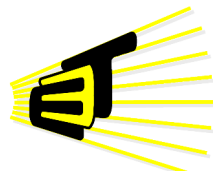


EVALUACIÓN DE MÉTRICAS EN DOMINIOS TRANSFORMADOS USADAS EN LAS METAHEURÍSTICAS PARA GENERAR UN PUNTO DE PARTIDA FAVORABLE A LA INVERSIÓN DE ONDA COMPLETA 2D

JHEYSTON OMAR SERRANO LUNA



Escuela de Ingenierías
Eléctrica, Electrónica y
de Telecomunicaciones



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

EVALUACIÓN DE MÉTRICAS EN DOMINIOS TRANSFORMADOS
USADAS EN LAS METAHEURÍSTICAS PARA GENERAR UN
PUNTO DE PARTIDA FAVORABLE A LA INVERSIÓN DE ONDA
COMPLETA 2D

Autor:
JHEYSTON OMAR SERRANO LUNA

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

Director:
MSc. Sergio A. Abreo Carrillo

Codirector:
PhD. Ana B. Ramirez S.

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

A mis padres y todas aquellas personas que me apoyan en cada una de mis etapas.

AGRADECIMIENTOS

Gracias Dios por darme la oportunidad para alcanzar este logro. A mis padres Jaime y Elizabeth por creer en mis metas. A Paula Andrea, a mis hermanos, primos y amigos que me apoyan en cada una de mis etapas. Agradezco al profesor Sergio Abreo y la Profesora Ana Beatriz Ramirez por sus valiosos consejos y sus continuas enseñanzas en mi formación académica. Agradezco a todos los integrantes del grupo de investigación CPS. A mi tío Jaime y mi madrina Esperanza por sus continuas enseñanzas y apoyo incondicional.

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

CONTENIDO

	Pág.
INTRODUCCIÓN	15
1 PROBLEMAS INVERSOS Y MÉTODOS DE OPTIMIZACIÓN GLOBAL	17
1.1. El método sísmico	17
1.2. El Problema directo e inverso de identificación	18
1.3. Full Waveform Inversion (FWI)	23
1.4. Métodos de optimización global	24
1.5. Función de costo	32
2 SELECCIÓN DE LOS DOMINIOS TRANSFORMADOS USADOS EN LAS MÉTRICAS	34
2.1. Dominio del tiempo	34
2.2. Dominio de la frecuencia	36
2.3. Dominio de Hilbert	38
3 METODOLOGÍA PARA LA CONSTRUCCIÓN DE LA MÉTRICA	43
3.1. Metodología propuesta	43
3.1.1. Método de diferencias finitas en el dominio tiempo	44
3.1.2. Construcción de la métrica	45
3.2. Modelos sintéticos	46
3.3. Resultados modelos sub-muestreados	48
3.4. Resultados de modelos a resolución de pixel	58
4 ANÁLISIS ESTADÍSTICO DE LOS DATOS	67
4.1. Seleccionar el número de experimentos	67
4.2. Selección de la función de distribución acumulada	69
4.2.1. Función de densidad de probabilidad Gaussiana	70
4.2.2. Función de densidad de probabilidad uniforme	70
4.2.3. Función de densidad de probabilidad Laplaciana	72
4.3. Estimadores	73
4.4. Matriz de varianzas <i>a-priori</i>	75
5 CONCLUSIONES	80
6 TRABAJO FUTURO	83

6.1. Ajuste de parámetros PSO	83
6.2. Otros dominios transformados	83
6.3. Construcción de la métrica	84
6.4. Costo computacional	85
6.5. Independencia entre píxeles	86
6.6. Puntos de partida para datos reales 2D y sintéticos 3D	87
ACRÓNIMOS	89
REFERENCIAS	90
BIBLIOGRAFÍA	92
ANEXOS	95

LISTA DE FIGURAS

	Pág.
Figura 1. Adquisición sísmica	18
Figura 2. Sismograma sintético	19
Figura 3. Problema directo e inverso de identificación.	20
Figura 4. Problema directo e inverso de identificación en el contexto geofísico.	21
Figura 5. Modificación del problema directo e inverso de identificación.	22
Figura 6. Problema del punto de partida en la FWI.	24
Figura 7. Puntos de partida en la FWI	25
Figura 8. Concepto del <i>cycle skipping</i>	27
Figura 9. Modelo empleado para ajustar w , c_1 y c_2	28
Figura 10. Modelo marmousi de dimensionalidad 11×3	32
Figura 11. Corrección de amplitud traza sísmica	35
Figura 12. Ejemplos del coeficiente de correlación	36
Figura 13. Espectro de la wavelet principal.	38
Figura 14. Método usando para computar el CS en el dominio de la frecuencia.	38
Figura 15. Representación de un sismograma en la traza compleja.	39
Figura 16. Representación de la traza compleja.	40
Figura 17. Cálculo del CS en la traza compleja.	41
Figura 18. Esquema solución del punto de partida.	44
Figura 19. Esquema en cascada de los dominios transformados.	47
Figura 20. Espectros de las fuentes PSO y FWI.	48
Figura 21. Modelos originales usados.	49
Figura 22. Uso del filtro sobre el modelo de velocidades	51
Figura 23. Modelo aleatorio generado en la primera iteración.	52
Figura 24. Modelo obtenido después de su filtrado.	52
Figura 25. Comportamiento de las iteraciones en función de la métrica \hat{C}_{corr}	53
Figura 26. Comportamiento de las iteraciones en función del tiempo.	54
Figura 27. Modelos obtenidos con cada dominio (Tiempo, frecuencia y Hilbert).	55
Figura 28. Cubo de los mejores globales de 100 experimentos.	56
Figura 29. Diagrama empleado para las combinaciones de métricas.	57
Figura 30. Zona de iluminación generada por las cinco fuentes.	57
Figura 31. Uso del operador FT_{angle} sobre la traza 50 con $\hat{C}S_{TC} + \hat{C}S_{FT} + \hat{C}_{corr}$ (\mathbf{m}_{ori} -cuadrado difractor).	58
Figura 32. Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{sub} -cuadrado difractor)	59
Figura 33. Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{sub} -cuadrado difractor).	59

Figura 34. Diagrama en cascada usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{sub} -cuadrado difractor).	60
Figura 35. \mathbf{m}_{sub} obtenido de \widehat{CS}_{TC} .	61
Figura 36. FWI usando \mathbf{m}_{sub} con métricas \widehat{CS}_{TC} .	61
Figura 37. \mathbf{m}_{sub} obtenido de $\widehat{CS}_{FT} + \widehat{CS}_{TC}$.	61
Figura 38. FWI usando \mathbf{m}_{sub} con métricas $\widehat{CS}_{FT} + \widehat{CS}_{TC}$.	61
Figura 39. \mathbf{m}_{sub} obtenido de $\widehat{CS}_{FT} + \widehat{C}_{corr} + \widehat{CS}_{TC}$.	62
Figura 40. FWI usando \mathbf{m}_{sub} con métricas $\widehat{CS}_{FT} + \widehat{C}_{corr} + \widehat{CS}_{TC}$.	62
Figura 41. \mathbf{m}_{sub} obtenido de $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$.	62
Figura 42. FWI usando \mathbf{m}_{sub} con métricas $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$.	62
Figura 43. Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full})-cuadrado difractor pixel.	63
Figura 44. Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -cuadrado difractor pixel).	63
Figura 45. Diagrama en cascada usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).	64
Figura 46. \mathbf{m}_{full} obtenido de \widehat{CS}_{TC} .	65
Figura 47. FWI usando \mathbf{m}_{full} con métricas \widehat{CS}_{TC} .	65
Figura 48. \mathbf{m}_{full} obtenido de $\widehat{CS}_{TC} + \widehat{C}_{corr}$.	65
Figura 49. FWI usando \mathbf{m}_{full} con métricas $\widehat{CS}_{TC} + \widehat{C}_{corr}$.	65
Figura 50. \mathbf{m}_{full} obtenido de $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$.	66
Figura 51. FWI usando \mathbf{m}_{full} con métricas $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$.	66
Figura 52. \mathbf{m}_{full} obtenido de $\widehat{CS}_{FT} + \widehat{C}_{corr}$.	66
Figura 53. FWI usando \mathbf{m}_{full} con métricas $\widehat{CS}_{FT} + \widehat{C}_{corr}$.	66
Figura 54. Ubicación del pixel para el análisis estadístico.	67
Figura 55. Varianzas normalizadas de las métricas.	68
Figura 56. Ajuste entre función de distribución acumulada empírica con la función de distribución acumulada normal.	71
Figura 57. Ajuste entre función de distribución acumulada empírica con la función de distribución uniforme acumulada.	72
Figura 58. función de distribución acumulada usando una distribución Laplaciana.	74
Figura 59. Construcción de la función de densidad de probabilidad de cada estimador.	77
Figura 60. Varianza con métrica \widehat{C}_{corr} .	78
Figura 61. Varianzas con métrica \widehat{CS}_{FT} .	78
Figura 62. Varianzas con métrica \widehat{CS}_{TC} .	78
Figura 63. $20 \cdot \log(\cdot)$ varianzas con métrica \widehat{C}_{corr} .	79
Figura 64. $20 \cdot \log(\cdot)$ varianzas con métrica \widehat{CS}_{FT} .	79
Figura 65. $20 \cdot \log(\cdot)$ varianzas con métrica \widehat{CS}_{TC} .	79
Figura 66. Esquema para activar varios dominios transformados.	85
Figura 67. Diagrama trabajo futuro.	86
Figura 68. Matriz de correlación usando la métrica \widehat{C}_{corr} .	88

Figura 69.	Matriz de correlación usando la métrica \widehat{CS}_{FT} .	88
Figura 70.	Matriz de correlación usando la métrica \widehat{CS}_{TC} .	88
Figura 70.	Uso del operador $FT_{angle}\{\cdot\}$ sobre la traza 50 obtenida con las métricas $\widehat{CS}_{FT}+\widehat{CS}_{TC}$ (\mathbf{m}_{ori} -2 capas).	95
Figura 71.	Uso del operador $FT_{angle}\{\cdot\}$ sobre la traza 50 obtenida de las métricas $\widehat{CS}_{TC}+\widehat{CS}_{FT}+\widehat{C}_{corr}$ (\mathbf{m}_{ori} -marmousi).	95
Figura 72.	Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -cuadrado difractor).	96
Figura 73.	Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).	96
Figura 74.	Diagrama multi-métrica usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).	96
Figura 75.	Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -2 capas).	96
Figura 76.	Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -2 capas).	97
Figura 77.	Diagrama multi-métrica usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -2 capas).	97
Figura 78.	Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -marmousi).	97
Figura 79.	Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -marmousi).	97
Figura 80.	Diagrama multi-métrica usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -marmousi).	98
Figura 81.	\mathbf{m}_{sub} cuadrado difractor estimado por \widehat{C}_{corr} .	103
Figura 82.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{FT}$.	103
Figura 83.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{FT}+\widehat{CS}_{TC}$.	103
Figura 84.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{TC}$.	103
Figura 85.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{TC}+\widehat{CS}_{FT}$.	104
Figura 86.	\mathbf{m}_{sub} cuadrado difractor estimado por \widehat{CS}_{FT} .	104
Figura 87.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{FT}+\widehat{C}_{corr}$.	104
Figura 88.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{FT}+\widehat{C}_{corr}+\widehat{CS}_{TC}$.	104
Figura 89.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{FT}+\widehat{CS}_{TC}$.	104
Figura 90.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{FT}+\widehat{CS}_{TC}+\widehat{C}_{corr}$.	105
Figura 91.	\mathbf{m}_{sub} cuadrado difractor estimado por \widehat{CS}_{TC} .	105
Figura 92.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC}+\widehat{C}_{corr}$.	105
Figura 93.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC}+\widehat{C}_{corr}+\widehat{CS}_{FT}$.	105
Figura 94.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC}+\widehat{CS}_{FT}$.	105
Figura 95.	\mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC}+\widehat{CS}_{FT}+\widehat{C}_{corr}$.	106
Figura 96.	\mathbf{m}_{full} cuadrado difractor estimado por \widehat{C}_{corr} .	107
Figura 97.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{FT}$.	107
Figura 98.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{FT}+\widehat{CS}_{TC}$.	107
Figura 99.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{TC}$.	107
Figura 100.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{C}_{corr}+\widehat{CS}_{TC}+\widehat{CS}_{FT}$.	108
Figura 101.	\mathbf{m}_{full} cuadrado difractor estimado por \widehat{CS}_{FT} .	108
Figura 102.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{FT}+\widehat{C}_{corr}$.	108
Figura 103.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{FT}+\widehat{C}_{corr}+\widehat{CS}_{TC}$.	108

Figura 104.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC}$	108
Figura 105.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$	109
Figura 106.	\mathbf{m}_{full} cuadrado difractor estimado por \widehat{CS}_{TC}	109
Figura 107.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$	109
Figura 108.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$	109
Figura 109.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$	109
Figura 110.	\mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$	110
Figura 111.	\mathbf{m}_{full} dos capas estimado por \widehat{C}_{corr}	111
Figura 112.	\mathbf{m}_{full} dos capas estimado por $\widehat{C}_{corr} + \widehat{CS}_{FT}$	111
Figura 113.	\mathbf{m}_{full} dos capas estimado por $\widehat{C}_{corr} + \widehat{CS}_{FT} + \widehat{CS}_{TC}$	111
Figura 114.	\mathbf{m}_{full} dos capas estimado por $\widehat{C}_{corr} + \widehat{CS}_{TC}$	111
Figura 115.	\mathbf{m}_{full} dos capas estimado por $\widehat{C}_{corr} + \widehat{CS}_{TC} + \widehat{CS}_{FT}$	112
Figura 116.	\mathbf{m}_{full} dos capas estimado por \widehat{CS}_{FT}	112
Figura 117.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{FT} + \widehat{C}_{corr}$	112
Figura 118.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{FT} + \widehat{C}_{corr} + \widehat{CS}_{TC}$	112
Figura 119.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC}$	112
Figura 120.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$	113
Figura 121.	\mathbf{m}_{full} dos capas estimado por \widehat{CS}_{TC}	113
Figura 122.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$	113
Figura 123.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$	113
Figura 124.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$	113
Figura 125.	\mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$	114
Figura 126.	\mathbf{m}_{full} marmousi estimado por \widehat{C}_{corr}	115
Figura 127.	\mathbf{m}_{full} marmousi estimado por $\widehat{C}_{corr} + \widehat{CS}_{FT}$	115
Figura 128.	\mathbf{m}_{full} marmousi estimado por $\widehat{C}_{corr} + \widehat{CS}_{FT} + \widehat{CS}_{TC}$	115
Figura 129.	\mathbf{m}_{full} marmousi estimado por $\widehat{C}_{corr} + \widehat{CS}_{TC}$	115
Figura 130.	\mathbf{m}_{full} marmousi estimado por $\widehat{C}_{corr} + \widehat{CS}_{TC} + \widehat{CS}_{FT}$	116
Figura 131.	\mathbf{m}_{full} marmousi estimado por \widehat{CS}_{FT}	116
Figura 132.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{FT} + \widehat{C}_{corr}$	116
Figura 133.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{FT} + \widehat{C}_{corr} + \widehat{CS}_{TC}$	116
Figura 134.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC}$	116
Figura 135.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$	117
Figura 136.	\mathbf{m}_{full} marmousi estimado por \widehat{CS}_{TC}	117
Figura 137.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$	117
Figura 138.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$	117
Figura 139.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$	117
Figura 140.	\mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$	118

RESUMEN

TÍTULO: Evaluación de métricas en dominios transformados usadas en las metaheurísticas para generar un punto de partida favorable a la inversión de onda completa 2D*

AUTOR: Jheyston Omar Serrano Luna**

PALABRAS CLAVE: Datos sísmicos, Dominios transformados, Trazas sísmicas, GPU, PSO.

DESCRIPCIÓN:

La industria del petróleo está interesada en la búsqueda de mejores métodos para estimar los parámetros del subsuelo a partir de los datos obtenidos durante una exploración sísmica. Los parámetros del subsuelo se utilizan para encontrar imágenes de reflectividad de la tierra que indica la existencia de un yacimiento de petróleo o gas. Estos parámetros del subsuelo deben estimarse correctamente de tal manera que la imagen resultante sea correcta. La técnica de inversión de onda completa (FWI) es una herramienta que permite obtener una imagen de alta resolución del subsuelo. Actualmente se está realizando un gran esfuerzo para industrializar esta técnica, que tiene dos desventajas: su elevado costo computacional y su sensibilidad a la elección del punto de partida. Algunas estrategias son usadas para establecer el punto de partida: métodos analíticos, tomografías y métodos de optimización global. Los métodos de optimización global son técnicas bioinspiradas que ofrecen muy buenos resultados cuando no es viable implementar el método óptimo, son de fácil implementación y necesitan de alguna métrica para su avance en su búsqueda sobre la función objetivo.

El objetivo principal de este trabajo de maestría es explorar distintos dominios transformados y generar una metodología para la construcción de una métrica que junto con el método de optimización global genere un punto de partida favorable para la FWI. Con el fin de medir si el modelo obtenido por la metaheurísticas es favorable, nosotros estudiamos el cycle skipping (CS) en 3 dominios: tiempo, frecuencia y traza compleja.

* Trabajo de investigación.

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Maestría en Ingeniería Electrónica. Director: MSc. Sergio A. Abreo Carrillo.

ABSTRACT

TITLE: Evaluation of metrics in transformed domains used in metaheuristics to generate a starting point to the Full Wave Inversion 2D*

AUTHOR: Jheyston Omar Serrano Luna**

KEYWORDS: Seismic data, Transformed domains, Seismic traces, GPU, PSO.

DESCRIPTION: Full-waveform inversion (FWI) is a state-of-the-art method used to estimate subsurface parameters, such as the underground seismic velocity. A good starting velocity model (SVM) is required in order to find adequate solutions of the FWI. SVM is a key element since the FWI is formulated as a local optimization problem with a non-unique solution. Currently, some strategies used to build SVMs are: analytical methods, reflection tomography, and global optimization methods. In particular, global optimization techniques need some metric to measure the distance between the modeled and observed data to build adequate SVMs.

A model is considered adequate for the FWI when the low frequency components of the model are correctly estimated or cycle-skipping is not present. In this work, we study the cycle-skipping problem in three different domains: time, frequency and complex trace domain. In this work, we propose three different metrics in the transformed domains, which are used in a particle swarm optimization meta-heuristic, to generate good starting velocity models.

* Trabajo de investigación.

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Maestría en Ingeniería Electrónica. Director: MSc. Sergio A. Abreo Carrillo.

INTRODUCCIÓN

El presente proyecto está enmarcado dentro del programa de investigación “Migración sísmica pre-apilado en profundidad por extrapolación de campos de onda utilizando computación de alto desempeño para datos masivos en zonas complejas”. Dicha investigación pretende aumentar la resolución de las imágenes del subsuelo requeridas por la industria del petróleo para identificar posibles reservas de hidrocarburos así como reducir el tiempo de ejecución de las aplicaciones sísmicas usando arquitecturas de cómputo diferentes a las CPU.

El petróleo brinda la mayoría de energía consumida por la sociedad. Los problemas de transporte, la escasez de alimentos y apagones son algunas consecuencias que podría traer la insuficiencia de este recurso energético. Debido a su demanda y a no ser recursos renovables, es necesaria la disponibilidad continua de ellos, obligando su búsqueda permanente. Actualmente se ha generado un interés en encontrar yacimientos de petróleo más profundos y complejos. Esto requiere la búsqueda y perfeccionamiento de técnicas de inversión que permitan minimizar la incertidumbre en la interpretación de las imágenes sísmicas de estas zonas. En Colombia hay muchas zonas de alta complejidad topográfica que aún no han sido explotadas ni exploradas, estas zonas son importantes para ECOPEPETROL S.A ya que aumenta sus reservas petroleras, generando un impacto económico en el territorio Colombiano.

La sísmica es una estrategia usada para la exploración de hidrocarburos. Sin embargo la exploración sísmica de zonas complejas es complicada y por esta razón se hace necesario el uso de técnicas que permitan llevar a cabo un estudio para determinar si es o no es viable excavar estas zonas y así aumentar las reservas económicas del país. En este proyecto se desarrolla una metodología que permita estimar un modelo de velocidades que es utilizado como punto de partida para la técnica de inversión de onda completa (*full waveform inversion*, FWI). Con esta metodología se selecciona un punto de partida con suficiente información de baja frecuencia o sin problemas de cycle skipping (CS) permitiendo obtener modelos de velocidades más acordes a la geología real del subsuelo Colombiano. Las técnicas de optimización local como la FWI tratan de encontrar un mínimo global a partir de un punto de partida cercano al vecindario de la solución. Sin embargo pierden el mínimo global si se parte de una solución más cercana a un mínimo local. Estas técnicas tienen el problema de la dependencia del punto de partida además de su alto costo computacional.

Existen diversos métodos para poder estimar los puntos de partida para las técnicas de optimización local: Métodos tradicionales (Tomografía, Gradiente de velocidades, análisis de semblanza) o métodos de optimización global. A diferencia de los métodos

de optimización local, los métodos de optimización global intentan encontrar el mínimo global de la función objetivo. La mayoría de los algoritmos de optimización global son de naturaleza estocástica y usan la información global para su búsqueda sobre la función objetivo, sin embargo la convergencia de estos métodos no está garantizada.

En este trabajo se propone obtener un punto de partida adecuado usando la técnica de optimización por enjambre de partículas (particle swarm optimization, PSO) junto con la medida de cycle skipping en tres diferentes dominios transformados: tiempo, frecuencia y traza compleja. Los métodos de optimización global presentan un alto costo computacional y por tanto se plantea su implementación en paralelo utilizando el lenguaje de programación CUDA-C (Compute Unified Device Architecture) desarrollado por NVIDIA. La alternativa de solución no es completa sin asignar incertidumbre a la solución obtenida y por esta razón en este trabajo se aborda un análisis estadístico que sirva como información a priori en una etapa posterior denominada cuantificación de la incertidumbre.

El presente documento está organizado de la siguiente manera: En el capítulo 1 se presenta los problemas inversos y métodos de optimización global. En el capítulo 2 se seleccionan los dominios transformados. En el capítulo 3 se construye una metodología basados en los dominios establecidos para la construcción de la métrica y así mismo se exponen los resultados obtenidos con la metodología propuesta. En el capítulo 4 se muestra un análisis estadístico de los datos. En el capítulo 5 se presenta el trabajo futuro que deja este proyecto. Por último, en el capítulo 6 se presentan las conclusiones y recomendaciones.

1. PROBLEMAS INVERSOS Y MÉTODOS DE OPTIMIZACIÓN GLOBAL

En este capítulo se presenta en qué consiste el método sísmico, posteriormente el problema directo o hacia adelante así como el problema inverso. Seguidamente se establece que el problema de inversión que se desea solucionar en este proyecto es un problema mal puesto. Por otra parte se definirá la FWI presentando sus ventajas y desventajas. Finalmente se presenta la técnica de optimización global (PSO) usada para obtener el punto de partida en la FWI.

1.1. El método sísmico

El método sísmico consiste en liberar energía en el subsuelo. La energía se genera por una fuente, por ejemplo dinamita o camiones vibradores. En la explosión de la dinamita se genera un rango de frecuencias mientras que en los camiones solo se vibra a determinada frecuencia. Esta energía viaja a través del subsuelo donde es transmitida, reflejada o difractada por los cambios de impedancia entre las capas. En superficie esta energía es capturada por un arreglo de dispositivos llamados geófonos en el caso terrestre o hidrófonos en el caso marino. El método sísmico es presentado en la Figura 1.

Posterior a la adquisición se obtiene un conjunto de trazas o lo que también es conocido como un sismograma (ver Figura 2). Sobre el sismograma se aplica una serie de etapas: primero una etapa de pre-procesamiento, que busca eliminar toda aquella información que no es de interés (ruido). Esta etapa de pre-procesamiento permite resaltar la información de eventos sísmicos que son de interés [2]. Una vez concluida la etapa de pre-procesamiento se continúa con la etapa de procesamiento donde una de sus etapas es encontrar un modelo de velocidades por medio de las trazas sísmicas, esto es conocido como método de inversión. Existen tres tipos de inversión: inversión por tomografía de tiempo de viaje, inversión simultánea e inversión de onda completa (*Full waveform inversion*, FWI). Este modelo de velocidad que se obtiene a través de la inversión, es necesario para la etapa posterior llamada migración sísmica. La migración sísmica es la etapa final del procesamiento de datos sísmicos que permite obtener la imagen del subsuelo para establecer sus características internas [3]. Finalmente un geofísico experto realiza la interpretación del dato migrado.

En este documento sólo nos enfocaremos en la etapa de inversión usando la técnica FWI.

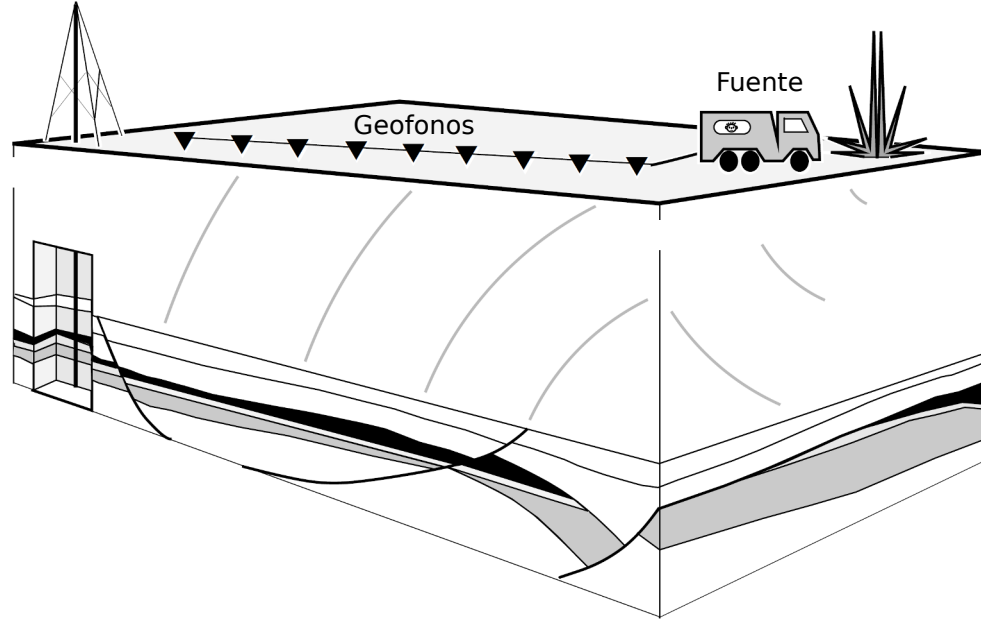


Figura 1: Adquisición sísmica, tomado de [1].

La FWI permite obtener modelos de velocidades de alta resolución. Sin embargo, a pesar de sus grandes ventajas la FWI tiene un elevado costo computacional al usar la ecuación de onda completa para obtener los datos modelados. Para mitigar el impacto del costo computacional se propone usar computación en paralelo para su implementación.

1.2. El Problema directo e inverso de identificación

El problema directo o problema hacia delante consiste en encontrar las observaciones \mathbf{d} a partir de los parámetros del modelo \mathbf{m} . El problema directo es definido por:

Sea $\mathbf{L}(\mathbf{m})(\cdot) : P \longrightarrow F$

\therefore dados $p \in P$, hallar $f \in F$ tal que $\mathbf{L}(\mathbf{m})(p) = f$.

El operador $\mathbf{L}(\mathbf{m})(\cdot)$ es conocido como el operador de d'Alembert y se define como $\mathbf{L}(c(x, z))(\cdot) = \nabla^2[\cdot] - \frac{1}{c^2(x, z)} \frac{\partial[\cdot]}{\partial t^2}$. En este caso $\mathbf{m} = c(x, z)$, $f = src(x, z)$, y $p(x, z)$ es el campo de presión. Si el operador de d'Alembert se aplica sobre el campo $p(x, z)$ y se introduce la fuente f obtenemos que

$$\mathbf{L}(c(x, z))(p(x, z)) \triangleq \frac{1}{c(x, z)^2} \frac{\partial^2 p(x, z)}{\partial t^2} - \frac{\partial^2 p(x, z)}{\partial x^2} - \frac{\partial^2 p(x, z)}{\partial z^2} = f. \quad (1.1)$$

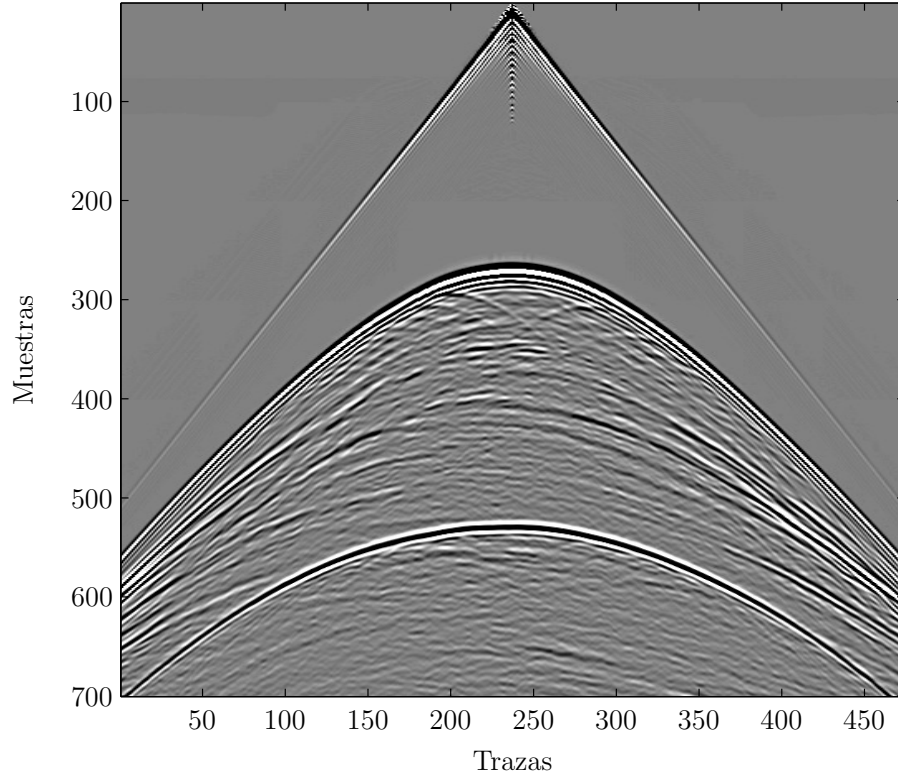


Figura 2: Sismograma sintético.

Reescribiendo (1.1) como una ecuación de onda se tiene que

$$\frac{1}{c(x, z)^2} \frac{\partial^2 p(x, z)}{\partial t^2} = \frac{\partial^2 p(x, z)}{\partial x^2} + \frac{\partial^2 p(x, z)}{\partial z^2} + src(x, z), \quad (1.2)$$

donde x y z representan las coordenadas espaciales, t la variable temporal, $c(x, z)$ el modelo de velocidades y $src(x, z)$ representa la fuente aplicada. En este caso la ecuación de onda es bidimensional, isotrópica y con densidad constante. El problema hacia delante se obtiene tras encontrar el campo de presión $p(x, z)$ en la ubicación de los receptores. Las condiciones de frontera utilizadas para solucionar el problema hacia delante son superficie libre en la parte superior del modelo y barreras absorbentes para todos los demás bordes. Por otra parte se tomaron como cero las primeras derivadas de $p(x, z)$ en las condiciones iniciales.

El problema inverso de identificación pretende estimar los parámetros del subsuelo ($c(x, z)$ que mide la velocidad de la onda sísmica) a partir de las observaciones (sismograma) conociendo $src(x, z)$ y $\mathbf{d}(x, t)$. En 1.2 se puede ver que la relación entre los parámetros ($c(x, z)$) y los datos \mathbf{d} es no lineal por el operador d'Alembert. En la Figura 3 se ilustra la definición sobre el problema inverso y directo.

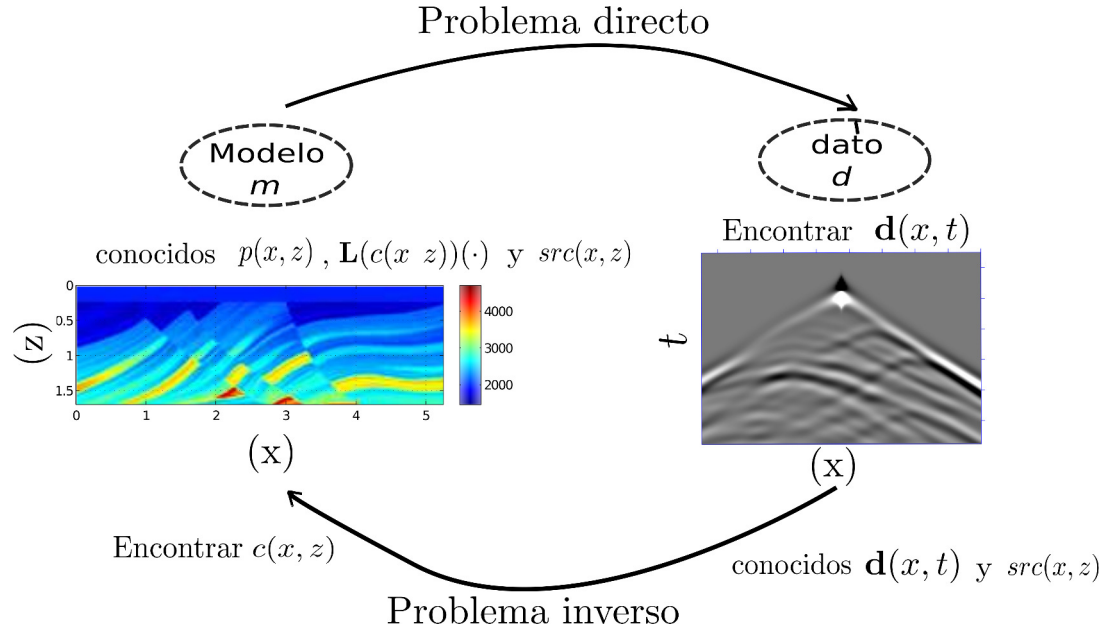


Figura 3: Definición del problema directo e inverso de identificación de parámetro.

Con la solución del problema inverso, se espera obtener unos parámetros (\mathbf{m}_{mod}) cercanos a \mathbf{m} , que sirvan como punto de partida para una técnica de optimización local. En la Figura 4 se presenta la solución propuesta al problema inverso. Se parte de unos parámetros \mathbf{m}_{sub} generados aleatoriamente y se aplica el operador $\mathbf{L}(\mathbf{m})(\cdot)$, conocida la fuente ($src(x, z)$). Posteriormente se aplica el operador $\mathbf{B}(\cdot)$ que discretiza el cubo del campo de presión $p(x, z, t)$, únicamente medido en una capa del cubo, lo que permite obtener los datos (d). Sobre los datos se aplica el operador $\mathbf{H}(\cdot)$ que transforma los datos (\mathbf{d}) al dominio de la traza compleja (TC). Finalmente el operador $\mathbf{C}(\cdot)$, mide las diferencias entre $\langle \mathbf{d}_{obs} \rangle$ y $\langle \mathbf{d}_{mod} \rangle$, obtenidos de \mathbf{m} y \mathbf{m}_{mod} respectivamente. El operador $\mathbf{C}(\cdot)$ planteado en la Figura 4 será la función de la metaheurística para determinar si los parámetros generados aleatoriamente están más cerca de la solución.

Se define el operador $L(\mathbf{m})(\cdot)$ como la matriz \mathbf{L} . Existen distintas técnicas exactas para solucionar los problemas inverso, sin embargo estas aplicaciones son limitadas por varias razones:

- La primera, muchas de estas técnicas son solo aplicables en situaciones ideales, por ejemplo la inversión de Herglotz-Wiechert donde plantea que la velocidad de la tierra depende únicamente de la profundidad y que la velocidad aumenta de forma monótona con la profundidad. La tomografía sísmica ha demostrado que estos dos principios no se cumplen en el manto de la tierra como le describe Nolet et al en [4].
- La segunda, se debe a que las técnicas de inversión son inestables desde la implementación. Esta inestabilidad se debe por el mal condicionamiento lo que genera valores

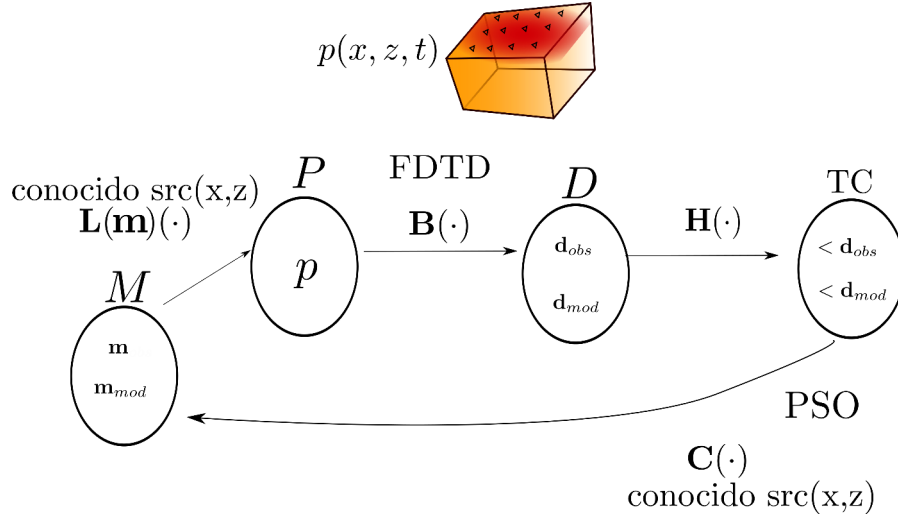


Figura 4: Definición del problema directo e inverso de identificación de parámetro en el contexto geofísico.

singulares de la matriz \mathbf{L} cercanos a cero. Esto se ha demostrado explícitamente por Darrent et al en [5].

- La tercera, que el número de observaciones difiere del número de parámetros del modelo. La matriz \mathbf{L} es usualmente una matriz no cuadrada y no existe su matriz inversa.

Esta última razón nos permite determinar que los datos no tienen suficiente información y por tanto no es posible estimar un modelo de forma única. El hecho de tener una cantidad finita de datos para reconstruir un modelo de infinitos grados de libertad, significa que el problema inverso es mal puesto y no tiene solución única en el sentido que varios modelos representan bien los datos [6].

