mathbf{C}(\mathbf{x}, t) = \delta\mathbf{C}(\mathbf{x})H(t). \quad (1.37)$$

Después de sustituir la Ec.(1.37) en la Ec.(1.34), se obtiene una versión simplificada de la derivada de la funcional objetivo

$$\nabla_{m\chi}\delta\mathbf{m} = - \int_T \int_G \delta\rho \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}} d^3\mathbf{x} dt + \int_T \int_G \nabla\mathbf{u}^\dagger(t) : \delta\mathbf{C} : \nabla\mathbf{u}(t) dt d^3\mathbf{x}. \quad (1.38)$$

Considerando un medio isotrópico, las componentes del tensor de constantes elásticas \mathbf{C} son dadas por Aki y Richards [2002]:

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}). \quad (1.39)$$

donde los símbolos λ y μ son las llamadas constantes de Lamé, por lo que la derivada de χ se compone de tres términos

$$\nabla_m \chi \delta \mathbf{m} = \nabla_\rho \chi \delta \rho + \nabla_\lambda \chi \delta \lambda + \nabla_\mu \chi \delta \mu \quad (1.40)$$

con

$$\nabla_\rho \chi \delta \rho = - \int_T \int_G \delta \rho \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}} \, dt d^3 \mathbf{x}, \quad (1.41)$$

$$\nabla_\lambda \chi \delta \lambda = \int_T \int_G \delta \lambda (\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{u}^\dagger) \, dt d^3 \mathbf{x}, \quad (1.42)$$

$$\nabla_\mu \chi \delta \mu = \int_T \int_G \delta \mu [(\nabla \mathbf{u}^\dagger) : (\nabla \mathbf{u}) + (\nabla \mathbf{u}^\dagger) : (\nabla \mathbf{u})^T] \, dt d^3 \mathbf{x}. \quad (1.43)$$

Se obtienen los kernels de sensibilidad considerando solo la integral sobre el tiempo, dando como resultado:

$$K_\rho = - \int_T \dot{\mathbf{u}}^\dagger \cdot \dot{\mathbf{u}} \, dt, \quad (1.44)$$

$$K_\lambda = \int_T (\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{u}^\dagger) \, dt, \quad (1.45)$$

$$K_\mu = \int_T [(\nabla \mathbf{u}^\dagger) : (\nabla \mathbf{u}) + (\nabla \mathbf{u}^\dagger) : (\nabla \mathbf{u})^T] \, dt. \quad (1.46)$$

Estos kernels corresponden a la parametrización fundamental $\mathbf{m} = (\rho, \lambda, \mu)$. Dependiendo de la preferencia, y por conveniencia numérica, es posible reescribir los kernels en función de otros parámetros, como por ejemplo en función de las velocidades de onda-P, α , y de onda-S, β , Fichtner [2010]:

$$K_\rho = K_\rho + (\alpha^2 - 2\beta^2)K_\lambda + \beta^2 K_\mu, \quad (1.47)$$

$$K_\beta = 2\rho\beta K_\mu - 4\rho\beta K_\lambda, \quad (1.48)$$

$$K_\alpha = 2\rho\alpha K_\lambda. \quad (1.49)$$

Uno de los retos en el cálculo de los kernels de sensibilidad recae en la parte computacional, donde se debe tener acceso simultáneo tanto al campo de desplazamientos regular como al adjunto.

Una de las soluciones es guardar el campo de desplazamiento regular en el disco duro para cada paso de tiempo y cuando se esté calculando el campo de desplazamientos adjunto, leer al mismo tiempo el campo regular del disco duro para posibilitar la construcción de los kernels de sensibilidad; Sin embargo, esta alternativa ralentiza severamente los códigos, ya que está demostrado que el acceso a disco duro es del orden de 100 veces más lento que el acceso a memoria RAM. Otra solución cuando no se tiene gran capacidad de almacenamiento es reconstruir el campo de desplazamientos regular a partir de un tiempo final. Esto es posible si se guarda el campo en su tiempo final y se considera como punto de partida en la retro-propagación, en otras palabras la integración hacia atrás en tiempo [Gauthier et al., 1986]. Esta integración del campo regular representa deshacer los efectos de la atenuación del campo, el cual es un proceso estable numéricamente hablando [Tarantola, 2005]. En el caso de que se utilicen condiciones de frontera absorbentes, es también necesario guardar las fronteras para cada paso de tiempo, para que en cada tiempo se sustituyan como condiciones de frontera en la retro-propagación [Gauthier et al., 1986].

1.1.4. Fuente adjunta y funcionales de desajuste

En la gran variedad de estudios encontrados en la literatura donde se utiliza la inversión de forma de onda completa, FWI, siempre se detecta que el uso correcto de una funcional de desajuste es crucial para el éxito de la inversión con datos reales. Por lo que existen una gran variedad de funcionales de desajuste dependiendo de la escala del problema (regional o global) y del detalle que se quiera obtener de las estructuras geológicas en el subsuelo. Para este trabajo se utilizó la norma L_2 , que mide la diferencia entre los datos observados y sintéticos y es muy eficiente para detectar cambios bruscos en el contraste de las propiedades físicas de la Tierra. También se utilizó la funcional de desajuste relacionada a la correlación cruzada entre datos sintéticos y observados, que nos define la diferencia entre tiempos de arribo de una fase dada. Para esta funcional se toma en cuenta el trabajo de Luo y Schuster [1991]; Fichtner [2010]

1.1.4.1. Función de Green Adjunta

Una de las ventajas del método adjunto es que puede ser usado como una herramienta para linealizar el problema directo. Lo anterior se puede lograr si definimos a χ como la componente

i -ésima del campo de desplazamiento \mathbf{u} , medido en la posición de los receptores $\mathbf{x} = \mathbf{x}^r$ en un tiempo $t = t^r$:

$$\chi(\mathbf{m}) = u_i(\mathbf{m}, \mathbf{x}^r, t^r). \quad (1.50)$$

Para derivar la fuente adjunta que corresponde a la funcional objetivo χ , Ec.(1.50), se reescribe la funcional objetivo en su forma integral, como se propone en la Ec.(1.15),

$$\chi(\mathbf{m}) = \langle \chi_1(\mathbf{m}) \rangle = \int_T \int_G \chi_1(\mathbf{m}) dt d^3\mathbf{x} = \int_T \int_G \mathbf{e}_i \cdot \mathbf{u}(\mathbf{m}; x, t) \delta(\mathbf{x} - \mathbf{x}^r) \delta(t - t^r) dt d^3\mathbf{x}, \quad (1.51)$$

donde \mathbf{e}_i es el vector unitario en la dirección- i . Aplicando la Ec.(1.14) al integrando que define χ_1 en la Ec.(1.51), obtenemos la fuente adjunta \mathbf{f}^\dagger , correspondiente a la funcional objetivo definida por la Ec.(1.50)

$$\mathbf{f}^\dagger = -\nabla_u \chi_1 = -\mathbf{e}_i \delta(\mathbf{x} - \mathbf{x}^r) \delta(t - t^r). \quad (1.52)$$

Como se puede observar en la Ec.(1.52), la fuente adjunta \mathbf{f}^\dagger es localizada en un tiempo y posición dada, y actúa en la ubicación del receptor \mathbf{x}^r y en el tiempo observado t^r . Sustituyendo Ec.(1.52) en la Ec.(1.31) obtenemos la ecuación adjunta

$$\mathbf{L}^\dagger(\mathbf{u}^\dagger) = -\mathbf{e}_i \delta(\mathbf{x} - \mathbf{x}^r) \delta(t - t^r), \quad (1.53)$$

y considerando que la función de Green adjunta se encuentra definida como:

$$\mathbf{L}^\dagger(\mathbf{g}_i^\dagger) = \mathbf{e}_i \delta(\mathbf{x} - \mathbf{x}^r) \delta(t - t^r), \quad (1.54)$$

por lo que comparando la Ec.(1.54) y Ec.(1.53), se observa que el campo de desplazamientos adjunto \mathbf{u}^\dagger es igual a menos la función de Green:

$$\mathbf{u}^\dagger(\mathbf{x}, t) = -\mathbf{g}_i^\dagger(\mathbf{x}^r, t^r; \mathbf{x}, t). \quad (1.55)$$

Por lo tanto la derivada de χ , se define como

$$\nabla_m \chi \delta \mathbf{m} = \nabla_m u_i(\mathbf{x}^r, t^r) \delta \mathbf{m} = -\langle \mathbf{g}_i^\dagger(\mathbf{x}^r, t^r) \cdot \nabla_m \mathbf{L} \delta \mathbf{m} \rangle \quad (1.56)$$

Este ultimo resultado va ser de utilidad en las siguientes subsecciones, para obtener la fuente adjunta de otras funcionales de desajuste de una manera más rápida, ya que la derivada de la funcional objetivo nos queda en términos de la función de Green, Ec.(1.56).

1.1.4.2. Norma L_2

La funcional de desajuste clásica usada en FWI es la norma L_2 , que mide la distancia entre los sismogramas observados \mathbf{u}^0 , y los sismogramas sintéticos \mathbf{u} en la posición de los receptores \mathbf{x}^r :

$$\chi(\mathbf{m}) = \frac{1}{2} \int_T [\mathbf{u}^0(\mathbf{x}^r, t) - \mathbf{u}(\mathbf{m}; \mathbf{x}^r, t)]^2 dt. \quad (1.57)$$

El integrando χ_1 que corresponda a χ es

$$\chi_1(\mathbf{m}) = \frac{1}{2} [\mathbf{u}^0(\mathbf{x}^r, t) - \mathbf{u}(\mathbf{m}; \mathbf{x}^r, t)]^2 \delta(\mathbf{x} - \mathbf{x}^r). \quad (1.58)$$

Recordando la sección 1.1.1, en específico la Ec.(1.14), donde se define el operador adjunto con su fuente adjunta, sustituyendo la Ec.(1.51), obtenemos la fuente adjunta para la norma L_2

$$\mathbf{f}^\dagger(\mathbf{x}, t) = -\nabla_u \chi_1 = -[\mathbf{u}^0(\mathbf{x}^r, t) - \mathbf{u}(\mathbf{m}; \mathbf{x}^r, t)] \delta(\mathbf{x} - \mathbf{x}^r). \quad (1.59)$$

Como se puede observar en la Ec.(1.59), la fuente adjunta está ubicada en la posición de los receptores y su evolución en tiempo es determinada por los residuales entre los datos sintéticos y los datos observados. Uno de los problemas de la norma L_2 es que no es una medida de desajuste robusta, ya que si existen *outliers* en los datos, estos se pueden volver dominantes en la funcional de desajuste. Otro problema es que la funcional de desajuste es controlada por las ondas de mayor amplitud, ignorando información invaluable que tiene una pequeña amplitud en el sismograma. Por último, otro aspecto de importancia es que agrega no-linealidad al problema, que es inherente al problema directo. Todos los factores mencionados anteriormente dan como resultado la existencia de muchos mínimos locales, los cuales son un problema al utilizar un método de optimización local, como el gradiente conjugado.

Debido a estas desventajas, se han diseñado varias funcionales de desajuste para contrarrestar estos problemas, las cuales se construyen en forma tal que se puedan extraer explícitamente la información de los tiempos de arribo entre fases de onda y las amplitudes.

1.1.4.3. Correlación cruzada - Tiempos de arribo $\Delta\mathcal{T}$

Debido a los problemas inherentes mencionados en la sección anterior de la norma \mathbf{L}_2 , aún es necesario extraer la mayor información posible de la diferencia entre los datos observados y sintéticos. Una de las formas de hacerlo es con el método propuesto por Luo y Schuster [1991]. Este método se basa en la estimación de tiempos de retraso por medio de la correlación entre los datos observados y sintéticos para una fase dada; de esta manera es posible incluir la fase explícitamente en la funcional objetivo, la cual depende cuasi-linealmente de los parámetros estructurales del modelo.

En el desarrollo se sigue muy de cerca el concepto desarrollado por Luo y Schuster [1991]. Considerando la componente- i de una forma de onda observada $\mathbf{u}_i^0(\mathbf{x}^r, t)$ y su sintética correspondiente $\mathbf{u}_i(\mathbf{m}; \mathbf{x}^r, t)$, en la posición del receptor \mathbf{x}^r , se asume implícitamente que se separa la fase correspondiente para ambos casos, como por ejemplo la onda directa P o S. La *correlación cruzada de tiempos de desfase* \mathcal{T} se define como el tiempo en el cual la correlación cruzada tiene un máximo global

$$\mathcal{C}(u_i^0, u_i)(\tau) = \int_T \mathbf{u}_i^0(\mathbf{x}^r, t) \mathbf{u}_i(\mathbf{m}; \mathbf{x}^r, t + \tau) dt. \quad (1.60)$$

Por lo tanto tenemos que $\mathcal{T} > 0$ cuando la forma de onda sintética llega después de la observada y $\mathcal{T} < 0$ cuando la forma de onda sintética llega antes que la observada, siguiendo la convención de Fichtner [2010]. Se considera que las formas de onda tanto en los datos sintéticos como los observados son filtradas y separadas apropiadamente. La funcional de desajuste que se intenta minimizar es la que se usa comúnmente en la tomografía de tiempos de arribo,

$$\chi(\mathbf{m}) = \frac{1}{2} \mathcal{T}^2(\mathbf{m}). \quad (1.61)$$

La definición en la Ec.(1.60) no se relaciona explícitamente con \mathcal{T} en la Ec.(1.61), por lo que es necesario derivar una función que relacione u_i y u_i^0 con la diferencia de tiempos de arribo \mathcal{T} .

Para esto, se puede ver que la Ec.(1.60) tiene un máximo cuando $\tau = \mathcal{T}$, por lo tanto satisface la condición de la primera derivada para valores extremos,

$$0 = \frac{d}{d\tau} \mathcal{C}(u_i^0, u_i)(\tau)|_{\tau=\mathcal{T}} = \int_T \mathbf{u}_i^0(\mathbf{x}^r, t) \dot{\mathbf{u}}_i(\mathbf{m}; \mathbf{x}^r, t + \mathcal{T}) dt. \quad (1.62)$$

La Ec.(1.62) define \mathcal{T} implícitamente, cuando la función de correlación cruzada tiene un máximo global, el cual es el único para todo el dominio de la correlación. Aplicando la diferenciación implícita se obtiene la derivada de la funcional de desajuste χ :

$$\nabla_m \chi \delta \mathbf{m} = \mathcal{T} \nabla_m \mathcal{T}(\mathbf{m}) \delta \mathbf{m} = \frac{\mathcal{T} \int_T \dot{u}_i^0(\mathbf{x}^r, t - \mathcal{T}) \nabla_m u_i(\mathbf{m}; \mathbf{x}^r, t) \delta \mathbf{m} dt}{\int_T \ddot{u}_i^0(\mathbf{x}^r, t - \mathcal{T}) u_i(\mathbf{m}; \mathbf{x}^r, t) \delta \mathbf{m} dt}. \quad (1.63)$$

Asumiendo que las formas de onda de u_i^0 y u_i solo se encuentran desfasadas y no distorsionadas con respecto una de la otra, se tiene que $u_i^0(\mathbf{x}^r, t - \mathcal{T}) = u_i(\mathbf{m}; \mathbf{x}^r, t)$, por lo que la Ec.(1.63) se reduce a

$$\nabla_m \chi \delta \mathbf{m} = - \frac{\mathcal{T}}{\|\dot{u}_i\|_2^2} \int_T \dot{u}_i(\mathbf{m}; \mathbf{x}^r, t) \nabla_m u_i(\mathbf{m}; \mathbf{x}^r, t) \delta \mathbf{m} dt, \quad (1.64)$$

con

$$\|\dot{u}_i\|_2^2 = \int_T \dot{u}_i^2(\mathbf{m}; \mathbf{x}^r, t) dt. \quad (1.65)$$

Para el termino $\nabla_m u_i(\mathbf{m}; \mathbf{x}^r, t) \delta \mathbf{m}$ en la Ec.(1.65) se sustituye el resultado obtenido en la Ec.(1.56):

$$\nabla_m \chi \delta \mathbf{m} = \frac{\mathcal{T}}{\|\dot{u}_i\|_2^2} \int_T \dot{u}_i(\mathbf{m}; \mathbf{x}^r, t) \mathbf{g}_i^\dagger(\mathbf{x}^r, t; \mathbf{x}, \tau) \cdot \nabla_m \mathbf{L} \delta \mathbf{m} dt. \quad (1.66)$$

Definiendo el campo adjunto,

$$\mathbf{u}^\dagger(\mathbf{x}, t) = \frac{\mathcal{T}}{\|\dot{u}_i\|_2^2} \int_T \dot{u}_i(\mathbf{m}; \mathbf{x}^r, t) \mathbf{g}_i^\dagger(\mathbf{x}^r, t; \mathbf{x}, \tau) dt, \quad (1.67)$$

nos permite escribir la derivada de χ en su forma canónica Ec.(1.15)

$$\nabla_m \chi \delta \mathbf{m} = \int_T \int_G \mathbf{u}^\dagger \cdot \nabla_u \mathbf{L} \mathbf{u} d^3 \mathbf{x} dt. \quad (1.68)$$

Revisando detalladamente la Ec.(1.67), se puede observar que el campo de desplazamientos adjunto \mathbf{u}^\dagger , se expresa en términos de la integral de la fuente adjunta por la función de Green adjunta, por lo tanto la fuente adjunta correspondiente a la funcional de desajuste χ viene dada por

$$\mathbf{f}^\dagger(\mathbf{x}, t) = \frac{\mathcal{T} \mathbf{e}_i}{\|\dot{u}_i\|_2^2} \dot{u}_i(\mathbf{m}; \mathbf{x}^r, t) \delta(\mathbf{x} - \mathbf{x}^r), \quad (1.69)$$

con \mathbf{e}_i como vector unitario de la dirección de observación. Como resultado la fuente adjunta se encuentra ubicada en la posición de los receptores y su evolución en tiempo es determinada con la derivada de los desplazamientos sintéticos. El factor de normalización permite que los kernels de sensibilidad no dependan de la amplitud del campo regular de desplazamientos [Tromp et al., 2005].

Los kernels de sensibilidad son cuasi-independientes de los datos observados debido a que la fuente adjunta no contiene ningún valor de los datos observados, excepto por el factor \mathcal{T} . Lo anterior se debe a que se asume similitud en la forma de onda entre los datos observados y sintéticos y solo existe desfase entre ellos. Existen muchas aplicaciones usando correlaciones cruzadas de tiempos de arribo [Tromp et al., 2005; Tape et al., 2010; Luo y Schuster, 1991], en los cuales se obtiene una mayor información usando diferentes bandas de frecuencia.

1.2. Optimización: Gradiente conjugado

En las secciones previas se ha mostrado como calcular la derivada de la funcional objetivo sin importar que funcional de desajuste se utilice. También se describió la forma de construir los kernels de sensibilidad, los cuales son un paso intermedio para el cálculo del gradiente de la funcional de desajuste y donde éste último es usado en el algoritmo de gradiente conjugado no-lineal. En esta sección se describe a grandes rasgos como funciona este método de optimización para mejorar el modelo de parámetros iterativamente y disminuir el error entre los datos observado y sintéticos Fletcher y Reeves [1964].

1.2.1. Gradiente Conjugado

Esencialmente el algoritmo de gradiente conjugado para el caso específico de la FWI consta de lo siguiente Tape [2009]: dado un modelo inicial \mathbf{m}^0 se estima $\chi(\mathbf{m}^0)$, $\mathbf{g}^0 = \nabla_m \chi(\mathbf{m}^0)$, se establece la dirección de búsqueda como el gradiente negativo de la funcional de desajuste, $\mathbf{p}^0 = -\mathbf{g}^0$. Si $\|\mathbf{p}^0\| < \epsilon$, donde ϵ es el error entre los datos observados y sintéticos que se intenta alcanzar, entonces \mathbf{m}^0 es el modelo que se está buscando y es el que minimiza la funcional de desajuste, sino es el caso:

1. Se realiza una búsqueda lineal para obtener un escalar ν^k , tal que minimice la funcional

$\tilde{\chi}^k(\nu)$ donde

$$\tilde{\chi}^k(\nu) = \chi(\mathbf{m}^k + \nu \mathbf{p}^k), \quad (1.70)$$

$$\tilde{g}^k(\nu) = \frac{\partial \tilde{\chi}^k}{\partial \nu} = \nabla_m \chi(\mathbf{m}^k + \nu \mathbf{p}^k) \cdot \mathbf{p}^k. \quad (1.71)$$

- Se elige un parámetro de prueba $\nu_t^k = -2\tilde{\chi}^k(0)/\tilde{g}^k(0)$ basado en una extrapolación cuadrática.
 - Cálculo del modelo de prueba $\mathbf{m}_t^k = \mathbf{m}^k + \nu_t^k \mathbf{p}^k$.
 - Se calcula la funcional de desajuste para el modelo de prueba $\chi(\mathbf{m}_t^k)$.
 - Se interpola la función de $\tilde{\chi}^k(\nu)$ con un polinomio cuadrático y se obtiene ν^k que produce el mínimo analítico del polinomio interpolado.
2. Se actualiza el modelo $\mathbf{m}^{k+1} = \mathbf{m}^k + \nu^k \mathbf{p}^k$ y se calcula $\mathbf{g}^{k+1} = \nabla_m \chi(\mathbf{m}^{k+1})$.
 3. Se actualiza la dirección de búsqueda el gradiente conjugado: $\mathbf{p}^{k+1} = -\mathbf{g}^{k+1} + \beta^{k+1} \mathbf{p}^k$, donde $\beta^{k+1} = \mathbf{g}^{k+1} \cdot (\mathbf{g}^{k+1} - \mathbf{g}^k) / (\mathbf{g}^k \cdot \mathbf{g}^k)$.
 4. Si $\|\mathbf{p}^{k+1}\| < \epsilon$, entonces \mathbf{m}^{k+1} es el modelo buscado; si no es el caso se repite el proceso a partir del paso numero 1.

Como se puede observar en el algoritmo de optimización, un punto clave es la elección del parámetro de prueba ν_t^k , el cual determina que tan lejos se debe de ir en la dirección de búsqueda

(inicialmente el gradiente negativo) para obtener un modelo de prueba. En este trabajo se asume una forma cuadrática de la funcional de desajuste y se determina el parámetro de prueba tomando en cuenta esta suposición Tape [2009].

1.2.2. Selección del modelo de prueba ν_t

Debido a que la solución numérica del modelo directo es demandante computacionalmente hablando, es necesario limitar el numero de simulaciones por iteración en el problema inverso, por lo que es muy importante decidir que parámetro de prueba ν_t es el correcto para cada iteración por el algoritmo de gradiente conjugado.

1.2.2.1. Selección del paso de prueba : Polinomio cuadrático

El paso de prueba, o parámetro de prueba ν_t , determina en el algoritmo de gradiente conjugado que tan lejos del modelo actual se debe de ir en la dirección de búsqueda, para obtener un modelo de prueba \mathbf{m}_t . Dada una funcional de desajuste $\chi(\mathbf{m})$ y su gradiente $\nabla_m \chi(\mathbf{m})$ en el modelo actual, el usuario enfrenta como determinar que tan lejos se debe de ir en la dirección de búsqueda para encontrar un modelo de prueba, para el cual es necesario estimar la funcional de desajuste del modelo. El vector gradiente $\mathbf{g} = \nabla_m \chi(\mathbf{m})$ es representado en el algoritmo de gradiente conjugado como la pendiente $\tilde{g}(0)$, y la funcional de desajuste en la dirección de búsqueda por $\tilde{\chi}(\nu)$. En el método de optimización, se selecciona el parámetro de prueba por medio de la interpolación de $\tilde{\chi}(\nu)$ usando un polinomio cuadrático, $Q(\nu)$:

$$Q(\nu) = a\nu^2 + b\nu + c, \quad (1.72)$$

donde a, b , y c son determinadas usando los valores (r_1) y la pendiente (g_1) para el modelo actual y para la ubicación del modelo de prueba se seleccionan los parámetros (r_2) y la pendiente (g_2) de tal manera que en ν_t son iguales a cero. Por lo que los cuatro valores son definidos de la siguiente manera:

$$\begin{aligned} r_1 &\equiv Q(0) = \tilde{\chi}(0) = \chi(\mathbf{m}), \\ g_1 &\equiv Q'(0) = \tilde{g}(0) = \nabla_m \chi(\mathbf{m}), \end{aligned}$$

$$\begin{aligned}
 r_2 &\equiv Q(\nu_t) = 0, \\
 g_2 &\equiv Q'(\nu_t) = 0.
 \end{aligned}
 \tag{1.73}$$

Usando la Ec.(1.73) se pueden determinar los coeficientes de $Q(\nu)$:

$$\begin{aligned}
 a &= -g_1/(2\nu_t) = g_1^2/(4r_1), \\
 b &= g_1, \\
 c &= r_1,
 \end{aligned}
 \tag{1.74}$$

y por lo tanto el parámetro de prueba queda determinado como:

$$\nu_t = \frac{-2r_1}{g_1} = \frac{-2\chi(\mathbf{m})}{\tilde{g}(0)}.
 \tag{1.75}$$

La *parabola de prueba* $Q(\nu)$ es elegida tal que su vértice se encuentre en $\chi = 0$; sin embargo, uno puede usar un vértice diferente a cero, el cual se puede determinar de la iteración pasada del algoritmo, permitiendo una convergencia más rápida del método.

Capítulo 2

Zona de Subducción Cocos-Rivera: Estudios Previos

La zona de subducción de Centroamérica es probablemente uno de los márgenes convergentes más complejos en la Tierra. A lo largo de la Trinchera de América Central, la placa de Cocos y la placa de Rivera subducen por debajo de la placa de Norteamérica y el Caribe, causando una distribución irregular de la sismicidad, eventos de deslizamiento lento y tremores no volcánicos [Kostoglodov et al., 2003; Pardo y Suárez, 1995].

Recientemente se ha realizado un esfuerzo multidisciplinario y multi-institucional con el objetivo de definir la geometría y el modelo de velocidades del slab, y obtener un modelo geodinámico del proceso de subducción.

2.1. Configuración tectónica

El ambiente tectónico de la subducción de la placa de Cocos y la placa de Rivera bajo la placa de Norteamérica y la placa del Caribe, tiene una historia de al menos 30 Ma aproximadamente, cuando la placa de Farallon se fragmentó gradualmente en pequeños fragmentos. A partir de uno de estos fragmentos surge la placa de Cocos y que esta a su vez da origen a la placa de Rivera [Atwater y Stock, 1998]. La placa de Cocos y Rivera tienen como límite físico, a lo largo de la

Trinchera de América Central, el graben *El Gordo*, el cual ha sido tomado para marcar el extremo suroeste de una zona de extensión activa entre las dos placas [Bandy et al., 2000].

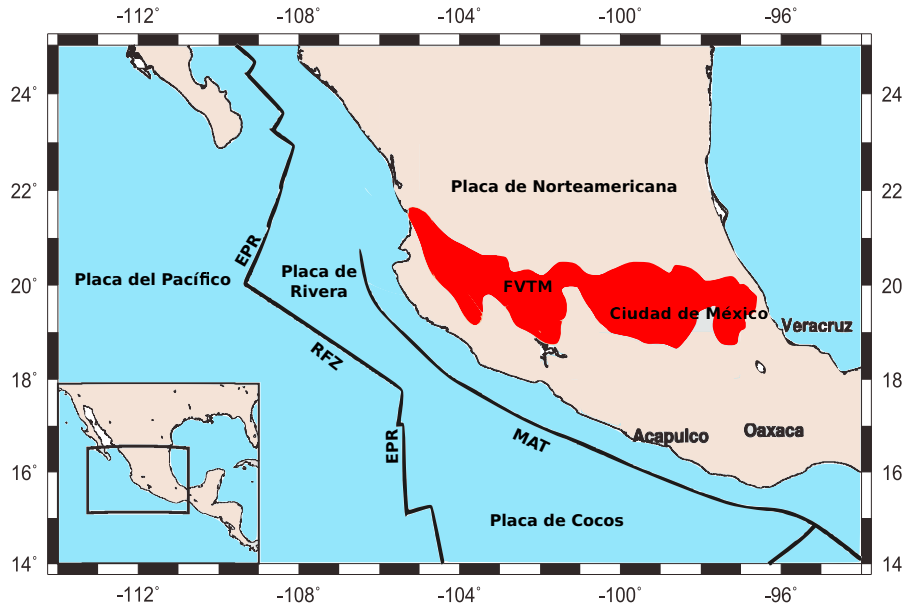


Figura 2.1: Mapa de la región que incluye las cuatro placas: Placas del Pacífico, Rivera, Cocos y Norteamérica. Las abreviaciones mostradas en el mapa son: EPR *Dorsal del Pacífico Este*, MAT *Trinchera de América Central* y RFZ *Zona de Fractura de Rivera*. La zona de color rojo es la Faja Volcánica Trans Mexicana FVTM.

2.1.1. Placa de Cocos

La placa de Cocos se formó hace aproximadamente 23 Ma cuando la placa de Farallón se dividió en dos, dando origen a la placa de Cocos de lado norte del ecuador y Nazca en la parte sur [Lonsdale, 2005]. A su vez la triple unión entre las placas de Cocos, Norteamérica y Pacífico migraron hacia el sureste, causando que la parte norte de la placa de Cocos se fragmentara en varias placas pequeñas, de las cuales terminaron siendo parte de la placa del Pacífico (Monterey, Arguello, Guadalupe y Magnadalena) [Lonsdale, 2005]. La placa de Rivera fue el último fragmento que perdió la placa de Cocos y la cual empezó a actuar independiente hace aproximadamente 10 Ma [DeMets y Traylen, 2000].

La placa de Cocos se encuentra delimitada: al noroeste por la placa de Norteamérica y la placa del Caribe, al oeste por la placa del Pacífico y al sur por la placa de Nazca (ver Fig.2.1). La placa Cocos es relativamente joven; con una edad de entre 10 a 25 Ma, ésta subduce a lo largo de la Trinchera de América Central con una tasa de convergencia variable que va desde 50 hasta 90 mm/año [DeMets, 2001]. La tasa de convergencia entre la placa de Cocos y la Norteamérica incrementa progresivamente hacia el suroeste de la trinchera, de 5 cm/año del centro de México hasta 9 cm/año al sur de Costa Rica [DeMets, 2001]. La edad de la placa Cocos aumenta gradualmente, en la parte oeste de México 10 Ma, a 15 Ma en el centro de México hasta 25 Ma en Guatemala.