Usualmente \mathbf{L} no es cuadrada y es necesario de un operador que permita estimar los parámetros del modelo, este operador se denota \mathbf{L}^{-g} que es llamada la inversa generalizada de la matriz \mathbf{L} y es calculada en (1.3) como

$$\mathbf{L}^{-g} = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T, \quad (1.3)$$

donde T es la matriz transpuesta y $()^{-1}$ la matriz inversa [6]. El modelo estimado ($\tilde{\mathbf{m}}$) es obtenido de

$$\mathbf{L}^g \tilde{\mathbf{m}} = \mathbf{d}. \quad (1.4)$$

Para tener un definición más cercana a la realidad se presenta

$$\mathbf{d} = \mathbf{Lm} + e, \quad (1.5)$$

donde e esta asociado al error generado por la adquisición de los datos. Reemplazando \mathbf{d} de (1.5) en (1.4) obtenemos que

$$\tilde{\mathbf{m}} = (\mathbf{L}^{-g}\mathbf{L})\mathbf{m} + \mathbf{L}^{-g}e, \quad (1.6)$$

y reagrupando términos tenemos que

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{L}^{-g}\mathbf{L} - I)\mathbf{m} + \mathbf{L}^{-g}e. \quad (1.7)$$

Cuando la aproximación de la matriz inversa \mathbf{L}^{-g} es la matriz \mathbf{L}^{-1} , el producto de $\mathbf{L}^{-g}\mathbf{L}$ es la matriz identidad logrando anular el segundo termino de (1.7). Esto muestra que aún teniendo \mathbf{L}^{-1} y en presencia de errores, los parámetros del modelo nunca serán iguales a los originales. En (1.7) podemos ver la no unicidad en la solución por lo ya mencionado. El diagrama de la Figura 5 ilustra lo presentado en (1.7).

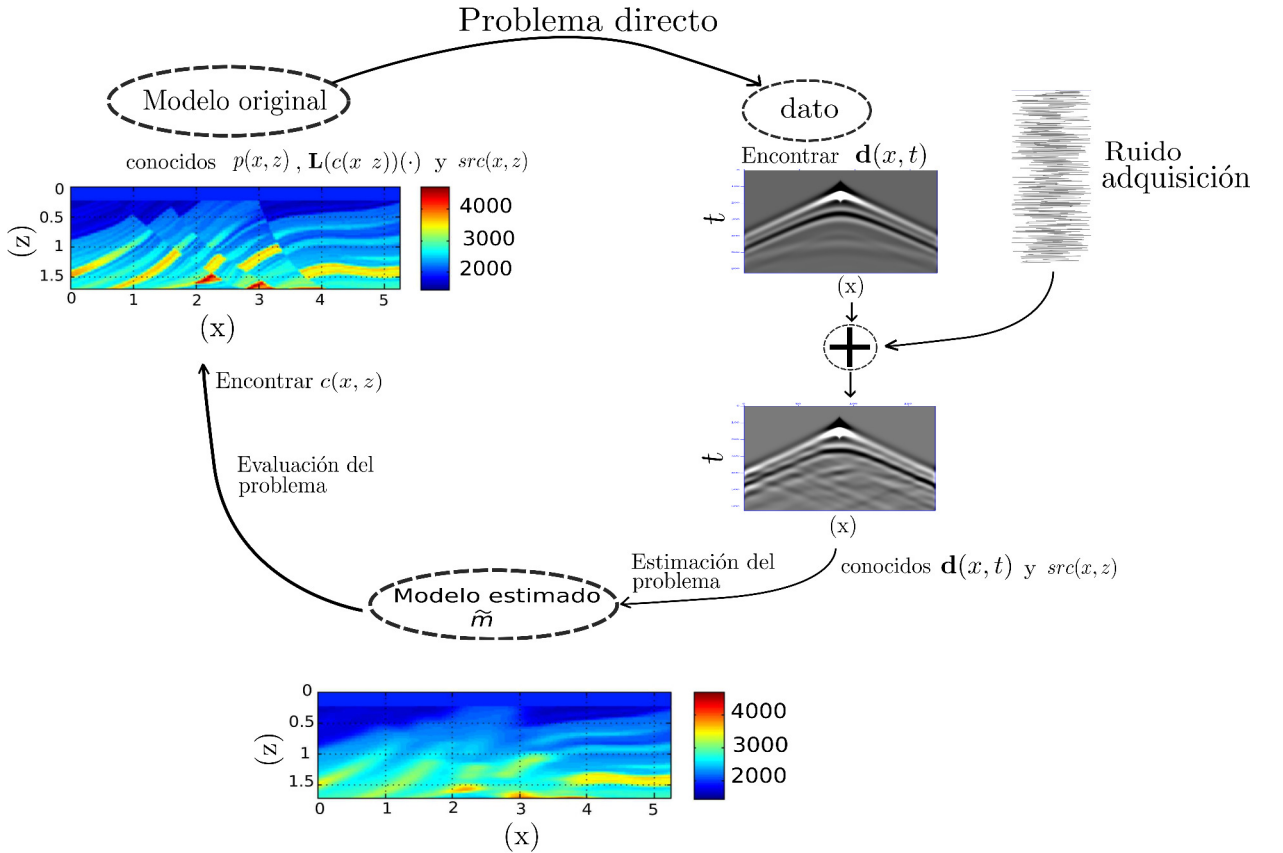


Figura 5: Definición más cercana a la realidad sobre el problema directo e inverso de identificación.

1.3. Full Waveform Inversion (FWI)

La FWI es una técnica de inversión que permite obtener parámetros del subsuelo como velocidades ($c(x, z)$). La FWI es un método de optimización local y lineal que minimiza el error cuadrático medio entre el dato observado (\mathbf{d}_{obs}) y el dato modelado $\mathbf{d}_{mod}(\mathbf{m})$ [7], como

$$\arg \min_{\mathbf{m}^k} \|\mathbf{d}_{mod}(\mathbf{m}^k) - \mathbf{d}_{obs}\|_2^2, \quad (1.8)$$

donde \mathbf{m}^k son los parámetros desconocidos en la k -ésima iteración. Como se describió previamente los parámetros en nuestro problema de inversión por identificación esta asociado a un modelo de velocidades del subsuelo. Este problema inverso se puede solucionar iterativamente seleccionando un punto partida \mathbf{m}^0 y actualizándolo aplicando el método de Newton penalizado [8]:

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \alpha_k \Delta \mathbf{m}^k, \quad (1.9)$$

donde α_k es el paso de avance y $\Delta \mathbf{m}^k$ en la k -ésima iteración esta dado por

$$\Delta \mathbf{m}^k = -[\mathbf{H}(\mathbf{m}^k)]^{-1} \mathbf{g}(\mathbf{m}^k), \quad (1.10)$$

siendo $[\mathbf{H}(\mathbf{m}^k)]^{-1}$ la inversa de la matriz Hessiana de la función de costo evaluada en el punto \mathbf{m}^k y $\mathbf{g}(\mathbf{m}^k)$ es el gradiente de la función de costo presentada en (1.8) y evaluada en el punto \mathbf{m}^k . En la primera iteración, la inversa de la matriz Hessiana es la matriz identidad y para las siguientes iteraciones se estima el producto de $[\mathbf{H}(\mathbf{m}^k)]^{-1} \mathbf{g}(\mathbf{m}^k)$ mediante el algoritmo de L-BFGS [4]. El cálculo del gradiente es presentado por Plessix en [9] Es importante resaltar que la FWI presenta dos grandes inconvenientes, su alto costo computacional y alta dependencia del punto de partida \mathbf{m}^0 . En la Figura 6, se presenta una función objetivo con múltiples mínimos locales y dos modelos iniciales. El modelo inicial 1 se ajusta al mínimo local mientras que el modelo inicial 2 se ajustará al mínimo global, esto deja ver la dependencia de su punto de partida al ser un método de optimización local.

Para mostrar la dependencia del punto de partida para la FWI, tomaremos como ejemplo el modelo Marmousi presentado en la Figura 7 a). Se aplican dos modelos de velocidades como puntos de partida para mostrar la importancia de obtener un punto de partida \mathbf{m}^0 cerca del vecindario del mínimo global. Los resultados después de aplicar FWI sobre estos dos puntos iniciales son presentados en la Figura 7.

Para poder obtener un \mathbf{m}^0 adecuado es decir cerca del vecindario del mínimo global, se deberá tener la suficiente información de baja frecuencia, es decir, no caer en problemas de *cycle skipping* (CS). El CS ocurre cuando existe un desfase mayor que la mitad del periodo de la fuente [11]. Por ejemplo en la Figura 8 a) se presenta un desfase mayor a $T/2$ con respecto a la Figura 8 b) es decir con CS. El lóbulo n de la Figura 8 (a) se ajusta al lóbulo $n-1$ de la Figura 8 b). Por otra parte la Figura 8 (c) presenta un desfase menor a $T/2$ con respecto a la Figura 8 (b) es decir sin CS. El lóbulo n de la Figura 8 c) se ajusta al lóbulo n de la 8 b).

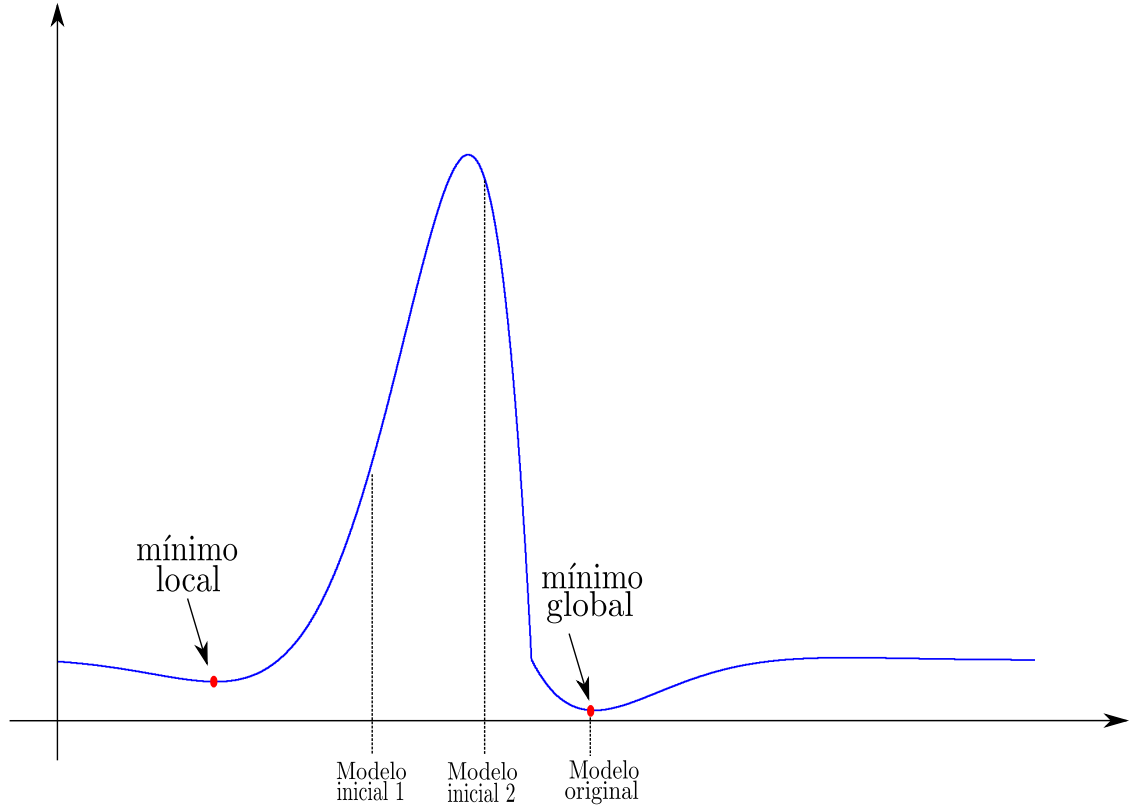


Figura 6: Problema del punto de partida en la FWI, adaptado de [10].

1.4. Métodos de optimización global

Los métodos de optimización global se clasifican en: determinísticos y probabilísticos. Estos métodos requieren un elevado costo computacional y para mitigar su efecto se aborda desde una implementación en paralelo mediante GPU (*Graphics Processor Unit*). Los métodos de optimización global son estrategias numéricas usando conceptos heurísticos, con el fin de intensificar sus búsquedas en zonas promisorias del mínimo global y así logran escapar de mínimos locales. En el contexto geofísico no podemos aplicar un método determinístico, dado que no conocemos la expresión matemática de la función objetivo. Además de no conocer si esta expresión es dos veces derivable para determinar si el punto crítico es un máximo o un mínimo. aplicar un método de optimización global que permita explorar la función de costo y así obtener un punto de partida (modelo de velocidades) cercano del vecindario del mínimo global necesario para que la FWI ajuste la solución.

PSO consta de un grupo de partículas que exploran un espacio solución de un problema. Las partículas se mueven de forma aleatoria y cada una de ellas recuerda la mejor posición local y global del enjambre. Del conjunto de mejores posiciones locales se calcula el menor valor sobre la función objetivo, siendo está la mejor posición global

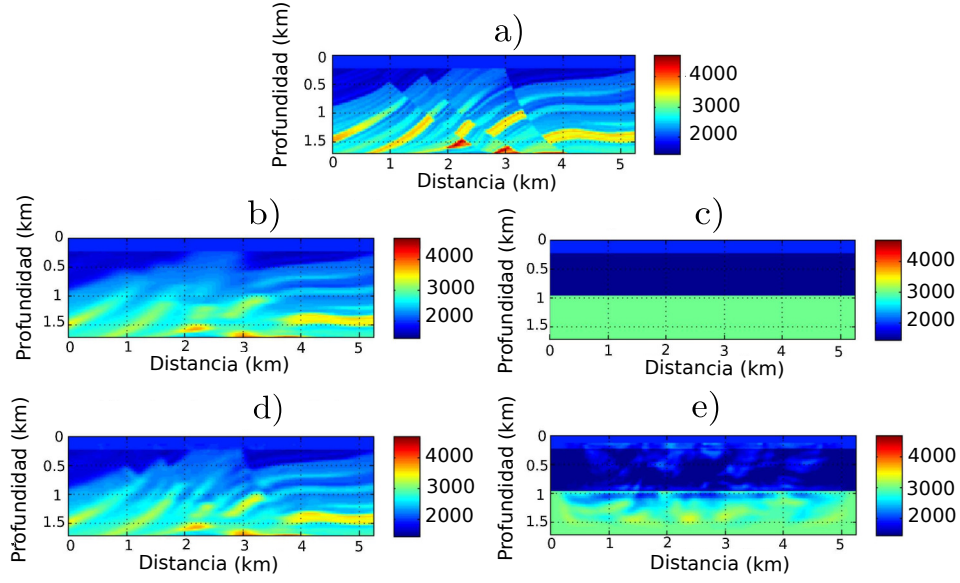


Figura 7: a) Modelo original Marmousi; b) Modelo inicial obtenido de suavizar 20 veces el modelo original; c) Modelo inicial de dos capas 1500-3000 m/s; d) Modelo obtenido después de aplicar FWI sobre el modelo b; e) Modelo obtenido después de aplicar FWI sobre el modelo c.

[12]. La ecuación de PSO en un espacio j -dimensional es

$$\mathbf{x}_{i,j}^{k+1} = \mathbf{x}_{i,j}^k + \mathbf{v}_{i,j}^{k+1} \Delta t, \quad (1.11)$$

donde $\mathbf{x}_{i,j}^{k+1}$ es el nuevo vector posición asociado a un modelo de velocidades del subsuelo de la i -ésima partícula, $\mathbf{x}_{i,j}^k$ el actual vector posición de la i -ésima partícula en la k -ésima iteración, $\mathbf{v}_{i,j}^{k+1}$ es el nuevo vector velocidad y Δt es el paso de avance dentro de la zona de búsqueda. Las velocidades del subsuelo en la primera iteración para todas las partículas son generadas en base a una distribución normal entre el límite inferior de velocidad del modelo V_{min} y el límite superior de velocidad del modelo V_{max} . El vector velocidad se obtiene de

$$\mathbf{v}_{i,j}^{k+1} = (w \mathbf{v}_{i,j}^k) + c_1 \text{ rand}_1 \left(\frac{(\mathbf{p}_i^k - \mathbf{x}_{i,j}^k)}{\Delta t} \right) + c_2 \text{ rand}_2 \left(\frac{(\mathbf{g}_j^k - \mathbf{x}_{i,j}^k)}{\Delta t} \right), \quad (1.12)$$

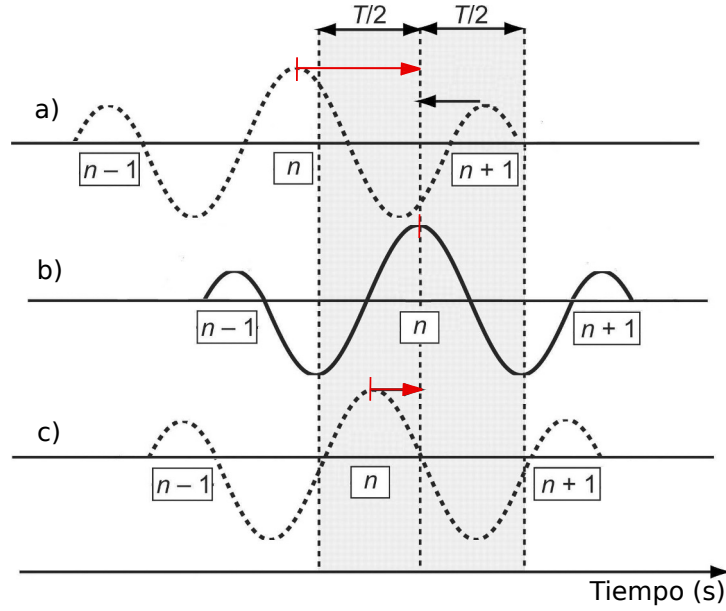
donde w es el peso inercial, encargada de controlar el movimiento de las partículas; c_1 y c_2 son las constantes asociadas con el comportamiento local y global del enjambre respectivamente, $\text{rand}()$ es una función que multiplica su argumento por un número aleatorio entre $[0,1]$ con distribución normal, \mathbf{x}_i^k es la mejor posición local y \mathbf{g}_j^k es la mejor posición global lograda por el enjambre. En el pseudocódigo 1 se presenta la adaptación de PSO. Para la selección de los parámetros w , Δt , c_1 y c_2 se hizo una revisión en el estado-del-arte con el fin de llegar a un “consenso científico” sobre

Algoritmo 1 Pseudocódigo de la adaptación PSO.

```

1: Crear e inicializar las partículas en el espacio  $j$ -dimensional
2:  $w, c_1, c_2$  ▷ inicializar las constates
3:  $V_{min}$  ▷ limite inferior de velocidad del modelo
4:  $V_{max}$  ▷ limite superior de velocidad del modelo
5:  $k = 1$  ▷ primera iteración
6:  $\mathbf{x}_{i,j}^k$  ▷ valores aleatorios según una distribución normal entre  $V_{min}$  y  $V_{max}$ 
7:  $\mathbf{v}_{i,j}^k$  ▷ inicializadas en cero.
8:  $\mathbf{g}_j^k = \mathbf{x}_{1,j}^k$ 
9: for  $i = 1$ :Partícula do
10:   for  $j = 1$ :Dimensiones do
11:      $\mathbf{p}_i^k = \mathbf{x}_{i,j}^k$ 
12:   end for
13:   if  $m(\mathbf{x}_{i,j}^k) < m(\mathbf{g}_j^k)$  then ▷ donde  $m$  es el funcional que mide la diferencia
     entre el dato observado y modelado.
14:      $\mathbf{g}_j^k = \mathbf{x}_{i,j}^k$ 
15:   end if
16: end for
17: for  $k = 1$ :Iteraciones do
18:   for  $i = 1$ :Partícula do
19:     for  $j = 1$ :Dimensiones do
20:        $\mathbf{v}_{i,j}^{k+1} = (w \mathbf{v}_{i,j}^k) + c_1 \text{ rand}_1 \left( \frac{(\mathbf{p}_i^k - \mathbf{x}_{i,j}^k)}{\Delta t} \right) + c_2 \text{ rand}_2 \left( \frac{(\mathbf{g}_j^k - \mathbf{x}_{i,j}^k)}{\Delta t} \right)$ 
21:        $\mathbf{x}_{i,j}^{k+1} = \mathbf{x}_{i,j}^k + \mathbf{v}_{i,j}^{k+1} \Delta t$ 
22:       if  $\mathbf{x}_{i,j}^{k+1} < V_{min}$  then
23:          $\mathbf{x}_{i,j}^{k+1} = V_{min}$ 
24:       end if
25:       if  $\mathbf{x}_{i,j}^{k+1} > V_{max}$  then
26:          $\mathbf{x}_{i,j}^{k+1} = V_{max}$ 
27:       end if
28:     end for
29:     if  $m(\mathbf{x}_{i,j}^{k+1}) < m(\mathbf{p}_i^k)$  para minimización then
30:        $\mathbf{p}_i^k = \mathbf{x}_{i,j}^{k+1}$ 
31:       if  $m(\mathbf{x}_{i,j}^{k+1}) < m(\mathbf{g}_j^k)$  then
32:          $\mathbf{g}_j^k = \mathbf{x}_{i,j}^{k+1}$ 
33:       end if
34:     end if
35:   end for
36: end for

```


 Figura 8: Concepto del *cycle skipping*.

esta selección en términos de convergencia de PSO. En la actualidad se han realizado distintos trabajos la tasa de convergencia de PSO. Shi y Eberhart en [13] asociaron el peso de inercia w como se describió en (1.12). Ellos investigaron sobre el efecto de w en un rango entre $[0,1.2]$. Los resultados muestran que valores de w entre $[0.8,1.2]$ presentan una convergencia lenta, mientras que los valores de w entre $[0.1, 0.4]$ presentan una convergencia rápida.

En [14] los trabajos de Clerc y [15] proponen un parámetro de restricción el cual ayuda a elegir las constantes w , c_1 y c_2 de tal manera que garantice la convergencia absoluta del algoritmo. Clerc en [14] propone trabajar con $\Delta t = 1$. Reescribiendo (1.12) tenemos

$$\mathbf{v}_{i,j}^k = K(\mathbf{v}_{i,j}^k + c_1 \text{rand}_1(\mathbf{p}_i^k - \mathbf{x}_{i,j}^k) + c_2 \text{rand}_2((\mathbf{g}_j^k - \mathbf{x}_{i,j}^k))), \quad (1.13)$$

donde K es el parámetro de restricción definido como

$$K = \frac{2}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|}, \quad (1.14)$$

en el cual $\varphi = c_1 + c_2$ y Clerc propone que $\varphi > 4$. Tomando $c_1 = c_2 = 2,05$ entonces $\varphi = 4,1$ y $K = 0,7298$.

Holland, en [16] discute sobre dos términos importantes en la selección de los parámetros w , c_1 y c_2 . Un algoritmo debe tener un balance entre la exploración y la explotación. La exploración es la tendencia del algoritmo de examinar nuevas regiones de búsqueda, mientras que la explotación es la tendencia a buscar el mínimo más cercano a ese punto inicial. Como se describió antes con w pequeño el algoritmo tiene una convergencia

rápida, sin embargo sería un algoritmo de poca exploración. Por otra parte con c_1 menor que c_2 , pone un punto de inercia muy alto tal que todas las partículas se vean atraídas y solo explorando en la región cercana al mejor global. Se debe buscar un equilibrio entre estos dos términos para el adecuado funcionamiento del algoritmo según su contexto.

Analizando el problema en el contexto geofísico, cada dimensión en el espacio de búsqueda de las partículas esta asociada a una velocidad desconocida del modelo del subsuelo $\mathbf{x}_{i,j}^k$. Cada partícula que recorre las j -ésima-dimensiones y tendrá asociado un modelo de velocidades del subsuelo. A partir del estado-del-arte se tomaron las constantes $\Delta t = 1$, $w = 0,729$, $c_1 = 2,8$, $c_2 = 3,1$, 300 partículas y 100 iteraciones propuestas por [12]. Se emplea un modelo de 18 dimensiones para ajustar los parámetros w , c_1 y c_2 , como se presenta en la Figura 9.

v_1	v_2	v_3	v_4	v_5	v_6
v_7	v_8	v_9	v_{10}	v_{11}	v_{12}
v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}

Figura 9: Modelo empleado para ajustar w , c_1 y c_2 .

donde $v_1 = v_2 = v_5 = v_6 = 1500$, $v_3 = v_4 = 2500$, $v_7 = v_8 = v_9 = v_{10} = v_{11} = v_{12} = 2000$ y $v_{13} = v_{14} = v_{15} = v_{16} = v_{17} = v_{18} = 2800$. Se propone ajustar inicialmente w . Se realizan 10 experimentos tomando los 5 mejores resultados con w de 1,2 (ver Tabla 1), 1 (ver Tabla 2) 0,729 (ver Tabla 3), 0,7 (ver Tabla 4) y 0.2 (ver Tabla 5) manteniendo constante c_1 y c_2 en 2,8 y 3,1 respectivamente. De estas Tablas podemos tomar $w = 0,7$.

$w=1,2$; $c_1=2,8$; $c_2=3,1$	m
Exp 1	0,6488
Exp 2	0,6867
Exp 3	0,6192
Exp 4	0,7280
Exp 5	0,7398
Promedio	0,6845

Tabla 1: parámetros $w=1,2$; $c_1=2,8$; $c_2=3,1$.

Seguidamente se sintoniza la constante c_2 , tomando valores de 2,1 (ver Tabla 6) y 1,1 (ver Tabla 7). Seleccionando un valor de $c_2 = 1,1$.

$w=1;$ $c_1=2,8;$ $c_2=3,1$	m
Exp 1	0,7039
Exp 2	0,7042
Exp 3	0,6814
Exp 4	0,7219
Exp 5	0,6730
Promedio	0,6969

Tabla 2: parámetros $w=1$; $c_1=2,8$; $c_2=3,1$.

$w=0,729;$ $c_1=2,8;$ $c_2=3,1$	m
Exp 1	0,6538
Exp 2	0,7378
Exp 3	0,7341
Exp 4	0,7825
Exp 5	0,7689
Promedio	0,7354

Tabla 3: parámetros $w=0,729$; $c_1=2,8$; $c_2=3,1$.

$w=0,7;$ $c_1=2,8;$ $c_2=3,1$	m
Exp 1	0,7365
Exp 2	0,7693
Exp 3	0,8043
Exp 4	0,7968
Exp 5	0,6943
Promedio	0,7602

Tabla 4: parámetros $w=0,7$; $c_1=2,8$; $c_2=3,1$.

Finalmente se varía c_1 en valores de 1,8 (ver Tabla 8), 1 (ver Tabla 9) y 0,8 (ver Tabla 10). Seleccionando $c_1 = 0,8$. En la sección de anexos se presentan todas las velocidades obtenidas por experimento para esta selección (ver Tablas 18 a 21).

Inicialmente se variaron las dimensiones con el fin de reducir el tiempo computacional. En la Figura 10 se presenta una malla de 11×3 dimensiones para el modelo marmousi de 211×68 dimensiones. El costo computacional con estas dimensiones alcanzaba las 72 horas en una implementación en lenguaje C en un procesador Intel(R) Xeon(R) CPU

$w=0,2;$ $c_1=2,8;$ $c_2=3,1$	m
Exp 1	0,6297
Exp 2	0,6730
Exp 3	0,6041
Exp 4	0,6643
Exp 5	0,6122
Promedio	0,6367

Tabla 5: parámetros $w=0,2$; $c_1=2,8$; $c_2=3,1$.

$w=0,7;$ $c_1=2,8;$ $c_2=2,1$	m
Exp 1	0,7845
Exp 2	0,8545
Exp 3	0,7807
Exp 4	0,7386
Exp 5	0,7418
Promedio	0,7800

Tabla 6: parámetros $w=0,7$; $c_1=2,8$; $c_2=2,1$.

$w=0,7;$ $c_1=2,8;$ $c_2=1,1$	m
Exp 1	0,7741
Exp 2	0,7710
Exp 3	0,9950
Exp 4	0,9315
Exp 5	0,7559
Promedio	0,8455

Tabla 7: parámetros $w=0,7$; $c_1=2,8$; $c_2=1,1$.

E5-2620 v3 @ 2.40GHz.

En PSO se acepta un nuevo modelo cuando una métrica obtenida entre \mathbf{d}_{obs} y \mathbf{d}_{mod} indique que este es un mejor modelo. La métrica puede estar asociada a una maximización por ejemplo en el caso de una correlación (entre \mathbf{d}_{obs} y \mathbf{d}_{mod}) o una minimización por ejemplo en caso del error cuadrático medio (entre $||\mathbf{d}_{obs}$ y $\mathbf{d}_{mod}||^2$).

Lo que se busca con el uso del método de optimización global no es obtener la solución exacta del modelo original, si no tener una solución suavizada o con la suficiente infor-

$w=0,7;$ $c_1=1,8;$ $c_2=1,1$	m
Exp 1	0,6989
Exp 2	0,8080
Exp 3	0,9281
Exp 4	0,9036
Exp 5	0,7314
Promedio	0,8140

Tabla 8: parámetros $w=0,7$; $c_1=1,8$; $c_2=1,1$.

$w=0.7;$ $c_1=1;$ $c_2=1.1$	m
Exp 1	0,8765
Exp 2	0,9885
Exp 3	0,9931
Exp 4	0,9070
Exp 5	0,8810
Promedio	0,9292

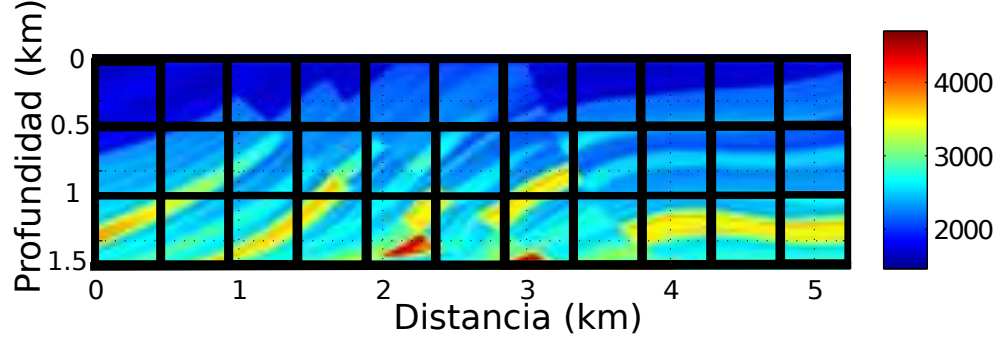
Tabla 9: parámetros $w=0,7$; $c_1=1$; $c_2=1,1$.

$w=0,7;$ $c_1=0,8;$ $c_2=1,1$	m
Exp 1	0,9995
Exp 2	0,9919
Exp 3	0,9992
Exp 4	0,9999
Exp 5	0,9663
Promedio	0,9914

Tabla 10: parámetros $w=0,7$; $c_1=0,8$; $c_2=1,1$.

mación de baja frecuencia. Se espera que mediante PSO se tenga un punto de partida adecuado en términos del porcentaje de CS (*cycle skipping*) presente en los datos. Por otra parte, la selección de PSO se debe en su bioinspiración hacia un enjambre de abejas, lo que permite obtener mas muestras de la función objetivo, logrando así ser un algoritmo de alta exploración y a su vez permitir escapar de mínimos locales.

Se trabajó con 500 partículas y 100 iteraciones sobre todos los experimentos. En el Capítulo 3 se dará la justificación de esta selección. Sin embargo al aumentar el núme-


 Figura 10: Modelo marmousi de dimensionalidad 11×3

ro de partículas a 500 conlleva un elevado costo computacional debido que sobre cada partícula se tiene asociado un modelo de velocidades sobre el cual se obtiene un sismograma y luego se estima una métrica. Para mitigar este efecto se plantea programar en paralelo tanto de la obtención del sismograma como la dimensionalidad utilizando unidades de procesamiento gráfico (GPU, por sus siglas en ingles).

Las industrias petroleras hoy en día sólo hacen tomografías para estimar los puntos de partida. Sin embargo, las tomografías requieren de expertos geofísicos y no existe un consenso de la solución. En este trabajo, por el contrario, se plantea usar un método cuantitativo basado en una metodología para dar respuesta del punto de partida en el problema de inversión.

1.5. Función de costo

Como se describió anteriormente cada i -ésima partícula en la k -ésima iteración, $\mathbf{x}_{i,j}^k$, de PSO representa un modelo de velocidades $\mathbf{c}(\mathbf{x}, \mathbf{z})$. El tamaño de la dimensionalidad j será $N_x \times N_z$, donde N_x representa el número de puntos en la dirección x y N_z el número de puntos en la dirección z . La evaluación de la función de costo $\mathbf{m}(\mathbf{x}_{i,j}^k)$ representa un operador matemático aplicado sobre el modelo de velocidades $F\{\mathbf{x}_{i,j}^k\}$. Este operador es definido como

$$F\{\mathbf{x}_{i,j}^k\} = \mathbf{m}(T\{R\{L(\mathbf{c}_{i,j}^k)[P(x, z)]\}, T\{\mathbf{d}_{obs}\}) \quad (1.15)$$

donde

- \mathbf{m} es el funcional que mide las diferencias entre el dato observado y modelado.
- $P(x, z)$ es el campo de presión en (x, z) ;

- $L(\mathbf{c}_{i,j}^k)(\cdot) = \nabla^2[\cdot] - \frac{1}{\mathbf{c}^2(x,z)} \frac{\partial^2[\cdot]}{\partial t^2}$ es el operador de d'Alembert;
- R es el operador de diferencias finitas en el dominio del tiempo;
- T es el operador del espacio transformado;

2. SELECCIÓN DE LOS DOMINIOS TRANSFORMADOS USADOS EN LAS MÉTRICAS

En el presente capítulo se exponen los criterios considerados para escoger los dominios transformados. Los dominios transformados son bases en donde podemos representar la información de otras formas, permitiendo obtener atributos de los datos \mathbf{d}_{obs} y \mathbf{d}_{mod} . Se establecieron como dominios transformados el tiempo, la frecuencia y la traza compleja. Para todos los dominios transformados se empleó el concepto de *cycle skipping* (CS). En el dominio del tiempo se empleó la métrica de correlación mientras que para el dominio de la frecuencia y el dominio de Hilbert se utilizó el CS. Todas las pruebas se llevaron a cabo utilizando la técnica de optimización global PSO tomando 5 disparos. Además en este capítulo se busca dar claridad sobre el operador \mathbf{m} presentado en el algoritmo 1, mostrando sus beneficios al ser aplicado en los distintos dominios transformados.

2.1. Dominio del tiempo

El primer dominio explorado es el tiempo donde se usa la métrica de correlación entre \mathbf{d}_{obs} y \mathbf{d}_{mod} . Antes de aplicar esta métrica, es necesario aplicar una corrección de amplitud ya que por naturaleza el frente de onda, este decae a razón de $1/R$, siendo R el radio de avance del frente de onda. La correlación de amplitud busca que la onda sísmica directa sea comparable con respecto a las reflexiones en la mayoría de las trazas. Esta corrección es aplicada tanto en el dato observado como para el dato modelado. El factor de corrección aplicado es

$$\mathbf{d}_{obs}^{s,r}(t) = \mathbf{d}_{obs}^{s,r}(t) \times q^2(t), \quad (2.1)$$

donde s es el número de fuentes, r es el número de receptores, t es tiempo y $q^2(t)$ es una función que corrige el decaimiento exponencial. En la Figura 11 se presenta un ejemplo de una traza sísmica con corrección de amplitud (azul) y sin corrección de amplitud (roja) obtenidas sobre un modelo de 2 capas con velocidades de 2000 [m/s] y 3000 [m/s], siendo esta última velocidad la de mayor profundidad. Se puede observar que la onda directa presenta gran amplitud con respecto a la onda reflejada. No corregir esta amplitud podría indicar que hemos llegado a la solución del modelo aún sin contemplar que hay un cambio de impedancia entre las capas.

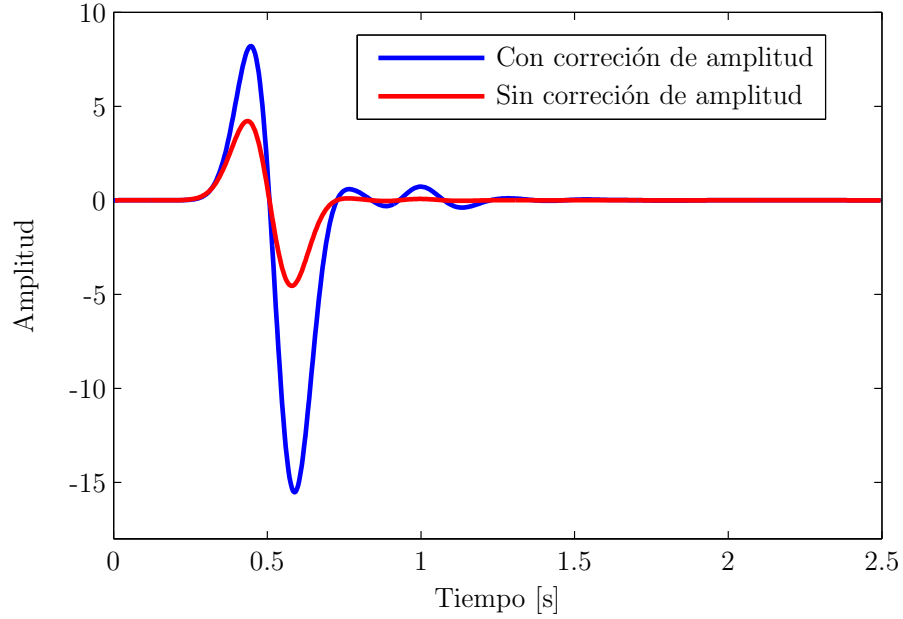


Figura 11: Traza sísmica con corrección de amplitud (azul) y sin corrección de amplitud (roja).