2.1.1.1. Configuración del slab de subducción

La placa de cocos subducida bajo México se caracteriza por su gran variabilidad y su geometría inusual a lo largo de su rumbo, lo cual es inconsistente con la variación gradual en los parámetros de subducción, como por ejemplo su edad y su ritmo de convergencia a lo largo de la Trinchera de América Central. En tan sólo unos cientos de kilómetros el ángulo de subducción cambia de 50° grados aproximadamente en el límite con la placa de Rivera a 0° en la parte central de México. Esta variabilidad en el ángulo de subducción del slab es una de las características más estudiadas y de mayor debate de la zona de subducción de centroamérica. Debido a esta gran incógnita y a la necesidad de tener mayor información de esta región, se realizó una colaboración entre la Universidad de California en los Angeles, UCLA, *California Institute of Thecnology*, CALTECH y la UNAM, a la cual se le llamó experimento MASE (*Meso-América Subduction Experiment*). El experimento MASE fue diseñado con la idea central de estudiar la dinámica de la placa de Cocos bajo el centro de México. Entre sus objetivos es determinar la geometría y estructura de velocidades del slab de subducción.

Dentro de los estudios que destacan, se encuentra el análisis de funciones receptor Pérez-Campos et al. [2008] y la tomografía de onda P y S Husker y Davis [2009]. Estos estudios mostraron que la placa inicialmente se inclina con un ángulo de 18° hasta 80 [km] de distancia de la trinchera y después se vuelve horizontal a una profundidad de 50 [km]. Estos estudios refinaron los parámetros obtenidos por Pardo y Suárez [1995], en donde se ubicaron los hipocentros de sismos regionales para ubicar en profundidad la placa de Cocos, Fig.(2.2). La subducción plana se extiende 200 [km] tierra

adentro sólo por debajo del moho de la placa superior. Además en estudios más recientes utilizando el arreglo MASE, se ha identificado una capa con un grosor de 3 a 5 [km], cuya característica principal es tener una baja velocidad y alta presión de fluido Kim et al. [2010]; Song et al. [2009], la cual podría ayudar al desacoplamiento entre la placa de cocos y placa superior dando origen a la subducción plana. Esta última observación es confirmada por la simulación numérica de Manea y Gurnis [2007], y explica que la evolución del slab de tener un ángulo de inclinación a una subducción horizontal en el centro de México pudo haber atrapado un canal de material de baja viscosidad. La interpretación sobre la naturaleza de la capa de baja velocidad no está clara, sin embargo se ha propuesto que representa un remanente de cuña del manto que experimentó una serpentización significativa ya que la subducción del slab se estableció desde el Mioceno medio Ferrari et al. [2012].

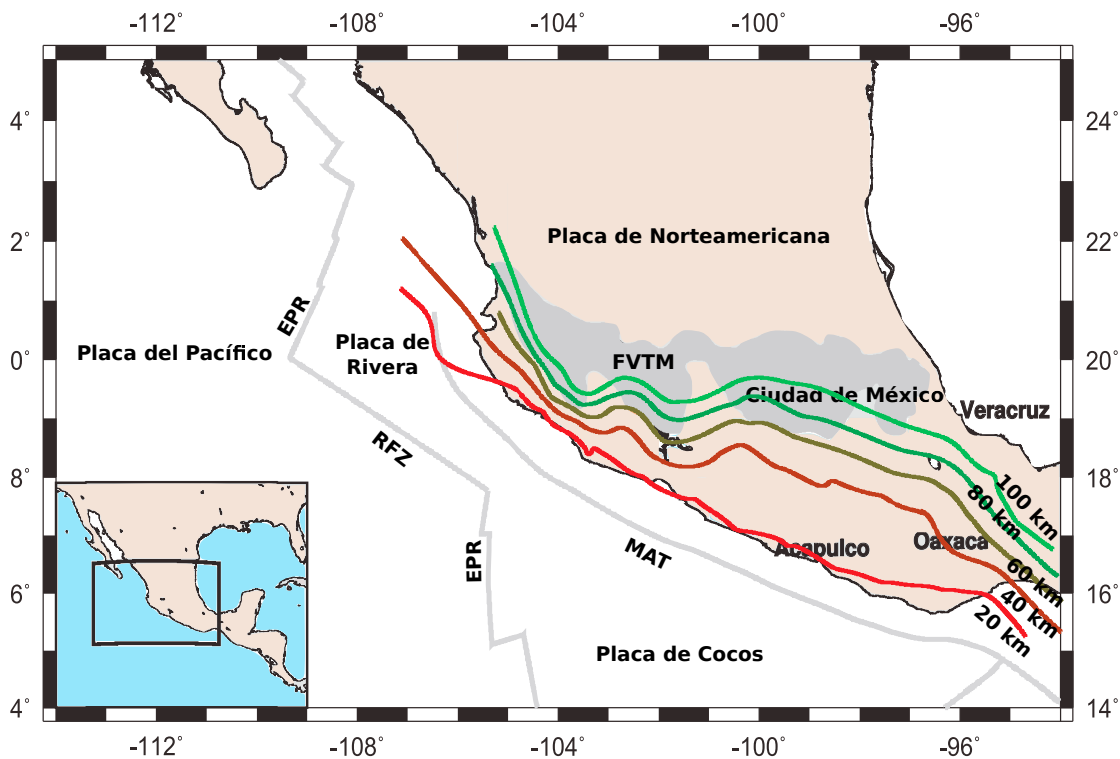


Figura 2.2: Mapa de la región de estudio que muestra las isóbatas de la subducción de la placa de Cocos bajo la placa de Norteamérica obtenidas por Pardo y Suárez [1995]. Modificado de Kim et al. [2010].

2.1.2. Placa de Rivera

La existencia de la placa de Rivera fue propuesta primeramente por Atwater y Stock [1998]. Esta actualmente subduce por debajo de la parte oeste de México primeramente por debajo del estado de Jalisco. El movimiento relativo de la placa de Rivera adyacente a las placas de Pacífico, Cocos y Norteamérica ha sido muy controversial [Bandy et al., 1998; DeMets y Traylen, 2000]. Se piensa que la placa de Rivera se separó de la placa de Cocos hace 5 a 10 Ma, pero la naturaleza entre las dos placas aún es incierta.

La litosfera de Rivera consumida en la trinchera es de finales del Mioceno (9 Ma). A pesar de que la sismicidad relacionada con la subducción de la placa de Rivera bajo el bloque de Jalisco es baja, al menos seis grandes sismos han ocurrido ($M_s > 7,0$) desde 1837, incluyendo el gran sismo de 1932 ($M_s = 8,2$) [Eissler y McNally, 1984; Singh et al., 1985]; lo anterior contradice la hipótesis de que la placa de Rivera subduce asísmicamente Nixon [1982]. Con base en los sismos localizados por Pardo y Suárez [1995] se concluye que la placa de Rivera subduce con un ángulo 50° aproximadamente a una profundidad de 40 [km].

2.1.2.1. Configuración del slab de subducción

En México, la sismicidad producida por la subducción de la placa de Rivera está limitada a profundidades no mayores de los 100 [km]; debido a esto no era posible determinar la geometría del slab más allá de esta profundidad. Usando los hipocentros de eventos locales y regionales Pardo y Suárez [1995] propusieron que la placa de Rivera subduce con una inclinación de 50° aproximadamente, Fig.(2.2). Sin embargo, la geometría del slab a grandes profundidades, como también sus variaciones laterales hasta hace unos años se mantenían desconocidas. Debido a la necesidad de tener mayor información de esta zona de subducción, se realizó una colaboración entre las Universidades de Texas en Austin, Nuevo México y la UNAM, definiendo el experimento MaRS (*Mapping the Rivera Subduction Zone*). Este experimento consistió en un arreglo de 50 estaciones de banda ancha ubicadas en el estado de Jalisco y Michoacán, cubriendo la región entre la costa y la parte noroeste de la Faja Volcánica Trans-Mexicana, FVTM.

Algunos de los estudios usando los datos del arreglo MARS revelaron la estructura tridimen-

sional del manto superior hasta 400 [km] de profundidad bajo el bloque Jalisco [Yang et al., 2009]. La tomografía de onda P reveló características importantes: 1) existe un espacio entre la placa de Rivera y Cocos que aumenta con la profundidad y 2) que a partir de una profundidad de 100 [km], tanto el slab de Rivera como el de Cocos aumentan su inclinación, a más de 200 km por debajo de la parte trasera de la FVTM. Además el slab de Rivera subduce de manera abrupta ($60-65^\circ$) comparada con la parte adyacente del slab de Cocos, aún a pesar de que esta es más antigua. Este hallazgo es bastante sorprendente porque la subducción de la placa de Rivera es sólo 10 Ma y se entendería que el slab es menos denso, por lo tanto un ángulo de subducción más pequeño.

Capítulo 3

Metodología

En este capítulo se detalla la metodología que se siguió para alcanzar los objetivos propuestos de este trabajo. Iniciando desde el problema directo, el algoritmo numérico empleado para la solución de la ecuación de onda, hasta su validación con una solución analítica del problema. Siguiendo después con la base de datos utilizada y las restricciones que se definieron para la selección de los eventos a utilizar para la FWI. En una segunda etapa, se detalla la solución del problema inverso, la cual consta de la generación del modelo inicial a partir de la información recopilada de estudios previos de la zona, un análisis de sensibilidad del método de inversión (*Checkerboard sensitivity test*), el procesamiento de los datos sintéticos y observados para la construcción de las fuentes adjuntas para cada evento seleccionado y por último el post-procesamiento para cada kernel de sensibilidad antes de la construcción del gradiente de la funcional de desajuste.

3.1. Problema directo

La solución de la ecuación de onda en este trabajo se resolvió por el método de diferencias finitas, usando la formulación de [Virieux, 1986]. Para reducir el número de puntos por longitud de onda y , por ende, disminuir la dispersión espacial del método, se utilizó diferencias finitas centradas de octavo orden en espacio, y donde los coeficientes de peso fueron generados siguiendo la metodología de [Fornberg, 1988]. Para el caso del tiempo se usaron diferencias finitas centradas de

segundo orden debido a que la dispersión numérica puede ser mitigada reduciendo el paso de tiempo y lo cual no representa en términos de memoria ningún incremento significativo. Conociendo estas características fue necesario calcular la condición de estabilidad (*Courant–Friedrichs–Lewy (CFL) condition*) del método, la cual fue determinada siguiendo el desarrollo matemático de [Moczo et al., 2000], quedando definida como

$$\frac{\Delta t \alpha \sqrt{3}[c_1 + c_2 + c_3 + c_4]}{\Delta h} \leq 1, \quad (3.1)$$

donde Δt es el paso de tiempo, α la velocidad de onda P, C_i los coeficientes de peso de las diferencias finitas centradas en espacio y Δh la discretización espacial. Para definir la discretización en tiempo Δt y en espacio Δh , se realizaron varias simulaciones para definir el valor óptimo de estos parámetros, dando como resultado que son necesarios 6 nodos por longitud de onda mínima $\Delta h = \frac{1}{6}L$ y un paso de tiempo de $\Delta t = 0,06\frac{\Delta h}{\alpha}$; ésta discretización produce un equilibrio entre el número de nodos por longitud de onda y el paso de tiempo para cumplir la condición de estabilidad Ec.(3.1), lo que se puede observar en la Fig.(3.1), en la cual se comparan la solución analítica y numérica para distintas discretizaciones.

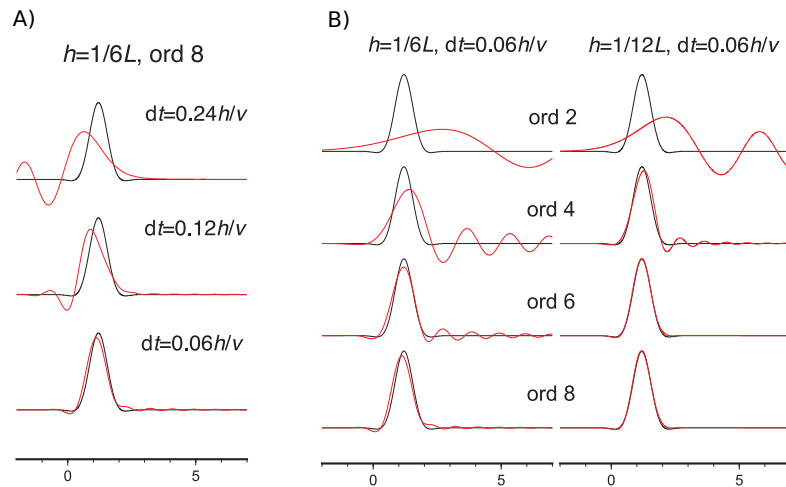


Figura 3.1: Simulación numérica para distintas discretizaciones tanto en tiempo y en espacio, con diferente orden de aproximación en las diferencias finitas.

Dado que al simular numéricamente la propagación de ondas en un medio finito, toda la energía

que llega a las fronteras adyacentes e inferiores del modelo sintético es reflejada, lo cual no pasa en la naturaleza ya que toda la energía se propaga en toda la tierra, es necesario agregar al modelo numérico fronteras absorbentes para simular un medio infinito y no aparezcan reflexiones espurias en la solución numérica de la ecuación de onda. En este trabajo se implemento en el algoritmo numérico la CPML *Convolutional Perfectly Matched Layer*, desarrollada por Komatitsch y Martin [2007], la cual es fácil de implementar, ya que sólo es necesario incrementar el modelo original con un máximo de diez nodos en cada dirección para generar los perfiles de amortiguamiento. Otra característica que es de importancia mencionar, es que su evolución en tiempo es estable y amortigua ondas con ángulos de incidencia muy bajos, el cual era un problema con la formulación tradicional de la PML [Berenger, 1996; Komatitsch y Martin, 2007].

Al haber ya definido los valores óptimos en la discretización y la implementación de las fronteras absorbentes solo nos queda por definir la implementación de la frontera libre y la simulación de una dislocación puntual en el algoritmo de diferencias finitas. Para el caso de la condición de frontera libre se utilizo la formulación de Mittet [2002], la cual se basa en definir las propiedades físicas del medio en la frontera como un medio isótropo transversal, de tal forma que la implementación es transparente para el método numérico y no es necesario más que agregar una condición en $z = 0$ para las propiedades del modelo. Por último se realizo la implementación de una dislocación puntual en el método de diferencia finitas siguiendo a Graves [1996], en donde define la aplicación del tensor de momento por medio de fuerzas de cuerpo distribuidas, que son agregadas para cada componente de la velocidad.

El código de la simulación numérica y de la inversión se realizo en C++, se paralelizo el código siguiendo la estrategia propuesta por Michéa y Komatitsch [2010], la cual consta de dividir el volumen de solución en la dirección z, para que cada procesador resuelva una parte de todo el dominio y solo sea necesario comunicarse entre procesadores en las fronteras de cada subdominio local a cada procesador, de esta forma se elimina el cuello de botella que se podría generar si se transferiera la información de cada subdominio a cada procesador. Al paralelizar el código se disminuye el tiempo de simulación para cada evento modelado y por ende la FWI es más rápida.

En la Fig.(3.3) se muestra la propagación en un medio homogéneo con $\rho = 2000.0 [kg/m^3]$; $\alpha = 2500.0 [m/s^2]$; $\beta = \alpha/\sqrt{3}$, aplicando una fuerza vertical a una profundidad de $z = 25.0 [m]$, también

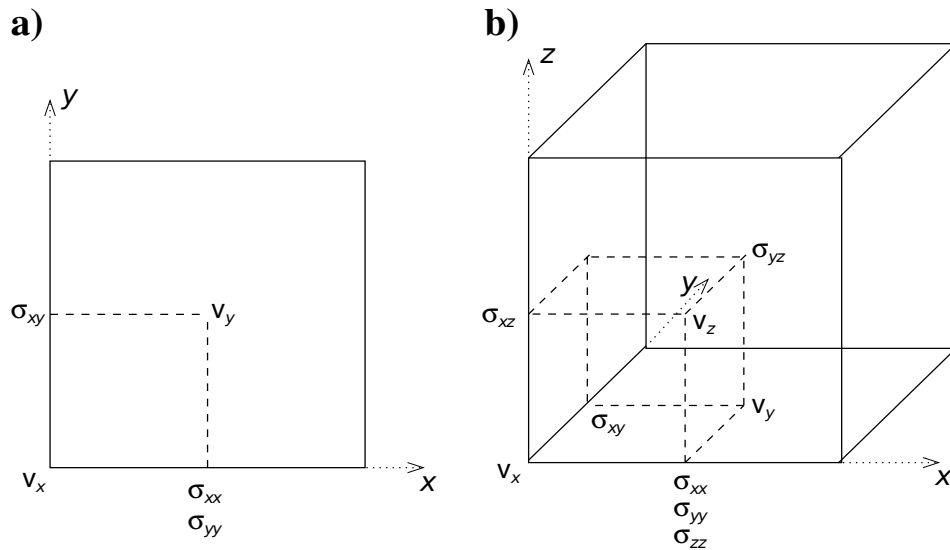


Figura 3.2: Celda elemental del método de diferencias finitas escalonadas en 3D [Madariaga, 1976]. En la cual se muestra la discretización en función de las componentes de velocidad \mathbf{v} y el tensor de esfuerzos σ .

se muestran los sismogramas calculados numéricamente y analíticamente y donde se puede observar que la discrepancia entre ellos es muy baja, por lo que el problema directo se valida, y puede ser utilizado para el método de inversión FWI. La solución analítica se calculó por el método de Cagniard–De Hoop, utilizando el programa de Per Berg del laboratorio de Matemáticas Aplicadas de Dinamarca (<http://www.spice-rtn.org/library/software/EX2DDIR.html>).

3.2. Base de Datos

Los datos utilizados en este trabajo fueron los sismos registrados en el arreglo MASE y la red permanente de monitoreo sísmico del Servicio Sismológico Nacional (SSN). El arreglo permanente del SSN consta de 23 estaciones colocadas en toda la República Mexicana. Por otro lado el arreglo MASE consta de 100 estaciones de banda ancha, las cuales se encuentran ubicadas perpendicularmente a la Trinchera de América Central, pasando por la Ciudad de México, el Estado de México, Guerrero, Hidalgo, Morelos y Veracruz, como se muestra en la Fig.(3.4). La geometría

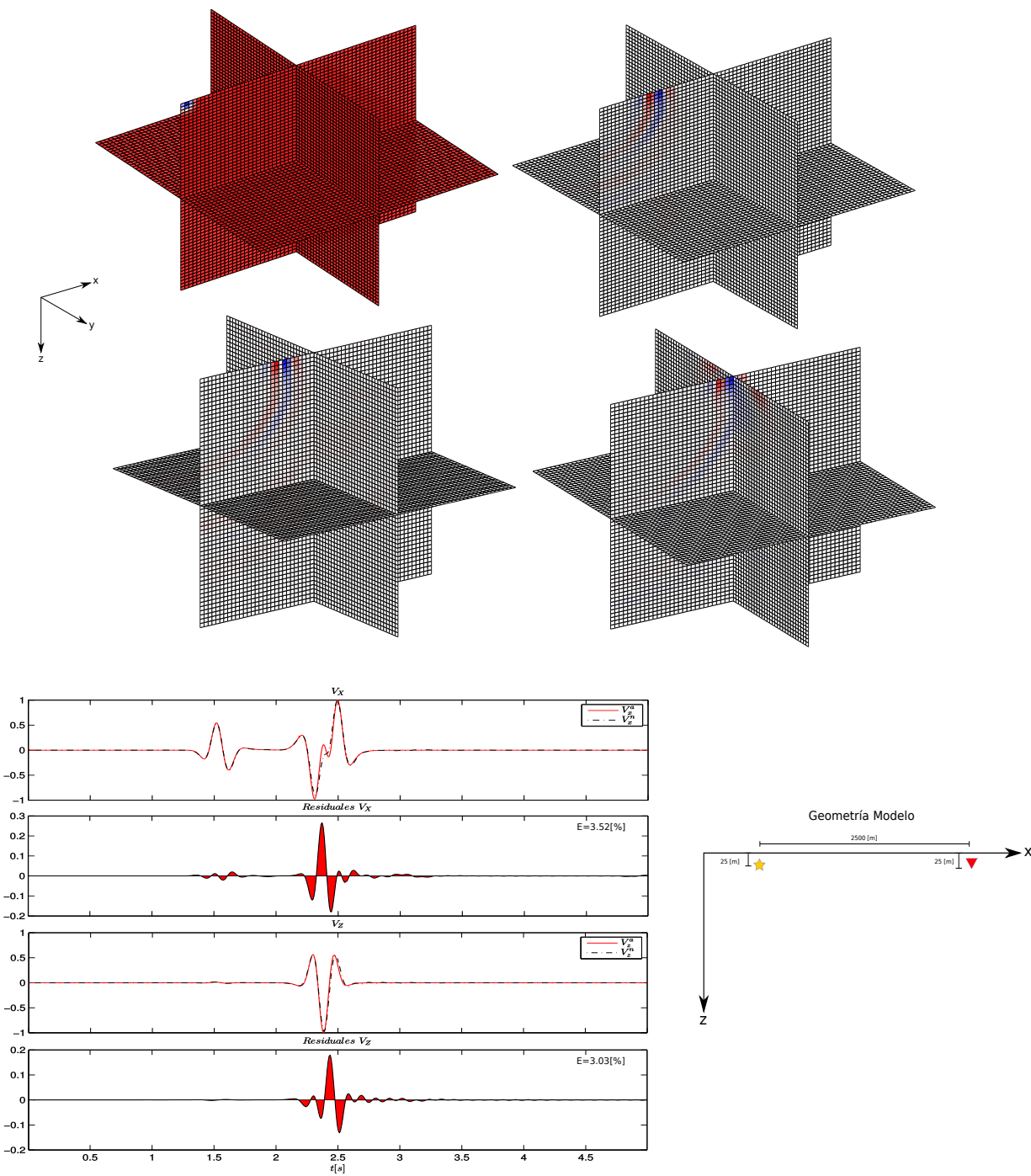


Figura 3.3: Propagación de ondas en un medio homogéneo. a) *Snapshots* de la simulación numérica componente v_z , b) Comparación entre la solución analítica v^a y numérica v^n

del arreglo considera una separación entre estaciones de cada 5 [km] aproximadamente. El periodo de operación del arreglo MASE fue de Diciembre de 2004 a Julio de 2007. Los datos registrados corresponden a eventos sísmicos (locales, regionales y telesísmicos), ruido sísmico y tremores no volcánicos.

El catálogo original consistió con 68 eventos sísmicos (locales y regionales), con al menos un registro en alguna de las 123 estaciones del arreglo MASE y SSN (100-MASE y 23-SSN). Se utilizó el catálogo ya depurado por Ortiz-Alemán et al. [2017], en el cual se seleccionaron todas las estaciones que tuviera un relación señal-ruido menor a 2. Resultando que solo 41 estaciones del arreglo MASE cumplen esas características y 14 de SSN, quedando con un total de 55 estaciones para el uso de sus datos durante la FWI. Finalmente de los 68 eventos del catálogo original se seleccionaron 12 con $4,5 \leq M_w \leq 5,5$. Se eligió este rango de magnitudes porque cuentan con una solución confiable del tensor de momentos y la suposición de una fuente puntual para su modelación es válida. La localización y parámetros de la fuente para cada evento fueron tomados del catálogo *Global-CMT* [Dziewonski et al., 1981; Ekström et al., 2012], los cuales se muestran en la Tabla (3.1).

Se realizó un procesamiento básico de los registros seleccionados, se corrigieron apropiadamente eliminando la respuesta del instrumento, la media y la tendencia. En la Fig.(3.5) se muestra la distribución de los eventos seleccionados, las estaciones seleccionadas y el área a modelar por el método de diferencias finitas.

3.3. Modelo inicial M_0

El modelo inicial M_0 para este trabajo fue construido con base en diversos estudios previos realizados de la zona, tomando en cuenta la estructura de velocidades en la corteza y el manto superior en México y la geometría de la subducción de la placa de Cocos bajo la placa de Norteamericana. El modelo inicial se construyó a partir de las curvas de dispersión obtenidas por Iglesias et al. [2010], las curvas se interpolaron para obtener un modelo 2D de velocidades de onda S. El modelo 2D obtenido se suavizó con un promedio móvil para evitar un sesgo a la hora de realizar la inversión y también para ampliar el espacio de soluciones y no volver quedar restringidos en

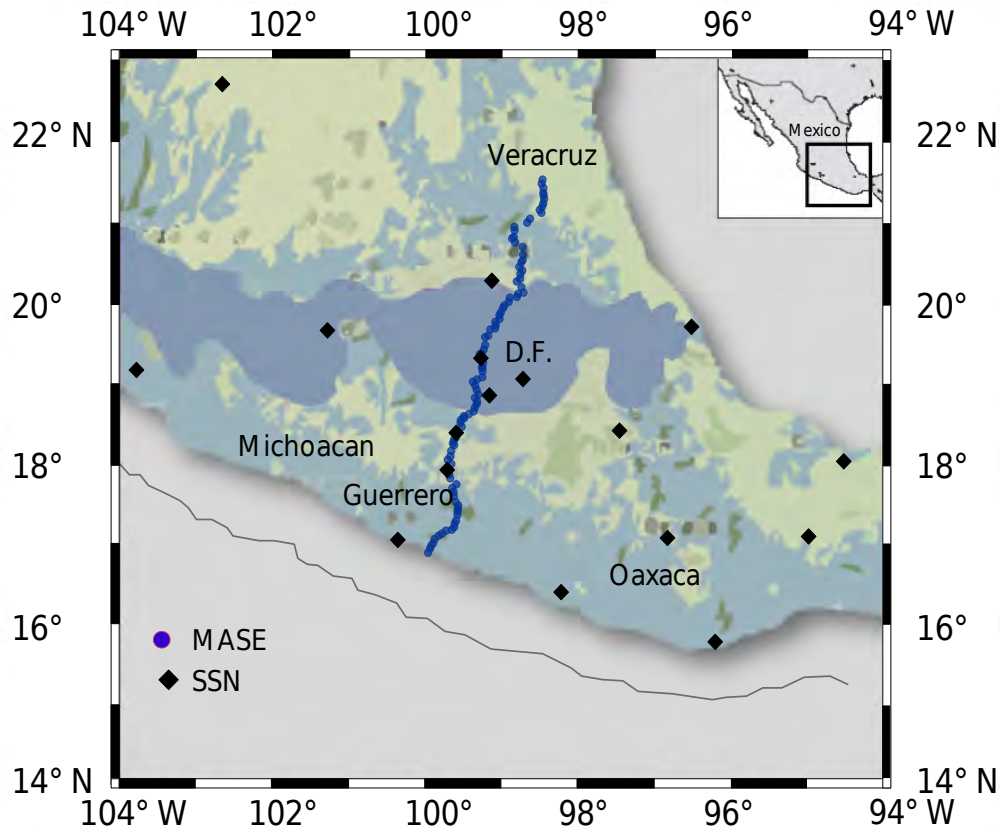


Figura 3.4: Localización de las estaciones del arreglo MASE (círculos azules) y la red permanente del SSN (rombos negros).

mínimo local al cual pertenecía este modelo del estudio realizado por Iglesias et al. [2010]. En un segundo paso, el modelo bidimensional se continuo hacia una tercera dirección para la construcción de un modelo tridimensional de velocidad de onda S, el cual se utilizo para la primera iteración de la FWI. Para obtener el modelo de velocidades de onda P, se considero una relación de Poisson de $\nu = 0,25$ y para el modelo de densidades se uso la relación $\rho = 0,32\alpha + 0,77[g/cm^3]$, donde la velocidad de onda S α es en $[km/s]$ [Berteussen, 1977]. En la Fig.(3.6) se muestra el modelo de velocidades de onda S obtenido en un corte paralelo al arreglo MASE, el cual es usado como modelo inicial M_0 .

Las dimensiones del modelo tridimensional son de $560 \times 500 \times 100 [km]$, con una resolución

Cuadro 3.1: Lista de eventos y sus parámetros de fuente utilizados en este estudio.

Num.	Evento	$t_0[s]$	M_0	Strike[°]	Dip[°]	Slip[°]	$t_{half}[s]$
1	20050814	4.6	1.56E+24	278	16	61	1.2
2	20051202	4.2	3.43E+23	309	27	96	0.7
3	20051214	1.8	2.78E+23	101	31	-78	0.7
4	20060220	1.9	9.13E+23	121	34	-86	1.0
5	20060811	3.3	1.26E+25	97	33	-94	2.4
6	20060819	2.9	2.78E+23	277	16	72	1.5
7	20061110	6.3	1.72E+23	310	40	102	0.6
8	20061118	2.6	2.88E+23	308	29	109	0.7
9	20070315	2.0	1.27E+24	63	8	35	1.1
10	20070330	5.2	9.90E+23	287	19	75	1.1
11	20070413A	3.5	1.07E+25	98	17	84	2.3
12	20070413B	2.0	1.26E+24	28	29	0	1.1

vertical y horizontal de 2 [km], dando como resultado 3,597,081 nodos en total del modelo tridimensional. La resolución de la propagación de ondas se limitó a una frecuencia de 0,25 [Hz], debido a que es la máxima frecuencia que se puede alcanzar con la discretización definida anteriormente sin que exista dispersión numérica en la solución. Los parámetros computacionales para cada iteración de la FWI fueron los siguientes: 6 nodos por longitud de onda, 180 segundos de simulación a 10 muestras por segundo y una frecuencia máxima de la función fuente $S(t)$ de 0.25 [Hz].