La métrica de correlación actúa en cada traza sísmica sobre el dato observado y el dato modelado, estimando un vector de coeficientes de correlación definido por

$$C_{corr}^{s,r} = \frac{\sum_t (\mathbf{d}_{obs}^{s,r}(t) - \hat{\mathbf{d}}_{obs}^{s,r})(\mathbf{d}_{mod}^{s,r}(t) - \hat{\mathbf{d}}_{mod}^{s,r})}{\sqrt{(\sum_t (\mathbf{d}_{obs}^{s,r}(t) - \hat{\mathbf{d}}_{obs}^{s,r})^2)(\sum_t (\mathbf{d}_{mod}^{s,r}(t) - \hat{\mathbf{d}}_{mod}^{s,r})^2)}} + 1, \quad (2.2)$$

donde $\hat{\mathbf{d}}_{obs}^{s,r}$ o $\hat{\mathbf{d}}_{mod}^{s,r}$ será el promedio en la dirección de las muestras; el promedio de todos los coeficientes de correlación en la dirección de los receptores \hat{C}_{corr}^s produce un coeficiente de correlación por cada fuente sísmica [17]; Finalmente, se promedian los coeficientes de corrección en la dirección de la fuente \hat{C}_{corr} produciendo un coeficiente de correlación por modelo de velocidad. Un nuevo modelo de velocidades es aceptado cuando el coeficiente de correlación es mayor que el anterior.

La métrica de correlación en el dominio del tiempo es de las métricas más utilizadas en el estado-del-arte hasta el momento en las técnicas de optimización global [17]. La correlación nos indica qué tanto se asemejan dos señales en fase. Para entender cómo funciona esta métrica se plantea el siguiente ejemplo: En la Figura 12 a) tenemos en rojo la señal $\sin(t)$ y en azul la señal $0,5 \cdot \sin(t)$ obteniendo un coeficiente de correlación de 1; por otra parte en la Figura 12 b) tenemos en rojo la señal $\sin(t)$ y en azul $0,5 \cdot \sin(t + \pi/4)$ obteniendo un coeficiente de correlación de 0.7071. Note que la amplitud siempre fue diferente para las dos señales (roja y azul), sin embargo cuando ellas se encontraban en fase obtuvimos una correlación alta mientras que con un desfase obtenemos un

coeficiente de correlación menor. Esto muestra que la correlación indica que tanto se parecen $\mathbf{d}_{obs}^{s,r}$ y $\mathbf{d}_{mod}^{s,r}$ en fase más no en amplitud.

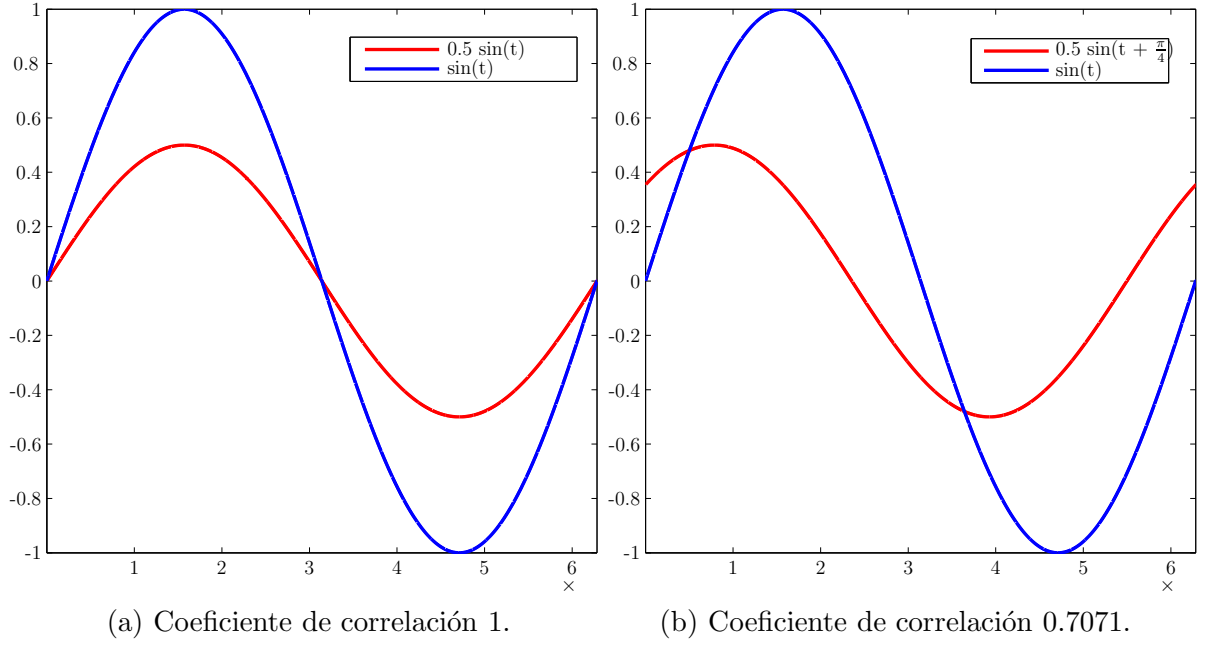


Figura 12: Ejemplos del coeficiente de correlación

2.2. Dominio de la frecuencia

La transformada de Fourier (FT) es un caso particular de la transformada de Laplace, donde se reemplaza s por $j\omega$. Mediante la transformada de Fourier se puede representar una señal $u(t)$ mediante la sumatoria de funciones exponenciales complejas de la forma $e^{j\omega t}$, permitiendo ir del dominio temporal al dominio de la frecuencia. La transformada de Fourier está definida como

$$\mathbf{U}(\omega) = \int_{-\infty}^{\infty} u(t)e^{-j\omega t} dt, \quad (2.3)$$

donde $u(t)$ debe ser derivable e integrable (en el sentido de Lebergue), ω es la frecuencia angular. En nuestro caso $u(t)$ es el sismograma $\mathbf{d}_{obs}^{s,r}$ o $\mathbf{d}_{mod}^{s,r}$. Como la FT es una representación mediante funciones de exponenciales complejas, tendrá una magnitud

$$|\mathbf{U}(\omega)| = \sqrt{\text{Re}\{\mathbf{U}(j\omega)\}^2 + \text{Im}\{\mathbf{U}(j\omega)\}^2}, \quad (2.4)$$

y una fase definida por

$$\theta(\omega) = \arctan\left(\frac{\text{Im}\{\mathbf{U}(j\omega)\}}{\text{Re}\{\mathbf{U}(j\omega)\}}\right). \quad (2.5)$$

Se selecciona el dominio de Fourier debido que hay eventos sísmicos que poseen características especiales de frecuencia que no es posible caracterizar sólo con el dominio del tiempo no es posible caracterizar. La reflexión sísmica, la difracción sísmica, el ruido coherente y el ruido incoherente son algunos ejemplos de eventos sísmicos que pueden ser diferenciados mediante el dominio de Fourier para su posterior procesado.

De esta manera, los datos observados y modelados son transformados al dominio de la frecuencia mediante el operador FT . Los datos en el dominio de la frecuencia son un número complejo donde sólo nos interesa estudiar su fase para el análisis del CS. $\theta_{obs}^{s,r}(\omega)$ será la fase obtenida tras aplicar FT sobre $\mathbf{d}_{obs}^{s,r}$ y $\theta_{mod}^{s,r}(\omega)$ será la fase obtenida tras aplicar FT sobre $\mathbf{d}_{mod}^{s,r}$. El coeficiente de CS en el dominio de Fourier, $CS_{FT}^{s,r}$ cuenta las ocurrencias donde $\theta_{obs}^{s,r}(\omega)$ y $\theta_{mod}^{s,r}(\omega)$ es mayor que π . El algoritmo 2 y la Figura 14 muestran cómo calcular el $CS_{FT}^{s,r}$. Se tomaron 30 componentes frecuenciales por traza con un $\Delta f = 0,2441$ Hz, sólo tomando las frecuencias menores a 10 Hz. Existen dos razones por las cuales solo se analiza las frecuencias menores a 10 Hz: la primera de ellas es por la frecuencia central de la fuente, en todos los experimentos se toma una fuente con frecuencia central 3 Hz y su espectro se presenta en la Figura 13, donde vemos que el ancho de banda es alrededor de 8 Hz; la segunda razón es debido que el CS solo ocurre para frecuencias bajas menores a 5 Hz [18], además es muy común en una adquisición real no tener suficiente información de baja frecuencia ya que por la propia electrónica de los instrumentos las frecuencias menores a 5 Hz son consideradas ruido. Una vez obtenido $CS_{FT}^{s,r}$ se calcula la norma l^2 en la dirección de los receptores \hat{C}_{FT}^s produciendo un coeficiente de FT por fuente; Finalmente, se promedia los coeficientes de FT en la dirección de la fuente \hat{C}_{FT} produciendo un coeficiente de FT por modelo de velocidad. Un nuevo modelo de velocidades es aceptado cuando el coeficiente de FT es menor que el anterior.

Algoritmo 2 CS en el dominio de la frecuencia

1: $\theta_{obs}^{s,r}(\omega)$	▷ Fase obtenida de la señal analítica $\mathbf{d}_{obs}^{s,r}(t)$
2: $\theta_{mod}^{s,r}(\omega)$	▷ Fase obtenida de la señal analítica $\mathbf{d}_{mod}^{s,r}(t)$
3: for $s = 1, \dots, S$ do	▷ S, número de fuentes
4: for $r = 1, \dots, R$ do	▷ R, número de receptores
5: $Counter = 0$	▷ reiniciar contador de CS
6: for $\omega = 0, \dots, f_c$ do	▷ f_c , es la frecuencia corte (10 Hz)
7: $diff = \theta_{mod}^{s,r}(\omega) - \theta_{obs}^{s,r}(\omega) $	
8: if $diff > \pi$ then	
9: $Counter = Counter + 1$	▷ Contador CS
10: end if	
11: end for	
12: $CS_{FT}^{s,r} = Counter$	▷ Salvar contador CS
13: end for	
14: end for	

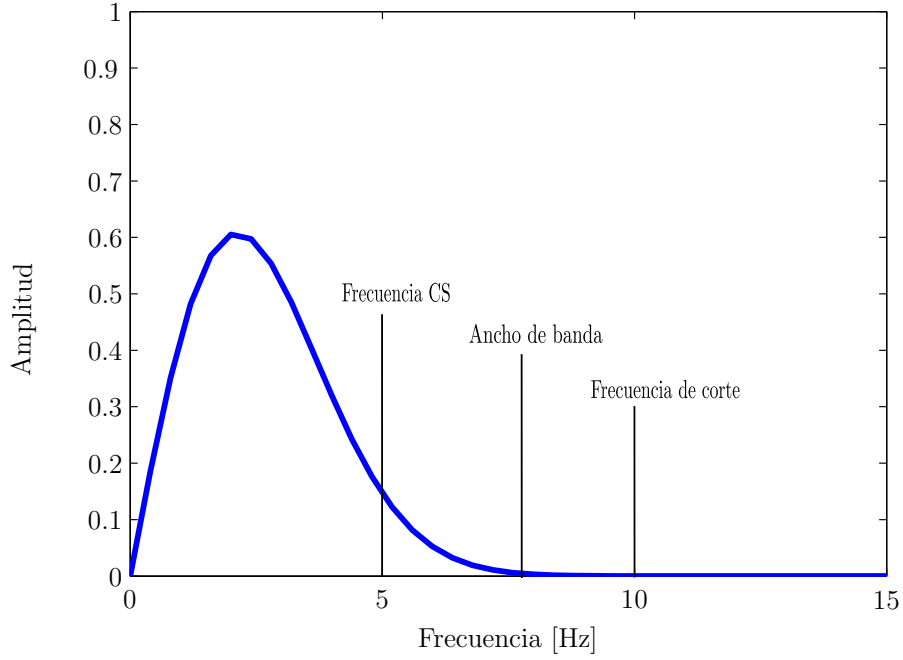


Figura 13: Espectro de la fuente.

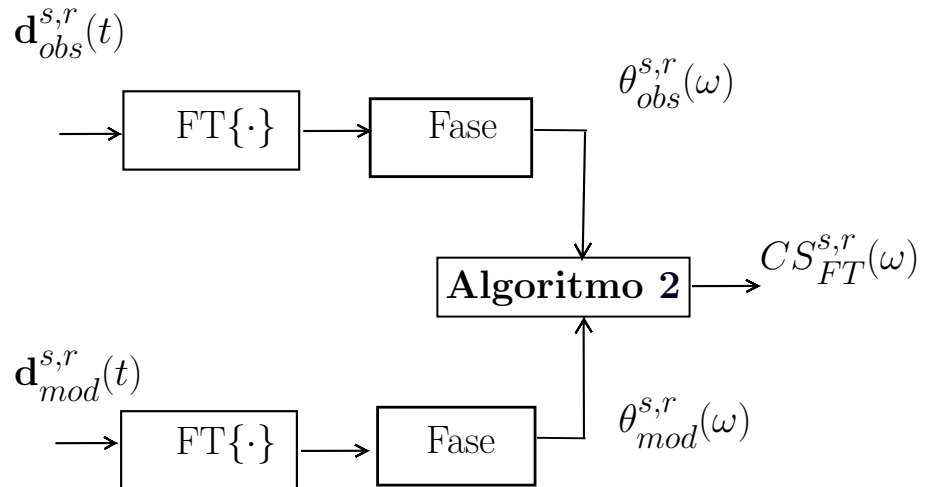


Figura 14: Método usado para computar el CS en el dominio de la frecuencia.

2.3. Dominio de Hilbert

En la actualidad, la transformación de la traza compleja es muy utilizada ya que permite identificar eventos sísmicos que no pueden observarse con facilidad en el dominio del tiempo. En la Figura 15 a) se presenta un registro sísmico en tiempo donde existe una fractura en el terreno alrededor de 2 [seg] y en la traza 80. No es fácil identificar esta fractura y por tanto se propone aplicar la transformación de la traza compleja para

estimar una magnitud (envolvente). Este resultado es presentado en la Figura 15 b).

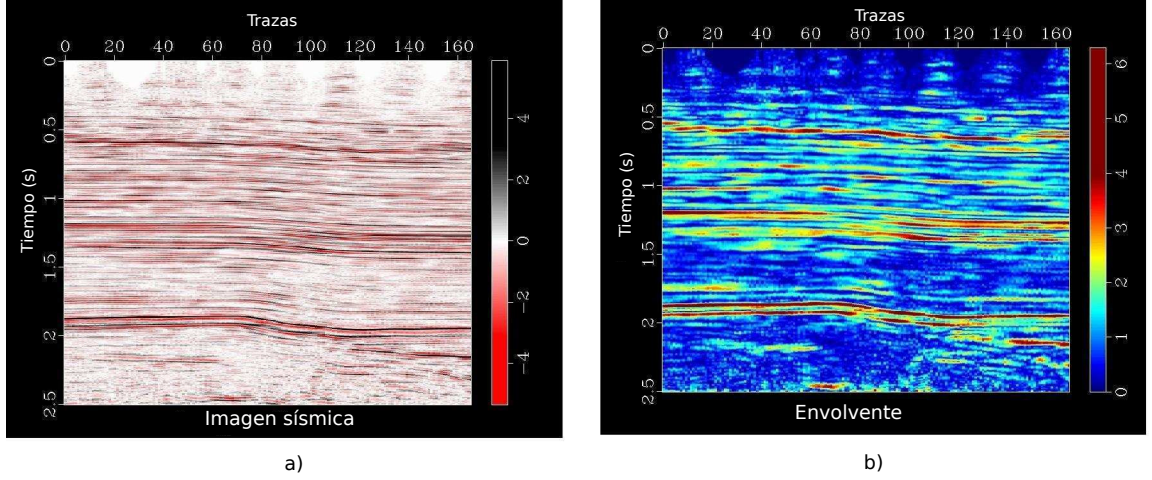


Figura 15: Representación de un sismograma en la traza compleja, a) Imagen sísmica, b) envolvente.

La traza compleja (TC) permite extraer atributos sísmicos tales como la intensidad de los reflectores o la frecuencia instantánea. Esta transformación permite separar la información proporcionada por la amplitud de la proporcionada por la fase. La traza compleja se puede interpretar como un cambio en la fase de -90° , es decir ondas cosenoidales serian transformadas en ondas senoidales. La traza compleja es definida por

$$\tilde{g}(t) = g(t) + iH\{g(t)\}, \quad (2.6)$$

donde $\tilde{g}(t)$ es conocida como la señal analítica y es obtenida sumando la traza sísmica en el dominio del tiempo ($g(t)$) y la transformada de Hilbert de la traza sísmica ($H\{g(t)\}$). La representación de la traza compleja es presentada en color rojo en la Figura 16. La traza compleja es obtenida de dos proyecciones: Una proyección en color azul denominada traza real y obtenida de los geófonos, y una proyección en color negro denominada traza imaginaria, obtenida tras aplicar la transformada Hilbert sobre el sismograma. La definición de la transformada de Hilbert es

$$H\{f(t)\} = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{t - \tau} d\tau, \quad (2.7)$$

donde el operador H se puede interpretar como una convolución en tiempo de $g(t) * (\frac{-1}{\pi t})$ o un desfase de -90° en tiempo [19].

Sobre $\mathbf{d}_{obs}^{s,r}(t)$ y $\mathbf{d}_{mod}^{s,r}(t)$ se aplica un filtro IRR (por sus siglas en ingles, Infinite Impulse Response). Este filtro IRR es un filtro pasa bajas (LP, por sus siglas en ingles low pass

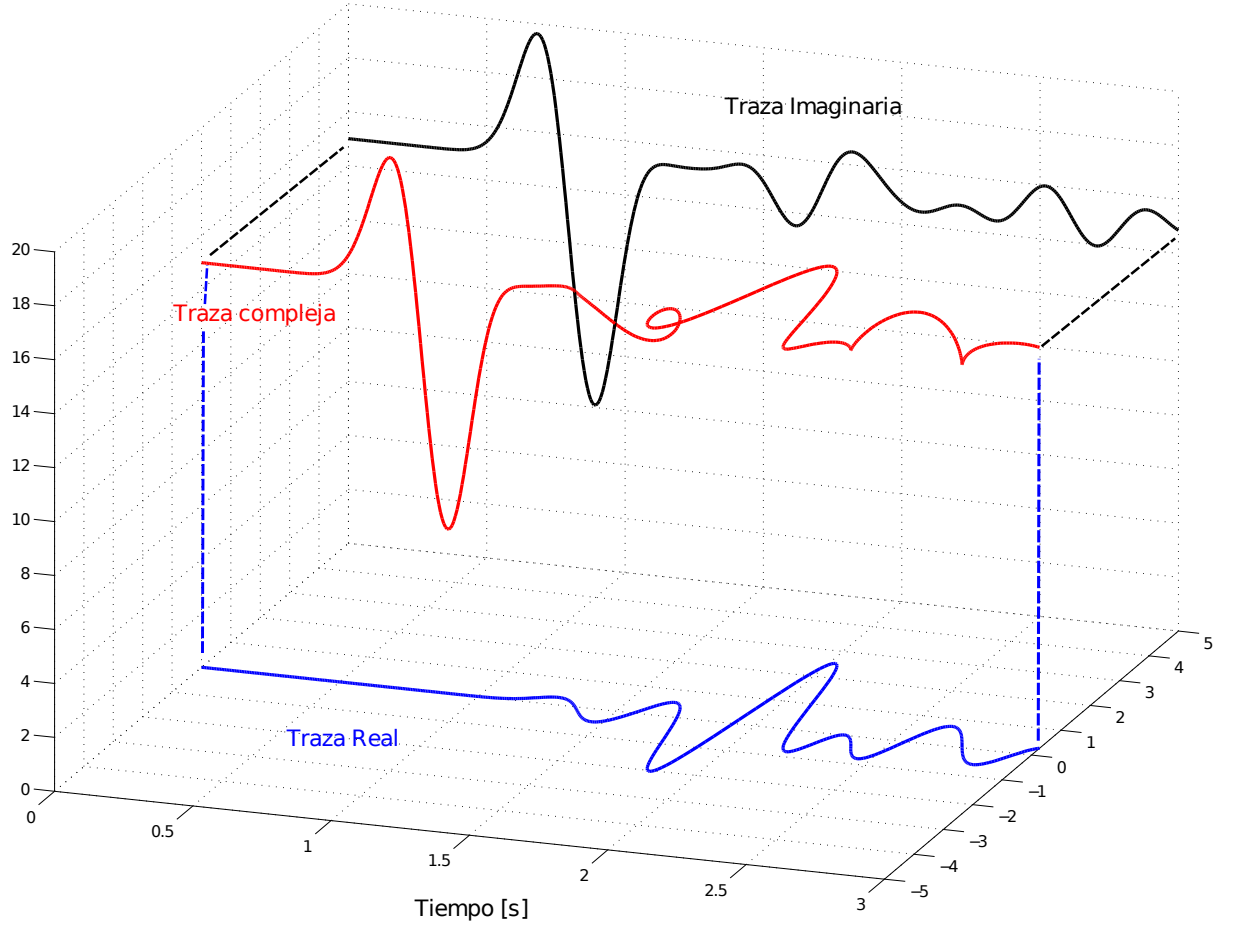


Figura 16: Representación de la traza compleja, compuesta por la traza sísmica (Traza real) y la transformada Hilbert (Traza imaginaria) las cuales son proyección de la traza compleja.

filter) y tiene como objetivo solo analizar las frecuencias menores a 10 Hz de $\mathbf{d}_{obs}^{s,r}$ y $\mathbf{d}_{mod}^{s,r}$. La frecuencia de corte seleccionada para el filtro es 10 Hz . Al aplicar el filtro pasa bajas IRR sobre $\mathbf{d}_{obs}^{s,r}$ se obtiene $\mathbf{g}_{obs}^{s,r}(t)$ y sobre $\mathbf{d}_{mod}^{s,r}$ se obtiene $\mathbf{g}_{mod}^{s,r}(t)$. La fase de las señales analíticas se calcula de la siguiente manera

$$\theta_{obs}^{s,r}(t) = \arctan \left(\frac{H\{g_{obs}^{s,r}(t)\}}{g_{obs}^{s,r}(t)} \right), \quad (2.8)$$

$\theta_{obs}^{s,r}(t)$ es la fase analítica estimada de $\mathbf{g}_{obs}^{s,r}(\mathbf{m}_{true})$, donde \mathbf{m}_{true} es el modelo original y $\theta_{mod}^{s,r}(t)$ es la fase analítica estimada por $\mathbf{g}_{mod}^{s,r}(\mathbf{m}^k)$, donde \mathbf{m}^k son los parámetros del modelo en la iteración k-ésima. El coeficiente de la traza compleja $CS_{TC}^{s,r}$ cuenta las

ocurrencias donde $|\theta_{obs}^{s,r}(t) - \theta_{mod}^{s,r}(t)| > \pi$. La Figura 17 y el algoritmo 3 muestran cómo calcular el $CS_{TC}^{s,r}$. Una vez obtenido $CS_{TC}^{s,r}$ se calcula la norma l^2 en la dirección de los receptores $\widehat{C_{TCs}}$ produciendo un coeficiente de TC por fuente; Finalmente, se promedian los coeficientes de la traza compleja en la dirección de la fuente $\widehat{C_{TC}}$ produciendo un coeficiente de TC por modelo de velocidad. Un nuevo modelo de velocidades es aceptado cuando el coeficiente de TC es menor que el anterior.

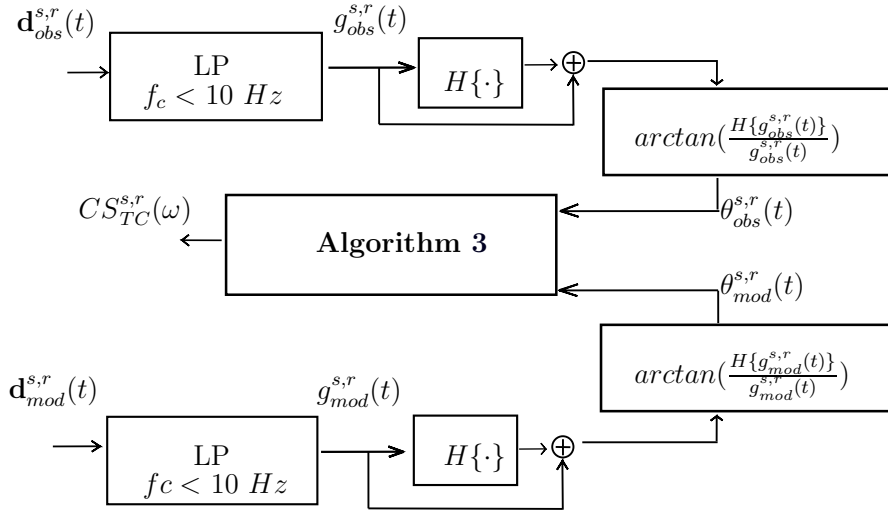


Figura 17: Diagrama empleado para calcular el CS en la traza compleja (CS TC).

Algoritmo 3 CS en la traza compleja

<pre> 1: $\theta_{obs}^{s,r}(t)$ 2: $\theta_{mod}^{s,r}(t)$ 3: for $s = 1, \dots, S$ do 4: for $r = 1, \dots, R$ do 5: $Counter = 0$ 6: for $t = 0, \dots, N_t$ do 7: $dif = \theta_{mod}^{s,r}(t) - \theta_{obs}^{s,r}(t)$ 8: if $dif > \pi$ then 9: $Counter = Counter + 1$ 10: end if 11: end for 12: $CS_{TC}^{s,r} = Counter$ 13: end for 14: end for </pre>	<p>▷ Fase obtenida de la señal analítica $\mathbf{g}_{obs}^{s,r}(t)$</p> <p>▷ Fase obtenida de la señal analítica $\mathbf{g}_{mod}^{s,r}(t)$</p> <p>▷ S, número de fuentes</p> <p>▷ R, número de receptores</p> <p>▷ Reiniciar contador de CS</p> <p>▷ N_t, tiempo de adquisición</p> <p>▷ contador CS</p> <p>▷ Salvar el contador CS</p>
--	---

Se aplican los tres dominios explorando todas sus posibles combinaciones. En el capítu-

lo 3 mostraremos cual es la mejor combinación pero vale la pena justificar el porqué de estos dominios. Como se mencionó anteriormente se busca encontrar suficiente información de baja frecuencia o lo que también es conocido como no tener problemas de CS. Todas las configuraciones de dominio y métrica buscan minimizar el desfase entre trazas. En el dominio del tiempo mediante la correlación buscamos ajustar las trazas en fase mas no en amplitud. En el dominio de la frecuencia en términos del CS, buscamos comparar las componentes frecuenciales en fase del dato observado y el dato modelado. Sin embargo en el dominio de la frecuencia se pierde la componente temporal y esto es un problema ya que el CS busca el desfase en cada tiempo. Se intento resolver este problema mediante el uso de la STFT (short time Fourier transform) pero fue descartado por su elevado costo computacional, alcanzando un incremento en tiempo de hasta 10 veces.

Finalmente la traza compleja permite traer la información de frecuencia al tiempo, lo que permite obtener una fase instantánea siendo ideal para identificar el CS. Nosotros probamos algunos otros dominios transformados como la transformada radon, τ -p y f-k pero los mejores dominios sin elevar mucho su tiempo computacional han sido los presentado en este documento.

3. METODOLOGÍA PARA LA CONSTRUCCIÓN DE LA MÉTRICA

En este capítulo se presenta la metodología propuesta con respecto a la construcción de las tres métricas descritas anteriormente (correlación en el dominio del tiempo, CS en el dominio de la frecuencia y CS en el dominio de la Hilbert). La metodología fue validada usando tres modelos (cuadrado difractor, dos capas paralelas y marmousi); Inicialmente, en el capítulo se presentan los resultados en un esquema en cascada para modelos sub-muestreados y seguidamente los resultados a resolución de pixel. Todas las pruebas fueron implementadas bajo computación en paralelo usando el lenguaje CUDA-C.

3.1. Metodología propuesta

En primera medida se necesita plantear una alternativa de solución para generar el punto de partida necesario para la FWI tal como se presenta en la Figura 18.

En el caso de una adquisición real, la parte superior del diagrama de la Figura 18 empezaría desde (\mathbf{d}_{obs}), el cual es adquirido en campo por un arreglo de geófonos. Sin embargo en este proyecto no están contemplados datos reales y para poder obtener \mathbf{d}_{obs} es necesario propagar un campo de presión sobre el modelo original. Se usa diferencias finitas para solucionar (1.2) con el fin de obtener el campo de presión $P(x, z)$ y así el dato observado. En la Figura 18 este modelado de la ecuación de onda es llamado propagador de onda.

Un primer modelo estimado puede ser obtenido de la información *a-priori* del terreno, sin embargo para nuestro caso en todas las pruebas se tomó aleatoriedad en las velocidades de los modelos iniciales de la metaheurística. Al igual que con el modelo original, sobre el modelo estimado también aplicamos el propagador de onda y obtenemos los datos modelados (\mathbf{d}_{mod}).

Se propone transformar estos datos en tres dominios (tiempo, frecuencia y Hilbert), con el objetivo de que estas transformaciones permitan obtener características del subsuelo que no son claras en un dominio pero sí lo son en otros. Seguidamente es necesario realizar una comparación de los datos en esos dominios usando el operador \mathbf{m} que mide el grado de similitud entre \mathbf{d}_{obs} y \mathbf{d}_{mod} . Finalmente la metaheurística se encargará de generar un nuevo modelo de velocidades de acuerdo a las dimensiones y las restricciones del problema.

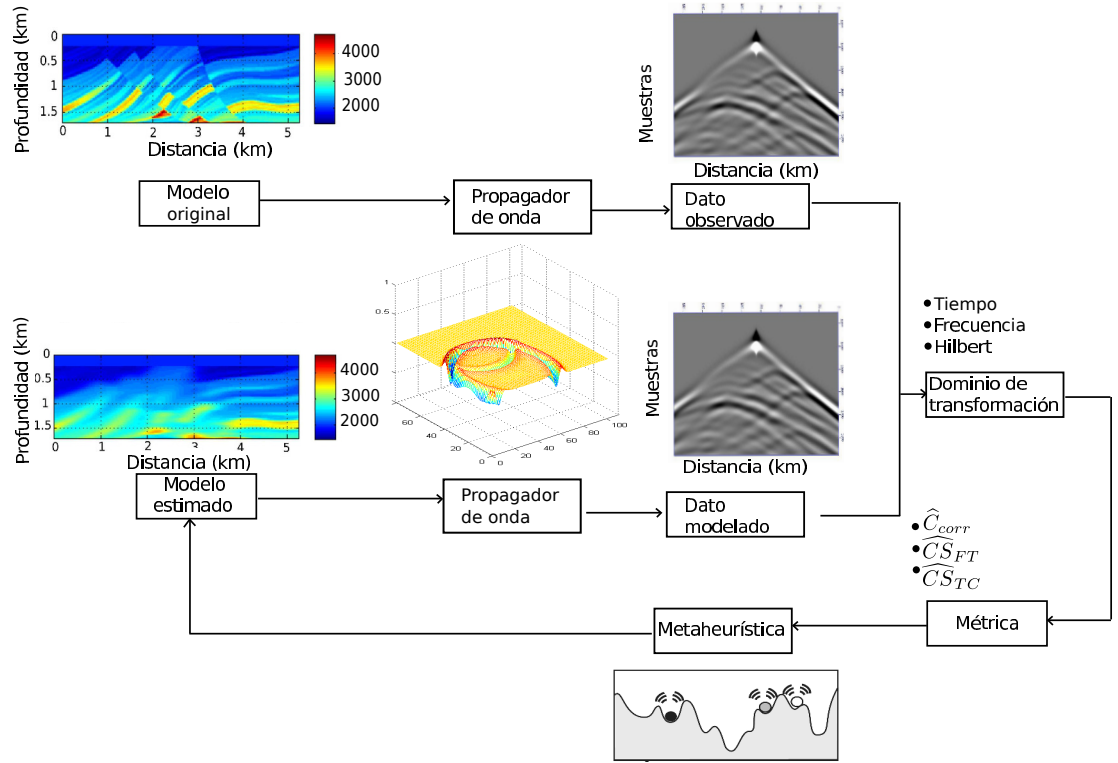


Figura 18: Diagrama de flujo de la solución para obtener el punto de partida.

3.1.1. Método de diferencias finitas en el dominio tiempo

Para resolver computacionalmente (1.2) es necesario su discretización que permita una posterior etapa de implementación. Existen distintas alternativas para resolver (1.2) entre las cuales tenemos: Diferencias finitas en el dominio del tiempo (FDTD) y elementos finitos. Se utilizan diferencias finitas en el dominio del tiempo (FDTD) debido que comparada con elementos finitos, permite corto tiempo computacional, fácil entendimiento y naturaleza explícita. La FDTD se solucionará para un rango de frecuencias según el ancho de banda definido por la fuente.

En (1.2) puede ser reescrita en términos de las diferencias finitas [20] expandiendo mediante series de Taylor's de segundo orden en el dominio temporal como

$$\frac{\partial^2 P(x, z, t)}{\partial t^2} \simeq \frac{P_{i,j}^{n+1} - 2P_{i,j}^n + P_{i,j}^{n-1}}{\Delta t^2}, \quad (3.1)$$

donde i , j y n , representan la discretización de las variables x , z y t respectivamente, Δx y Δz . Para la expansión en series de Taylor's de octavo orden de P en el dominio

espacial tenemos

$$\frac{\partial^2 P}{\partial x^2} \simeq \frac{-\frac{1}{560}P_{i+4,j}^n + \frac{8}{315}P_{i+3,j}^n - \frac{1}{5}P_{i+2,j}^n + \frac{8}{5}P_{i+1,j}^n - \frac{205}{72}P_{i,j}^n + \frac{8}{5}P_{i-1,j}^n - \frac{1}{5}P_{i-2,j}^n + \frac{8}{315}P_{i-3,j}^n - \frac{1}{560}P_{i-4,j}^n}{\Delta x^2}. \quad (3.2)$$

La misma expansión es utilizada para la derivada en z . Reemplazando términos en (1.2) para obtener los campos tenemos

$$\begin{aligned} \frac{1}{c_{i,j}^2} \frac{P_{i,j}^{n+1} - 2P_{i,j}^n + P_{i,j}^{n-1}}{\Delta t^2} &= \frac{1}{\Delta x^2} \left(-\frac{1}{560}P_{i+4,j}^n + \frac{8}{315}P_{i+3,j}^n - \frac{1}{5}P_{i+2,j}^n + \frac{8}{5}P_{i+1,j}^n \right. \\ &\quad \left. - \frac{205}{72}P_{i,j}^n + \frac{8}{5}P_{i-1,j}^n - \frac{1}{5}P_{i-2,j}^n + \frac{8}{315}P_{i-3,j}^n - \frac{1}{560}P_{i-4,j}^n \right) \\ &\quad + \frac{1}{\Delta z^2} \left(-\frac{1}{560}P_{i,j+4}^n + \frac{8}{315}P_{i,j+3}^n - \frac{1}{5}P_{i,j+2}^n + \frac{8}{5}P_{i,j+1}^n \right. \\ &\quad \left. - \frac{205}{72}P_{i,j}^n + \frac{8}{5}P_{i,j-1}^n - \frac{1}{5}P_{i,j-2}^n + \frac{8}{315}P_{i,j-3}^n - \frac{1}{560}P_{i,j-4}^n \right) \end{aligned} \quad (3.3)$$

Resolviendo el campo de onda $P_{i,j}^{n+1}$, con $\Delta x = \Delta z = \Delta h$, (1.2) podrá ser reescrita como

$$\begin{aligned} P_{i,j}^{n+1} &= 2P_{i,j}^n - P_{i,j}^{n-1} + \frac{c_{i,j}^2 \Delta t^2}{\Delta h^2} \left(-\frac{1}{560}P_{i+4,j}^n + \frac{8}{315}P_{i+3,j}^n - \frac{1}{5}P_{i+2,j}^n + \frac{8}{5}P_{i+1,j}^n - \frac{205}{72}P_{i,j}^n \right. \\ &\quad \left. + \frac{8}{5}P_{i-1,j}^n - \frac{1}{5}P_{i-2,j}^n + \frac{8}{315}P_{i-3,j}^n - \frac{1}{560}P_{i-4,j}^n - \frac{1}{560}P_{i,j+4}^n + \frac{8}{315}P_{i,j+3}^n - \frac{1}{5}P_{i,j+2}^n \right. \\ &\quad \left. + \frac{8}{5}P_{i,j+1}^n - \frac{205}{72}P_{i,j}^n + \frac{8}{5}P_{i,j-1}^n - \frac{1}{5}P_{i,j-2}^n + \frac{8}{315}P_{i,j-3}^n - \frac{1}{560}P_{i,j-4}^n \right) + src_{i,j}. \end{aligned} \quad (3.4)$$

En (3.4) fue implementada en CUDA-C para tomar ventaja de la computación en paralelo para resolver la ecuación de onda y así el problema directo [21].

3.1.2. Construcción de la métrica

El operador B permite obtener el campo de presión p de (3.4) en la ubicación de los receptores. Se define como

$$\mathbf{d}_{mod} = B\{p\}. \quad (3.5)$$

Debemos aplicar tres dominios transformados que facilitarán la interpretación de los datos, para esto se plantea

$$T_{\mathfrak{x}}\{\mathbf{d}_{mod}\} \quad (3.6)$$

donde $\mathfrak{x} = 1, 2, 3$, son las transformaciones de las trazas modeladas (\mathbf{d}_{mod}) en los tres dominios transformados ya mencionados. Por otra parte estas mismas transformaciones deben aplicarse sobre el dato observado teniendo que

$$T_{\mathfrak{x}}\{\mathbf{d}_{obs}\}. \quad (3.7)$$

Sobre estos dominios se aplica el operador $\mathbf{m}_{\mathfrak{x}}(T_{\mathfrak{x}}\{\mathbf{d}_{mod}\}, T_{\mathfrak{x}}\{\mathbf{d}_{obs}\})$ que es el funcional que mide que tanto se parecen \mathbf{d}_{obs} y \mathbf{d}_{mod} en los respectivos dominios transformados $T_{\mathfrak{x}}$. Como se propone evaluar cada una de las métricas \mathbf{m} de cada dominio, nos conduce a la ecuación

$$\begin{aligned} \mathbf{m}_{\mathfrak{T}} = \mathbf{m}_1(T_1\{\mathbf{d}_{mod}\}, T_1\{\mathbf{d}_{obs}\}) + \mathbf{m}_2(T_2\{\mathbf{d}_{mod}\}, T_2\{\mathbf{d}_{obs}\}) + \\ \mathbf{m}_3(T_3\{\mathbf{d}_{mod}\}, T_3\{\mathbf{d}_{obs}\}). \end{aligned} \quad (3.8)$$

Sin embargo estos dominios transformados deben irse ajustando con unos pesos según la importancia de cada evento sísmico, estos pesos se denotan como W_1, W_2 y W_3 . Por otra parte se usan banderas de activación para cada dominio C_1, C_2 y C_3 donde sus valores son de 0 o 1. Reescribiendo (3.8) tenemos que

$$\begin{aligned} \mathbf{m}_{\mathfrak{T}} = C_1 W_1 \mathbf{m}_1(T_1\{\mathbf{d}_{mod}\}, T_1\{\mathbf{d}_{obs}\}) \\ + C_2 W_2 \mathbf{m}_2(T_2\{\mathbf{d}_{mod}\}, T_2\{\mathbf{d}_{obs}\}) \\ + C_3 W_3 \mathbf{m}_3(T_3\{\mathbf{d}_{mod}\}, T_3\{\mathbf{d}_{obs}\}). \end{aligned} \quad (3.9)$$

En todos los experimentos los pesos W_x serán iguales a uno. En la Figura 19 se presenta un ejemplo del esquema en cascada, primero activamos el dominio T_1 (Figura 19 a)), en otro instante activamos el dominio T_2 (Figura 19 b)) y finalmente el dominio T_3 (Figura 19 c)).