Para todas las simulaciones se utilizó un servidor con las siguientes características:

- 4 procesadores Xeon Intel ®Xeon ®E7-4850
- 10 unidades de procesamiento por procesador
- 512 GB de memoria RAM
- *Hyperthreading Technology* desactivada

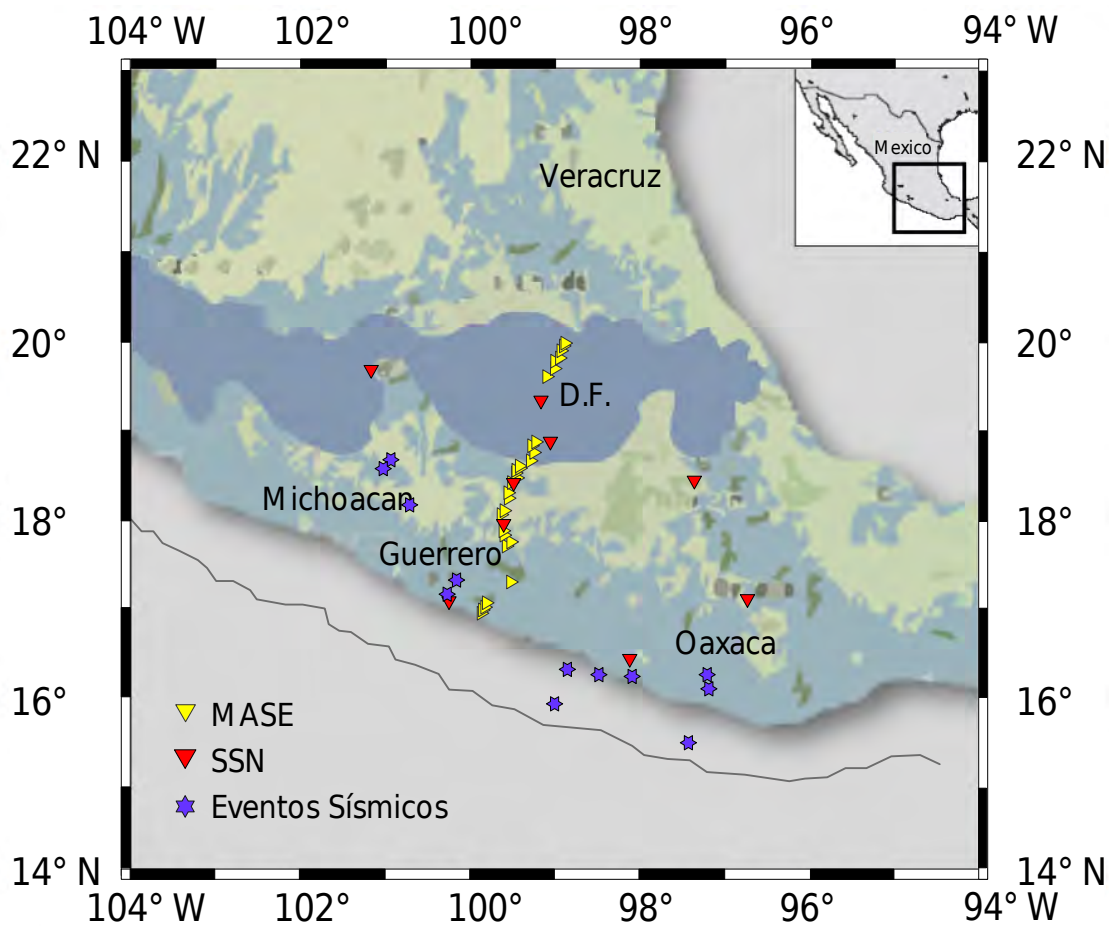


Figura 3.5: Distribución de eventos sísmicos y estaciones sísmicas seleccionadas.

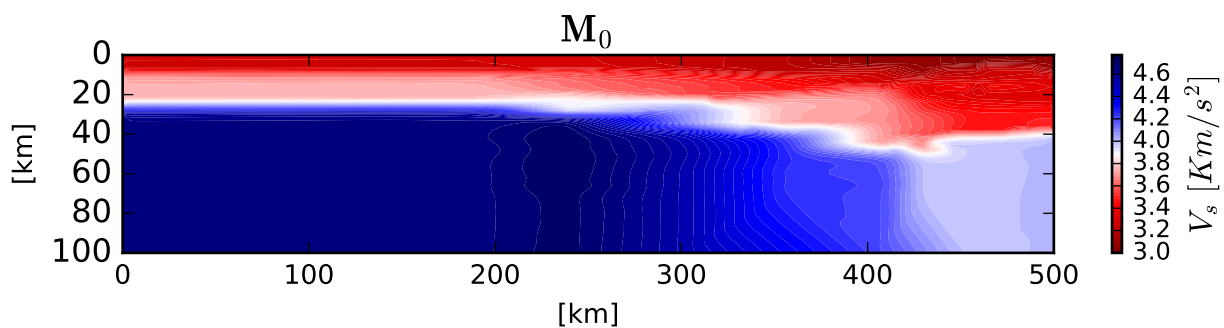


Figura 3.6: Sección del modelo inicial de velocidades de onda S.

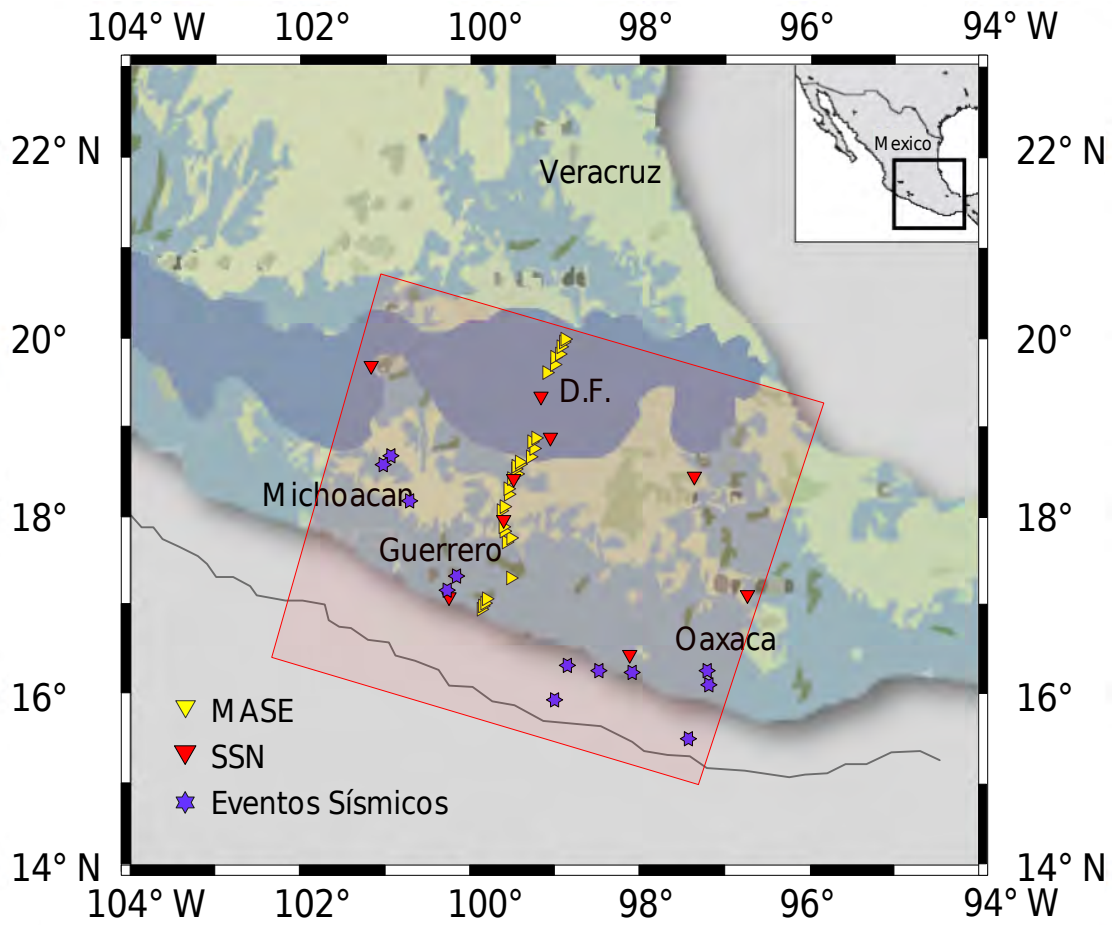


Figura 3.7: Región de color rojo utilizada para el modelado directo y la FWI.

- *Scientific Linux* como sistema operativo

3.4. Análisis de sensibilidad

Un paso importante antes de realizar la inversión, es analizar la sensibilidad que se tiene en el modelo debida a la geometría definida por la fuente y receptor para cada evento. Una forma de realizar este análisis es obtener la matriz Hessiana \mathbf{H} de la funcional de desajuste del problema y observar los elementos de la diagonal principal de la matriz \mathbf{H} ; si estos valores para cada fila

de la matriz son dominantes se tiene buena resolución, y por lo tanto los parámetros son bien restringidos e independientes; por otro lado si los elementos fuera de la diagonal son en magnitud aproximadamente iguales no se tiene buena resolución, y por lo tanto existe dependencia entre los parámetros [Fichtner, 2010]. Una de las desventajas de hacer de esta forma el análisis de sensibilidad es la necesidad de altos recursos computacionales para calcular la matriz Hessiana \mathbf{H} .

Tomando en cuenta el alto consumo computacional en la construcción de la matriz Hessiana, una alternativa para realizar el análisis de sensibilidad es realizar un *checkerboard sensitivity test*, el cual consta de los siguientes pasos:

1. Generar un modelo homogéneo \mathbf{M}_0 del cual se obtienen los datos sintéticos \mathbf{u} .
2. Se genera un modelo objetivo \mathbf{M}_t , el cual consiste en la generación de perturbaciones al modelo homogéneo. Estas perturbaciones siguen un patrón sinusoidal con una longitud de onda y amplitud (porcentaje de perturbación al modelo homogéneo) específicas definidas por el usuario.
3. Se obtienen los datos observados \mathbf{u}^0 a partir del modelo objetivo \mathbf{M}_t .
4. Se obtiene la derivada de la funcional con respecto a los parámetros $\nabla_m \chi$.
5. Se interpreta la derivada de la funcional comparada con el modelo objetivo, observando donde se obtiene mayor sensibilidad en los parámetros.

Para este trabajo se propuso el modelo inicial \mathbf{M}_0 usado en la iteración inicial de la FWI (Fig.(3.6)) y un modelo objetivo \mathbf{M}_t con perturbaciones sinusoidales del 10 % con una longitud de onda de $\lambda = 40 [km]$, resultando en cubos de $20 \times 20 \times 20 [km]$, con signo alternado. En la Fig.(3.8) se puede observar el modelo objetivo \mathbf{M}_t , en un corte transversal paralelo al arreglo MASE.

En la Fig.(3.9) se muestra un perfil del gradiente de la velocidad de onda S para la primera iteración. Este perfil se encuentra ubicado en la Fig.(4.3), como el perfil B-B', en el cual se observa que la sensibilidad del método de inversión se concentra en la parte central de la imagen. Esta área de mayor sensibilidad está delimitada en el tramo de entre 100 y 400 [km] en distancia horizontal y en profundidad entre 10 y 90 [km].

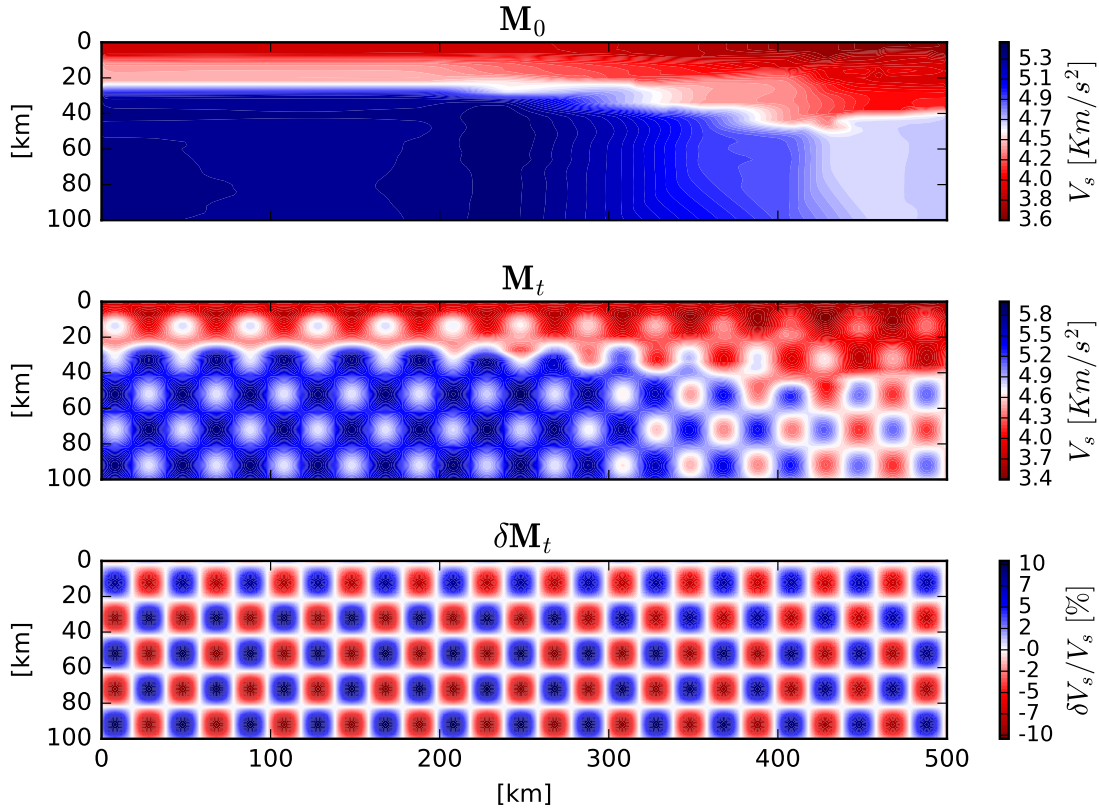


Figura 3.8: Modelo inicial M_0 , modelo objetivo M_t con perturbaciones máximas y mínimas del $\pm 10\%$ y modelo de perturbaciones δM_t .

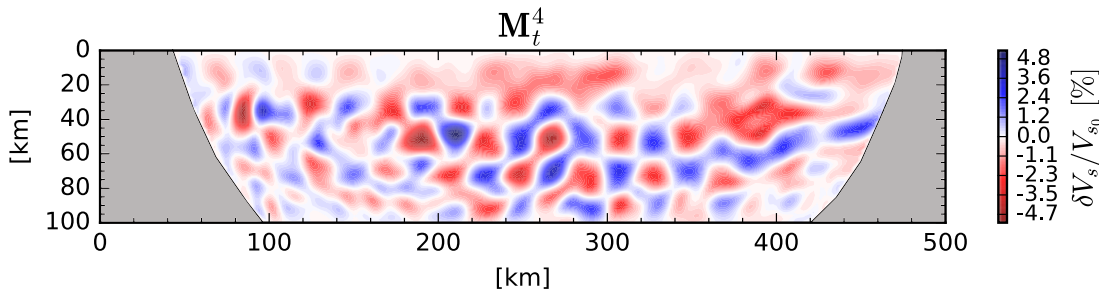


Figura 3.9: Gradiente de velocidad de onda S para la primera iteración de la prueba de sensibilidad del método de inversión (*checkerboard sensitivity test*).

3.5. Procesamiento de datos

En la sección anterior 3.2, se mencionó que a los datos observados se les realizó un procesamiento básico, eliminando la respuesta del instrumento, la media y la tendencia, pero para poder

comparar los datos observados con los sintéticos y construir las funciones de la fuente adjunta es necesario aplicar un procesamiento a ambos datos, para que tengan el mismo intervalo de muestreo y frecuencia.

El procesamiento que se siguió para los datos observados consistió simplemente en decimar a 10 muestras por segundo los sismogramas, aplicar un filtro pasabajos 0.1 [Hz] y ventanear hasta 180 segundos, que es lo que dura la simulación de propagación de ondas. Para el caso de los sismogramas sintéticos, se siguió el procedimiento propuesto por Komatitsch y Tromp [2002]. Este consiste en simular la propagación de ondas con una función fuente *Heaviside*, esto es una función tasa de momento \dot{M} que es igual a una delta de Dirac $\delta(t)$. En un segundo paso se realiza la convolución con cualquier función fuente que ajuste mejor a los sismogramas observados. Este procedimiento permite usar una gran variedad de funciones sin tener que resolver el modelo directo para cada función, y de esta forma se disminuye el tiempo computacional en la creación de las fuentes adjuntas.