3.2. Modelos sintéticos

Para todos los experimentos se busca que los parámetros de las pruebas no tuvieran problemas de dispersión y se garantizará estabilidad numérica [22]. Los parámetros

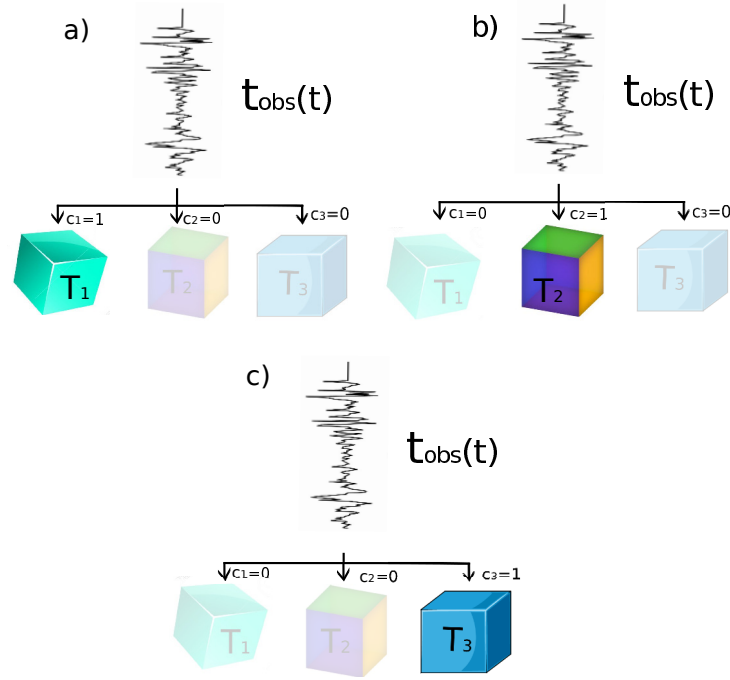


Figura 19: Esquema en cascada de los dominios transformados: a) Activo dominio transformado T_1 ; b) Activo dominio transformado T_2 ; c) Activo dominio transformado T_3 .

de las pruebas empleados son: $\Delta t = 4 \times 10^{-3}$ s, un tiempo de adquisición de 2.5 s, un espaciado $\Delta x = 25$ m, $\Delta z = 25$ m. Las dimensiones del modelo son de 5275×1700 m² con cinco fuentes ubicadas en 525 m, 1600 m, 2625 m, 3725 m y 4775 m a una profundidad de 125 m. Para la FWI se utiliza la fuente de excitación descrita por

$$g_s = 1 - 2(\pi f)^2(t - t_0)e^{-(\pi f)^2(t-t_0)^2}, \quad (3.10)$$

donde f es la frecuencia central (3Hz) y t_0 es el retardo de la fuente (0.5 s). Para PSO la fuente de excitación empleada es

$$g_s = -2(\pi f)^2(t - t_0)e^{-(\pi f)^2(t-t_0)^2}. \quad (3.11)$$

En la Figura 20, se presentan los espectro para la fuente de excitación de PSO a 3Hz y la fuente de excitación de FWI a 3Hz, 6Hz y 9Hz. Note que (3.11) es una modificación de la utilizada en (3.10). Se selecciona la fuente presentada en (3.11) para PSO, porque su ancho de banda contiene las componentes frecuenciales más bajas.

La metodología es validada usando tres modelos de velocidades: Cuadrado difractor Figura 21 a), modelo de capas paralelas Figura 21 b) y modelo Marmousi Figura 21

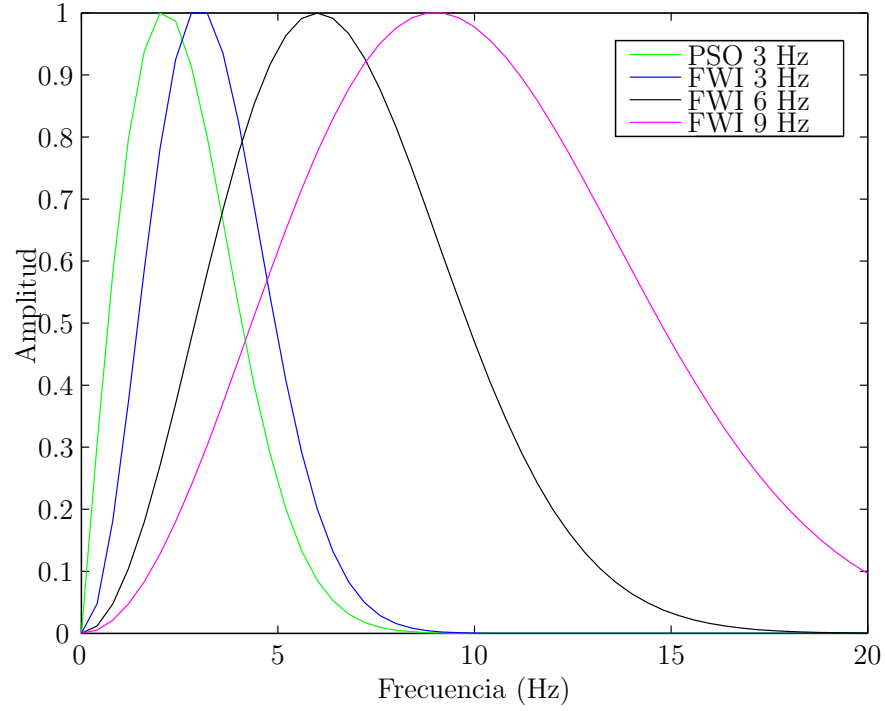


Figura 20: Espectro de las fuentes PSO y FWI.

c). Sin embargo, sólo emplearemos el modelo del cuadrado difractor de la Figura 21 a) como modelo original para todo el procedimiento de la metodología, en la sección de anexos se presentan los resultados de los modelos restantes.

3.3. Resultados modelos sub-muestreados

En primera medida se prueban las tres métricas (\widehat{C}_{corr} , \widehat{CS}_{FT} y \widehat{CS}_{TC}), partiendo de un modelo aleatorio con distribución normal y moviéndonos en un rango de velocidades entre V_{min} y V_{max} . Estos rangos de velocidad pueden ser establecidos por conocimientos geofísicos *a-priori*. Sin embargo, es necesario determinar V_{max} se usa el criterio de Courant para un malla de segundo orden temporal y octavo orden espacial. $V_{max} = 3450m/s$ es la velocidad máxima del modelo donde se logra mantener controlada la dispersión numérica y estabilidad [22]. El criterio de Courant establece que

$$C \frac{\Delta t}{\Delta x} \leq \sqrt{\frac{b_2}{b_1}} \leq 0,5546, \quad (3.12)$$

donde b_1 y b_2 representan la sumatoria de los valores absolutos de los coeficientes espaciales y temporales presentados en (3.3). Como se utilizó un malla de segundo

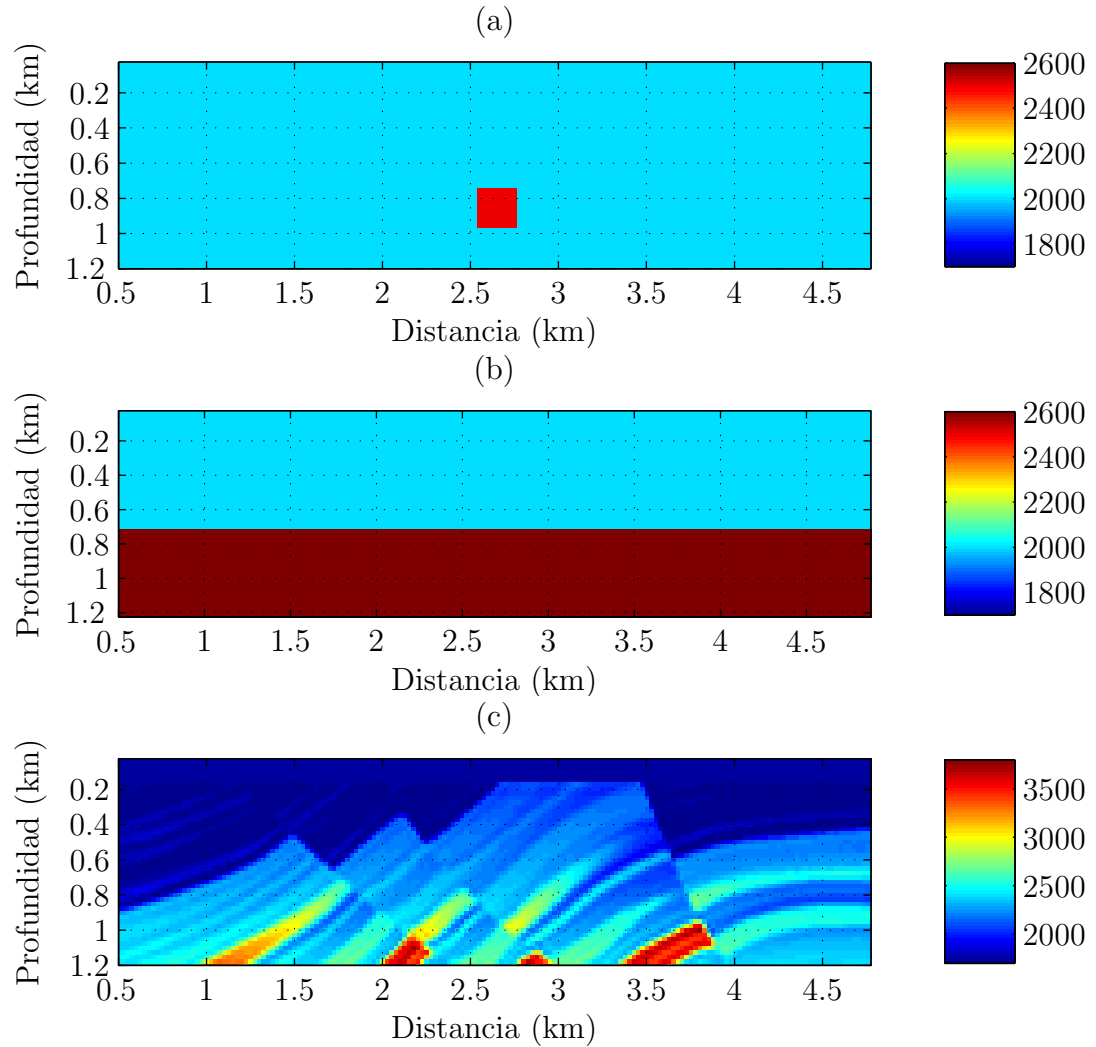


Figura 21: Modelos originales usados: (a)Cuadrado difractor velocidad fuera del cuadrado de 2000 m/s y en el cuadrado de 2500 m/s dimensiones de 9×9 , (b) Modelo 2 capas $2000\text{-}2800 \text{ m/s}$, (c) Modelo marmousi.

orden espacial tenemos que b_1 es

$$b_1 = |1| + |-2| + |1| = 4, \quad (3.13)$$

y como utilizamos coeficientes de octavo orden temporal tenemos que b_2 es

$$b_2 = 4 \cdot (|\frac{-1}{560}| + |\frac{8}{315}| + |\frac{-1}{5}| + |\frac{8}{5}|) + 2 \cdot (|\frac{-205}{72}|) = 13,0032. \quad (3.14)$$

En la Figura 23 se muestra un modelo submuestreado (\mathbf{m}_{sub}), generado aleatoriamente por la metaheurística en su primera iteración. \mathbf{m}_{sub} tiene dimensiones de 25×7 , donde 25 es el número de píxeles en la dimensión de x y 7 es el número de píxeles en la dimensión de z . Se observó que este modelo tiene contrastes fuertes debido a los cambios de velocidades entre píxeles adyacentes. Con el fin de mitigar el efecto de estos contrastes fuertes, se propone aplicar un filtro con el objetivo de suavizar este modelo y así ayudar en la búsqueda de la metaheurística. Se usa un filtro uniforme definido por

$$\mathbf{h} = \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (3.15)$$

donde \mathbf{h} tiene dimensiones de 5×5 con valores de $1/25$, de tal forma que la suma de todos sus elementos sea 1. En la Figura 22 se presenta el uso del operador \mathbf{h} sobre el modelo de velocidades para obtener un pixel. El proceso de la Figura 22 se repetirá hasta recorrer todos los píxeles del modelo de velocidades. El nuevo pixel es obtenido como

$$N_{pixel} = \sum_{y=0}^{24} h_y \times m_y. \quad (3.16)$$

\mathbf{h} es empleado para todos los modelos obtenidos de la metaheurística siempre actuando antes de la propagación. El resultado de aplicar \mathbf{h} sobre el modelo de la Figura 23 es presentado en la Figura 24.

En el Capítulo 1 se definieron los parámetros $w = 0,7$, $c_1 = 0,8$ y $c_2 = 1,1$ según el contexto geofísico. Sin embargo, es necesario determinar la cantidad de iteraciones y partículas por experimento. Se propone variar el número de iteraciones en valores de 10, 50 y 100; y el número de partículas en valores de 10, 50, 200, 400 y 500. El objetivo es encontrar un equilibrio entre costo computacional y avance de la métrica. En la Figura 25 se muestra que con 500 partículas y 100 iteraciones se obtiene el mejor avance en

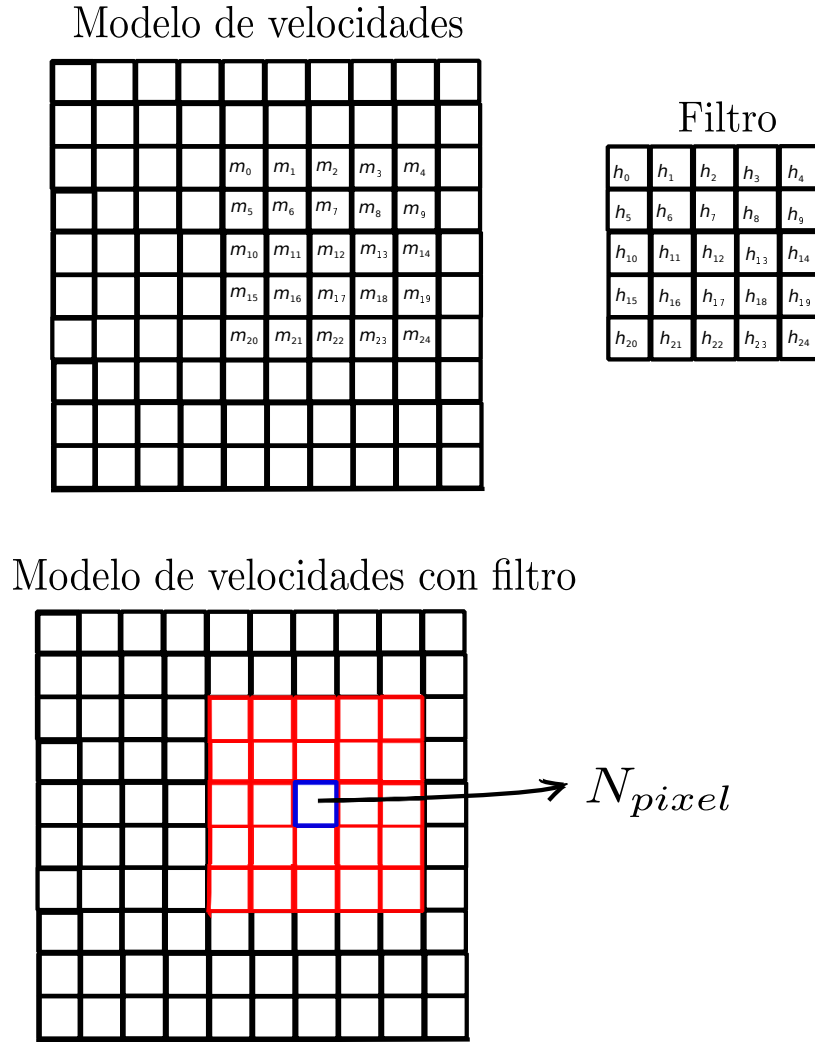


Figura 22: Uso del filtro sobre el modelo de velocidades

la métrica con un costo computacional de 3.85 horas como se presenta en la Figura 26. Para obtener los tiempos obtenidos en la Figura 26 se tuvo en cuenta: programación en paralelo tanto la dimensionalidad como la propagación de onda 2D, acústica y de densidad constante usando con una GPU Tesla K40c de 12Gb de RAM; ejecutar las métricas de manera serial bajo un procesador Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40GHz. La selección del número de partículas y dimensiones está estrechamente relacionado con las dimensiones del modelo.

Los resultados del mejor global de un experimento para cada una de las métricas se presenta en la Figura 27: a) \hat{C}_{corr} , b) \hat{CS}_{FT} y c) \hat{CS}_{TC} . Al ser un proceso de optimización probabilístico, es necesario definir el número de experimentos (n). Es importante recordar que cada experimento cuenta con 100 partículas, 500 iteraciones y solo guardando el mejor global. Se obtiene 100 experimentos de cada métrica por dos razones:

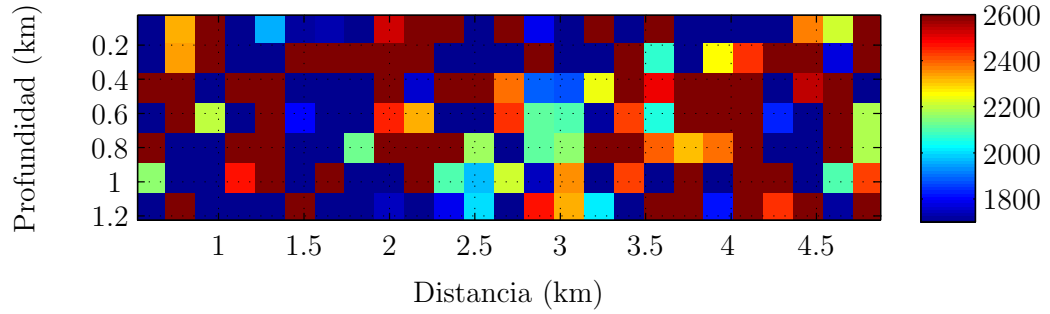


Figura 23: Modelo aleatorio generado en la primera iteración de la metaheurística.

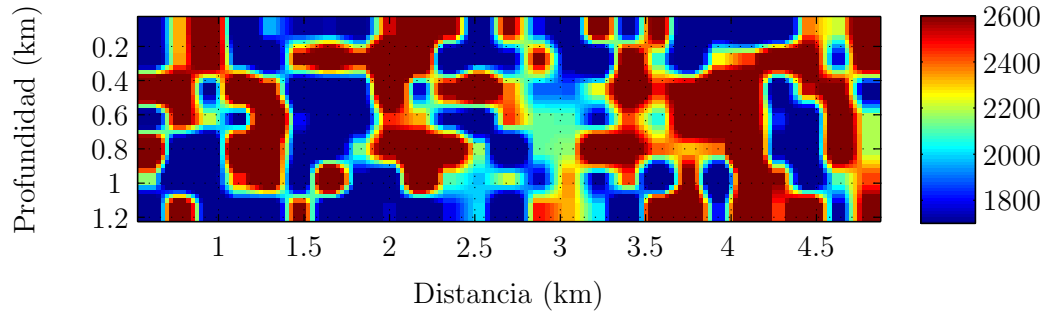


Figura 24: Modelo obtenido tras aplicar el operador $h\{\cdot, \cdot\}$.

La primera es por el tiempo de ejecución de las pruebas, dependiendo de la métrica y de la resolución del modelo este tiempo puede ser de 96 horas. La segunda razón es por el porcentaje de error de las varianzas con respecto al número de experimentos, siendo 100 experimentos la cantidad de experimentos necesarios para garantizar un porcentaje de error menor del 1 %. En el capítulo 4 se presentará detalladamente el fundamento estadístico para la selección de n .

De cada experimento se obtendrá un mejor global por prueba lanzada. Usando los 100 mejores globales de 100 experimentos, se construye un cubo de experimentos como se presenta en la Figura 28. Este cubo de experimentos es promediado en la dirección de los experimentos con el fin de generar un único modelo (\mathbf{m}_{avg}) que resalte los eventos que se repiten y disminuir el impacto de aquellos valores que no tienen coherencia.

En un esquema en cascada donde tenemos en cascada más de una métrica, \mathbf{m}_{avg} de la métrica actual es tomado como mejor global para la primera iteración en la nueva exploración de la siguiente métrica. Esto genera un punto de gravedad que es usado como referencia para la exploración aleatoria de las siguientes métricas. Se probaron todas las posibles combinaciones de métricas como se describió en la Figura 29, con

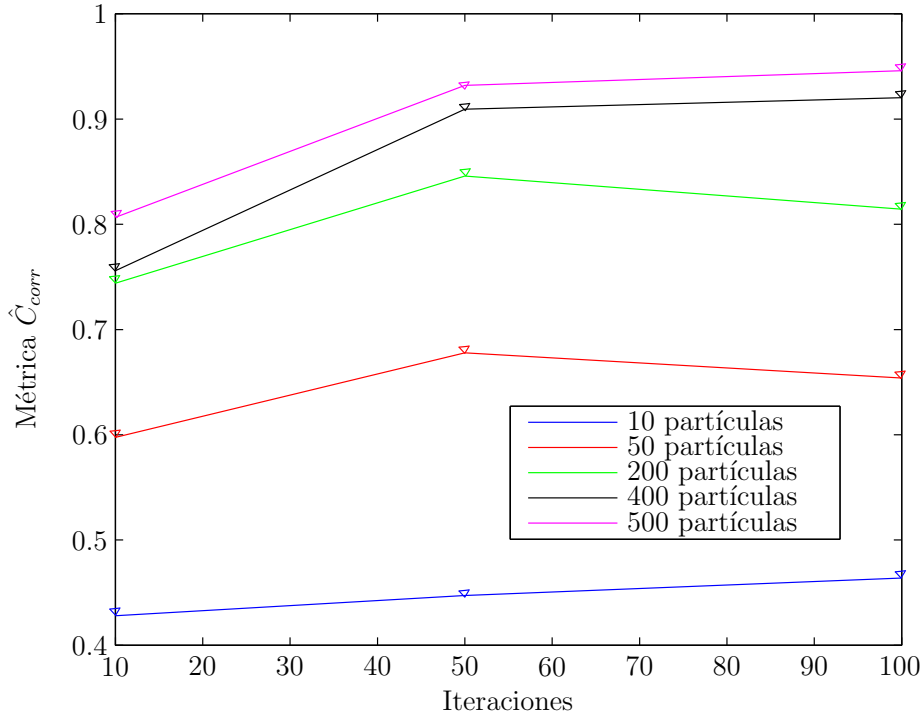


Figura 25: Comportamiento de las iteraciones en función de la métrica \hat{C}_{corr} .

el fin de encontrar la mejor combinación de ellas. Los resultados se presentan desde la Figura 81 a) hasta la Figura 95 a).

\hat{C}_{corr} y el \hat{CS}_{TC} son las métricas que mejor solucionan la zona de iluminación (ver zona sombreada de la Figura 30), obteniendo mas valores cercanos a 2000 m/s . Las velocidades en el cuadrado para todas las métricas siempre estuvieron por encima de las velocidades originales, con un error de $\pm 300 \text{ m/s}$. El error generado por la longitud de onda utilizada en la fuente, implica una mala ubicación del cuadrado difractor. La metaheurística aumenta la velocidad en esta zona, con el fin de que los tiempos de llegada de \mathbf{d}_{mod} se ajusten a \mathbf{d}_{obs} . Para lograr su ubicación exacta necesitamos ampliar el rango de frecuencias, sin embargo no se plantea aumentar el rango de frecuencias ya que no pretendemos solucionar todo el problema de inversión, sino que garantiza la información de baja frecuencia.

Sin aplicar FWI, es necesario definir cuáles de las combinaciones presentadas en la Figura 29 son las mejores. Se propone aplicar el operador $FT\{\cdot\}$ sobre las trazas obtenidas de \mathbf{m}_{avg} y \mathbf{m}_{true} . Se escoge una traza aleatoria y se propone el operador

$$FT_{abs}\{\mathbf{d}_{obs}, \mathbf{d}_{avg}\} = ||\text{abs}\{FT\{\mathbf{d}_{obs}\}\} - \text{abs}\{FT\{\mathbf{d}_{avg}\}\}||_2, \quad (3.17)$$

donde abs es la función magnitud de la transformada de Fourier y \mathbf{d}_{avg} son las trazas ob-

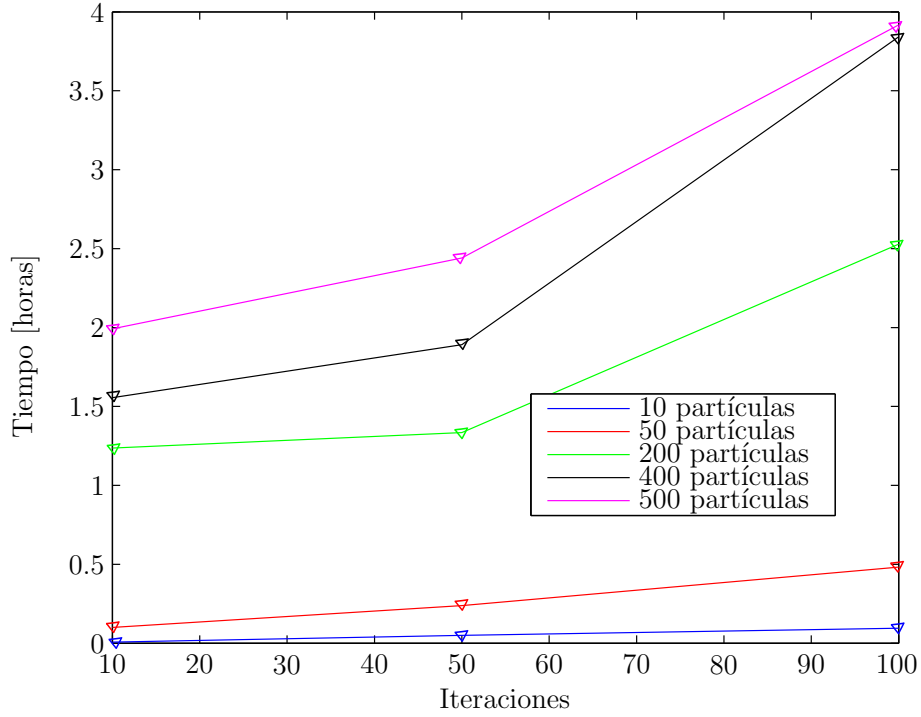


Figura 26: Comportamiento de las iteraciones en función del tiempo.

tenidas de \mathbf{m}_{avg} . También se obtuvo la fase de la transformada de Fourier, proponiendo el siguiente operador

$$FT_{angle}\{\mathbf{d}_{obs}, \mathbf{d}_{avg}\} = \|\angle FT\{\mathbf{d}_{obs}\} - \angle FT\{\mathbf{d}_{avg}\}\|_2, \quad (3.18)$$

donde \angle es la fase de la transformada de Fourier. En la Figura 31 se presenta el resultado de aplicar $\angle FT\{\mathbf{d}_{obs}\}$ menos $\angle FT\{\mathbf{d}_{avg}\}$ sobre la traza 50 obtenida de PSO con la métrica $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$. Como se aprecia en la Figura 31 las componentes frecuenciales menores a 6 Hz del modelo inicial obtenido PSO son casi cero, es decir la información de baja frecuencia esta siendo solucionada y es un punto adecuado para la FWI. Por otra parte la FWI actúa sobre las frecuencias altas (>6 Hz).

En la sección de anexos se presenta el operador $\angle FT$ sobre la traza 50 obtenida del modelo de 2 capas (ver Figura 70) y marmousi (ver Figura 71) respectivamente.

Se aplica el operador FT_{abs} sobre cada conjunto de trazas obtenidas del esquema en cascada presentado en la Figura 29, los resultados se presentan en la Figura 32. La convención utilizada en todos los diagramas multi métrica a partir de la Figura 34 es: Para una sola métrica el punto verde, para dos métricas la secuencia de puntos rojos y finalmente para tres métricas la secuencia de puntos azules.

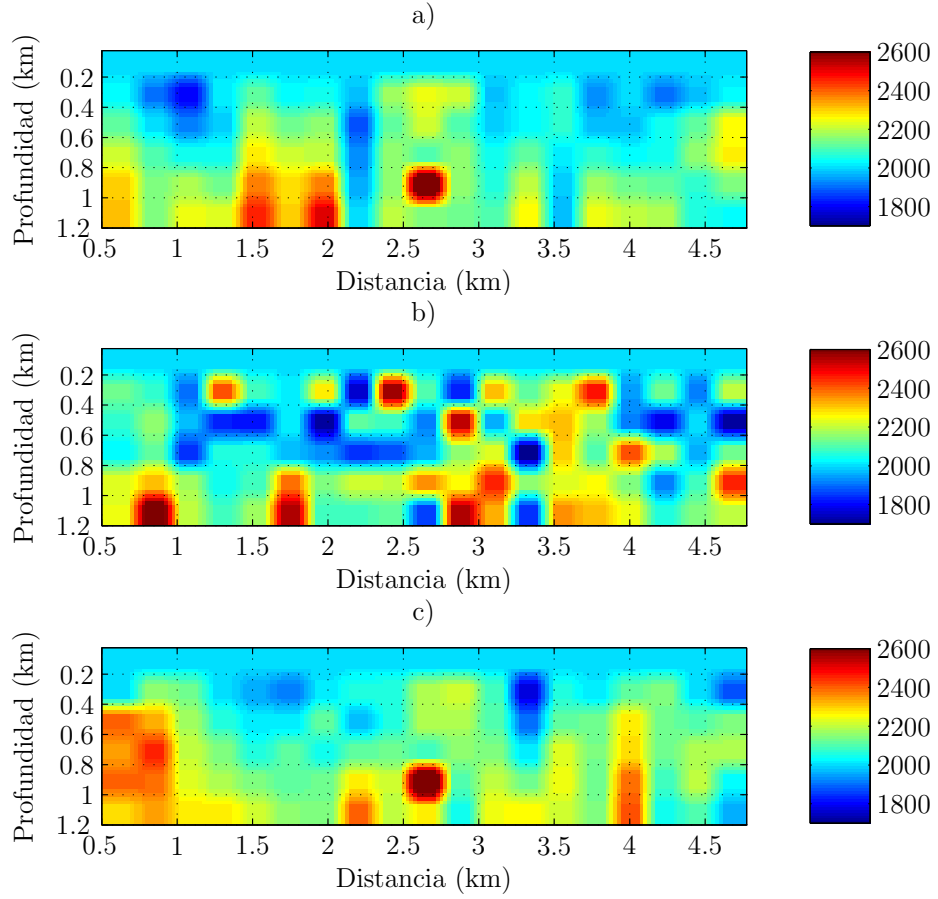


Figura 27: Modelos obtenidos con cada dominio a) Tiempo \widehat{C}_{corr} , b) Frecuencia \widehat{CS}_{FT} , c) Hilbert \widehat{CS}_{TC} .

Los 15 modelos obtenidos del esquema en cascada se emplean como puntos de partida para la FWI, con el objetivo de encontrar cual de ellos es el mejor. Es importante aclarar que primero aplicamos una FWI para una frecuencia de 3 Hz con 40 iteraciones y 21 fuentes. El resultado es usado como punto de partida para una siguiente FWI de 6 Hz con igual número de iteraciones y fuentes. Finalmente el resultado es usado como punto de partida para una tercera FWI de 9 Hz con 40 iteaciones y 21 fuentes. Este proceso es conocido en el estado-del-arte como una FWI multi escala [23]. Los puntos de partida con los modelos obtenidos al aplicar FWI son presentados desde la Figura 81 b) hasta la Figura 95 b). De los modelos obtenidos necesitamos estimar de manera cuantitativa cual es el mejor modelo. Para esto se plantea la siguiente métrica

$$L2^{s,r} = \|\mathbf{d}_{obs}^{s,r}(t) - \mathbf{d}_{mod}^{s,r}(t)\|_2, \quad (3.19)$$

donde s es el número de fuentes y r es el número de receptores. De (3.19) se obtiene la norma l^2 en la dirección de los receptores ($L2^s$) y finalmente la norma l^2 en la dirección

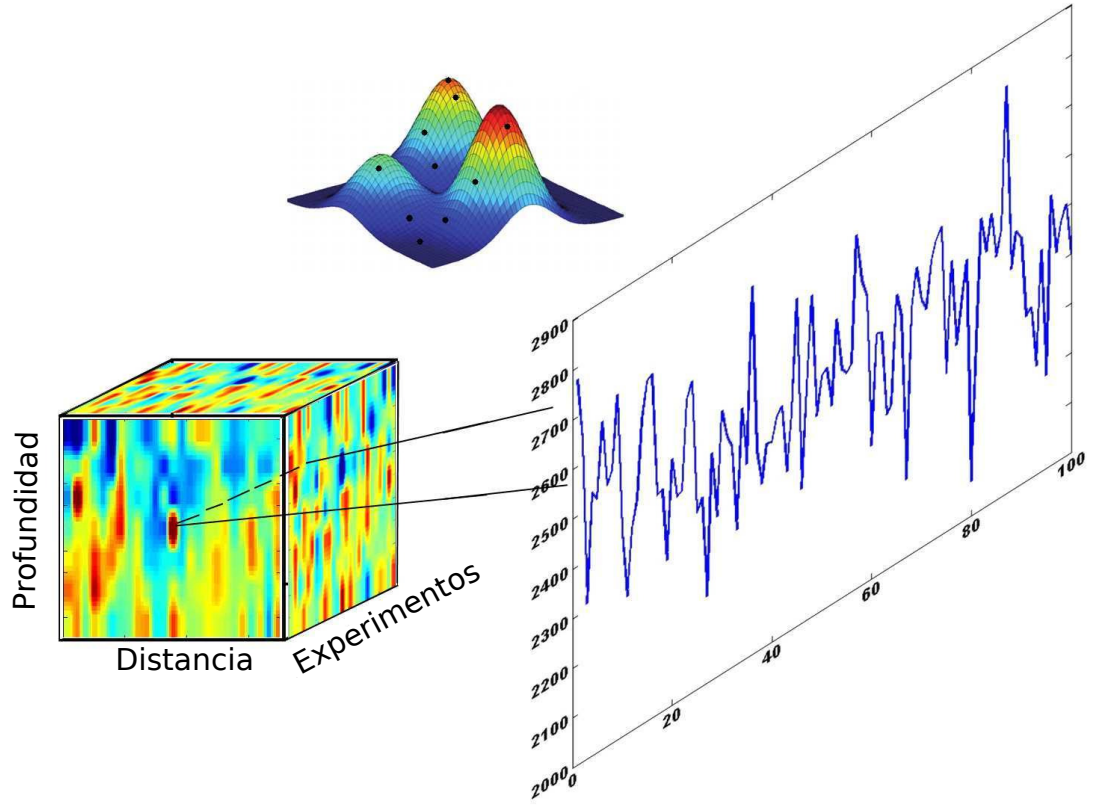


Figura 28: Cubo de los mejores globales de 100 experimentos.

de las fuentes ($L2$) para obtener un único escalar que relaciona todos los disparos. Se plantea la norma l^2 por que la FWI minimiza el error cuadrático medio entre las trazas como se describió en (1.8). Se espera que después de aplicar FWI, estemos cerca del vecindario del mínimo global y por tanto un comportamiento convexo de la función de costo. Los resultados de aplicar la norma l^2 entre las trazas del modelo obtenido de FWI y \mathbf{d}_{obs} , son presentados en la Figura 33.

El modelo obtenido de la métrica \widehat{CS}_{TC} es el mejor modelo al sólo aplicar una métrica. En los recuadros con puntos rojos las métricas $\widehat{CS}_{FT} + \widehat{CS}_{TC}$ son la mejor combinación para dos métricas en cascada y en el recuadro con puntos azules las métricas $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{CS}_{Corr}$ son la mejor combinación de tres métricas en cascada.

El tiempo de ejecución es importante al momento de seleccionar con cuántas métricas se generará el punto de partida. En la métrica de correlación el tiempo de ejecución fue de 22,34 horas, para la métrica \widehat{CS}_{FT} su tiempo de ejecución fue de 24,5 horas, mientras que para la métrica de \widehat{CS}_{TC} se necesitaron de 84.4 horas. Estos tiempos son calculados usando un servidor Intel(R), Xeon(R), CPU E5-2620 v3, 2.40GHz con 256 GB de RAM que incluye una GPU Tesla K40c de 12 GB de RAM siempre usando el

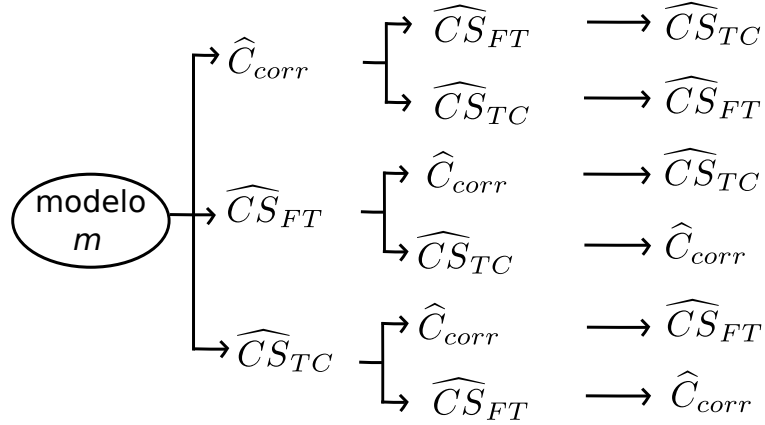


Figura 29: Diagrama empleado para las combinaciones de métricas.

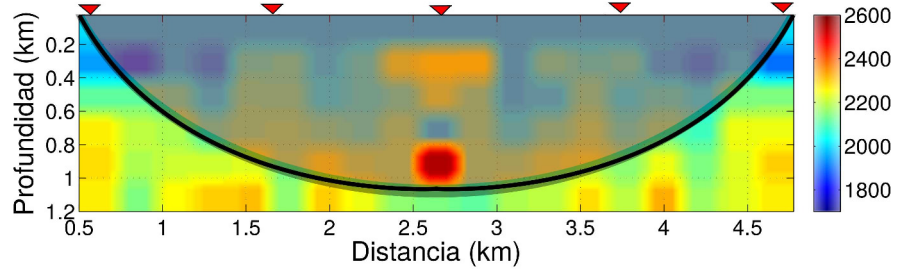


Figura 30: Zona de iluminación generada por las cinco fuentes.

sistema operativo Linux.

Se programo en paralelo para todos los experimentos, tanto para el propagador de onda como para la dimensionalidad de PSO para calcular el modelo. Con esta paralelización se logro una aceleración de seis veces comparada con una serial bajo lenguaje C en este mismo servidor.

Como se trabajó con datos sintéticos se puede corroborar esta metodología usando el modelo original. Se calculó la norma l^2 entre las soluciones del modelo original con las de modelos obtenidos de cada combinación de métricas. El objetivo es estimar cual es el mejor punto de partida. En el diagrama de la Figura 34 se presentan los respectivos valores de la norma l^2 corroborando los resultados de la Figura 33, cuando se usa una métrica o varias métricas. Se debe resaltar que el resultado de la Figura 34 será solo teórico ya que como se mencionó en una adquisición real el modelo original es desconocido.

Desde la Figura 35 hasta la Figura 42 se presentan los mejores modelos según el número de métricas. En las Figuras 37 y 41 se presentan el mejor y peor modelo inicial en

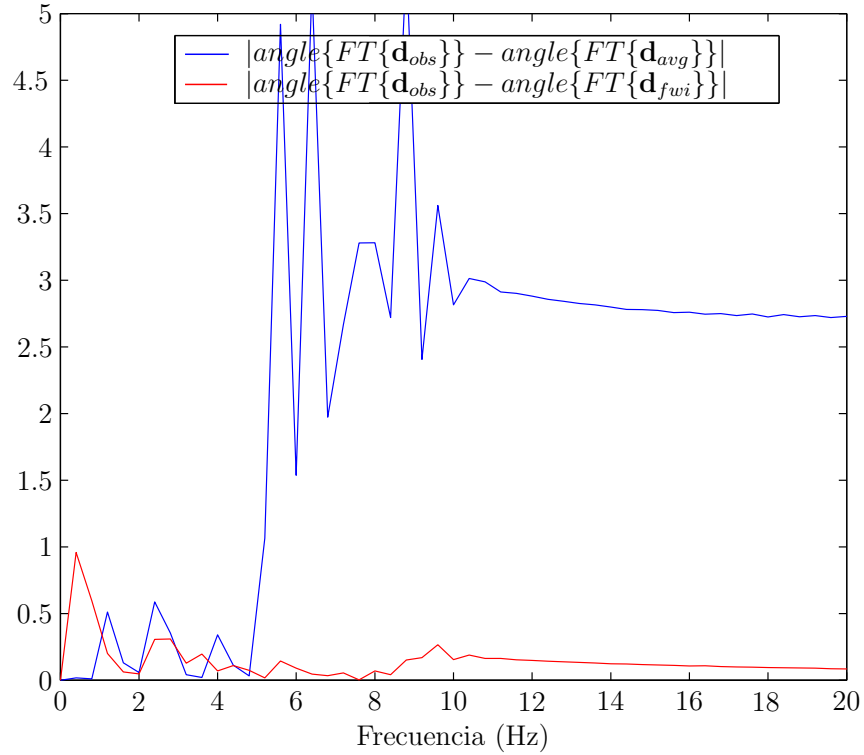


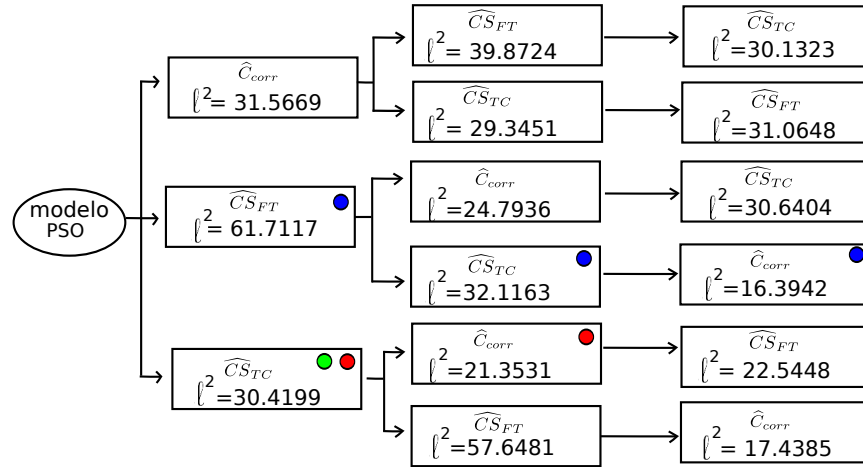
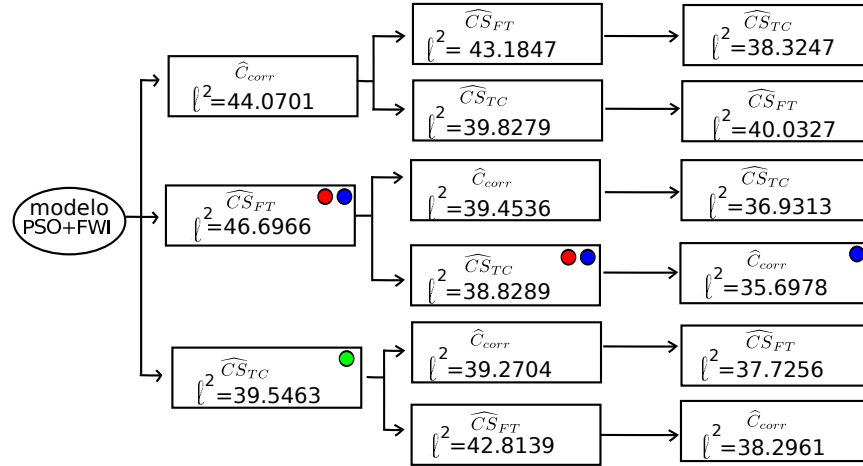
Figura 31: Uso del operador FT_{angle} sobre la traza 50 obtenida con las métricas $\widehat{C}S_{TC} + \widehat{C}S_{FT} + \widehat{C}_{corr}$ (\mathbf{m}_{ori} -cuadrado difractor).

términos de norma l^2 .

3.4. Resultados de modelos a resolución de pixel

Con el objetivo de acercarnos a la realidad se propone poder establecer modelos con dimensiones de pixel (\mathbf{m}_{full}). Para poder obtener los resultados de resolución de pixel fueron necesarios dos aspectos importantes: El primero programar en paralelo la propagación de onda acústica 2D sobre el modelo para obtener tanto los datos modelados como observados. La segunda es programar en paralelo la dimensionalidad de PSO, es decir la actualización de cada modelo \mathbf{x}_{i+1}^n para todas las partículas. Estas dos modificaciones aportan una disminución del tiempo computacional comparado a la serial de seis veces.

Al igual que con \mathbf{m}_{sub} se usa el diagrama es cascada presentado en la Figura 29, obteniendo los resultados presentados desde la Figura 96 a la Figura 110. Nuevamente obtuvimos FT_{abs} como se presentó en (3.17), buscando cual de estas combinaciones son los mejores puntos de partida para la FWI según el número de métricas. Los resultados se presentan en la Figura 43.


 Figura 32: Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{sub} -cuadrado difractor)

 Figura 33: Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{sub} -cuadrado difractor).

Aplicamos FWI sobre estos 15 modelos obtenidos del esquema en cascada a resolución de pixel. Es importante aclarar que en estas pruebas también aplicamos una FWI multi escala con 21 fuentes y frecuencias de 3 Hz, 6 Hz y 9 Hz. En la Figura 44 se presenta el diagrama multi-métrica en términos de la norma l^2 entre las trazas del modelo obtenido de FWI y \mathbf{d}_{obs} .

Cuando usamos una sola métrica podremos seguir usando \widehat{CS}_{TC} , cuando usamos dos métricas en cascada podremos usar \widehat{CS}_{TC} seguido de \widehat{C}_{corr} , mientras que si usamos tres métricas se recomienda usar $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$. Estos últimos ordenes para dos y tres métricas en cascada difieren del presentado en \mathbf{m}_{sub} . La razón se debe por tener un modelo sub-muestreado, ayuda a la métrica \widehat{CS}_{FT} .

La traza compleja permite obtener la información de fase en cada instante de tiempo,

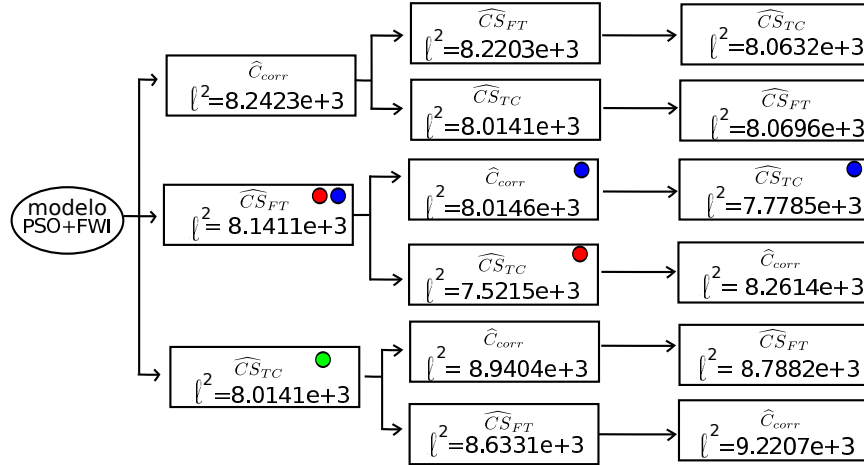


Figura 34: Diagrama en cascada usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{sub} -cuadrado difractor).

logrando caracterizar de una mejor forma los eventos sísmicos. Aplicando la métrica \widehat{CS}_{TC} se logra reducir la zona de búsqueda y genera un punto de inercia como mínimo global para la exploración de la siguiente métrica. Por otra parte al igual que con \mathbf{m}_{sub} existe una zona de iluminación tal como se presentó en la Figura 30, pero hay muchos menos valores cercanos a la solución. Para \mathbf{m}_{full} es más complejo alcanzar la solución debido que al tener un modelo sub-muestreado se está perdiendo información relevante. Finalmente se estima la norma l^2 entre cada modelo obtenido de la FWI multi escala con \mathbf{m}_{true} , los resultados son presentados en la Figura 45. De las Figuras 34 y 45, se corrobora la correcta selección de la primera métrica. Es importante mencionar que esta primera selección puede cambiar según la información de baja frecuencia proporcionada por el modelo.

Los mejores modelos según el número de métricas antes y después de aplicar FWI, son presentados desde la Figura 46 hasta la Figura 53. En las Figuras 50 y 52 se presentan el mejor y peor modelo inicial en términos de la norma l^2 .

Las secuencias que se presentan para los modelos a resolución de pixel del cuadrado difractor (ver Figura 74), marmousi (ver Figura 80) y modelo de dos capas (77) difieren. Esto se debe a que cada modelo ofrece información de baja frecuencia, siendo más fácil caracterizarse en un dominio que en otros. Algunos modelos obtenidos de la metaheurística como los de marmousi aunque cualitativamente no son parecidos al modelo original, contiene la información de baja frecuencia necesaria para la FWI y son puntos de partida adecuados desde el concepto de CS.

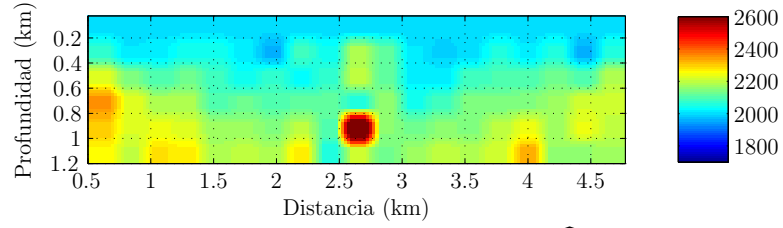


Figura 35: \mathbf{m}_{sub} obtenido de \widehat{CS}_{TC} .

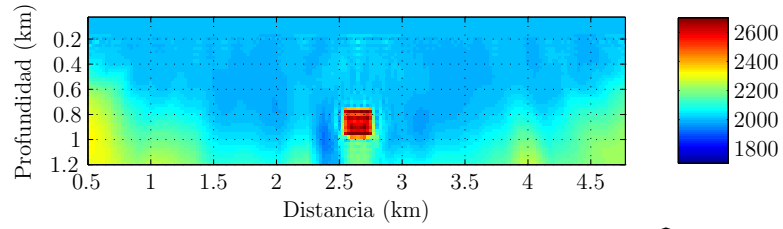


Figura 36: FWI usando \mathbf{m}_{sub} con métricas \widehat{CS}_{TC} .

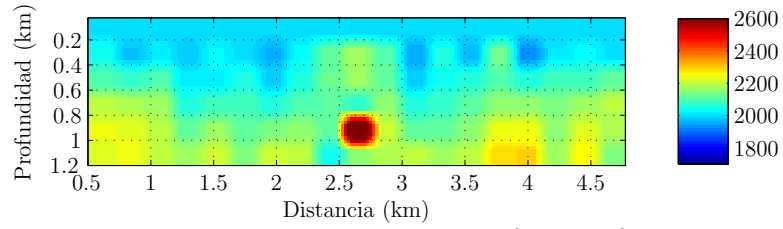


Figura 37: \mathbf{m}_{sub} obtenido de $\widehat{CS}_{FT} + \widehat{CS}_{TC}$.

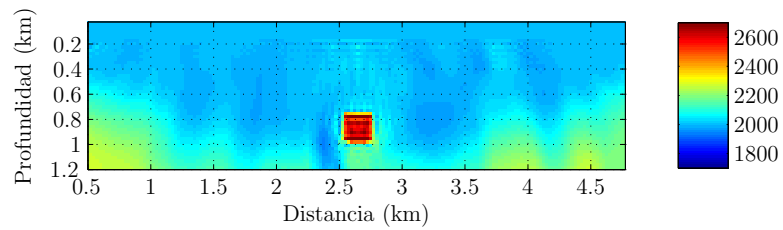


Figura 38: FWI usando \mathbf{m}_{sub} con métricas $\widehat{CS}_{FT} + \widehat{CS}_{TC}$.

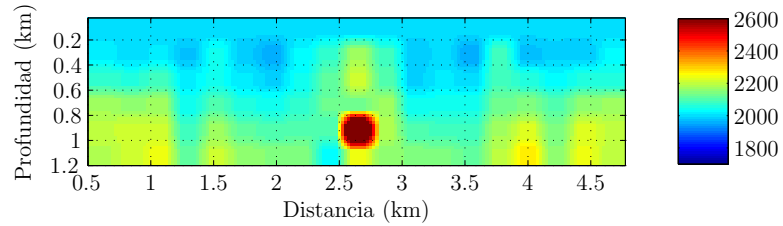


Figura 39: Modelo estimado mediante $\widehat{CS}_{FT} + \widehat{C}_{corr} + \widehat{CS}_{TC}$.

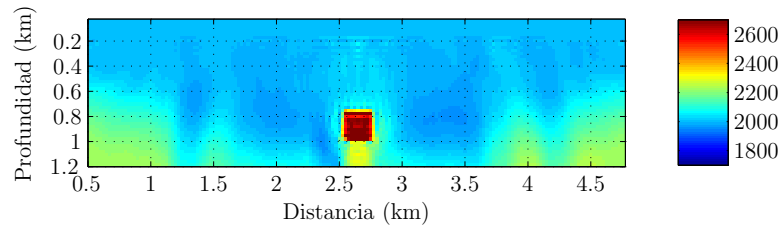


Figura 40: FWI usando \mathbf{m}_{sub} con métricas $\widehat{CS}_{FT} + \widehat{C}_{corr} + \widehat{CS}_{TC}$.

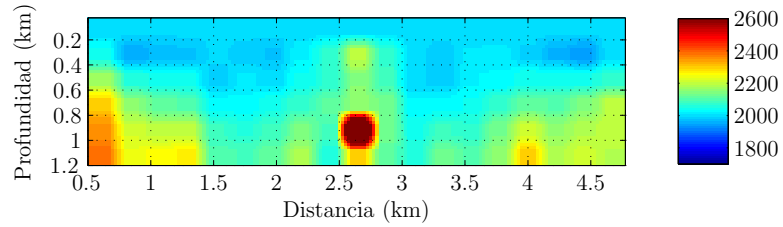


Figura 41: Modelo estimado mediante $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$.

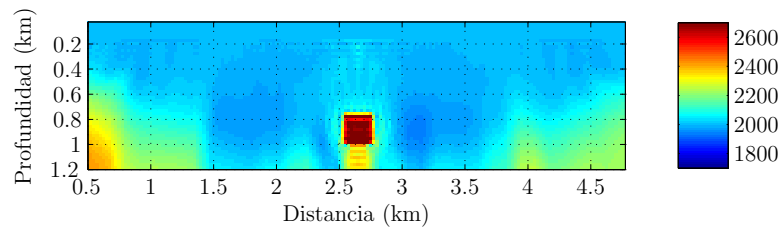


Figura 42: FWI usando \mathbf{m}_{sub} con métricas $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$.

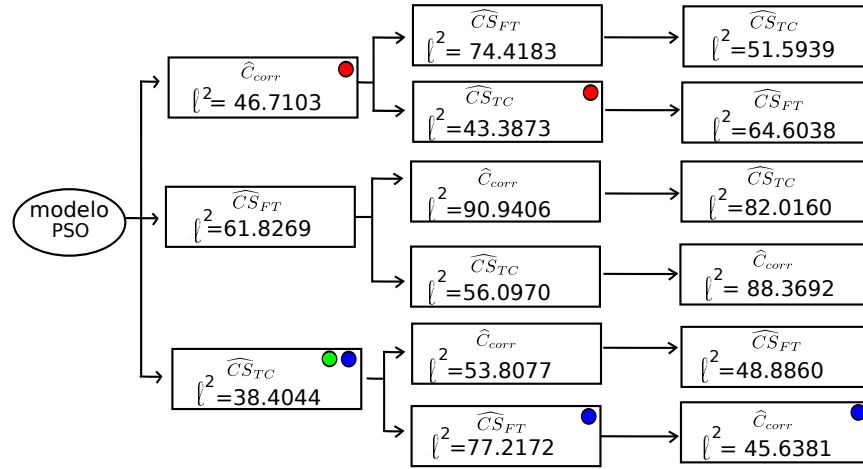


Figura 43: Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -cuadrado difractor)

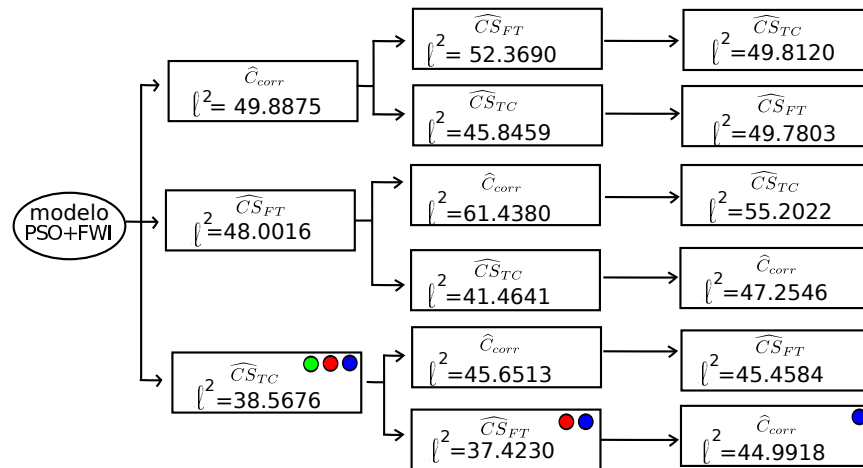


Figura 44: Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).

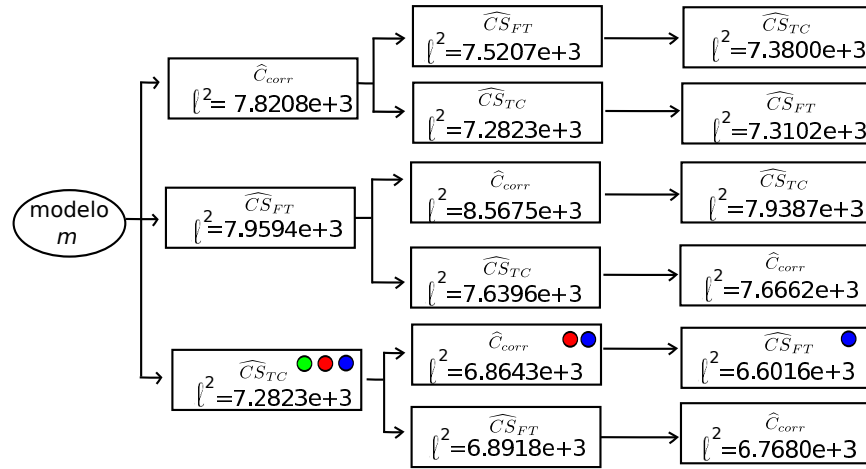


Figura 45: Diagrama en cascada usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).

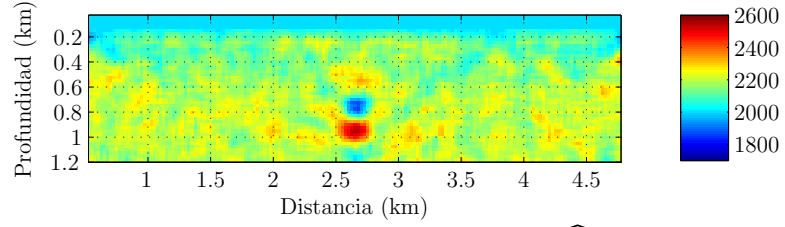


Figura 46: \mathbf{m}_{full} obtenido de \widehat{CS}_{TC}

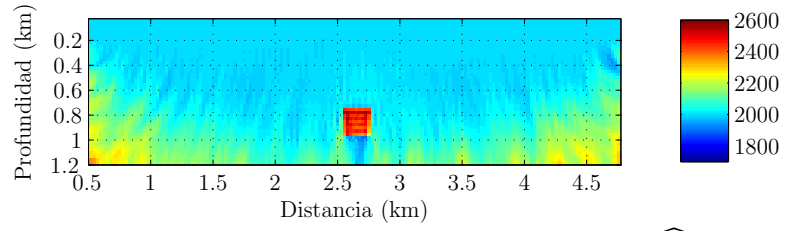


Figura 47: FWI usando \mathbf{m}_{full} con métricas \widehat{CS}_{TC}

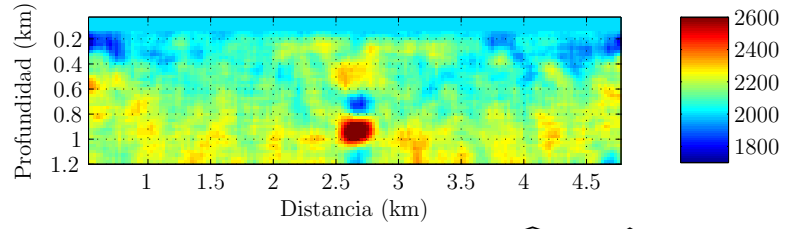


Figura 48: \mathbf{m}_{full} obtenido de $\widehat{CS}_{TC} + \widehat{C}_{corr}$

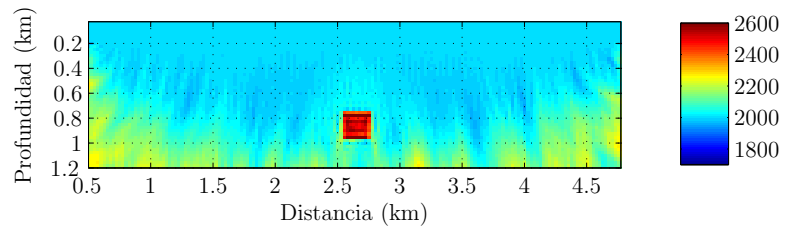


Figura 49: FWI usando \mathbf{m}_{full} con métricas $\widehat{CS}_{TC} + \widehat{C}_{corr}$

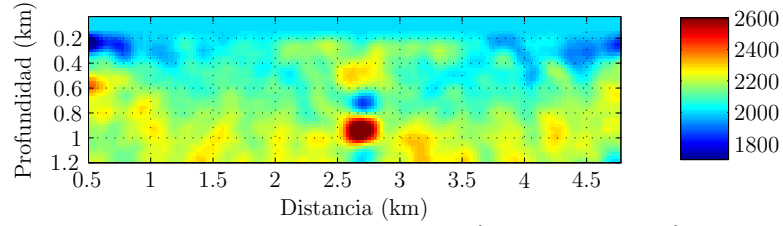


Figura 50: \mathbf{m}_{full} obtenido de $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$

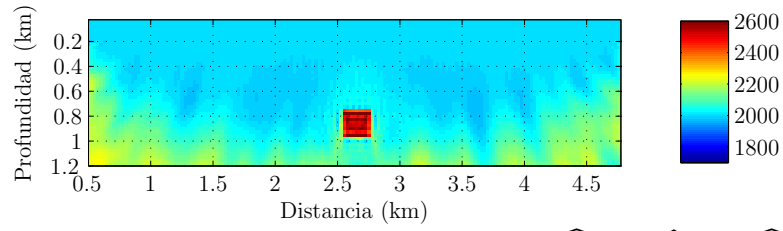


Figura 51: FWI usando \mathbf{m}_{full} con métricas $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$

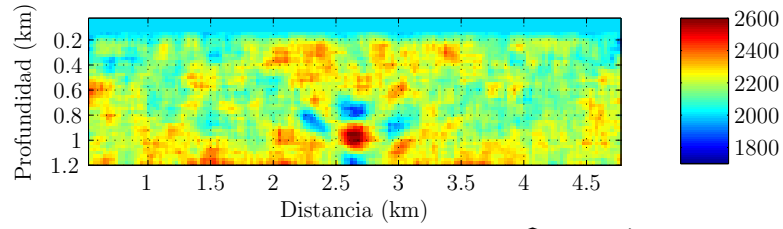


Figura 52: \mathbf{m}_{full} obtenido de $\widehat{CS}_{FT} + \widehat{C}_{corr}$

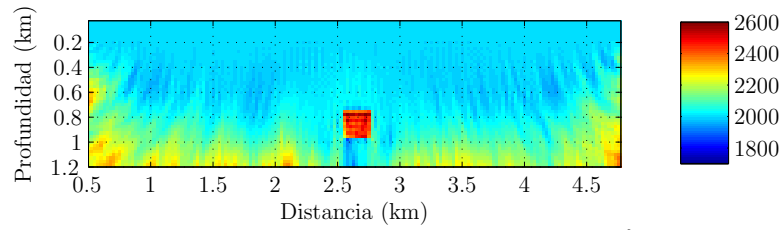


Figura 53: FWI usando \mathbf{m}_{full} con métricas $\widehat{CS}_{FT} + \widehat{C}_{corr}$

4. ANÁLISIS ESTADÍSTICO DE LOS DATOS

En este Capítulo se emplea el test de Kolmogorov-Smirnov la función de densidad de probabilidad que se ajuste a los datos obtenidos de cada métrica ???. Las funciones de densidad de probabilidad empleadas son: la Gaussiana, la uniforme y la Laplaciana. Una vez definida la función de densidad de probabilidad, se calculan los estimadores que mejor describen los datos. El modelo del cuadrado difractor presentado en la Figura 21 (a) será utilizado como ejemplo para presentar la metodología propuesta para el análisis estadístico.

4.1. Seleccionar el número de experimentos

El análisis estadístico requiere de una muestra significativa de la población para analizar la convergencia de los experimentos obtenidos de la metaheurística. La muestra \mathbf{x}_m será la proyección de un pixel sobre un cubo construido según el número de experimentos de cada métrica (ver Figura 28). Los pixeles en una vecindad no son independientes e idénticamente distribuidos(i.i.d), ya que existe una correlación espacial entre pixeles adyacentes, como se demuestra en el Capítulo 6. Sin embargo, se asumirá que todos los pixeles son (i.i.d). La ubicación del pixel seleccionado se presenta sombreado en la Figura 54.

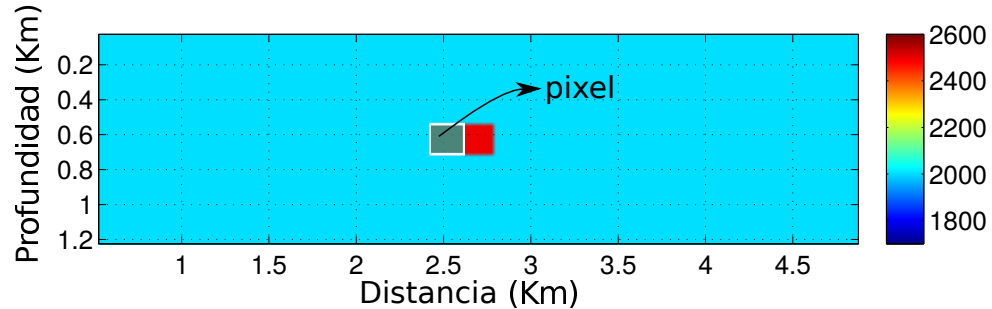


Figura 54: Ubicación del pixel para el análisis estadístico

Es importante recordar que para obtener un experimento se necesita de 500 partículas y 100 iteraciones. Para determinar el número de experimentos n requerido por el análisis

estadístico, se propone variar n en valores de 2, 10, 50, 100, 150 y 200. Sobre cada valor de n se construye un cubo donde se mide la varianza de la muestra \mathbf{x}_m , donde el subíndice m indica alguna de las métricas establecidas en el capítulo 2 (correlación (x_{corr}), CS en el dominio de la frecuencia (\mathbf{x}_{FT}) y CS en el dominio de la traza compleja (\mathbf{x}_{TC})). El objetivo es encontrar la convergencia del valor esperado (varianza) a medida que el número de experimentos crece [24]. En la Tabla 11 y la Figura 55 se presentan los valores de varianza obtenidos para las métricas en cada dominio. Se encontró de esta manera que $n = 100$ garantiza un cambio de varianza menor al 1 % en las tres métricas, tal como se presenta en la Figura 55 (b).

n° de experimentos	Varianzas con \mathbf{m} en el dominio del Tiempo ($\times 10^5$)	Varianzas con \mathbf{m} en el dominio de Hilbert ($\times 10^5$)	Varianza con \mathbf{m} en el dominio de la Frecuencia ($\times 10^5$)
2	2.4965	7.4995	9.4900
10	1.6400	1.7854	3.7168
50	1.1581	1.1811	1.5939
100	1.1230	1.1460	1.5381
150	1.0808	1.0386	1.5729
200	1.0486	1.0215	1.5797

Tabla 11: Varianzas de cada métrica en función del número de experimentos.

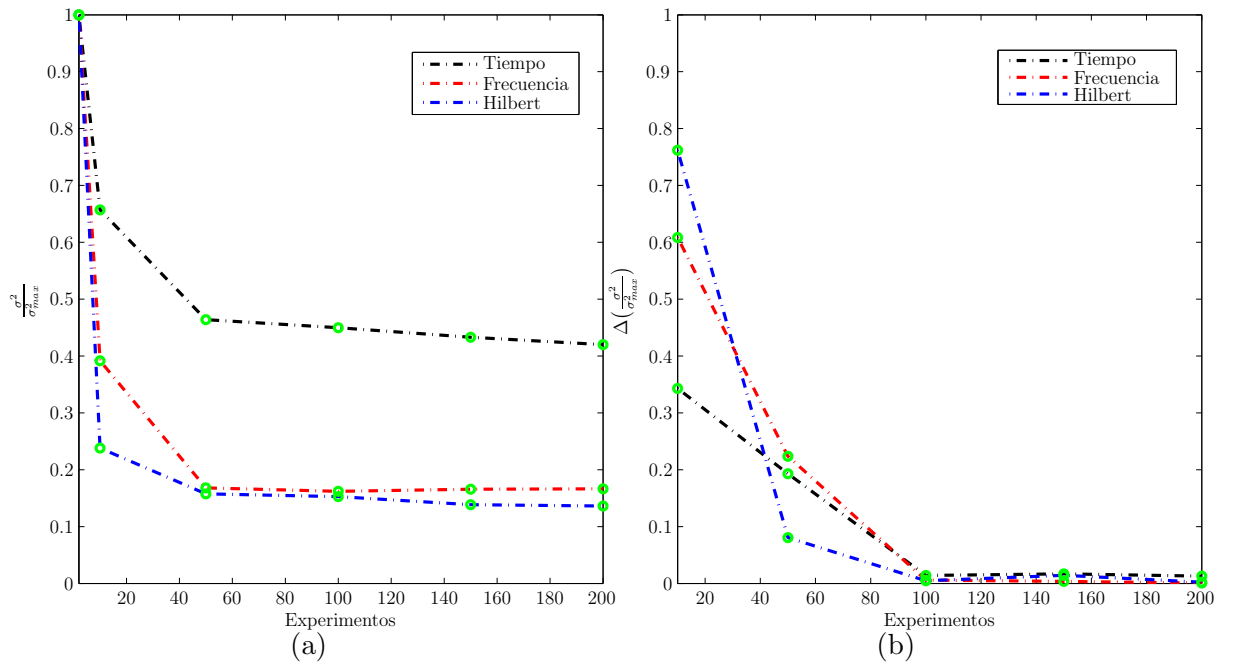


Figura 55: (a) Varianzas normalizadas de las métricas, (b) Tasa de cambio de las varianzas normalizadas.

Definiendo el número de experimentos $n = 100$, donde un experimento tiene 500 partículas y 100 iteraciones, se debe determinar sobre qué función de densidad de probabilidad se ajusta mejor cada muestra (\mathbf{x}_m).

4.2. Selección de la función de distribución acumulada

Se prueba el test no paramétrico de Kolmogórov-Smirnov (K-S) que consiste en comparar, dos distribuciones de probabilidad, siempre y cuando se tengan momentos de orden uno finitos [25]. Esta condición se cumple para las tres funciones de probabilidad seleccionadas en este documento (Gaussiana, uniforme y Laplaciana). La prueba K-S consiste en calcular la distancia entre la función de distribución acumulada empírica ($F_n(\mathbf{x}_m)$) con la función de distribución acumulada elegida ($F(\mathbf{x}_m)$). La distancia entre estas dos funciones de distribución es

$$D_m = \max\{|F_n(\mathbf{x}_m) - F(\mathbf{x}_m)|\}, \quad (4.1)$$

Es importante recordar que la muestra \mathbf{x}_m depende de la métrica y por tanto se obtendrán tres valores de D_m , D_{corr} , D_{TF} , D_{TC} . Por otra parte, $F_n(\mathbf{x}_m)$ es definida como

$$F_n(\mathbf{x}_m) = \frac{1}{n} \sum_{i=1}^n \begin{cases} 1 & \text{si } \mathbf{x}_i \leq \mathbf{x}_m \\ 0 & \text{en caso contrario,} \end{cases}$$

donde n es el tamaño de la muestra, \mathbf{x}_i es el i -ésimo valor observado en la muestra \mathbf{x}_m , cuyos valores han sido ordenados previamente de menor a mayor. Note que $F_n(\mathbf{x}_m)$ es una función escalón que incrementa en pasos de $1/n$. $F(\mathbf{x}_m)$ es obtenida a partir de la distribución de probabilidad que se especifica como hipótesis nula. La hipótesis nula no se rechaza si se cumple que

$$\sqrt{n}D_m < K_\alpha, \quad (4.2)$$

donde K_α es obtenido de [26]. Se debe elegir un nivel de significancia α para juzgar si los resultados son estadísticamente significativos. Se toma $\alpha = 0,05$, de tal forma que máximo el 5% de los datos estén fuera del máximo error permitido para aceptar la hipótesis nula y con este nivel de significancia obtenemos un $K_\alpha = 1,36$ tal como se presenta en [26]. Dado que de 4.2 se tiene que

$$D_m < \frac{K_\alpha}{\sqrt{n}} = 0,136. \quad (4.3)$$

de lo cual se sigue que la máxima diferencia tolerada para no rechazar la hipótesis nula en las tres métricas es $D_m = 0,136$.

4.2.1. Función de densidad de probabilidad Gaussiana

La primera función de densidad de probabilidad evaluada es la Gaussiana. Se denomina función de densidad de probabilidad normal o Gaussiana a una distribución de probabilidad de variable continua que frecuentemente ocurre en fenómenos reales. Su función de densidad de probabilidad es definida como

$$PDF = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_m - \mu_m}{\sigma_m}\right)^2} \quad (4.4)$$

donde μ_m y σ_m son los valores de media y desviación estándar obtenidos de la muestra de cada métrica [27]. La distribución normal presenta simetría con respecto a un parámetro estadístico y su función de distribución acumulada es definida por

$$F(\mathbf{x}_m) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\mathbf{x}_m - \mu_m}{\sigma_m \sqrt{2}} \right) \right] \quad (4.5)$$

donde erf es la función error. Aplicando el test de K-S a través de 4.3 y tomando la función de distribución acumulada normal para cada una de las métricas \mathbf{x}_m , se obtuvieron los resultados presentados en la siguiente tabla

Métrica	D_m
\widehat{C}_{corr}	0.10748
\widehat{CS}_{FT}	0.04181
\widehat{CS}_{TC}	0.09207

Tabla 12: K-S test sobre la función de distribución acumulada normal

En la Figura 56 se presentan los ajustes entre la función de distribución acumulada empírica contra la función de distribución acumulada normal para las muestras x_m de cada una de las métricas, no rechazándose la hipótesis nula en todos los casos.