La fuente a convolucionar se define como:

$$s(t) = \frac{\alpha}{\sqrt{\pi\tau}} e^{-(\frac{\alpha t}{\tau})^2}, \quad (3.2)$$

donde α y τ son el decaimiento y duración media de la fuente respectivamente, como se muestra en la Fig.(3.10). En la última parte del procesamiento de los datos sintéticos, se realiza un filtrado de los sismogramas a 0.1 [Hz] para tener tanto los datos observados como sintéticos en el mismo rango de frecuencias y poder realizar los ventaneos necesarios para generar las fuentes adjuntas. En la Fig.(3.11) se muestran los sismogramas sintéticos y observados para una estación, a una frecuencia de 0.1 [Hz], en los cuales se puede observar una gran similitud en la forma de onda y un pequeño desfase entre ellos. Por lo que el modelo elegido como modelo inicial se encuentra muy cerca del mínimo global de la solución que se busca.

3.5.1. Construcción fuentes adjuntas

Para la construcción de las fuentes adjuntas se seleccionaron tres ventanas por componente, considerando las ondas de cuerpo P y superficiales S, y una ventana de 20 segundos para la norma L_2 donde ya se tenía una similitud en la forma de onda. Por lo que se midieron 1,476 tiempos de arriba y 1,476 ventanas para el ajuste de la norma L_2 . Para las correlaciones entre los datos

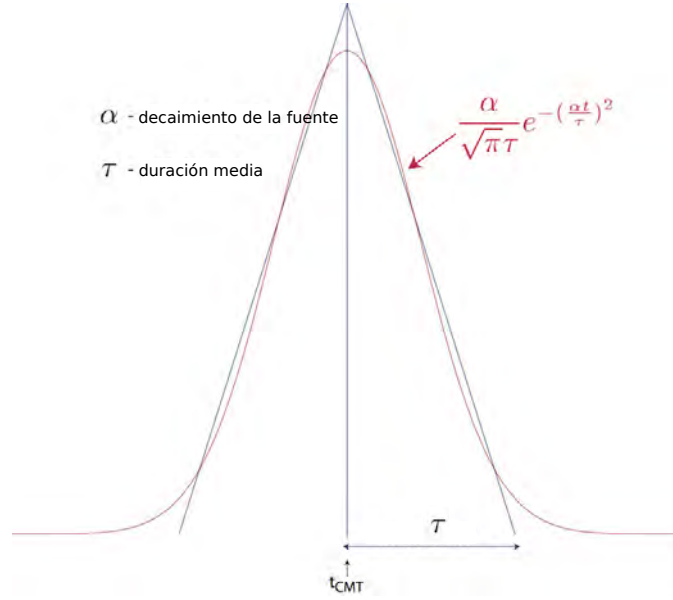


Figura 3.10: Comparación entre la función triángulo y la función gaussiana usada en este trabajo, modificada de [Komatitsch y Tromp, 2002].

sintéticos y observados se normalizaron en función de la energía promedio de las correlaciones de cada evento para no generar un sesgo en los eventos de mayor magnitud. El factor de normalización se definió siguiendo a Zhu [2013]

$$\mathbf{f}^\dagger(\mathbf{x}, t) = \frac{w_r \mathcal{T} \mathbf{e}_i \dot{u}_i(\mathbf{m}; \mathbf{x}^r, t) \delta(\mathbf{x} - \mathbf{x}^r)}{N_r}, \quad (3.3)$$

donde

$$N_r = \int_T w_r ||\dot{u}_i(\mathbf{m}; \mathbf{x}^r, t)||^2 dt, \quad \text{para } w_r = \frac{1}{\int_T ||\dot{u}_i^0(\mathbf{m}; \mathbf{x}^r, t)||^2 dt}, \quad (3.4)$$

de tal forma que la fuente adjunta y la construcción de los kernels de sensibilidad no dependan de la magnitud de los eventos, y de esta manera se evita el sesgo en la inversión de eventos de mayor magnitud.

3.5.2. Post-procesamiento kernels de sensibilidad

El post-procesamiento realizado a los kernels de sensibilidad consta de aplicar un operador que suavice los kernels para eliminar amplitudes espurias que se encuentran ubicadas en la periferia de



Figura 3.11: Comparación entre sismogramas sintéticos obtenidos con el modelo inicial propuesto \mathbf{M}_0 y los sismogramas observados, para las estaciones QUEM, XALI y KM67 para la componente z, del evento 20050814, filtrados a 0.1 [Hz].

las fuentes y receptores [Tape et al., 2010]. Lo anterior se puede lograr convolucionando el kernel de sensibilidad con una función Gaussiana de la forma

$$G(x, y, z) = \frac{4}{\pi\Gamma^2} e^{-4\left(\frac{x^2+y^2+z^2}{\Gamma^2}\right)}, \quad (3.5)$$

donde Γ es el ancho de la función Gaussiana, definida de tal forma que en coordenadas esféricas $r = \Gamma/2$, con una amplitud igual a $G(r) = G(0) e^{-1}$; de esta manera, Γ controla el suavizado del kernel. Una de las formas de definir el valor de este parámetro es a partir de la longitud de onda más pequeña que se este modelando en el problema directo [Tape et al., 2010]. La selección del parámetro Γ se puede interpretar análogamente al parámetro de amortiguamiento γ que usualmente se usa en las inversiones de tomografía sísmica para disminuir la dependencia entre los

parámetros del modelo a invertir.

Capítulo 4

Resultados

En cada iteración de la FWI se siguió la metodología propuesta en la sección anterior. Para la primera iteración se minimizaron los tiempos de arribo de las fases predominantes en los sismogramas observados filtrados a 0.1 [Hz]. En total se tomaron tres ventanas de duración variable para los datos observados y sintéticos con el fin de realizar las correlaciones y obtener los tiempos de desfase entre los datos, para posteriormente generar las funciones de fuente adjunta definidas por la Ec.(3.3). En esta primera iteración se minimizaron sólo los tiempos de arribo de acuerdo a la funcional objetivo Ec.(1.61), debido a que ésta depende linealmente de los parámetros estructurales presentes en el modelo a invertir. De esta manera, la funcional facilita la convergencia de la FWI, aun cuando carece de sensibilidad a pequeñas perturbaciones en el modelo de parámetros.

Para cada evento de cada iteración se obtuvieron los kernels de sensibilidad en función de la parametrización: densidad, \mathbf{K}_ρ , velocidad de onda S, \mathbf{K}_β , y velocidad de onda P, \mathbf{K}_α . El cálculo del kernel total para los 12 eventos ocupados en este trabajo se realizó con el post-procesado de los kernels individuales de cada evento, y sumando cada uno de ellos en función de su parametrización. Este post-procesamiento consistió en convolucionar el kernel de cada evento con una función gaussiana definida por la Ec.(3.5), con un ancho $\Gamma = 12 [km]$, el cual viene definido por la longitud de onda mínima alcanzada por el problema directo. En un segundo proceso se normalizó cada kernel de cada evento, para que al obtener la suma de los kernels de todos los eventos no existiera un evento predominante que sesgara el kernel total de cada iteración.

Con base en el algoritmo de gradiente conjugado, descrito en la subsección (1.2) para obtener

la actualización del modelo inicial, es necesario calcular el gradiente de la funcional para cada parámetro del modelo (ρ, α, β) , donde este gradiente se define en función de los kernels de sensibilidad totales de cada parámetro para cada iteración. Siguiendo a Tape et al. [2010] para calcular la actualización del modelo de parámetros es necesario obtener la funcional de desajuste $\chi(\mathbf{m})$, su gradiente respecto a cada parámetro $\nabla_m \chi(\mathbf{m})$ y obtener la funcional de desajuste de un modelo de prueba \mathbf{m}_t , de acuerdo al punto uno del algoritmo en la subsección (1.2.1). Tomando en cuenta lo anterior, es necesario hacer una simulación directa para el cálculo de la funcional de desajuste $\chi(\mathbf{m}^k)$, una simulación directa y adjunta para el cálculo del gradiente $\nabla_m \chi(\mathbf{m}^k)$ y por último una simulación directa para el cálculo de la funcional de desajuste para el modelo de prueba $\chi(\mathbf{m}_t)$. En total son necesarias 3 simulaciones por evento e iteración, para posteriormente realizar una interpolación polinomial de segundo grado de estos parámetros y obtener el valor del paso ν^k que produce el mínimo analítico del polinomio interpolado y la actualización del nuevo modelo \mathbf{m}^k . Para reducir el tiempo de cómputo en cada iteración y reducir el número de simulaciones a 2 por evento, se fijó el valor de paso ($\nu^k = 0.05$), de tal manera que la actualización en el modelo en cada iteración no sea tan drástica pero tampoco tan conservadora, asegurando la convergencia al mínimo de la funcional de desajuste.

La reducción a 2 simulaciones por evento impacta de gran manera en la reducción del tiempo de computo por iteración. Tomando en cuenta que son 12 eventos por iteración, se realizan 24 simulaciones para obtener los kernels totales de cada parámetro. Por otro lado, si no se fijara el parámetro de paso ν^k se realizarían 36 simulaciones por iteración, por lo que al hacer fijo el parámetro de paso en el algoritmo de gradiente conjugado se reduce el tiempo de cómputo en un 30%, comparado con el algoritmo tradicional de Tape et al. [2010].

El tiempo de cómputo de la simulación para el problema directo, para los 12 eventos, es de $T \approx 6 [h]$ aproximadamente y para la simulación adjunta es de $T \approx 12 [h]$ ya que esta consta de una simulación directa y una en tiempo reverso. Teniendo en cuenta lo anterior, cada iteración de la FWI tiene un tiempo de cómputo de $T \approx 18 [h]$, sin considerar el post-procesamiento de los kernels y la actualización del modelo entre cada iteración.

Todo el procedimiento anterior y los tiempos de computo mencionados son los mismos para cada iteración de la FWI, con la excepción de la funcional de desajuste que se utiliza para obtener

el modelo que la minimiza, ya sea la de tiempos de desfase \mathcal{T} o la norma \mathbf{L}_2 .

Para la segunda iteración de la FWI, se minimizaron los tiempos de arribo de las fases predominantes en los sismogramas observados filtrados a 0.1 [Hz]. Así mismo, se minimizó la norma \mathbf{L}_2 . En total, se tomaron tres ventanas de duración variable para los datos observados y sintéticos para realizar las correlaciones y obtener los tiempos de desfase entre los datos, y también para la diferencia entre los datos sintéticos y observados de acuerdo a la Ec.(1.59), derivada de la funcional de desajuste de la norma \mathbf{L}_2 . Fue necesario aplicar un procesamiento extra a la fuente adjunta producida por la normal \mathbf{L}_2 . Este consistió en normalizarla con su máximo en cada ventana y la energía del sismograma sintético y observado de acuerdo a la normalización en la Ec.(3.4). Lo anterior se realizó para que no existiera dependencia de la amplitud en las fuentes adjuntas y sólo de la forma de onda. Por lo que en esta iteración se aseguró una mayor sensibilidad en los parámetros del modelo respecto a la iteración anterior, de tal manera que se tuviera sensibilidad a estructuras más pequeñas presentes en los datos.

Para la tercera y cuarta iteración de la FWI sólo se minimizó la norma \mathbf{L}_2 . En el caso de la tercera iteración se procesaron los datos sintéticos y observados con un filtro pasabajos de 0.2 [Hz], y se obtuvieron las diferencias entre los datos sintéticos y observados para cada ventana seleccionada, obteniéndose la fuente adjunta para esta iteración. En cuanto a la cuarta iteración se realizó el mismo procedimiento que en la interacción anterior pero en ésta se filtraron los sismogramas a 0.25 [Hz], que es la frecuencia máxima alcanzable por el método de diferencias finitas para la discretización definida en este trabajo.

Para cada iteración se utilizó una banda de frecuencias distinta y esto se debe a que se aplicó una aproximación multiescala, en otras palabras, se ajustan primero los periodos largos de los sismogramas observados que están relacionados con las estructuras de gran tamaño, luego los periodos intermedios y por último los periodos cortos de los sismogramas, de tal manera que se va detallando el modelo a invertir en cada iteración. Esta forma de aproximación facilita la búsqueda del mínimo de la funcional de desajuste, ya que en la aproximación en bajas frecuencias se tiene un número menor de mínimos locales, por lo que es más fácil aproximarse al mínimo global de la funcional en bajas frecuencia y posteriormente ir detallándolo conforme se aumenta el contenido espectral de los datos. En la Fig.(4.1) se muestra de manera esquemática la aproximación multiescala para

inversión.

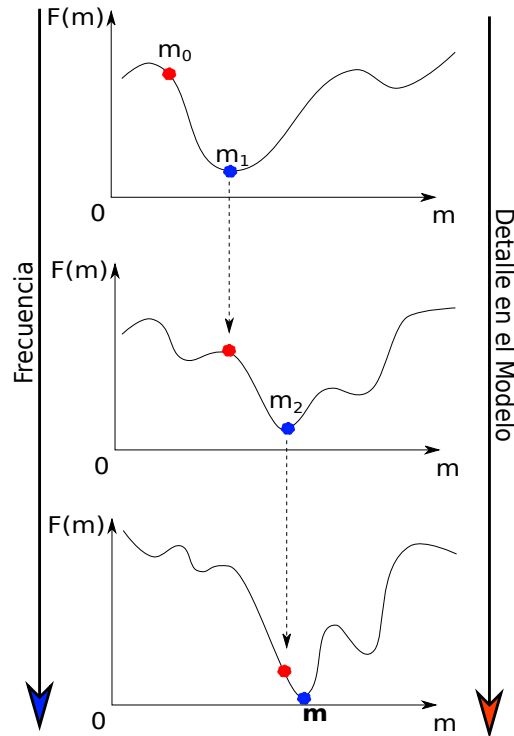


Figura 4.1: Esquema de la aproximación multiescala en tres etapas: La primera para periodos largos en los datos observados, encontrando un mínimo en m_1 , el cual es usado como modelo inicial para la siguiente etapa. En la segunda etapa se ajustan los datos de periodos intermedios, donde el algoritmo de minimización encuentra un mínimo en m_2 y éste a su vez es usado para la minimización en periodos cortos como el modelo inicial. Finalmente se encuentra el mínimo global de la funcional de desajuste m .

En la Fig.(4.2) se muestra el perfil del modelo 3D obtenido para cada iteración, ubicada en la Fig.(4.3) como la sección B-B'. En estos perfiles se puede observar la evolución del gradiente para cada iteración y la forma en que se van detallando las estructuras presentes en los datos. Otro aspecto de importancia en cada iteración es la aproximación multiescala que se utilizó, ya que para la primera iteración se detallan las estructuras de velocidad grandes, para la segunda las estructuras de tamaño medio, y por último en la iteración tres y cuatro se detallan estructuras

relativamente pequeñas presentes en la zona de estudio.