4.2.2. Función de densidad de probabilidad uniforme

La Segunda función de densidad de probabilidad evaluada es la uniforme. Es una familia de distribuciones de probabilidad para variables aleatorias continuas, en donde todas las muestras son igualmente probables. Su función de densidad de probabilidad está definida por

$$F_n(\mathbf{x}_m) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq \mathbf{x}_m \leq b. \\ 0 & \text{en caso contrario,} \end{cases}$$

donde a y b son el valor máximo y mínimo de cada muestra (\mathbf{x}_m) respectivamente [27]. La función de distribución acumulada es definida

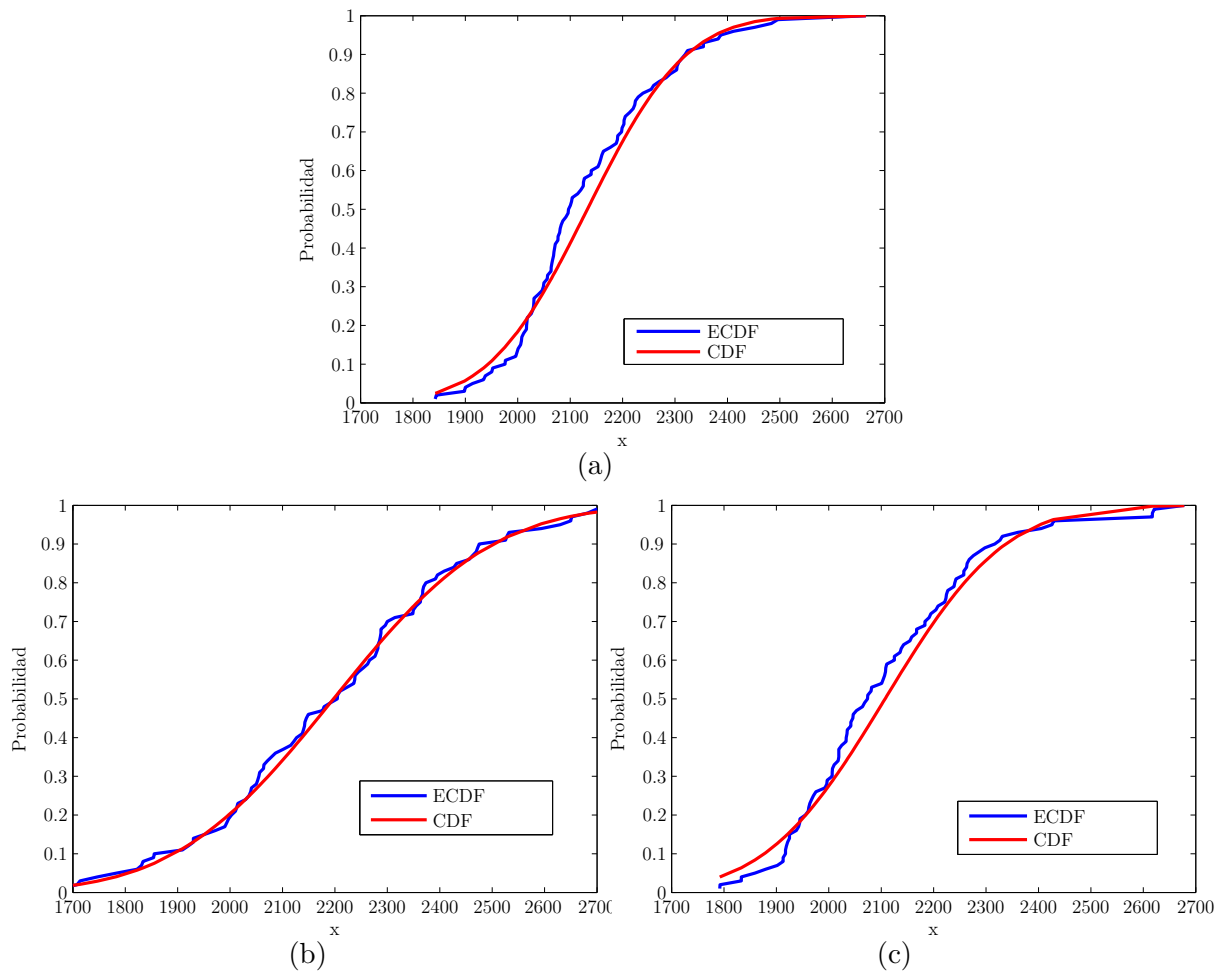


Figura 56: Ajuste entre función de distribución acumulada empírica con la función de distribución acumulada normal para las métricas: a) correlación, b) \widehat{CS}_{FT} , c) \widehat{CS}_{TC} .

$$F_n(\mathbf{x}_m) = \begin{cases} 0 & \text{si } \mathbf{x}_m < a \\ \frac{\mathbf{x}_m - a}{b - a} & \text{si } a \leq \mathbf{x}_m \leq b \\ 1 & \text{si } \mathbf{x}_m \geq b, \end{cases}$$

Al aplicar la prueba K-S con la función de distribución uniforme acumulada a cada muestra de cada métrica, se obtuvieron los siguientes resultados

Métrica	D_m
\widehat{C}_{corr}	0.15075
\widehat{CS}_{FT}	0.10357
\widehat{CS}_{TC}	0.14975

Tabla 13: K-S test sobre la función de distribución uniforme acumulada.

En la Figura 57 se presenta los ajustes entre la función de distribución acumulada empírica contra la función de distribución uniforme acumulada para las muestras \mathbf{x}_m de cada una de las métricas, no rechazándose la hipótesis nula para la métrica \widehat{CS}_{FT} .

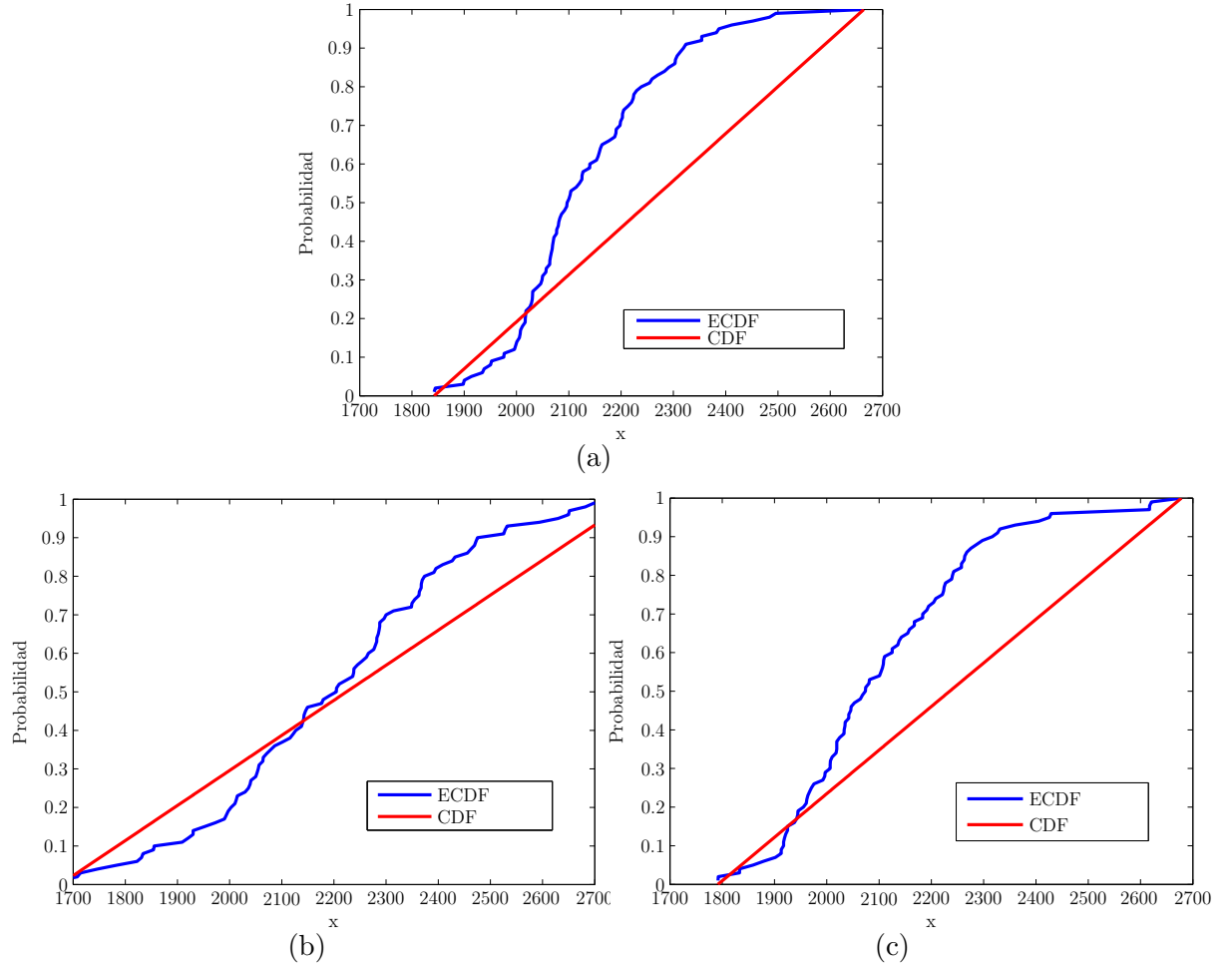


Figura 57: Ajuste entre función de distribución acumulada empírica con la función de distribución uniforme acumulada para las métricas: (a) correlación, (b) \widehat{CS}_{FT} , (c) \widehat{CS}_{TC} .

4.2.3. Función de densidad de probabilidad Laplaciana

La tercera función de densidad de probabilidad evaluada es la Laplaciana. La cual es una función de densidad de probabilidad continua, también conocida como distribución doble exponencial puesto que puede ser considerada como la relación de las densidades de dos distribuciones exponenciales. Su función de densidad de probabilidad es definida como

$$F(\mathbf{x}_m) = \frac{1}{2b} e^{(-\frac{|\mathbf{x}_m - \mu_m|}{b})}, \quad (4.6)$$

donde μ_m es la media de los datos para cada muestra y b es el parámetro de escala [27]. Su función de distribución uniforme acumulada esta definida por

$$F_n(\mathbf{x}_m) = \begin{cases} \frac{1}{2}e^{(-\frac{\mu_m - \mathbf{x}_m}{b})} & \text{si } \mathbf{x}_m < \mu_m \\ 1 - \frac{1}{2}e^{(-\frac{\mathbf{x}_m - \mu_m}{b})} & \text{si } \mathbf{x}_m \geq \mu_m, \end{cases}$$

Aplicando el test de K-S y tomando la función de distribución acumulada Laplaciana de cada muestra obtenida de las métricas, se obtuvieron los siguientes resultados

Métrica	D_m
\widehat{C}_{corr}	0.32406
\widehat{CS}_{FT}	0.17157
\widehat{CS}_{TC}	0.14974

Tabla 14: K-S test sobre la función de distribución uniforme acumulada Laplaciana.

En la Figura 58 se presentan los ajustes entre la función de distribución acumulada empírica contra la función de distribución acumulada Laplaciana para las muestras x_m de cada una de las métricas, rechazándose en todos los casos la hipótesis nula.

La Tabla 15 resume los parámetros obtenidos del test de K-S. Los recuadros verdes indican las pruebas donde no ha sido posible rechazar la hipótesis nula, los recuadros rojos indican que se rechaza la hipótesis nula. Es importante resaltar que después de aplicar el test de Kolmogorov-Smirnov sobre las tres funciones de distribución acumulada, donde no ha sido posible rechazar la hipótesis nula para la distribución normal en las tres métricas. Se asume que los datos tiene una distribución normal.

Dominio \ D_m	Normal	Uniforme	Laplaciana
Tiempo	0.10748	0.15075	0.32406
Frecuencia	0.04181	0.10357	0.17157
Hilbert	0.09207	0.14975	0.14974

Tabla 15: Resumen de la prueba K-S sobre cada métrica para las funciones de distribución acumulada

4.3. Estimadores

Una vez se define una función de densidad de probabilidad, el siguiente paso es encontrar los estimadores que mejor representen los datos. Un estimador es un estadístico que provee una completa caracterización del parámetro a través de una función de

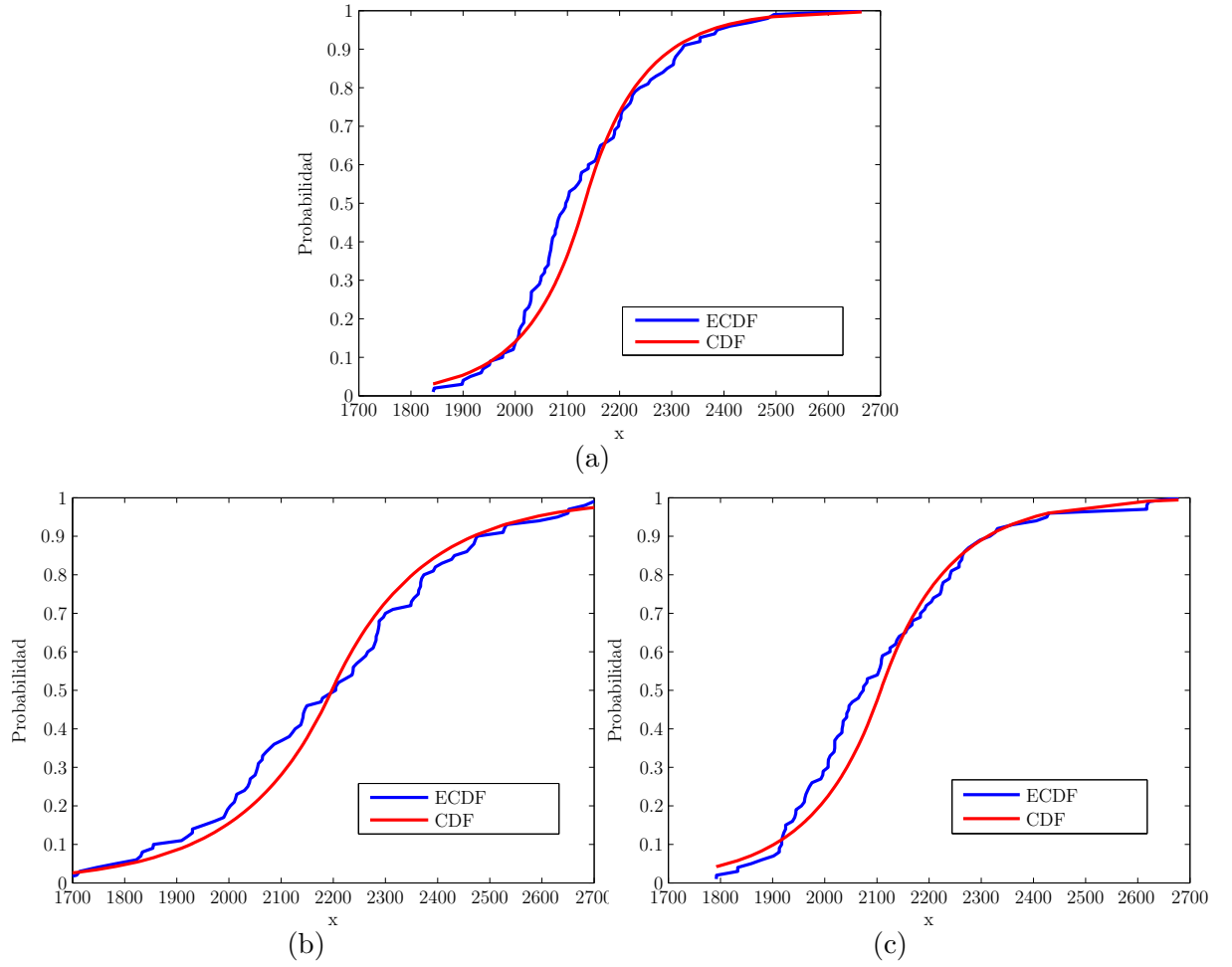


Figura 58: función de distribución acumulada usando una distribución Laplaciana para las métricas: a) correlación, b) \widehat{CS}_{FT} , c) \widehat{CS}_{TC} .

distribución. Se obtuvieron los siguientes estimadores: Primero, el estimador de máxima verosimilitud (Maximum Likelihood, ML), el cual brinda un procedimiento para calcular los parámetros de un modelo probabilístico de tal manera que sean los más probables a partir de los datos obtenidos. Esta definido como

$$\hat{\theta}_{ML} = \arg \max_{\theta} P(\mathbf{x}_m | \theta), \quad (4.7)$$

donde $P(\mathbf{x}_m | \theta)$ se llama la probabilidad condicional de θ aportada por los datos y θ son los parámetros de los valores aleatorios. El objetivo es encontrar un valor de θ entre todos los posibles que maximiza la función de probabilidad conjunta descrita en 4.7.

El siguiente estimador es el de máximo *a-posteriori* (Maximum A Posteriori, MAP), el cual es posible obtenerlo cuando tenemos información *a-priori* [27]. Partiendo de

inferencia bayesiana tenemos que

$$P(\theta|\mathbf{x}_m) = \frac{P(\mathbf{x}_m|\theta)P(\theta)}{P(\mathbf{x}_m)}, \quad (4.8)$$

donde $P(\theta|\mathbf{x}_m)$ es llamada la probabilidad *a-posteriori*, $P(\theta)$ es la probabilidad *a-priori* y $P(x)$ la probabilidad marginal [27]. De la inferencia bayesiana tenemos que el máximo *a-posteriori* esta definido como

$$\hat{\theta}_{MAP} = \arg \max_{\theta} P(\theta|\mathbf{x}_m) \propto \arg \max_{\theta} P(\mathbf{x}_m|\theta)P(\theta). \quad (4.9)$$

Finalmente, se tiene el estimador del mínimo error cuadrático medio (Minimum Mean Squared Error, MMSE), obtenido de

$$\hat{\theta}_{MMSE} = \int \theta P(\theta|\mathbf{x}_m) d\theta. \quad (4.10)$$

En la Tabla 16 se presentan los estimadores de media y varianza para ML, MAP y MMSE. Es importante recordar que estos estimadores son también obtenidos del pixel presentado en la Figura 54 para cada una de las métricas. Con los estimadores obtenidos (ML, MAP y MMSE) se construyen las funciones de densidad de probabilidad para determinar el grado de ajuste con los datos. Como se aprecia en la Figura 59 no existe mayores diferencias entre la selección del tipo de estimador.

Estimador \ Dominio	Tiempo	Frecuencia	Hilbert
μ_{ML}	2.1328e+03	2.1973e+03	2.1072e+03
μ_{MAP}	2.1328e+03	2.1973e+03	2.1072e+03
μ_{MMSE}	2.1328e+03	2.1973e+03	2.1072e+03
σ_{ML}^2	147.0682	236.7654	178.8374
σ_{MAP}^2	146.3383	235.5904	177.9499
σ_{MMSE}^2	147.8091	237.9582	179.7383

Tabla 16: Estimadores para las tres métricas presentadas

4.4. Matriz de varianzas *a-priori*

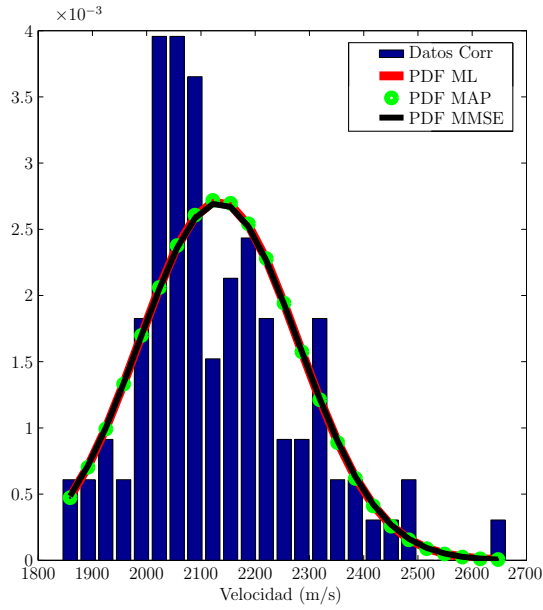
Como la hipótesis nula no fue rechazada para la función de distribución acumulada normal, se construye un cubo de 100 experimentos como se presentó en la Figura 28.

Se proyecta cada pixel en la dirección de los experimentos, permitiendo calcular las varianzas de la forma

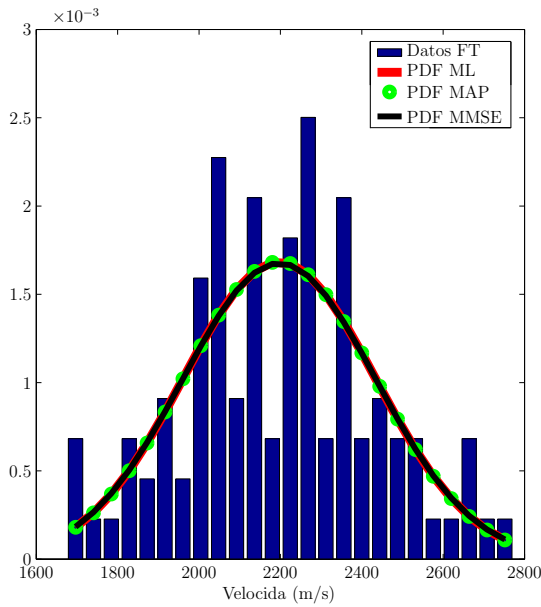
$$\sigma_n^2 = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_m^i - \bar{\mathbf{x}}_m)^2 \quad (4.11)$$

donde $\bar{\mathbf{x}}_m$ es la media aritmética de los datos, n el número de datos (100 experimentos) y \mathbf{x}_m^i cada dato. Los resultados se presentan en escala normal en las Figuras 60, 61 y 62; y en escala logarítmica en las Figuras 63, 64 y 65.

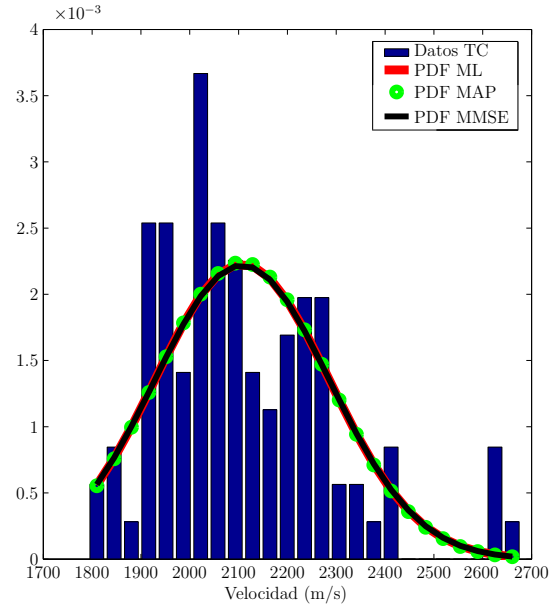
Se concluye que las técnicas metaheurísticas como PSO, permiten obtener los suficientes experimentos sobre los cuales hacer un análisis estadístico de la información. Todas las métricas se ajustaron a una distribución de probabilidad que junto con los estimadores y un análisis de varianzas, permite obtener una herramienta capaz de brindar la información *a-priori* requerida para procesos de cuantificación de la incertidumbre [28].



(a)



(b)



(c)

Figura 59: Construcción de la función de densidad de probabilidad de cada estimador para las métricas de: (a) correlación, (b) \widehat{CS}_{FT} , (c) \widehat{CS}_{TC}

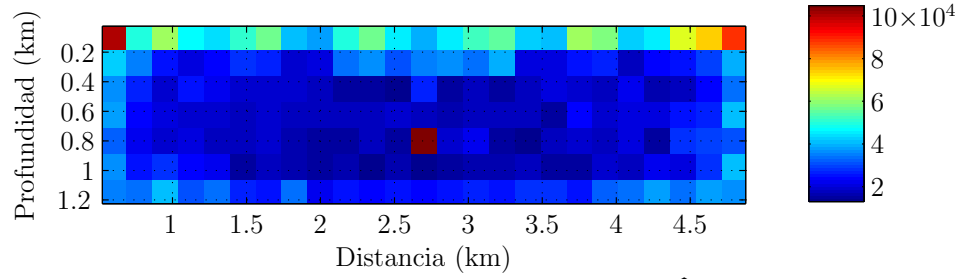


Figura 60: Varianzas con métrica \hat{C}_{corr} .

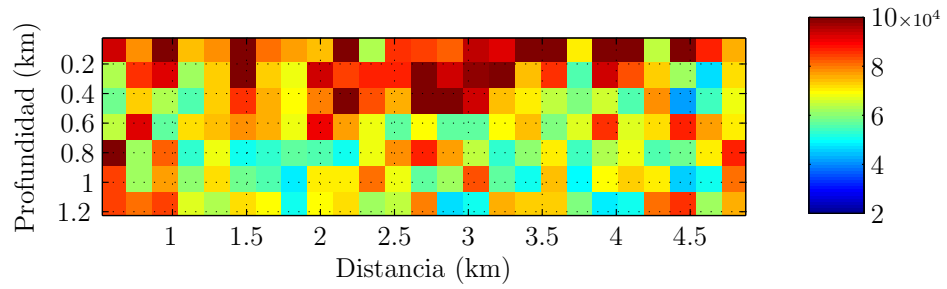


Figura 61: Varianzas con métrica \hat{CS}_{FT} .

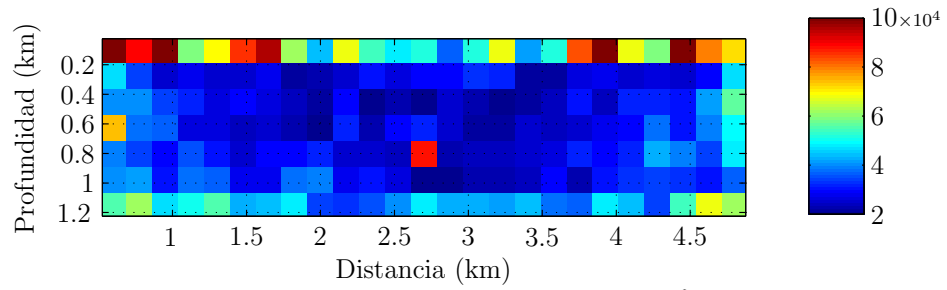


Figura 62: Varianzas con métrica \hat{CS}_{TC} .

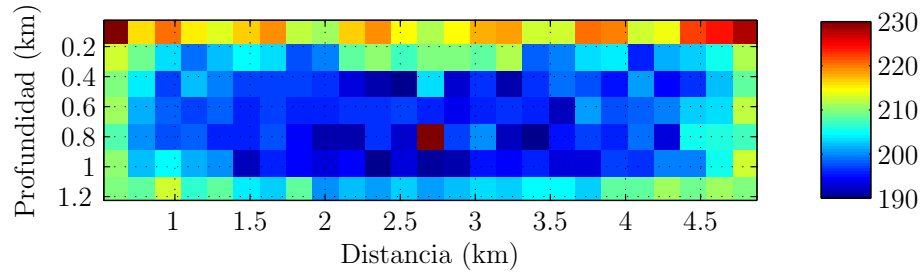


Figura 63: $20 \cdot \log(\cdot)$ varianzas con métrica \hat{C}_{corr} .

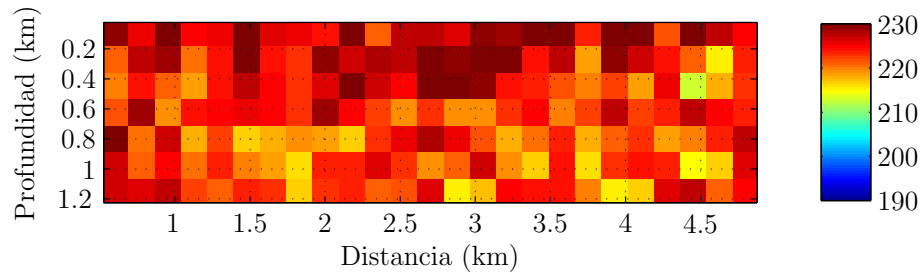


Figura 64: $20 \cdot \log(\cdot)$ Varianzas con métrica \hat{C}_{ST} .

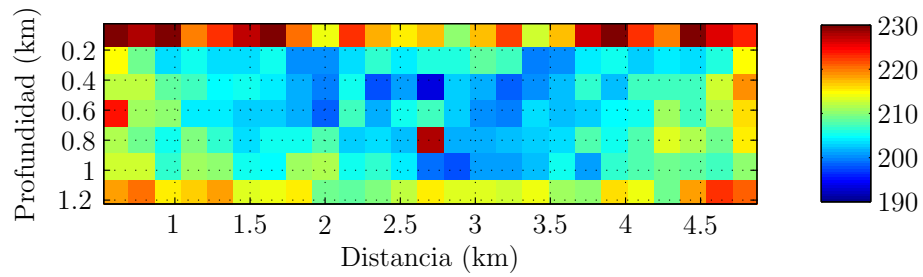


Figura 65: $20 \cdot \log(\cdot)$ varianzas con métrica \hat{C}_{STC} .

5. CONCLUSIONES

En este capítulo se presentan las conclusiones desarrolladas en este trabajo de investigación. La pregunta de investigación desarrollada en este trabajo de investigación era determinar si es viable obtener un punto de partida favorable en términos del CS, usando diferentes dominios transformados para una técnica de optimización global. Por otra parte, se logro establecer que las métricas se enfocan en solucionar las componentes de baja frecuencia (ver Figura 31).

Se pudo probar que es viable obtener un punto de partida favorable en términos de contener la información de baja frecuencia (en terminos del cycle skipping) usando una técnica de optimización global (PSO) con un esquema multi-métrica.

La metodología propone emplear tres posibles combinaciones según la cantidad de métricas a utilizar. Las combinaciones de métricas para los modelos del cuadrado difractor, 2 capas y marmousi difieren. Esto es debido a la información de baja frecuencia proporcionada por cada modelo. Algunos modelos pueden ser mejor caracterizados por alguna métrica que por las otras.

Se encontró que el CS visto desde el dominio de Hilbert es una mejor métrica que la correlación en dominio del tiempo en términos del porcentaje de CS en las trazas sísmicas. La métrica del CS, vista desde el dominio de Hilbert, da mejores resultados que la métrica de correlación debido que permite tener en tiempo tanto la información de magnitud como la información de fase; mientras que la correlación solo nos permite ajustar las fases. Por otra parte, se encontró que para proponer otras posibles métricas que midan el CS, los dominios transformados no deben perder el dominio del tiempo.

Inicialmente, se tomaron los parámetros de PSO revisando el estado-del-arte. En el contexto geofísico Ranjit y Shalichan en [12] proponen una inversión de reflectividad en una dimensión con los parámetros $w = 0,729$, $c_1 = 2,8$ y $c_2 = 3,1$. Sin embargo, se diseñó una metodología para estimar los parámetros w , c_1 y c_2 para invertir velocidad. De estos experimentos se obtuvieron los valores: $w = 0,7$, $c_1 = 0,8$ y $c_2 = 1,1$.

Se logró establecer una metodología para determinar el número de iteraciones y partículas, dando como resultado 100 iteraciones y 500 partículas para 14348 dimensiones. Los ajustes del número de iteraciones y partículas están sujetas a cambios según las dimensiones del problema. Por otra parte, al ser PSO una técnica heurística se encontró que con 100 experimentos se tienen de experimentos necesarios para obtener una varianza menor al 1 %, como se presentó en la Figura 55 (b).

En este trabajo de investigación sólo se emplea modelos sintéticos. El objetivo es llevar esta herramienta a datos reales y para esto se requiere modificar el operador que modela

la ecuación de onda 2D acústica y de densidad constante. Según el grado de complejidad del operador d' Alembert y en la medida que se pueda modelar más eventos sísmicos se podrá llegar a modelos más complejos y más cercanos a la realidad del subsuelo. Sin embargo, cambiar a un operador más cercano a la realidad implica elevar el costo computacional.

Al aplicar la FWI sobre el modelo inicial se desconoce si el resultado obtenido de la FWI es el correcto. Al tratarse de un problema inverso mal puesto, no existe unicidad en la solución y de varios modelos se pueden obtener las mismas trazas. Existen técnicas para estimar la incertidumbre de la solución pero como es un problema mal condicionado, se requiere de información *a-priori*. La metodología usada en el análisis estadístico de esta tesis puede ser usadas como información *a-priori* para la cuantificación de la incertidumbre y de esta forma hacer un análisis estadístico de los datos, brindando zonas de confianza para explotación.

Se generó una herramienta que permite obtener el punto de partida sin ningún conocimiento *a-priori* del subsuelo. Sin embargo, se puede guiar la búsqueda de la metaheurística en ciertas zonas con información *a-priori* suministrada por una tomografía o una migración. Esta adaptación podría generar mejoras significativas sobre las soluciones obtenidas hasta este momento. Sin embargo, se debe tener precaución de no estar sesgado por esta información *a-priori* ya que podría estar indicando un mínimo local y no el mínimo global.

La metodología se probó sobre tres modelos sintéticos de 14.348 dimensiones; sin embargo, este modelo es pequeño en comparación a un modelo real donde fácilmente alcanzamos dimensiones de 80 millones (10 km × 8 km). Para hacer viable el dimensionamiento en un dato real se propuso programar en paralelo la forma en como se sacudía cada dimensión tal como se presentó en (1.12).

Las propagaciones para obtener tanto el dato observado como el dato modelado son obtenidas usando una GPU en lenguaje CUDA-C. Con la mejora de la propagación y el dimensionamiento de las partículas se logró obtener una reducción del costo computacional de 6 veces en comparación con el código en lenguaje C.

Programar en paralelo el dimensionamiento de PSO permitirá utilizar esta metodología para generar puntos de partida sintéticos tridimensionales. En la Tabla 17 se presentan los recursos de RAM utilizados en GPU para una propagación usando (5.1). Los datos de RAM son obtenidos de

$$RAM_{3D} = \frac{\beta}{1024^2 \times 8} \times (N_s \times N_t \times N_x \times N_y + 11 \times N_x \times N_y \times N_z + N_t + 4 \times N_x + 4 \times N_y + 4 \times N_y + 4 \times N_z + 5 \times N_t \times N_x \times N_y), \quad (5.1)$$

donde N_x es la dimensiones del modelo en la dirección x , N_y es la dimensiones del modelo en la dirección y , N_z es la dimensiones del modelo en la dirección z , N_t es

el número de pasos en tiempo y β es la precisión (single o double). Las dimensiones del punto de partida tridimensionales sintéticos pueden crecer tanto, como quepa una propagación en RAM de GPU. Es importante mencionar que en la pruebas se uso una Tesla K40c con 12 GB de RAM.

Dimensiones	RAM GPU (MB)	% de RAM usada
175 (\mathbf{m}_{sub})	100.5467	8.3789e-04
14348 (\mathbf{m}_{full})	135.8234	0.0011
1442280 (\mathbf{m}_{3d}^0)	196.0315	0.0016

Tabla 17: Recursos de RAM en GPU en función de las dimensiones del modelo (MB).

En esta metodología se programo en paralelo tanto el dimensionamiento de PSO como la propagación de onda 2D acústica y de densidad constante. Sin embargo aún el algoritmo puede mejorar desde su programación en paralelo y para esto se proponen las siguientes alternativas: tomar las partículas de manera independiente por iteración, cada partícula estará haciendo al tiempo el ciclo de solución planteado en la Figura 18; programar en paralelo las métricas, de las tres métricas usadas, la métrica en el dominio de Hilbert es el caso más critico donde un experimento puede tardar hasta 96 horas en tiempo de ejecución; una última alternativa es usar un cluster multi-GPU usando un estándar de comunicación llamado MPI (Message Passing Interface). El tiempo de reducción del costo computacional de esta última alternativa es casi proporcional al numero de GPUs usadas en el cluster.

La metodología para obtener el punto de partida presentada en este trabajo puede ser empleada en cualquier contexto de inversión no lineal que utilice un operador lineal: por ejemplo, en el contexto geofísico la técnica GPR (por sus siglas en inglés, ground penetrating radar), permite caracterizar la zona somera para poder brindar más detalle de lo que hay a mayores profundidades.

Utilizar herramientas heurísticas como PSO proporciona un conjunto de experimentos que permiten obtener funciones de densidad de probabilidad y estimadores, sobre los cuales hacer una análisis estadístico. En nuestro caso se prueba que en todas las métricas se ajustan a un función de distribución acumulada normal como se presento en la tabla 15.

Se parte de la hipótesis que PSO garantiza la información de baja frecuencia (< 6 Hz) y la FWI las altas frecuencias (> 6 Hz). La Figura 31 para el modelo del cuadrado difractor, la Figura 70 para el modelo 2 capas y la Figura 71 para el modelo Marmousi, permiten corroborar que la información de baja frecuencia (< 6 Hz), son las componentes frecuenciales donde las métricas concentran su solución. Esto permite corroborar esta hipótesis.

6. TRABAJO FUTURO

En este capítulo se presentan algunas mejoras que se pueden desarrollar a partir de este trabajo de investigación. Desde los parámetros de PSO se ajustaron w , c_1 y c_2 , para un problema de inversión en velocidad en el contexto geofísico. Sin embargo, estos parámetros no son los óptimos y es necesario de otras herramientas para la sintonización de los parámetros. En el trabajo se seleccionaron 3 dominios transformados, pero aún es posible explorar otros dominios donde no se pierda la línea temporal. Por otra parte, se planteó activar un dominio al tiempo y se propone activar más de un dominio a la vez. Desde el punto de vista del costo computacional, aún es posible mejorar la independencia de cada partícula así como la métrica en el dominio de Hilbert. Finalmente, se demuestra la dependencia entre pixeles y por tanto estudios de probabilidad mas avanzados para analizar los experimentos obtenidos.

6.1. Ajuste de parámetros PSO

Inicialmente se seleccionaron las constantes c_1 , c_2 , w y Δt de PSO con el estado-del-arte para inversión de reflectividad. Sin embargo, los ajustes de w , c_1 y c_2 pueden cambiar según el contexto o la aplicación. En el contexto geofísico para inversión de velocidad, el mejor ajuste encontrado en los parámetros es $w = 0,7$, $c_1 = 0,8$ y $c_2 = 1,1$ tal como se describió en el capítulo 3. Sin embargo, estos parámetros no son los óptimos. Para ajustar los parámetros en el contexto geofísico se propone realizar un conjunto de experimentos con sustento estadístico. Algunos autores proponen iniciar con c_1 y c_2 igual a cero y variar w . Después de ajustar w empleando el principio de Clerc descrito en 1.14, ajustar c_1 y c_2 en términos de convergencia [16]. Se probaron algunos otros parámetros propuestos en el estado-del-arte pero no se obtuvieron mejores resultados a los presentados en este trabajo.

6.2. Otros dominios transformados

Se propusieron tres dominios transformados según el contexto geofísico tal como se describió en el capítulo 2. Sin embargo, esto abre numerosas preguntas sobre qué otros dominios transformados se puede emplear y si estos dominios transformados son los óptimos. Se encontró que no es conveniente perder el eje del tiempo cuando se mide el CS. Para mejorar la métrica \widehat{CS}_{FT} se propone utilizar la STFT (Short Time Fourier Transform). Sin embargo, el costo computacional de esta transformación puede incrementar dramáticamente los tiempos estimados hasta este momento. Por otra parte,

existen algunos retos al usar la STFT y es cómo seleccionar el tipo de ventana y el ancho de la misma.

Es posible mejorar las métricas propuestas hasta este momento. Para la métrica de correlación se calcula por traza un coeficiente de correlación por disparo ($\widehat{C}_{corr}^{s,r}$). Luego, se promedia en la dirección de los receptores obteniendo un único coeficiente parcial que relaciona todo el disparo (\widehat{C}_{corr}^s), finalmente se promedia en la dirección de las fuentes obteniendo un coeficiente que relaciona todos los disparos (\widehat{C}_{corr}) se propone sumar en la dirección de las fuentes obteniendo

$$\widehat{CS}_{corr}^r = \sum_{r=1}^R \sum_{s=1}^S CS_{corr}^{s,r}, \quad (6.1)$$

donde R es el número de receptores y S es el número de fuentes. Finalmente, se calculan la magnitud de $|\widehat{CS}_{corr}^r|$. Al no promediar, se esperan obtener mejores resultados con respecto a la métrica planteada en el capítulo 2. Se plantea usar este principio también en el cálculo de las métricas \widehat{CS}_{TC} y \widehat{CS}_{FT} . Se estima un vector asociado a contar los desfases mayores a π entre \mathbf{d}_{obs} y \mathbf{d}_{mod} , tal como se describió en el capítulo 2. Finalmente se calcula la norma l^2 por disparo y se promedia (\widehat{CS}_{FT} o \widehat{CS}_{TC}). Se propone considerar

$$\widehat{CS}_{FT}^r = \sum_{r=1}^R \sum_{s=1}^S CS_{FT}^{s,r}, \quad (6.2)$$

$$\widehat{CS}_{TC}^r = \sum_{r=1}^R \sum_{s=1}^S CS_{TC}^{s,r}, \quad (6.3)$$

y al igual que con la métrica \widehat{CS}_{corr} , se plantea calcular la magnitud de los vectores \widehat{CS}_{FT}^r y \widehat{CS}_{TC}^r .

6.3. Construcción de la métrica

Se propuso una métrica en cascada donde sólo se activa un dominio transformado al tiempo. Esto es logrado mediante el uso de las banderas C_1 , C_2 y C_3 tal como se describe en (6.4). Como trabajo futuro se propone activar más de un dominio transformado al tiempo. Los pesos W_1 , W_2 y W_3 se ajustarán según la importancia del evento sísmico. En la Figura 66 se presenta un ejemplo donde se activan los tres dominios transformados

poniendo en uno las banderas C_1, C_2 y C_3 . Mediante el uso de algún otro método de optimización se puede encontrar la combinación óptima para los pesos de W_x .

$$\begin{aligned} m_{\bar{x}} = & C_1 \times W_1 \times m_1(T_1\{\mathbf{d}_{\text{mod}}\}, T_1\{\mathbf{d}_{\text{obs}}\}) \\ & + C_2 \times W_2 \times m_2(T_2\{\mathbf{d}_{\text{mod}}\}, T_2\{\mathbf{d}_{\text{obs}}\}) \\ & + C_3 \times W_3 \times m_3(T_3\{\mathbf{d}_{\text{mod}}\}, T_3\{\mathbf{d}_{\text{obs}}\}). \end{aligned} \quad (6.4)$$

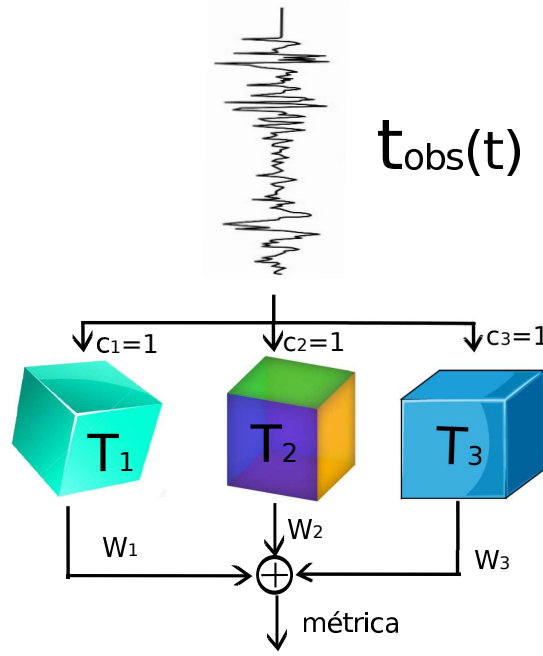


Figura 66: Esquema para activar varios dominios transformados

6.4. Costo computacional

El tiempo computacional para obtener un punto de partida usando esta metodología del esquema en cascada de métrica es:

Como se aprecia en la tabla 18 una configuración en cascada de métricas en los tres dominios transformados tarda aproximadamente 6 días. Sin embargo este algoritmo aún puede mejorar en su programación en paralelo. En la Figura 67 se presenta un diagrama de trabajo futuro del algoritmo desde el punto de vista del costo computacional. Aquí se puede ver que dos de las cuatro líneas de trabajo han sido abordadas y permite seguir avanzando en reducir el costo computacional de los resultados obtenidos hasta ahora. De las líneas de trabajo no abordadas hasta el momento vale la pena mencionar que es posible tomar cada partícula de forma independiente. Su independencia podría

Métricas	Tiempo (horas)
\widehat{C}_{corr}	24.8
$\widehat{C}\widehat{S}_{FT}$	26.3
$\widehat{C}\widehat{S}_{TC}$	84.4
$\widehat{C}_{corr} + \widehat{C}\widehat{S}_{FT}$	51.1
$\widehat{C}_{corr} + \widehat{C}\widehat{S}_{TC}$	109.2
$\widehat{C}\widehat{S}_{FT} + \widehat{C}\widehat{S}_{TC}$	110.7
$\widehat{C}_{corr} + \widehat{C}\widehat{S}_{TC} + \widehat{C}\widehat{S}_{FT}$	135.5

Tabla 18: Tiempo computacional esquema en cascada de métricas

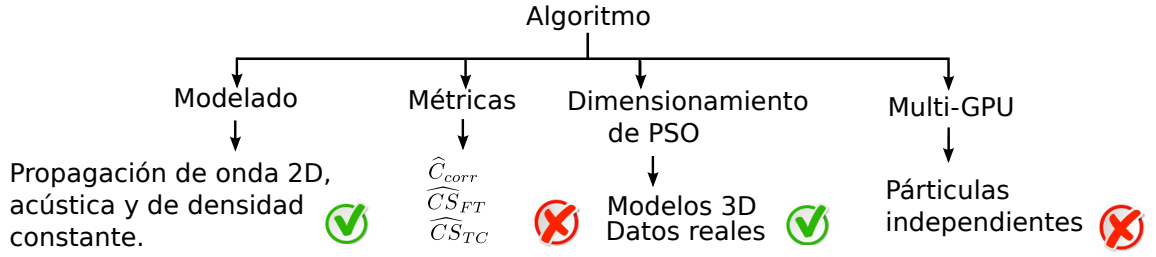


Figura 67: Diagrama trabajo futuro.

presentar el inconveniente de perder el mínimo global, ya que varias partículas pueden intentar escribir al mismo tiempo. Sin embargo por la dinámica de PSO esto no será de gran importancia. La independencia de las partículas de PSO será posible empleando arquitectura multi-GPU, donde es factible reducir su costo computacional proporcional al número de GPUs utilizadas. Otro aspecto a programar en paralelo son las métricas, sobre todo en el caso de la métrica $\widehat{C}\widehat{S}_{TC}$ donde su costo computacional es tres veces mayor que las otras dos métricas.

6.5. Independencia entre pixeles

Se asumió que todos los pixeles eran independientes e idénticamente distribuidos, con el fin de llevar a cabo la metodología del análisis estadístico sobre un solo pixel. Para comprobar la dependencia entre pixeles primero se estima la matriz de covarianza definida por

$$\sigma_{i,j} = \frac{\sum_{i=1}^N ((A_i - \mu_A)(B_i - \mu_B))}{n - 1}, \quad (6.5)$$

donde μ_A y μ_B son la media de A y B respectivamente. La diagonal de la matriz de covarianza son las varianzas. Por otra parte, de la matriz de covarianza es posible

estimar la matriz de correlación, su relación esta definida como

$$r_{i,j} = \frac{\sigma_{i,j}}{\sqrt{\sigma_{i,i}\sigma_{j,j}}}. \quad (6.6)$$

Desde las Figuras 68 a 70 se presentan la matriz de correlación sobre los modelos presentados en las Figuras 96, 101 y 106, respectivamente. Note que existe una alta correlación entre los pixeles, lo que demuestra que los pixeles no son independientes. Además, conociendo la naturaleza del frente de onda sabemos que la velocidad de un pixel altera a sus pixeles adyacentes.

6.6. Puntos de partida para datos reales 2D y sintéticos 3D

La exploración de PSO es viable para la metodología propuesta sobre un modelo de 211×68 con 500 partículas y una propagación que tarda 1×10^{-3} segundos usando una GPU Tesla K40c de Nvidia. En un dato real 2D las dimensiones podrían ser de 10000×8000 , por lo que se requeriría de 5 millones de partículas y una propagación tardaría 1 minuto. Dejando el número de iteraciones y experimentos constantes e iguales a 100, requeriríamos de 5×10^{10} minutos. Por otra parte para generar un punto de partida tridimensional de dimensiones $211 \times 68 \times 101$ se requeriría de 100 mil partículas y una propagación podría tardar 20 minutos. Si se deja el número de iteraciones y partículas nuevamente constantes e iguales a 100, requeriría de 2×10^{10} minutos. Sin embargo sobre estos tiempos es necesario sumarle el doble del tiempo de propagación, ya que es el tiempo necesario para obtener la métrica. Logrando propagaciones de al menos 10^{-6} es viable obtener un punto de partida para un dato real 2D o sintético 3D, donde los tiempos de ejecución son 41,67 horas y 50 minutos respectivamente.

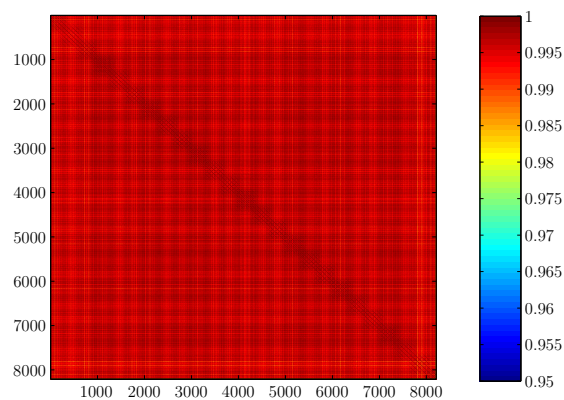


Figura 68: Matriz de correlación usando la métrica \hat{C}_{corr} .

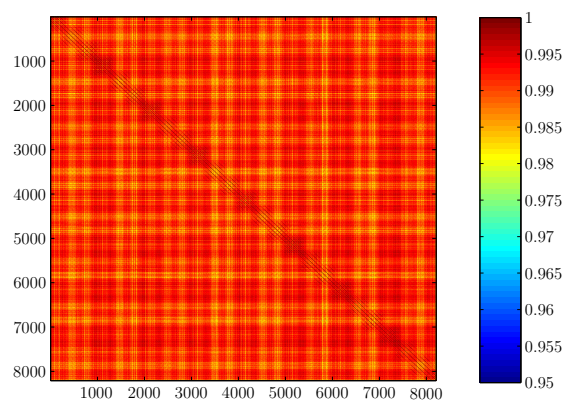


Figura 69: Matriz de correlación usando la métrica $\hat{C}_{S_{FT}}$.

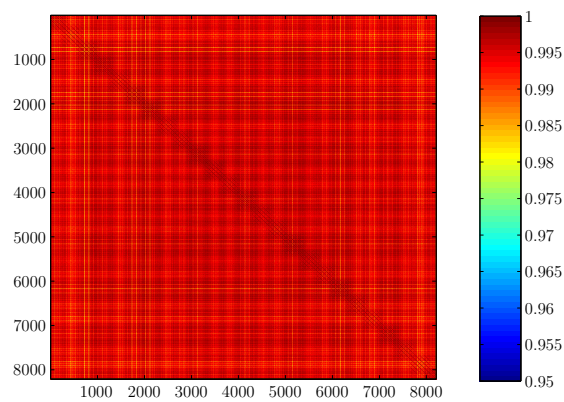


Figura 70: Matriz de correlación usando la métrica $\hat{C}_{S_{TC}}$.

ACRÓNIMOS

FWI- Full waveform inversión

TC- Complex trace

LP- Low pass filter

CS- Cycle skipping

$\widehat{CS}_{corr}^{s,r}$ - Métrica del cycle skipping en base a correlación

$\widehat{CS}_{FT}^{s,r}$ - Métrica del cycle skipping en base a Fourier

$\widehat{CS}_{TC}^{s,r}$ - Métrica del cycle skipping en base a la traza compleja

FDTD- Diferencias finitas en el dominio del tiempo

PSO- Particle swarm optimization

GPU- Graphics Processor Unit

\mathbf{m}_{sub} - Modelo submuestreado

\mathbf{m}_{ori} - Modelo original

\mathbf{m}_{full} - Modelo a resolución de pixel

SVM- Starting velocity model

ECDF- Empirical cumulative distribution function

CDF- Cumulative distribution function

MAP- Máxima verosimilitud

MMSE- Minimum mean square error

PDF- Función de densidad de probabilidad

REFERENCIAS

- [1] Drijkoningen G.G. and of Applied Earth Sciences TU-Delft. Faculty, *Seismic Data Acquisition*. Delft University of Technology, 1996.
- [2] Sanchez I, Serrano J, Niño C Sierra D y Agudelo W., “Svd polarization filter taking into account the planarity of ground roll energy,” *CTyF*, vol. 6, pp. 5–24, 2016.
- [3] Sercel. <http://www.sercel.com/about/Pages/what-is-geophysics.aspx>. Reviewed: June of 2015.
- [4] Nolet, G., S.P. Grand and Kennett. B.L.N., “Seismic heterogeneity in the upper mantle,” *Geophys.*, vol. 74, pp. 127–152, 1994.
- [5] Sneider R. and Trampert J., “The stability of one-dimensional inverse scattering,” *University of Utrech*, vol. 74, pp. 865–880, 1993.
- [6] Sneider R. and Trampert J., “Inverse problems in geophysics,” *Dept. of Geophysics*, vol. 74, pp. 127–152, 1999.
- [7] Virieux J. and Operto S., “An overview of full-waveform inversion in exploration geophysics,” *Geophysics*, vol. 74, pp. 127–152, 2009.
- [8] Goldstein Allen A, “On newton’s method,” *Numerische Mathematik*, vol. 7, no. 5, pp. 391–393, 1965.
- [9] Plessix R.-E., “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications,” *Geophysical Journal International*, vol. 167, no. 2, pp. 495–503, 2006.
- [10] Tarantola Albert, “Inversion of seismic reflection data in the acoustic approximation,” *Geophysics*, vol. 49, no. 8, pp. 1259–1266, 1984.
- [11] Y. Ma, *Waveform-based velocity estimation from reflection seismic data*. No. 1, Colorado School of Mines, 2012.
- [12] Ranjit S. and Shalichahan S., “Particle swarm optimization: A new tool to invert geophysical data,” *Society of Exploration Geophysicists*, vol. 54, pp. 772–775, 2007.
- [13] Shi Y. and Eberhart Russ C., “A modified particle swarm optimizer,” *IEEE International Conference of Evolutionary Computation*, vol. 54, pp. 772–775, 1998.

- [14] Clerc M., “The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization,” *In Proceedings of the Congress on Evolutionary Computation*, vol. 54, pp. 1951–1957, 1999.
- [15] Clerc M. and Kennedy J., “The particle swarm: Explosion, stability and convergence in a multi-dimensional complex space,” *IEEE International Conference of Evolutionary Computation*, vol. 54, pp. 60,61,77, 2001.
- [16] Holland J., *Adaption in Natural and Artificial Systems*. University of Michigan, 1975.
- [17] Mrinal K. Sen, University of Texas, Austin, *Global Optimization Methods in Geophysical Inversion*. Cambridge, 2013.
- [18] Hondori J., Mikada H., Asakawa E. and S. Mizohata, “A new initial model for full waveform inversion without cycle skipping,” *SEG*, vol. 1, no. 2, pp. 1446–1450, 2015.
- [19] F. Cl  rbout J., *EARTH SOUNDINGS ANALYSIS: Processing versus Inversion*. Stanford University, 2004.
- [20] Sullivan Dennis M, *Electromagnetic simulation using the FDTD method*. John Wiley & Sons, 2013.
- [21] Abreo-Carrillo S. A., Ramirez A. B., Reyes O. M., and Abreo-Carrillo D.L, “A practical implementation of acoustic full waveform inversion on graphical processing units,” *CTyF ecopetrol*, p. 1, 2015.
- [22] Cangellaris A., “Numerical stability and numerical dispersion of a compact 2-d/fdtd method used for the dispersion analysis of waveguides,” 1993.
- [23] Bunks Carey, Saleck Fatimetou M, Zaleski S, and Chavent G, “Multiscale seismic waveform inversion,” *Geophysics*, vol. 60, no. 5, pp. 1457–1473, 1995.
- [24] Grimmett Geoffrey and Stirzaker David, *Probability and Random Processes*. Oxford, 2001.
- [25] Ropero, E. Eleftheriou M., Gava L. and E. Romero, *Manual de Estad  stica Empresarial con ejercicios resueltos*. Delta publicaciones, 2008.
- [26] Lopes, R., Reid I. and P. Hobson, “The two-dimensional kolmogorov-smirnov test,” 2015.
- [27] G. Casell and R. Burger, *Statistical Inference*. Duxbury, 2002.
- [28] Fang Zhilong, Herrmann Felix J, and Lee Chia Ying, “A stochastic quasi-newton mcmc method for uncertainty quantification of full-waveform inversion,” 2015.

BIBLIOGRAFÍA

Abreo-Carrillo S. A., Ramirez A. B., Reyes O. M., and Abreo-Carrillo D.L, “*A practical implementation of acoustic full waveform inversion on graphical processing units*”, CTyF ecopetrol, p. 1, 2015.

Bunks Carey, Saleck Fatimetou M, Zaleski S, and Chavent G, “*Multiscale seismic waveform inversion*”, Geophysics, vol. 60, no. 5, pp. 1457–1473, 1995.

Clerc M., “*The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization*”, In Proceedings of the Congress on Evolutionary Computation, vol. 54, pp. 1951–1957, 1999.

Cangellaris A., “*Numerical stability and numerical dispersion of a compact 2d/fdtd method used for the dispersion analysis of waveguides*”, 1993.

Casell G. and Burger R., “*Statistical Inference*”, Duxbury, 2002.

Claerbout J., “*EARTH SOUNDINGS ANALYSIS: Processing versus Inversion*”, Stanford University, 2004.

Clerc M. and Kennedy J., “*The particle swarm: Explosion, stability and convergence in a multi-dimensional complex space*”, IEEE International Conference of Evolutionary Computation, vol. 54, pp. 60,61,77, 2001.

Drijkoningen G.G. and of Applied Earth Sciences TU-Delft. Faculty, “*Seismic Data Acquisition*”, Delft University of Technology, 1996.

Fang Zhilong, Herrmann Felix J, and Lee Chia Ying, “*A stochastic quasi-newton mcmc method for uncertainty quantification of full-waveform inversion*”, 2015.

Goldstein Allen A, “*On newton’s method*”, Numerische Mathematik, vol. 7, no. 5, pp. 391–393, 1965.

Grimmett Geoffrey and Stirzaker David, *“Probability and Random Processes. Oxford”*, 2001

Holland J., *“Adaption in Natural and Artificial Systems”*, University of Michigan, 1975.

Hondori J., Mikada H., Asakawa E. and S. Mizohata, *“A new initial model for full waveform inversion without cycle skipping”*, SEG, vol. 1, no. 2, pp. 1446–1450, 2015.

Lopes, R., Reid I. and P. Hobson, *“The two-dimensional kolmogorov-smirnov test”*, 2015.

Mrinal K. Sen, *“Global Optimization Methods in Geophysical Inversion ”*, University of Texas, Austin, Cambridge, 2013.

Nolet, G., S.P. Grand and Kennett. B.L.N., *“Seismic heterogeneity in the upper mantle ”*, Geophys., vol. 74, pp. 127–152, 1994.

Plessix R.-E., *“A review of the adjoint-state method for computing the gradient of a functional with geophysical applications”*, Geophysical Journal International, vol. 167, no. 2, pp. 495–503, 2006.

Ranjit S. and Shalichan S., *“Particle swarm optimization: A new tool to invert geophysical data ”*, Society of Exploration Geophysicists, vol. 54, pp. 772–775, 2007.

Roperio, E. Eleftheriou M., Gava L. and E. Romero, *“Manual de Estadística Empresarial con ejercicios resueltos”*, Delta publicaciones, 2008.

Sanchez I, Serrano J, Nino C, Sierra D. y Agudelo W., *“Svd polarization filter taking into account the planarity of ground roll energy ”*, CTyF, vol. 6, pp. 5–24, 2016.

Sercel, *“<http://www.sercel.com/about/Pages/what-is-geophysics.aspx> ”*, Reviewed: June of 2015.

Shi Y. and Eberhart Russ C., “*A modified particle swarm optimizer*”, IEEE International Conference of Evolutionary Computation, vol. 54, pp. 772–775, 1998.

Sneider R. and Trampert J., “*The stability of one-dimensional inverse scattering*”, University of Utrech, vol. 74, pp. 865–880, 1993.

Sneider R. and Trampert J., “*Inverse problems in geophysics*”, Dept. of Geophysics, vol. 74, pp. 127–152, 1999.

Sullivan Dennis M., “*Electromagnetic simulation using the FDTD method*”, John Wiley and Sons, 2013.

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

Virieux J. and Operto S., “*An overview of full-waveform inversion in exploration geophysics*”, Geophysics, vol. 74, pp. 127–152, 2009.

Y. Ma, “*Waveform-based velocity estimation from reflection seismic data*”, No. 1, Colorado School of Mines, 2012.

ANEXOS

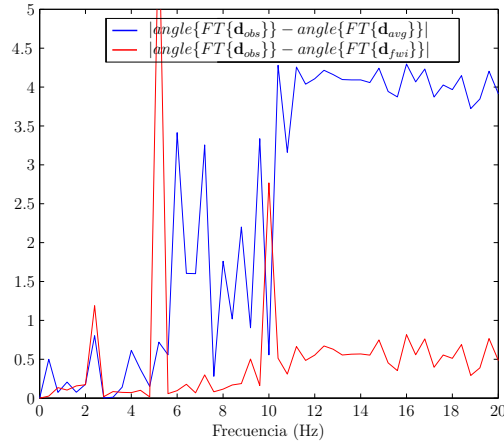


Figura 70: Uso del operador $FT_{angle}\{\cdot\}$ sobre la traza 50 obtenida con las métricas $\widehat{CS}_{FT} + \widehat{CS}_{TC}$ (\mathbf{m}_{ori} -2 capas)].

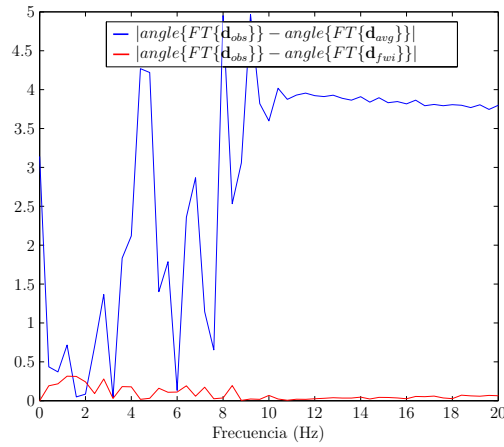


Figura 71: Uso del operador $FT_{angle}\{\cdot\}$ sobre la traza 50 obtenida de las métricas $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$ (\mathbf{m}_{ori} -marmousi).

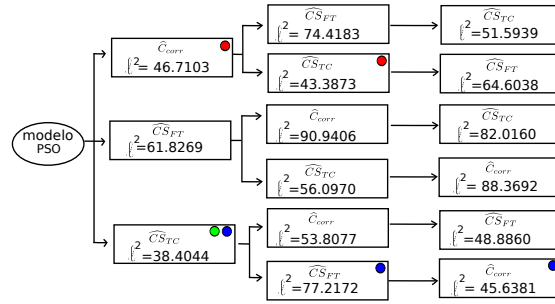


Figura 72: Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -cuadrado difractor).

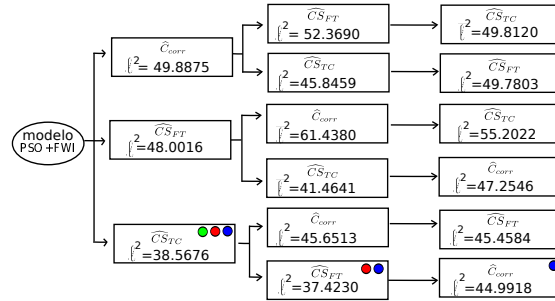


Figura 73: Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).

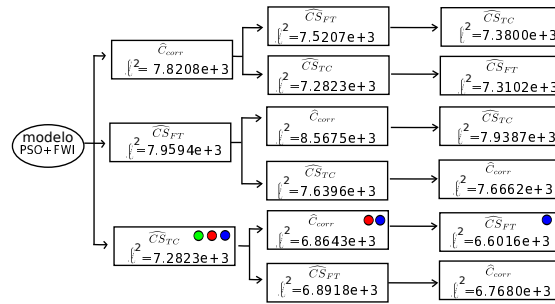


Figura 74: Diagrama multi-métrica usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -cuadrado difractor).

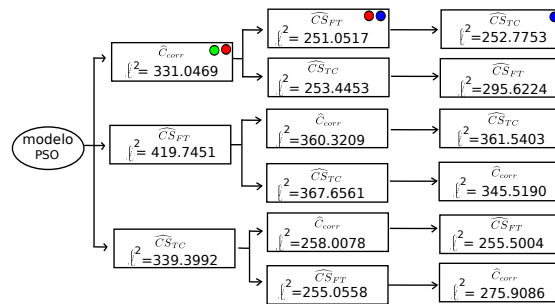
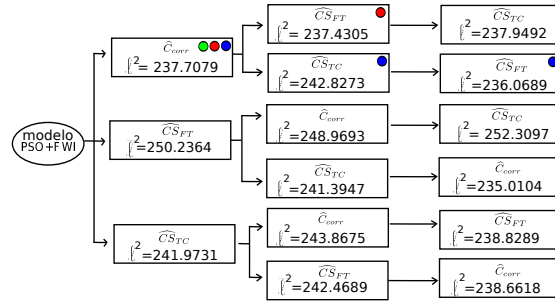
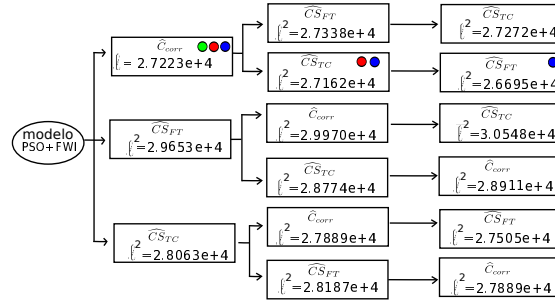
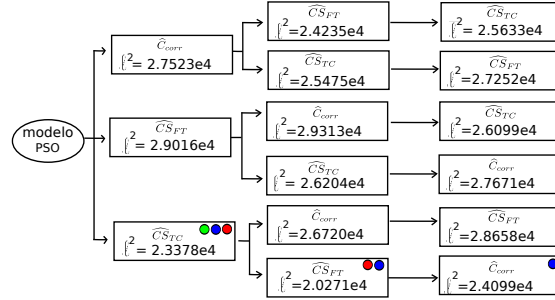
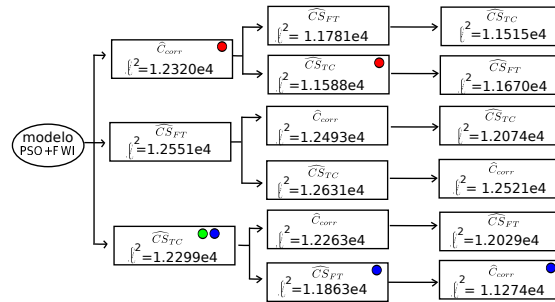


Figura 75: Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -2 capas).

Figura 76: Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -2 capas).Figura 77: Diagrama multi-métrica usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} (\mathbf{m}_{full} -2 capas).Figura 78: Norma l^2 entre los espectros \mathbf{d}_{obs} y \mathbf{d}_{avg} (\mathbf{m}_{full} -marmousi).Figura 79: Norma l^2 entre \mathbf{d}_{obs} y \mathbf{d}_{fwi} (\mathbf{m}_{full} -marmousi).

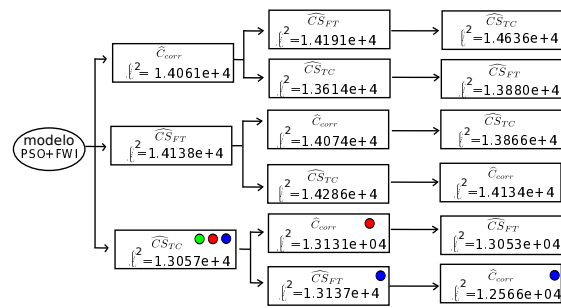


Figura 80: Diagrama multi-métrica usando norma l^2 entre \mathbf{m}_{ori} y \mathbf{m}_{fwi} ($\mathbf{m}_{full-marmousi}$).

$w=1.2;$ $c_1=2.8;$ $c_2=3.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	1
Exp 1	1243	1796	3474	2326	1522	1442	1965	1865	1713	2364	3082	2886	2730	2711	2550	3188	1313	1091	0.6488
Exp 2	1292	1500	2847	2604	1731	2257	1913	1864	2921	1734	1640	1522	2185	2647	2844	2001	3200	2258	0.6867
Exp 3	1332	1869	1815	3463	1521	1501	2082	1897	2497	2387	2303	2979	2200	2793	2672	2629	2494	3067	0.6192
Exp 4	2382	1747	2145	2687	1529	1476	1698	1617	1733	2495	2968	2668	1626	1811	2497	1508	3032	1667	0.7280
Exp 5	1415	1527	2757	2237	1690	1518	2083	2175	1778	2514	2293	2449	2359	2675	2579	1458	1357	2950	0.7398
$w=1;$ $c_1=2.8;$ $c_2=3.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	1
Exp 1	1691	1388	2287	3223	1601	1464	2328	2415	2125	1866	1996	2814	2190	2012	2609	2272	2575	2454	0.7039
Exp 2	1153	1459	1988	3031	1509	1380	1998	1982	2027	2680	1942	1493	1818	2452	2577	2609	2679	1672	0.7042
Exp 3	1283	1775	3446	2211	1440	1383	2015	1850	1771	2363	3045	2131	2731	2603	3306	1381	1295	1069	0.6814
Exp 4	1829	1736	2368	2216	1416	1375	1942	1810	1721	3128	2137	2098	2650	2620	2493	3132	3101	2987	0.7219
Exp 5	1415	1998	2051	2266	2369	1816	2297	1626	1534	1647	1906	1652	2063	1924	1861	2119	2691	1990	0.6730
$w=0.729;$ $c_1=2.8;$ $c_2=3.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	1
Exp 1	1455	1586	2622	2467	2350	2253	2111	2005	2156	2511	2426	3134	2972	2209	2075	2255	3452	2156	0.6538
Exp 2	1550	1528	3454	2199	1490	1404	2028	1940	1753	2453	2339	2828	2186	2734	2833	2481	3227	3045	0.7378
Exp 3	1432	1849	3344	2033	1528	1973	3259	1799	1883	2118	1843	1827	2561	2058	2730	2842	3414	3289	0.7341
Exp 4	2062	1517	2249	2472	1486	1470	2242	1981	2005	2072	1956	2308	2385	2169	2880	2528	2206	2314	0.7825
Exp 5	1430	1380	2472	2883	1774	1634	1473	2091	1994	1911	1795	1676	1391	2934	1978	2631	3250	3182	0.7689

Tabla 18: Ajuste del parámetro w , tomando valores de 1.2, 1 y 0.729, con $c_1 = 2,8$ y $c_1 = 3,1$ constante.

$w=0.2;$ $c_1=2.8;$ $c_2=3.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	\mathbf{m}	
	Ori	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	1	
	Exp 1	1497	1435	2610	2353	1767	1478	2533	1899	2130	1966	2292	3220	2783	2321	2756	2089	1922	2126	0.7365
	Exp 2	2035	1885	3013	2176	1756	1944	1515	1560	2002	1803	1826	1881	2283	2876	2967	2177	1839	2565	0.7693
	Exp 3	1586	1525	2164	2726	1529	1774	1845	1887	2627	1838	1785	1683	2192	2366	2481	2641	2029	1911	0.8043
	Exp 4	1473	1500	2316	2972	1736	1905	2310	2218	1768	2495	1595	1653	2618	2475	2455	2886	2224	1727	0.7968
Exp 5	2419	2148	2667	2369	1494	1409	1688	2174	2434	2287	2694	2238	2125	2400	3394	1831	2220	2227	0.6943	
$w=0.2;$ $c_1=2.8;$ $c_2=3.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	\mathbf{m}	
	Ori	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	2800	1
	Exp 1	1906	1810	2235	2665	2243	1748	2555	2210	2296	2417	1891	1634	2212	2311	1787	2708	2184	2754	0.6297
	Exp 2	1443	1462	1993	3048	1360	1460	2230	1989	1998	2645	1925	3201	1806	2445	2604	2588	2591	1693	0.6730
	Exp 3	1665	2577	3079	2377	1465	1511	1629	2929	2826	2512	2298	2027	2470	1480	2498	2679	1388	1724	0.6041
	Exp 4	2236	2043	2184	2305	2229	1365	1545	2092	1619	2033	2852	3126	1611	2127	1812	2337	1685	2024	0.6643
Exp 5	1700	1705	2545	3076	2240	1783	2072	1727	2266	2289	1945	2592	2884	2343	3104	1565	1565	2580	0.6122	

Tabla 19: Ajuste del parámetro w , tomando valores de 0.7 y 0.2, con $c_1 = 2,8$ y $c_1 = 3,1$ constante.

$w=0.7;$ $c_1=2.8;$ $c_2=2.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m	
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	2800	1
Exp 1	1581	1641	2805	2332	1631	1876	1935	1767	2473	2124	1760	2028	2685	2611	2294	1951	1917	2275	0.7845	
Exp 2	1422	1621	2599	2729	1540	1680	1936	1908	2146	2503	1759	1888	2553	2563	2005	1917	2621	2262	0.8545	
Exp 3	1543	1518	2064	2901	1840	1973	2126	1896	1991	1950	1943	2244	2575	2410	3073	2062	2092	2045	0.7807	
Exp 4	1789	2021	2896	2208	1527	1489	1965	1969	2612	2515	2491	2491	3354	1920	3068	1795	1797	2749	0.7386	
Exp 5	1476	1530	2268	2695	2185	2544	1901	2055	1893	2683	2339	2394	2619	3002	2011	2228	2350	1859	0.7418	

$w=0.7;$ $c_1=2.8;$ $c_2=1.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m	
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	2800	1
Exp 1	1581	1459	1501	2313	2924	1852	1582	2209	2399	2294	2993	1970	1834	3402	2420	1693	2487	1837	0.7741	
Exp 2	1453	1495	2381	2716	1909	1961	2181	2634	2310	2667	1945	1632	3127	2160	1626	3446	1858	2143	0.7710	
Exp 3	1516	1502	2447	2563	1507	1494	1929	1978	2007	1976	1981	2081	2642	2912	2762	2854	2400	2614	0.9950	
Exp 4	1526	1522	2182	2481	1461	1451	1916	1927	2065	2727	1890	1775	2764	2856	2789	2462	2974	1639	0.9315	
Exp 5	2248	1848	2346	2639	1771	1731	3020	2103	1974	1781	1667	1821	2727	1930	1963	2828	2012	1973	0.7559	

Tabla 20: Ajuste del parámetro c_2 , tomando valores de 2.1 y 1.1, con $w = 0,7$ y $c_1 = 2,8$ constante.

$w=0.7;$ $c_1=1.8;$ $c_2=1.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m	
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	2800	1
Exp 1	2026	2015	2684	2443	2007	1915	2101	2167	3106	1982	2354	2261	1897	3412	1836	1438	3211	2023	0.6989	
Exp 2	1513	1502	2419	2678	1939	2075	1914	1957	2152	2715	2067	1567	2951	3268	1796	3485	2585	3450	0.8080	
Exp 3	1491	1497	2363	2628	1495	1465	1707	1843	2415	2369	1946	1938	1592	2626	2519	2268	3266	2932	0.9281	
Exp 4	1465	1502	2524	2508	1538	1499	2177	2427	2449	1919	1925	1937	3499	1713	2068	2522	2745	2639	0.9036	
Exp 5	1442	1714	3016	2170	1772	1407	2078	1764	3209	1807	2238	1898	3093	2492	2513	2516	3088	3249	0.7314	
$w=0.7;$ $c_1=1;$ $c_2=1.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m	
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	2800	1
Exp 1	1423	1510	2559	2678	1547	1485	2134	2028	1861	2730	2060	2020	3399	2829	2394	1627	2378	3496	0.8765	
Exp 2	1470	1496	2520	2405	1512	1512	2043	2013	2029	1984	2033	1949	2849	2800	2970	2289	3088	2858	0.9885	
Exp 3	1497	1497	2523	2495	1509	1490	2067	1999	1991	1982	1978	2042	2790	2429	2864	2793	2586	2646	0.9931	
Exp 4	1468	1514	2635	2499	1529	1507	2039	2313	2667	2270	1832	1824	3497	2045	1804	1922	2921	2154	0.9070	
Exp 5	1467	1518	2474	2562	1523	1500	1984	2339	2860	2246	1893	1870	2815	1751	2658	1846	3387	2619	0.8810	
$w=0.7;$ $c_1=0.8;$ $c_2=1.1$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}	m	
Ori	1500	1500	2500	2500	1500	1500	2000	2000	2000	2000	2000	2000	2800	2800	2800	2800	2800	2800	2800	1
Exp 1	1499	1501	2470	2515	1498	1497	1999	1993	1999	2008	2015	1998	2743	2766	2790	2735	2909	2934	0.9995	
Exp 2	1493	1501	2508	2464	1509	1500	2010	2030	1979	2016	2001	1945	2631	2713	2577	2472	2918	2603	0.9919	
Exp 3	1503	1502	2465	2541	1497	1494	1999	1996	1983	2014	2006	2000	2693	2716	2717	2859	2886	2954	0.9992	
Exp 4	1500	1499	2497	2506	1500	1501	1999	2003	1998	1999	1992	1997	2805	2814	2776	2868	2785	2791	0.9999	
Exp 5	1467	1473	2539	2519	1536	1536	2154	2192	1957	1966	1848	1828	3362	2850	2466	3062	2256	1956	0.9663	

Tabla 21: Ajuste del parámetro c_1 , tomando valores de 1.8, 1 y 0.8, con $w = 0,7$ y $c_2 = 1,1$ constante.