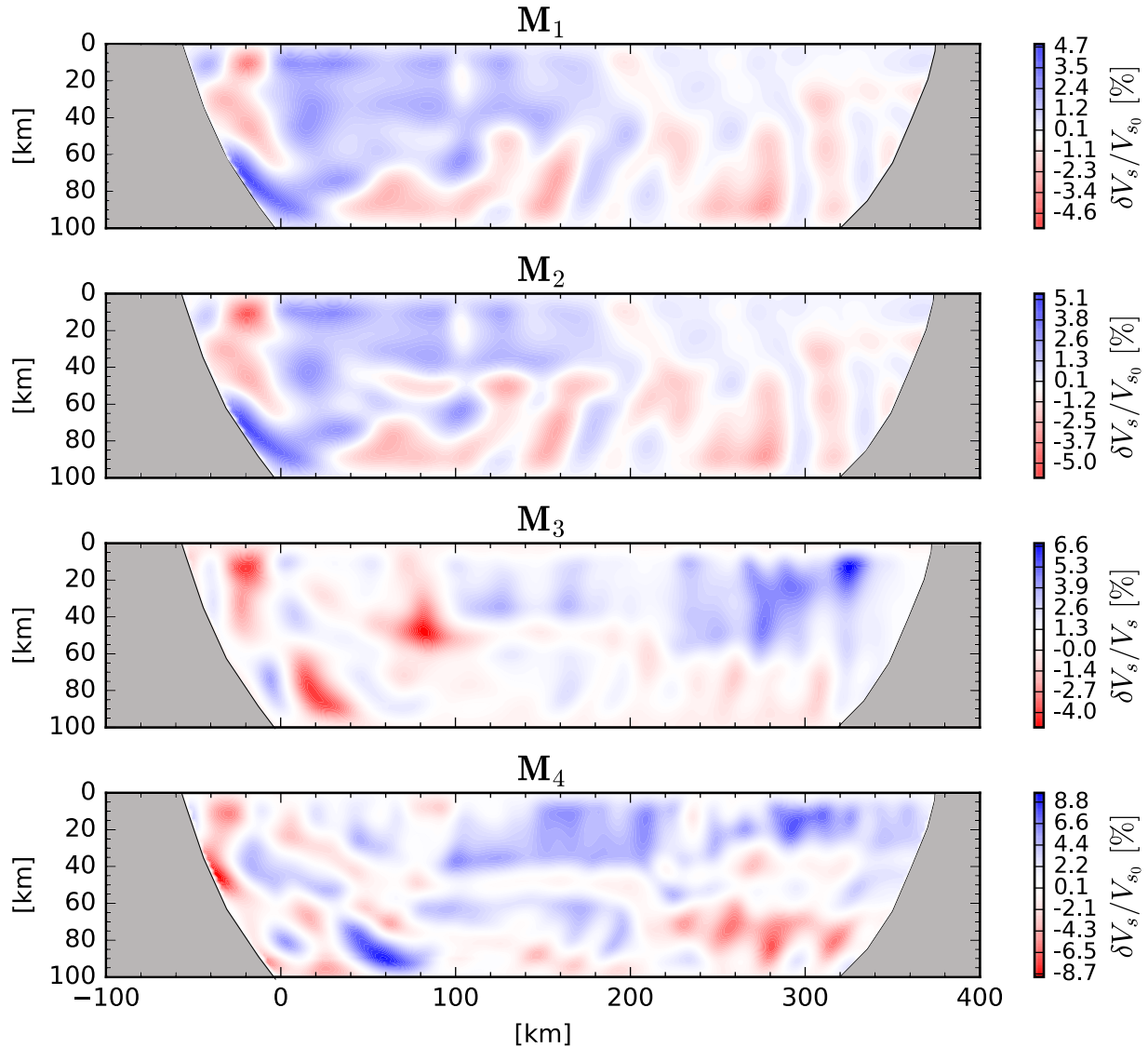


Figura 4.2: Evolución de los gradientes de las primeras cuatro iteraciones para el perfil B-B'.

Para la última iteración de la FWI se obtuvieron tres perfiles paralelos al arreglo MASE y perpendiculares a la trinchera, los cuales se indican en el mapa de la Fig.(4.3) como A-A', B-B' y C-C'. La ubicación de los perfiles se definió en función de la zona en donde se tiene una mayor sensibilidad al realizar la inversión; esta sensibilidad es función de la geometría de las fuentes y

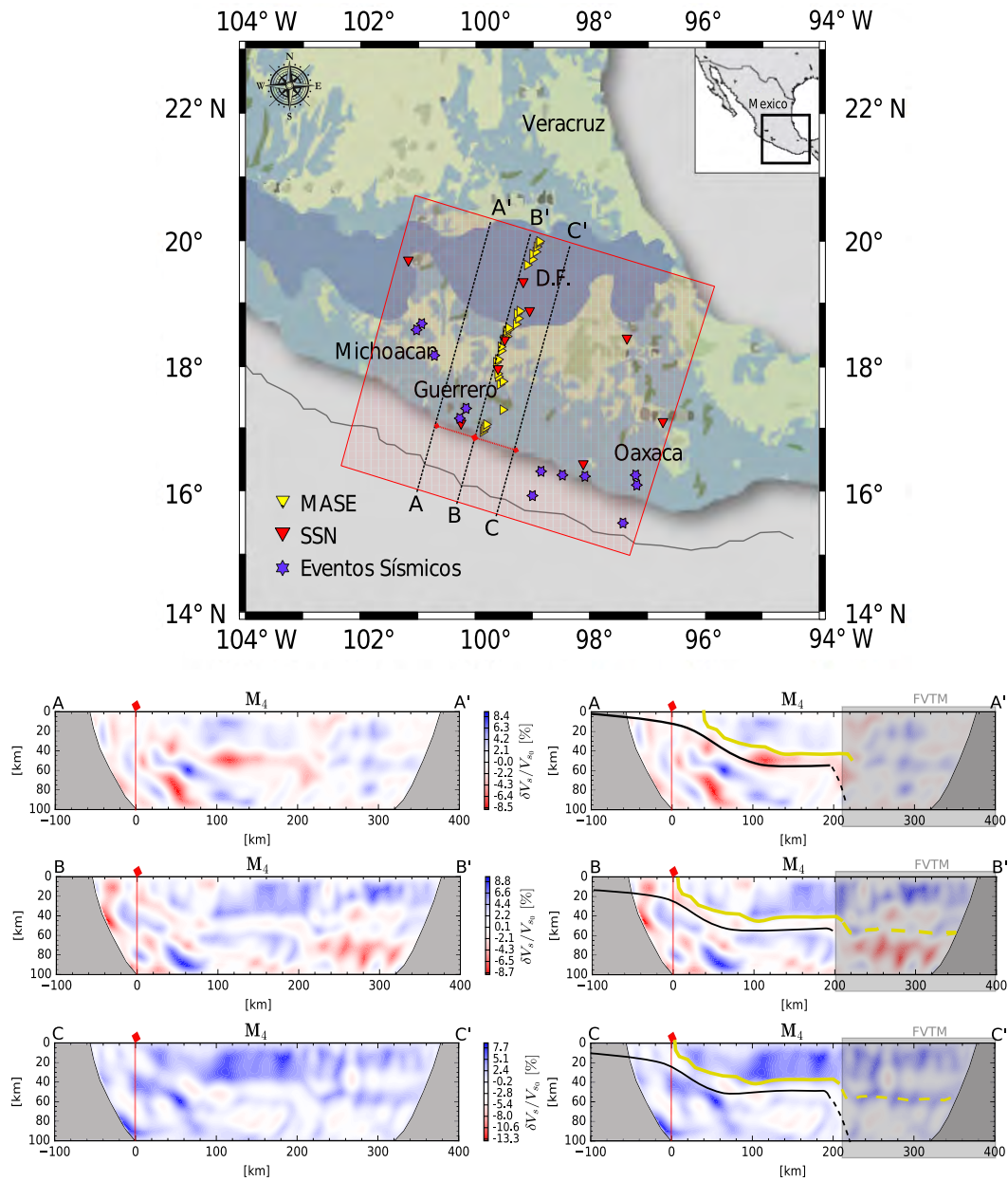


Figura 4.3: Gradiente de perturbaciones para la iteración M_4 calculado por la FWI, del cual se obtienen tres perfiles paralelos al arreglo MASE (A-A', B-B' y C-C'), con una separación entre cada perfil de 70 [km].

receptores con que se cuenta.

En el perfil A-A' en la Fig.(4.3), se puede observar claramente una anomalía de baja velocidad que va desde los -100 [km] hasta los 240 [km] en dirección hacia la FVTM. Esta anomalía es resaltada con una línea negra continua sólo donde se observa claramente, y con una línea punteada en los sitios en que no se tiene certeza que sea una anomalía o un artefacto generado por el método numérico. Esta anomalía se puede dividir en cuatro secciones principales. La primera a partir del punto de referencia de 0 [km] hacia la trinchera (-100 [km]), a una profundidad de 20-25 [km], se observa una estructura de baja velocidad casi horizontal, con un ángulo de 6° , una longitud media de 100 [km] y un espesor de entre 15-25 [km] aproximadamente. La segunda sección se define a partir del punto de referencia de 0 [km] a 120 [km] en dirección al arco volcánico, en donde se encuentra un cambio de pendiente en la anomalía; en este punto la pendiente tiene un valor aproximado de 20° con respecto a la horizontal. El espesor de la anomalía se conserva entre 5-10 [km] y la profundidad máxima que alcanza es aproximadamente de entre 50 a 60 [km], a una distancia de 120 [km] a partir del punto de referencia (línea roja punteada). En la tercera sección, localizada entre 120 [km] y 200 [km] existe otra transición en el ángulo de inclinación de la anomalía, pasando de una pendiente de 20° a un ángulo casi nulo; la longitud de la anomalía en este sector es de 100 [km], a una profundidad constante de 40 a 60 km en toda su extensión. En esta misma sección se puede observar otra anomalía de alta velocidad (línea amarilla continua) respecto al medio circundante; ésta presenta una longitud de 200 [km] a partir del punto de referencia, y un espesor de entre 40 y 50 [km] ubicada sobre la parte horizontal de la anomalía de baja velocidad, y un acuñaamiento entre el punto de referencia y 100 [km] a partir de este. En la cuarta y última sección de este perfil se tiene una imagen muy caótica de la cual no se puede extraer gran información, excepto en el tramo entre 200 y 240 [km], en el cual se puede observar un nuevo cambio en el ángulo de inclinación de la anomalía, con un valor aproximado de 68° con respecto la horizontal. En este tramo también se puede observar que a una profundidad de 80 [km] la anomalía de baja velocidad es trucada por una anomalía de alta velocidad y vuelve aparecer la anomalía de baja velocidad a una profundidad de entre 90 y 100 [km], por lo que no es posible saber con certeza si estas dos anomalías son la misma o es un artefacto generado por el método numérico y su sensibilidad a esta profundidad. Esta parte de la anomalía se marca con una línea

punteada como una de las posibles interpretaciones que se puede dar a este tramo, considerando que es continua la anomalía a estas profundidades.

En el perfil B-B' en la Fig.(4.3), en el cual se tiene la mayor sensibilidad de acuerdo a la configuración entre fuentes y receptores, se pueden observar claramente dos anomalías, una de baja y una de alta velocidad. La anomalía de alta velocidad (línea color amarillo) tiene una longitud de 400 [km] aproximadamente, con un espesor variable que va desde 60 [km] debajo de la FVTM (rectángulo sombreado) a 50 [km] en el sector ubicado entre 60 [km] y 200 [km] de distancia a partir del punto de referencia. Por último se observa un acuñamiento en el tramo entre el punto de referencia y 60 [km] a partir de éste. En el caso de la segunda anomalía, que es de baja velocidad y que va desde los -100 [km] hasta 240 [km] en dirección hacia la FVTM, la cual es delimitada con una línea negra continua, se dividió en cuatro secciones principales de igual manera que en el perfil A-A'. La primera sección queda definida a partir del punto de referencia de 0 [km] hacia la trinchera (-100 [km]) a una profundidad de 20 [km]; se observa una estructura de baja velocidad casi horizontal con un ángulo de 7 [°], una longitud de 100 [km] y un espesor entre 15-25 [km] aproximadamente. La segunda sección se define a partir del punto de referencia de 0 [km] a 60 [km] en dirección a la Ciudad de México, en donde se encuentra un cambio de pendiente en la anomalía, pasando de 7 [°] a 21 [°] de inclinación con respecto a la horizontal. El espesor de la anomalía se conserva entre 5 y 10 [km] y la profundidad máxima que alcanza es aproximadamente entre 50 y 55 [km], a una distancia de 60 [km] a partir del punto de referencia. En la tercera sección, localizada entre 60 [km] y 200 [km], existe otra transición en el ángulo de inclinación de la anomalía, pasando de 21 [°] a un ángulo casi nulo; la longitud de la anomalía en este sector es de 140 [km] a una profundidad constante de 45 a 55 km en toda su extensión. En la cuarta y última sección de este perfil se tiene una imagen más clara comparada con el perfil anterior; en esta se puede observar, entre 200 y 240 [km], un cambio en el ángulo de inclinación de la anomalía con un valor aproximado de 69 [°] con respecto a la horizontal. En este tramo también se puede observar que a una profundidad de 90 [km] la anomalía de baja velocidad es truncada por una anomalía de alta velocidad como se observó en el perfil A-A', pero a diferencia del perfil anterior se tiene una mejor delimitación de la anomalía, por lo que es posible dar una interpretación más confiable.

Por último en el perfil C-C', la sensibilidad del método es menor comparada con los dos perfiles

anteriores, pero se puede interpretar una anomalía de baja velocidad en color blanco, la cual también se divide en cuatro secciones principales como en los perfiles anteriores. La primera sección tiene una extensión de 100 [km] medidos desde el punto de referencia en dirección de la trinchera. En este tramo se puede observar una anomalía horizontal no muy bien definida, con una extensión de 100 [km] y un espesor de entre 15 y 25 [km] aproximadamente. En la segunda sección, que inicia en el punto de referencia y termina a una distancia de 70 [km] hacia la FVTM, se observa un cambio de pendiente en la anomalía, pasando de 7 [°] a 19 [°] aproximadamente; en este sector la anomalía tiene un grosor entre 5 a 10 [km] y alcanza una profundidad máxima de 50-55 [km] en el punto de transición del ángulo de inclinación. En la tercera sección ubicada entre 60 [km] y 200 [km] existe otra transición en el ángulo de inclinación de la anomalía, pasando a un ángulo casi nulo, con una longitud de 140 [km] a una profundidad constante de 45 a 55 km en toda su extensión. En la última sección de este perfil se tiene una imagen caótica del subsuelo debido a la baja sensibilidad del método de inversión en esta zona; lo que se puede observar con ayuda de los perfiles anteriores es un cambio en el ángulo de inclinación de la anomalía con un valor aproximado de 71 [°] con respecto a la horizontal. En este tramo también se puede observar que a una profundidad de 90 [km] la anomalía de baja velocidad es trucaada por una anomalía de alta velocidad como se observo en el perfil A-A' y B-B'.

Capítulo 5

Discusión y Conclusiones

En este trabajo se implementó la Tomografía de Forma de Onda Completa con su formulación adjunta para obtener un modelo de velocidades 3D que delimite la geometría del slab de subducción bajo el sector centro-oriental de la FVTM. Los resultados obtenidos en este estudio fueron interpretados con base en el marco geológico regional de la zona; éstos se discuten a lo largo de este capítulo.

Los resultados mostrados en el perfil B-B' de la cuarta iteración de la FWI, se dividen en cuatro secciones principales a partir del cambio en el ángulo de inclinación de la anomalía de baja velocidad encontrada en los tres perfiles descritos previamente. Las interpretaciones geológicas se realizaron sobre este perfil debido a que se encuentra en la zona de mayor sensibilidad de acuerdo a la geometría que se tiene entre fuentes y receptores.

La Fig.(5.1) representa el perfil B-B', el cual se divide en cuatro secciones. En la primera sección se encuentra una anomalía de baja velocidad con una longitud de 100 [km] aproximadamente, la cual se interpreta como el límite inferior de la placa de Cocos debido a que la profundidad observada en esta anomalía concuerda con el catálogo de localización de los hipocentros, realizado por Pardo y Suárez [1995]. Para la segunda sección, que va del punto de referencia a una distancia de 100 [km] en dirección a la FVTM, se observa un cambio de inclinación en la anomalía; esta se interpreta como una variación en el ángulo de subducción del slab, el cual presenta una longitud aproximada de 60 a 70 km y que pasa de tener una inclinación de 7-8 [°] a 18-23 [°], aproximadamente. En esta zona se encuentra la mayor concentración de sismos, los cuales definen el plano de Wadati-

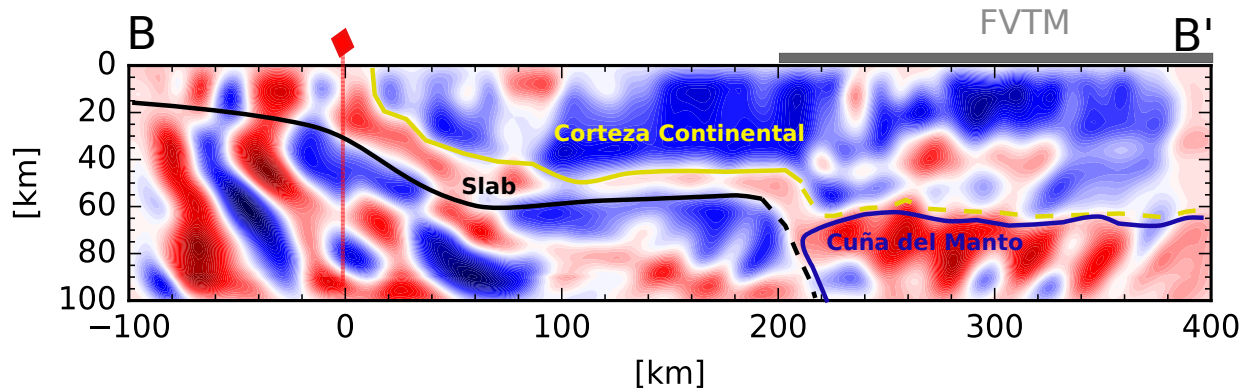


Figura 5.1: Gradiente de perturbaciones para la iteración M_4 . Se observa el slab de subducción asociado a la anomalía de baja velocidad delimitada con una línea negra continua, el límite entre la corteza continental y el manto superior delimitado con una línea amarilla continua y por último una anomalía de baja velocidad interpretada como la cuña del manto asociada al modelo de subducción, delimitada con una línea continua de color azul en la imagen.

Benioff para esta sección. La profundidad de la anomalía en este sector coincide con los parámetros reportados en los estudios realizados por Pardo y Suárez [1995] con la localización de hipocentros de sismos regionales y locales, y el estudio realizado con funciones de receptor por Pérez-Campos et al. [2008]. Para la tercera división del perfil B-B', que comprende el tramo entre 60 [km] y 200 [km] de distancia a partir del punto de referencia, se observa un decremento en el ángulo de subducción del slab, el cual subduce de manera semi horizontal a lo largo de 140 [km]; esto coincide con lo reportado en los estudios previos de la zona Pardo y Suárez [1995]; Pérez-Campos et al. [2008]; Kim et al. [2010, 2012], en los que se propone un modelo de subducción plana para este sector. Una de las hipótesis recurrentes en la literatura que trata de explicar la subducción plana del slab es la propuesta por Manea y Gurnis [2007]; Ferrari et al. [2012]. En esta hipótesis se plantea la existencia de una capa de baja velocidad (o baja viscosidad) que desacopla la parte superior del slab y la parte inferior de la placa cabalgante, explicando la ausencia de la zona de Waddati-Benioff y el bajo ángulo de subducción del slab. En la última y cuarta sección se observa un cambio abrupto en la inclinación del slab; este consiste en una variación drástica en el ángulo de inclinación, que pasa de ser semi horizontal a tener un ángulo de 65-70 [°] aproximadamente, a

una distancia del punto de referencia de 220 [km], lo cual concuerda con los estudios tomográficos realizados por Kim et al. [2012]; Pérez-Campos et al. [2008].

Por último, en el perfil B-B' se puede observar una anomalía de alta velocidad (línea amarilla) respecto al medio circundante, ubicada en el tramo definido entre el punto de referencia y un punto a una distancia de 400 [km], a partir de este. Esta se interpreta como la corteza continental y el engrosamiento de la misma. En la zona antearco presenta un espesor máximo de 40 [km], el cual llega a ser de hasta 55 [km] por debajo del arco volcánico (FVTM). También es posible observar en esta zona la discontinuidad de Mohorovičić localizada a una profundidad variable entre 50 y 55 [km], aproximadamente.

La aplicación de la técnica de Tomografía de Forma de Onda Completa en este trabajo, tuvo como principal objetivo obtener un modelo de velocidades del subsuelo que caracterizara la geometría del slab de subducción bajo el centro de México. Con los resultados obtenidos fue posible identificar las principales zonas de transición del slab en dirección perpendicular a la trinchera. Definiendo tres zonas de importancia, la primera definida en los primeros cien kilómetros en dirección a la FVTM, la cual es de gran interés científico ya que es donde se concentra la mayor cantidad de sismos ocurridos en México y también donde se originan los sismos de mayor magnitud. En la segunda zona definida, incluida en la zona de antearco, se tiene una subducción plana, la cual sigue siendo motivo de estudio debido a la geometría del slab y la ausencia de sismos en esta zona. Por último la tercera zona definida es aquella en donde se observa un cambio abrupto en el ángulo de subducción del slab; esto es de gran relevancia para comprender los procesos de vulcanismo y migración del arco volcánico en superficie.

Todos los resultados obtenidos en este trabajo concuerdan con la mayoría de los estudios realizados en la zona Kim et al. [2010]; Pérez-Campos et al. [2008]; Husker y Davis [2009] respecto a los valores de profundidad y el ángulo de subducción la placa de Cocos. La FWI nos permitió obtener un modelo más detallado de la geometría del slab de subducción, respecto a estudios previos similares. El mejoramiento de la imagen de la geometría del slab en profundidad contribuirá al mejoramiento del modelo de subducción que se acepta actualmente.

Las mejoras observadas son resultado de la implementación del método de la FWI ya que ésta, a diferencia de la tomografía sísmica clásica, permite el uso de toda la información contenida en los

datos observados, resultando en una imagen de alta calidad dependiendo de la banda de frecuencias que se utilice.

Referencias

- Aki, K. y Richards, P. G. (2002). *Quantitative seismology*, volume 1.
- Arfken, G. B. y Weber, H. J. (1999). *Mathematical methods for physicists*.
- Atwater, T. y Stock, J. (1998). Pacific-north america plate tectonics of the neogene southwestern united states: an update. *International Geology Review*, 40(5):375–402.
- Bandy, W., Hilde, T., y Yan, C. (2000). The rivera-cocos plate boundary: Implications for rivera-cocos relative motion and plate fragmentation. *Geological Society of America Special Papers*, 334:1–28.
- Bandy, W., Kostoglodov, V., y Mortera-Gutiérrez, C. (1998). Southwest migration of the instantaneous rivera-pacific euler pole since 0.78 ma. *GEOFISICA INTERNACIONAL-MEXICO-*, 37:153–170.
- Berenger, J.-P. (1996). Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 127(2):363–379.
- Berteussen, K.-A. (1977). Moho depth determinations based on spectral-ratio analysis of norsar long-period p waves. *Physics of the Earth and Planetary Interiors*, 15(1):13–27.
- Campos-Enriquez, J. y Sánchez-Zamora, O. (2000). Crustal structure across southern mexico inferred from gravity data. *Journal of South American Earth Sciences*, 13(6):479–489.
- Chavent, G., Dupuy, M., Lemmonier, P., et al. (1975). History matching by use of optimal theory. *Society of Petroleum Engineers Journal*, 15(01):74–86.

- DeMets, C. (2001). A new estimate for present-day cocos-caribbean plate motion: Implications for slip along the central american volcanic arc. *Geophysical Research Letters*, 28(21):4043–4046.
- DeMets, C. y Traylen, S. (2000). Motion of the rivera plate since 10 ma relative to the pacific and north american plates and the mantle. *Tectonophysics*, 318(1):119–159.
- Dziewonski, A., Chou, T.-A., y Woodhouse, J. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, 86(B4):2825–2852.
- Eissler, H. K. y McNally, K. C. (1984). Seismicity and tectonics of the rivera plate and implications for the 1932 jalisco, mexico, earthquake. *Journal of Geophysical Research: Solid Earth*, 89(B6):4520–4530.
- Ekström, G., Nettles, M., y Dziewoński, A. (2012). The global cmt project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200:1–9.
- Ferrari, L. (2004). Slab detachment control on mafic volcanic pulse and mantle heterogeneity in central mexico. *Geology*, 32(1):77–80.
- Ferrari, L., Orozco-Esquivel, T., Manea, V., y Manea, M. (2012). The dynamic history of the trans-mexican volcanic belt and the mexico subduction zone. *Tectonophysics*, 522:122–149.
- Fichtner, A. (2010). *Full seismic waveform modelling and inversion*. Springer Science & Business Media.
- Fichtner, A., Kennett, B. L., Igel, H., y Bunge, H.-P. (2008). Theoretical background for continental-and global-scale full-waveform inversion in the time–frequency domain. *Geophysical Journal International*, 175(2):665–685.
- Fletcher, R. y Reeves, C. M. (1964). Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154.
- Fornberg, B. (1988). Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of computation*, 51(184):699–706.

- Gauthier, O., Virieux, J., y Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms: Numerical results. *Geophysics*, 51(7):1387–1403.
- Graves, R. W. (1996). Simulating seismic wave propagation in 3d elastic media using staggered-grid finite differences. *Bulletin of the Seismological Society of America*, 86(4):1091–1106.
- Husker, A. y Davis, P. M. (2009). Tomography and thermal state of the cocos plate subduction beneath mexico city. *Journal of Geophysical Research: Solid Earth*, 114(B4).
- Iglesias, A., Clayton, R., Pérez-Campos, X., Singh, S., Pacheco, J., García, D., y Valdés-González, C. (2010). S wave velocity structure below central mexico using high-resolution surface wave tomography. *Journal of Geophysical Research: Solid Earth*, 115(B6).
- Jödicke, H., Jording, A., Ferrari, L., Arzate, J., Mezger, K., y Rüpke, L. (2006). Fluid release from the subducted cocos plate and partial melting of the crust deduced from magnetotelluric studies in southern mexico: Implications for the generation of volcanism and subduction dynamics. *Journal of Geophysical Research: Solid Earth*, 111(B8).
- Kim, Y., Clayton, R., y Jackson, J. (2010). Geometry and seismic properties of the subducting cocos plate in central mexico. *Journal of Geophysical Research: Solid Earth*, 115(B6).
- Kim, Y., Miller, M. S., Pearce, F., y Clayton, R. W. (2012). Seismic imaging of the cocos plate subduction zone system in central mexico. *Geochemistry, Geophysics, Geosystems*, 13(7).
- Komatitsch, D. y Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, 72(5):SM155–SM167.
- Komatitsch, D. y Tromp, J. (2002). Spectral-element simulations of global seismic wave propagation—i. validation. *Geophysical Journal International*, 149(2):390–412.
- Kostoglodov, V., Singh, S. K., Santiago, J. A., Franco, S. I., Larson, K. M., Lowry, A. R., y Bilham, R. (2003). A large silent earthquake in the guerrero seismic gap, mexico. *Geophysical Research Letters*, 30(15).

- Lions, J.-L. (1968). Contrôle optimal de systemes gouvernés par des équations aux dérivées partielles.
- Lonsdale, P. (2005). Creation of the cocos and nazca plates by fission of the farallon plate. *Tectonophysics*, 404(3):237–264.
- Luo, Y. y Schuster, G. T. (1991). Wave-equation travelttime inversion. *Geophysics*, 56(5):645–653.
- Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, 66(3):639–666.
- Manea, V. y Gurnis, M. (2007). Subduction zone evolution and low viscosity wedges and channels. *Earth and Planetary Science Letters*, 264(1):22–45.
- Melgar, D. y Pérez-Campos, X. (2011). Imaging the moho and subducted oceanic crust at the isthmus of tehuantepec, mexico, from receiver functions. *Pure and applied geophysics*, 168(8-9):1449–1460.
- Michéa, D. y Komatitsch, D. (2010). Accelerating a three-dimensional finite-difference wave propagation code using gpu graphics cards. *Geophysical Journal International*, 182(1):389–402.
- Mittet, R. (2002). Free-surface boundary conditions for elastic staggered-grid modeling schemes. *Geophysics*, 67(5):1616–1623.
- Moczo, P., Kristek, J., y Halada, L. (2000). 3d fourth-order staggered-grid finite-difference schemes: Stability and grid dispersion. *Bulletin of the Seismological Society of America*, 90(3):587–603.
- Nixon, G. T. (1982). The relationship between quaternary volcanism in central mexico and the seismicity and structure of subducted ocean lithosphere. *Geological Society of America Bulletin*, 93(6):514–523.
- Ortiz-Alemán, C., Reyes-Olvera, M., Iglesias-Mendoza, A., Orozco-Del-Castillo, M., y Hernández-Gómez, J. (2017). Estimation of source; path and site effects at mase array stations: a comprehensive study. *GEOFÍSICA INTERNACIONAL*, 56(1):71–85.

- Pardo, M. y Suárez, G. (1995). Shape of the subducted rivera and cocos plates in southern mexico: Seismic and tectonic implications. *J. geophys. Res*, 100(12):357–12.
- Pérez-Campos, X., Kim, Y., Husker, A., Davis, P. M., Clayton, R. W., Iglesias, A., Pacheco, J. F., Singh, S. K., Manea, V. C., y Gurnis, M. (2008). Horizontal subduction and truncation of the cocos plate beneath central mexico. *Geophysical Research Letters*, 35(18).
- Peter, D., Komatitsch, D., Luo, Y., Martin, R., Le Goff, N., Casarotti, E., Le Loher, P., Magnoni, F., Liu, Q., Blitz, C., et al. (2011). Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes. *Geophysical Journal International*, 186(2):721–739.
- Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Singh, S., Ponce, L., y Nishenko, S. (1985). The great jalisco, mexico, earthquakes of 1932: Subduction of the rivera plate. *Bulletin of the Seismological Society of America*, 75(5):1301–1313.
- Song, T.-R. A., Helmberger, D. V., Brudzinski, M. R., Clayton, R. W., Davis, P., Pérez-Campos, X., y Singh, S. K. (2009). Subducting slab ultra-slow velocity layer coincident with silent earthquakes in southern mexico. *Science*, 324(5926):502–506.
- Tape, C. (2009). *Seismic tomography of southern California using adjoint methods*. PhD thesis, California Institute of Technology.
- Tape, C., Liu, Q., Maggi, A., y Tromp, J. (2010). Seismic tomography of the southern califor-
nia crust based on spectral-element and adjoint methods. *Geophysical Journal International*, 180(1):433–462.
- Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.
- Tromp, J., Tape, C., y Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160(1):195–216.

- Urrutia-Fucugauchi, J. y Flores-Ruiz, J. H. (1996). Bouguer gravity anomalies and regional crustal structure in central Mexico. *International Geology Review*, 38(2):176–194.
- Virieux, J. (1986). P-sv wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 51(4):889–901.
- Virieux, J. y Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.
- Yang, T., Grand, S. P., Wilson, D., Guzman-Speziale, M., Gomez-Gonzalez, J. M., Dominguez-Reyes, T., y Ni, J. (2009). Seismic structure beneath the Rivera subduction zone from finite-frequency seismic tomography. *Journal of Geophysical Research: Solid Earth*, 114(B1).
- Zhu, H. (2013). *Seismic structure of the European crust and upper mantle based on adjoint tomography*. PhD thesis, Princeton University.