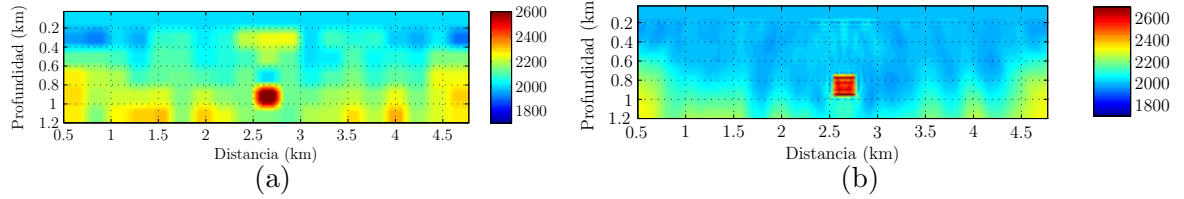


Figura 81: \mathbf{m}_{sub} cuadrado difractor estimado por \hat{C}_{corr} (a) PSO (b) FWI.

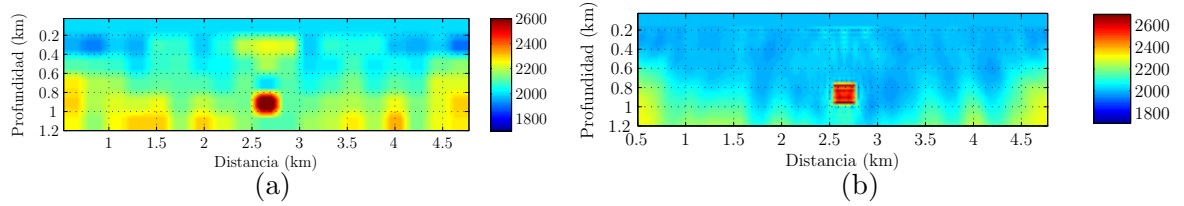


Figura 82: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}S_{FT}$ (a) PSO (b) FWI.

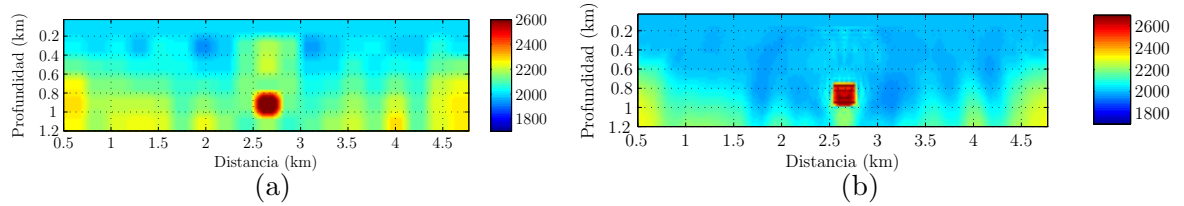


Figura 83: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}S_{FT} + \hat{C}S_{TC}$ (a) PSO (b) FWI.

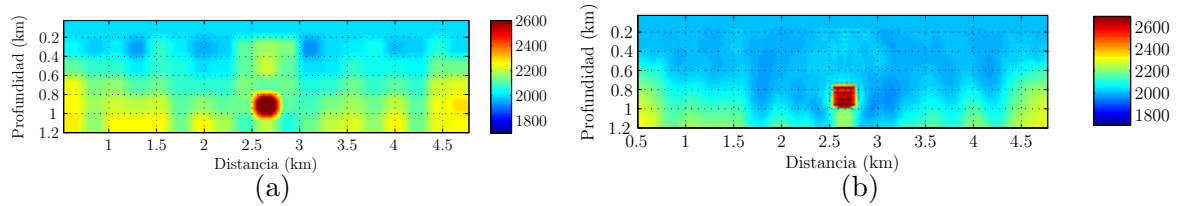


Figura 84: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}S_{TC}$ (a) PSO (b) FWI.

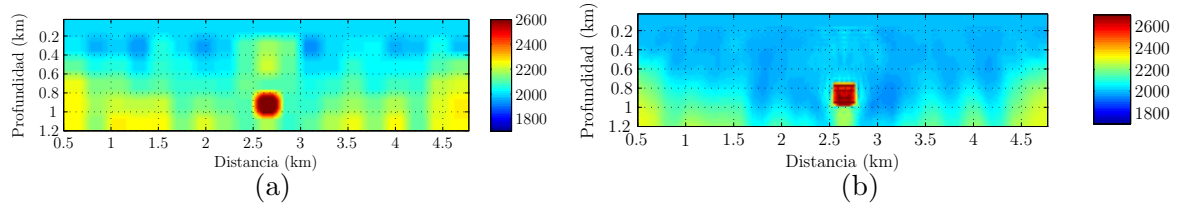


Figura 85: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{TC} + \hat{C}\hat{S}_{FT}$ (a) PSO (b) FWI.

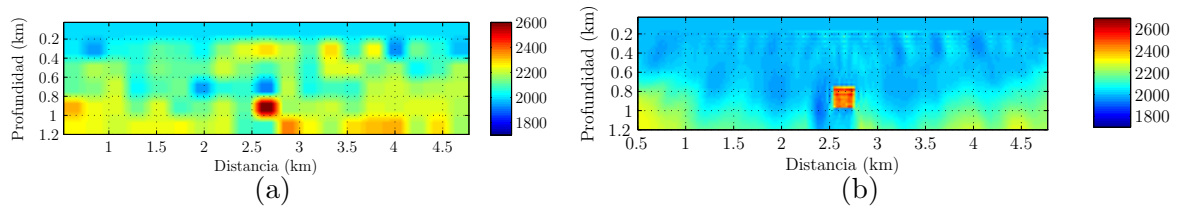


Figura 86: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}\hat{S}_{FT}$ (a) PSO (b) FWI.

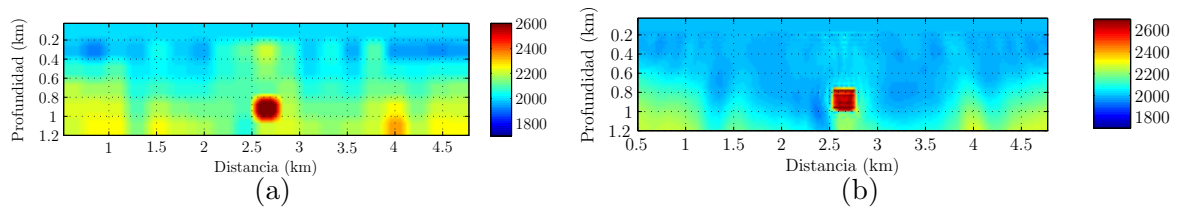


Figura 87: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}\hat{S}_{FT} + \hat{C}_{corr}$ (a) PSO (b) FWI.

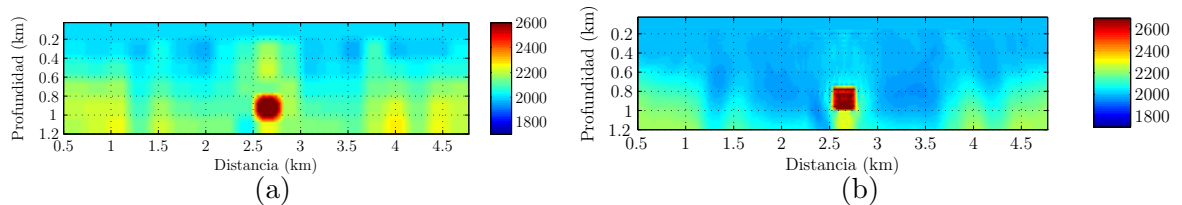


Figura 88: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}\hat{S}_{FT} + \hat{C}_{corr} + \hat{C}\hat{S}_{TC}$ (a) PSO (b) FWI.

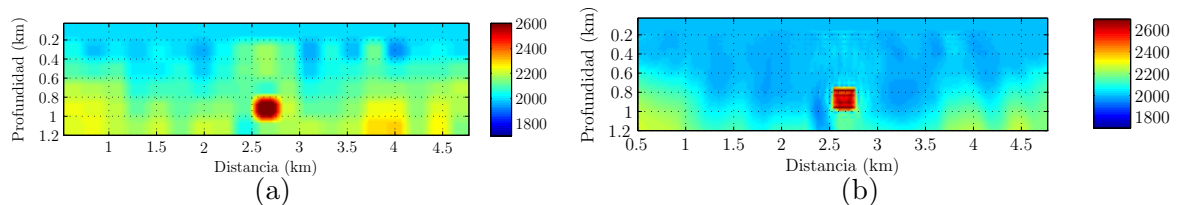


Figura 89: \mathbf{m}_{sub} cuadrado difractor estimado por $\hat{C}\hat{S}_{FT} + \hat{C}\hat{S}_{TC}$ (a) PSO (b) FWI.

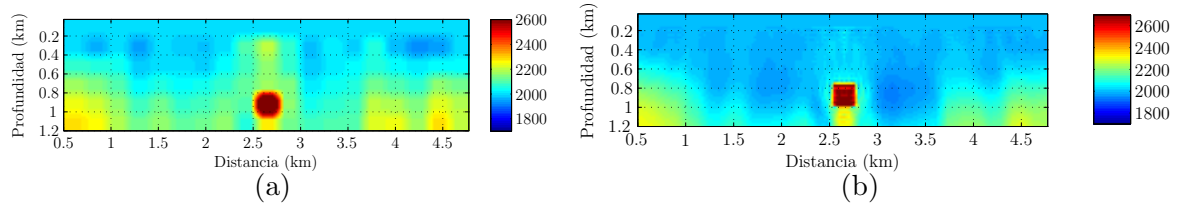


Figura 90: \mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

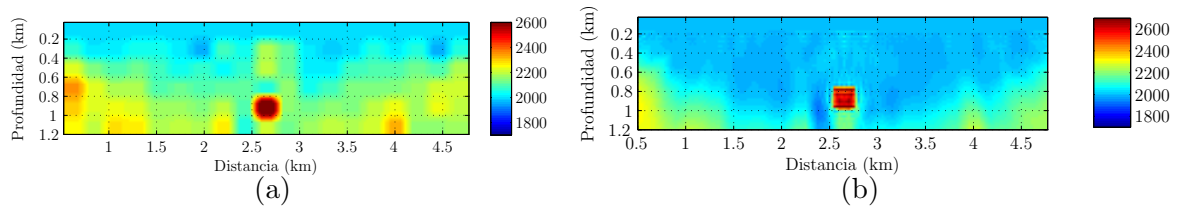


Figura 91: \mathbf{m}_{sub} cuadrado difractor estimado por \widehat{CS}_{TC} (a) PSO (b) FWI.

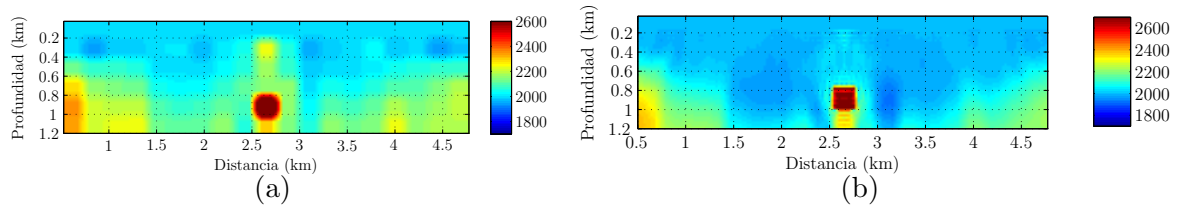


Figura 92: \mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

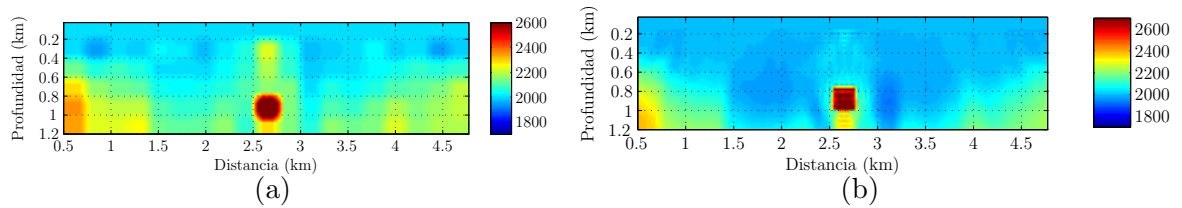


Figura 93: \mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

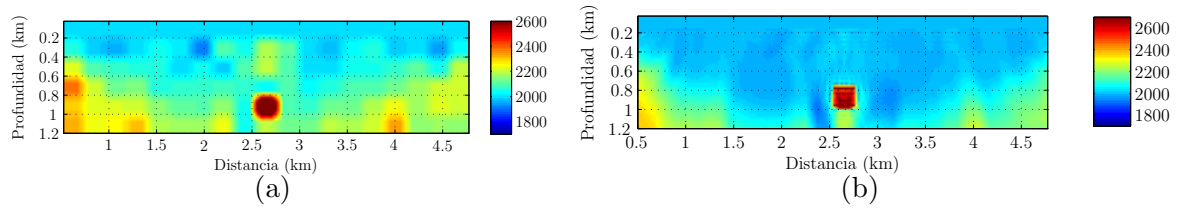


Figura 94: \mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

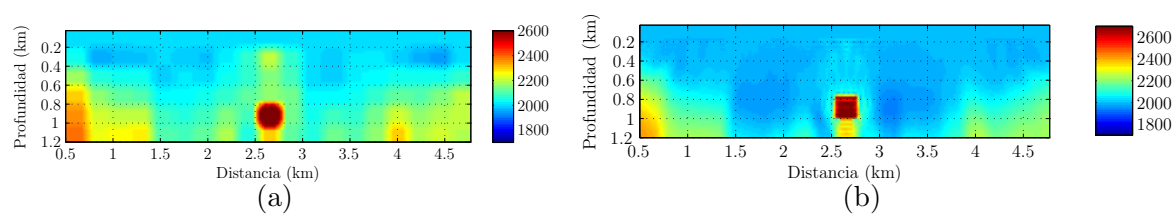


Figura 95: \mathbf{m}_{sub} cuadrado difractor estimado por $\widehat{C}\widehat{S}_{TC} + \widehat{C}\widehat{S}_{FT} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

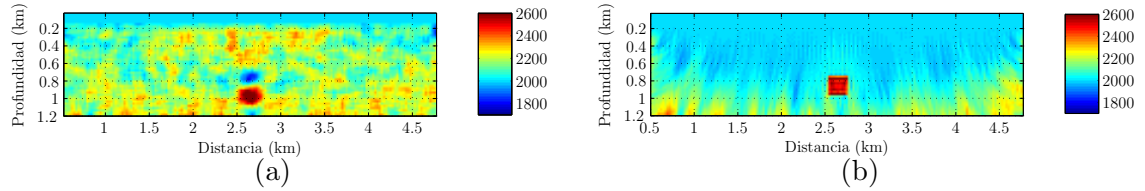


Figura 96: \mathbf{m}_{full} cuadrado difractor estimado por \hat{C}_{corr} (a) PSO (b) FWI.

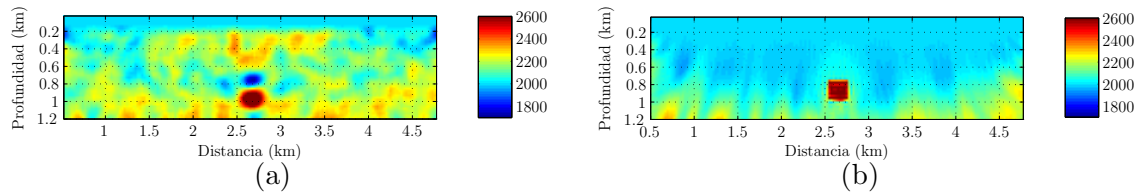


Figura 97: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{FT}$ (a) PSO (b) FWI.

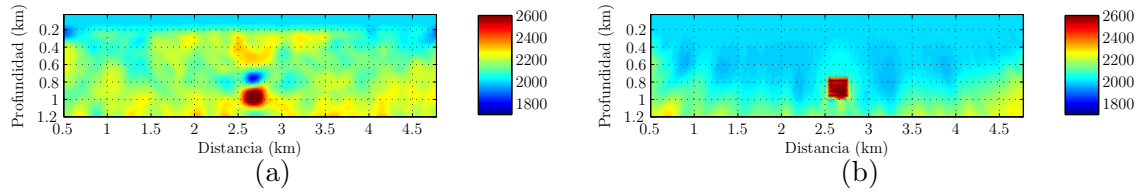


Figura 98: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{FT} + \hat{C}\hat{S}_{TC}$ (a) PSO (b) FWI.

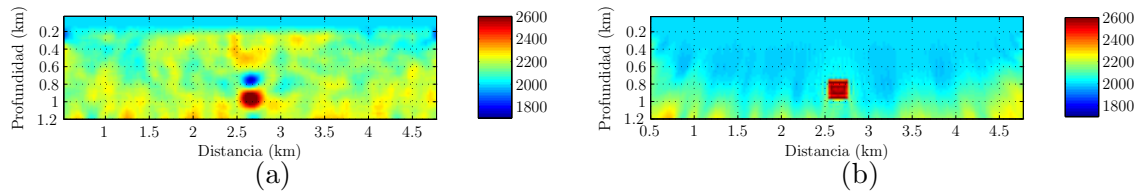


Figura 99: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{TC}$ (a) PSO (b) FWI.

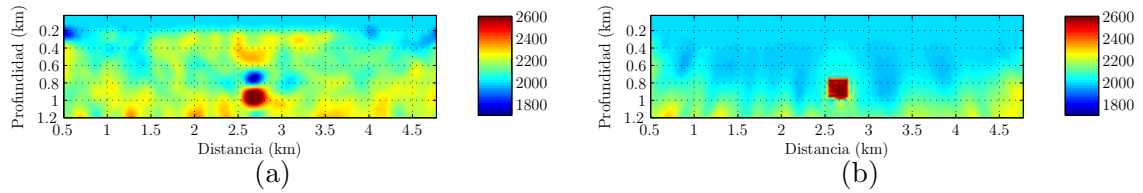


Figura 100: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}_{corr} + \hat{C}S_{TC} + \hat{C}S_{FT}$ (a) PSO (b) FWI.

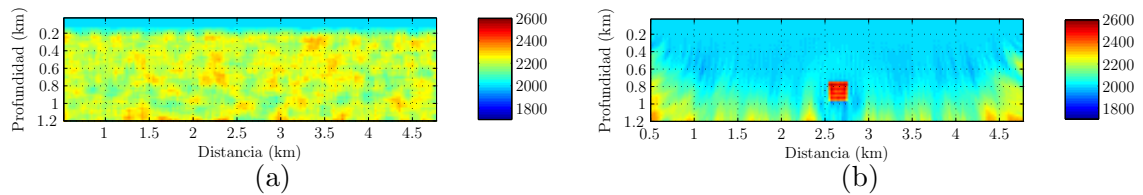


Figura 101: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}S_{FT}$ (a) PSO (b) FWI.

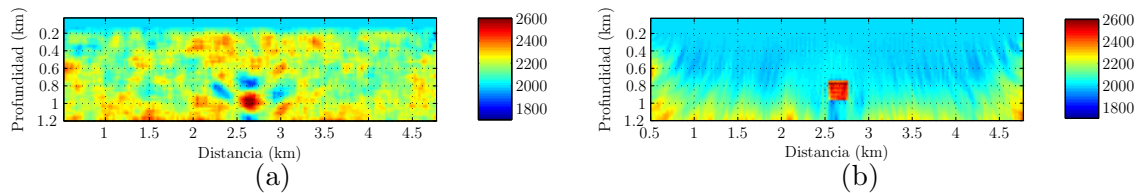


Figura 102: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}S_{FT} + \hat{C}_{corr}$ (a) PSO (b) FWI.

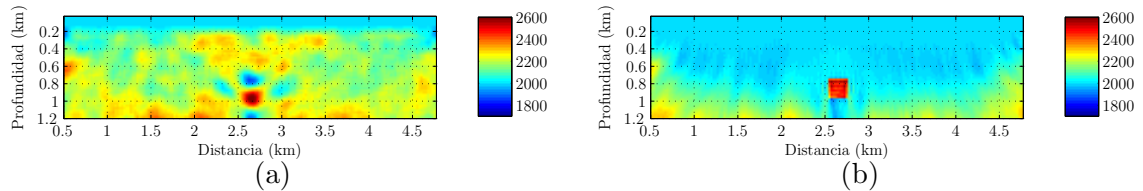


Figura 103: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}S_{FT} + \hat{C}_{corr} + \hat{C}S_{TC}$ (a) PSO (b) FWI.

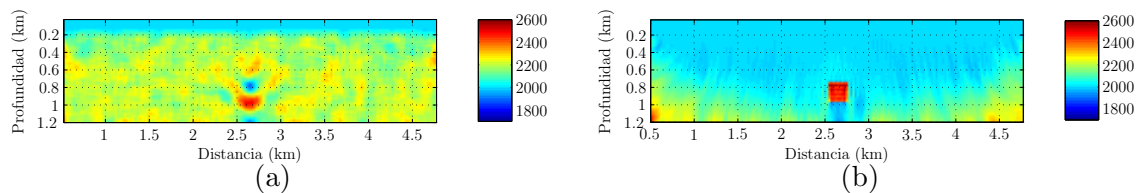


Figura 104: \mathbf{m}_{full} cuadrado difractor estimado por $\hat{C}S_{FT} + \hat{C}S_{TC}$ (a) PSO (b) FWI.

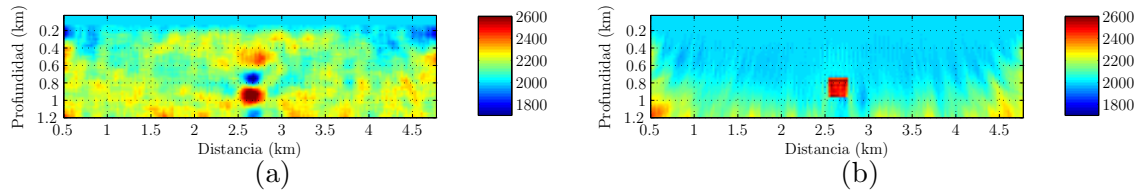


Figura 105: \mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

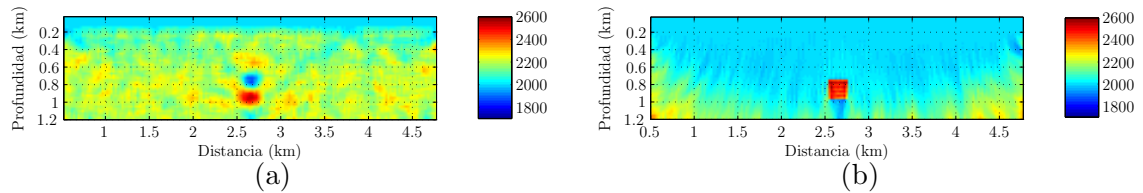


Figura 106: \mathbf{m}_{full} cuadrado difractor estimado por \widehat{CS}_{TC} (a) PSO (b) FWI.

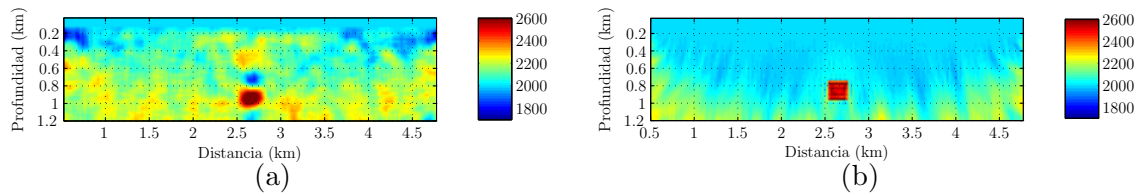


Figura 107: \mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

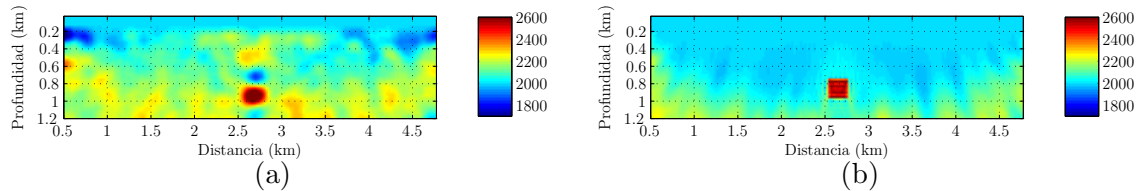


Figura 108: \mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

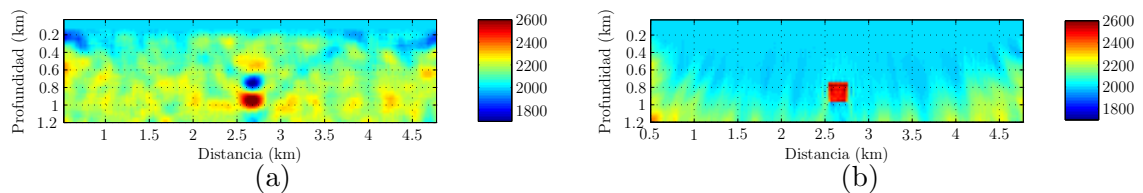


Figura 109: \mathbf{m}_{full} cuadrado difractor estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

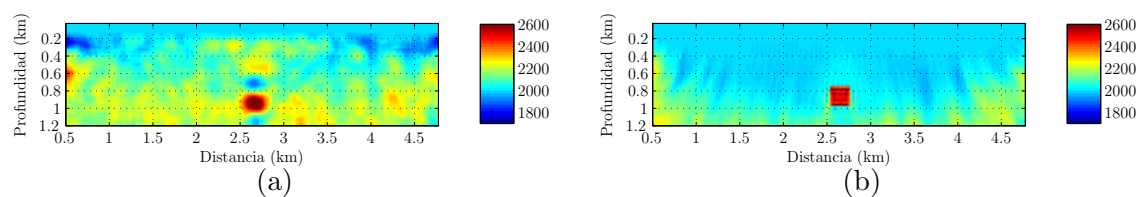


Figura 110: \mathbf{m}_{full} cuadrado difractor estimado por $\widehat{C}\widehat{S}_{TC} + \widehat{C}\widehat{S}_{FT} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

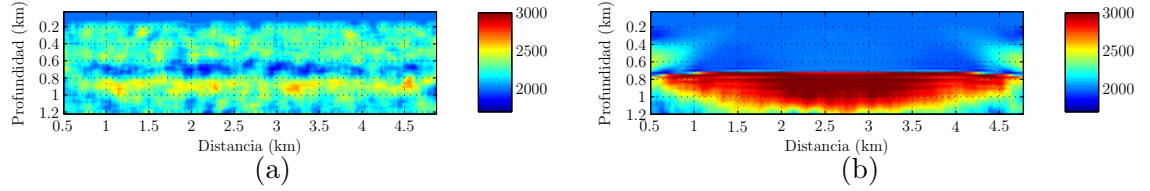


Figura 111: \mathbf{m}_{full} dos capas estimado por \hat{C}_{corr} (a) PSO (b) FWI.

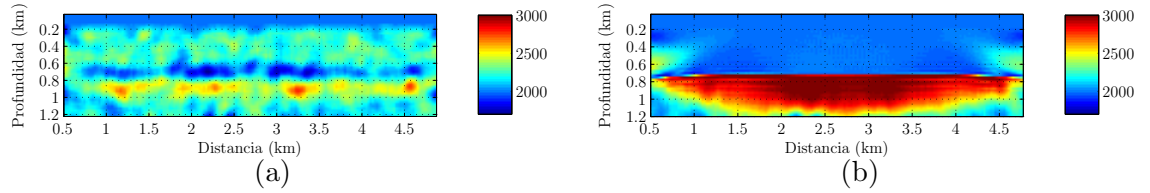


Figura 112: \mathbf{m}_{full} dos capas estimado por $\hat{C}_{corr} + \hat{CS}_{FT}$ (a) PSO (b) FWI.

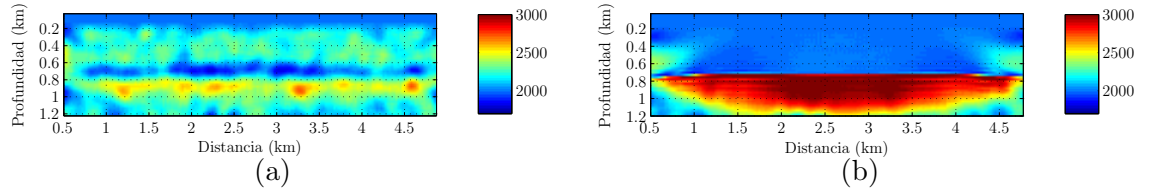


Figura 113: \mathbf{m}_{full} dos capas estimado por $\hat{C}_{corr} + \hat{CS}_{FT} + \hat{CS}_{TC}$ (a) PSO (b) FWI.

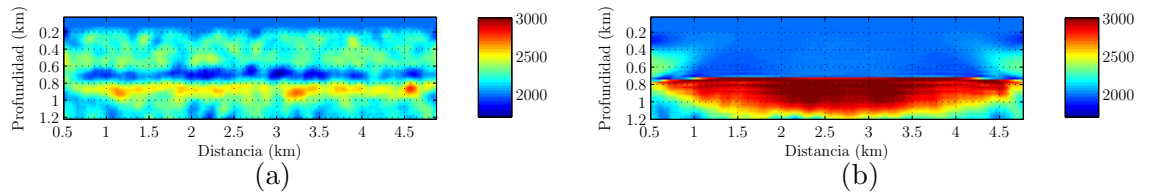


Figura 114: \mathbf{m}_{full} dos capas estimado por $\hat{C}_{corr} + \hat{CS}_{TC}$ (a) PSO (b) FWI.

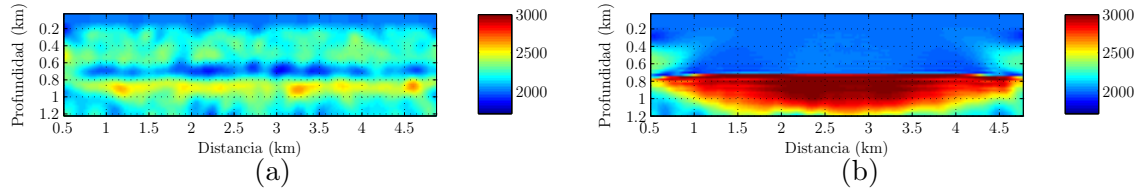


Figura 115: \mathbf{m}_{full} dos capas estimado por $\widehat{C}_{corr} + \widehat{C}S_{TC} + \widehat{C}S_{FT}$ (a) PSO (b) FWI.

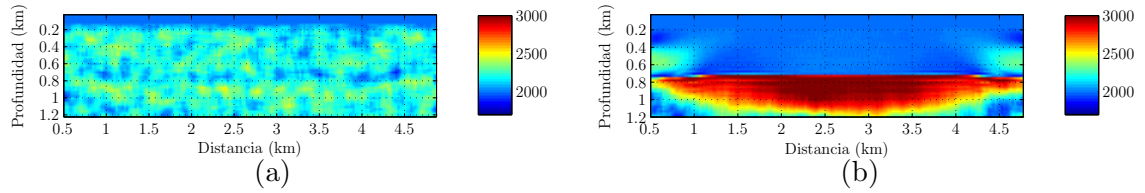


Figura 116: \mathbf{m}_{full} dos capas estimado por $\widehat{C}S_{FT}$ (a) PSO (b) FWI.

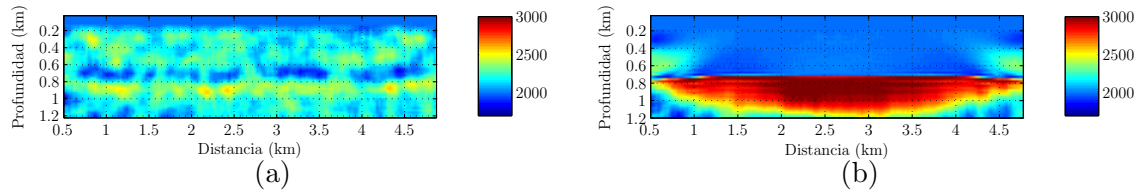


Figura 117: \mathbf{m}_{full} dos capas estimado por $\widehat{C}S_{FT} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

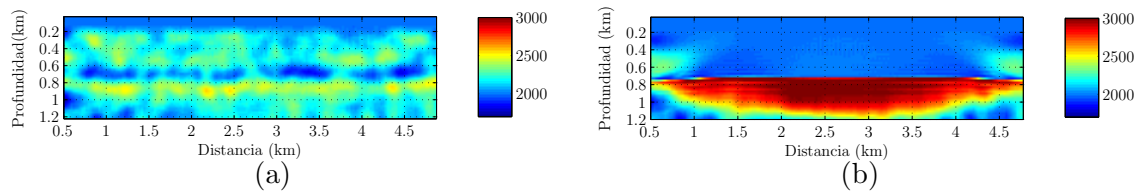


Figura 118: \mathbf{m}_{full} dos capas estimado por $\widehat{C}S_{FT} + \widehat{C}_{corr} + \widehat{C}S_{TC}$ (a) PSO (b) FWI.

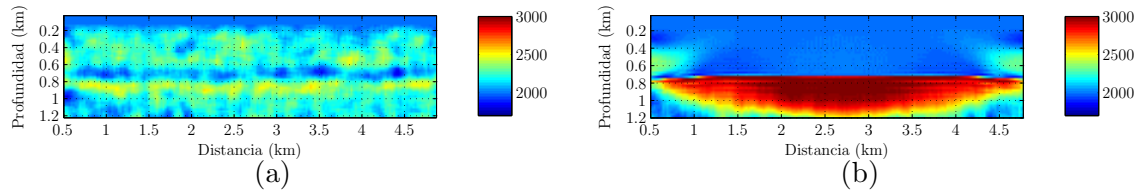


Figura 119: \mathbf{m}_{full} dos capas estimado por $\widehat{C}S_{FT} + \widehat{C}S_{TC}$ (a) PSO (b) FWI.

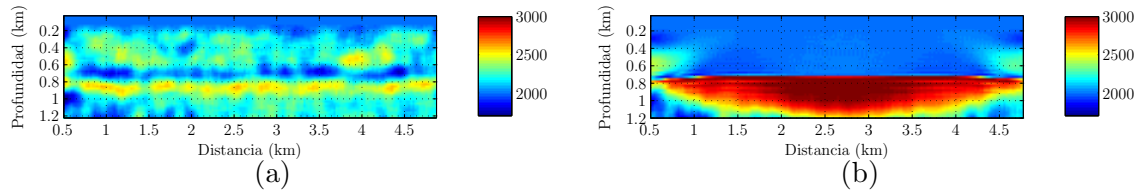


Figura 120: \mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

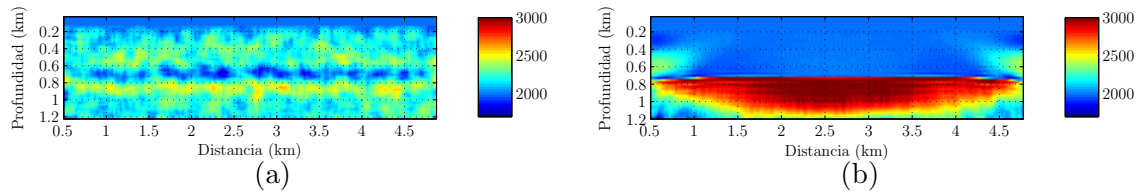


Figura 121: \mathbf{m}_{full} dos capas estimado por \widehat{CS}_{TC} (a) PSO (b) FWI.

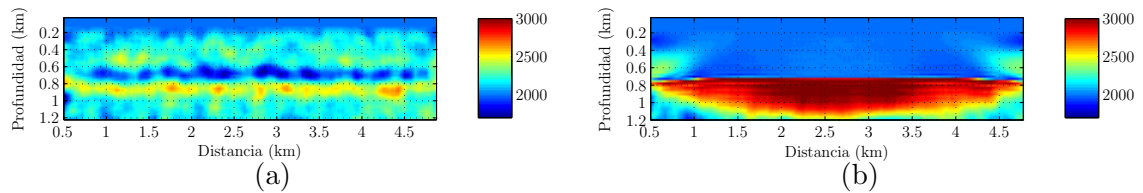


Figura 122: \mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

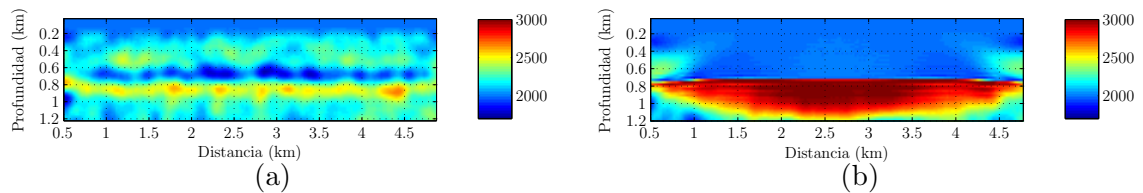


Figura 123: \mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

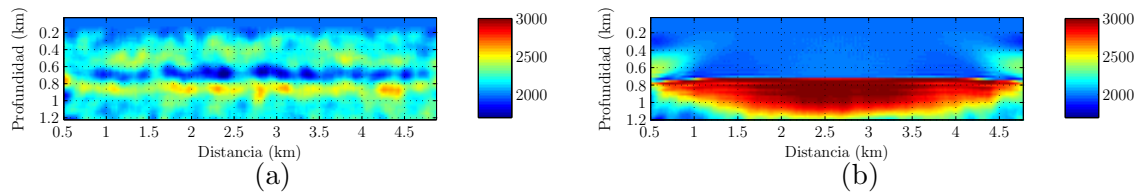


Figura 124: \mathbf{m}_{full} dos capas estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$ a) PSO b) FWI.

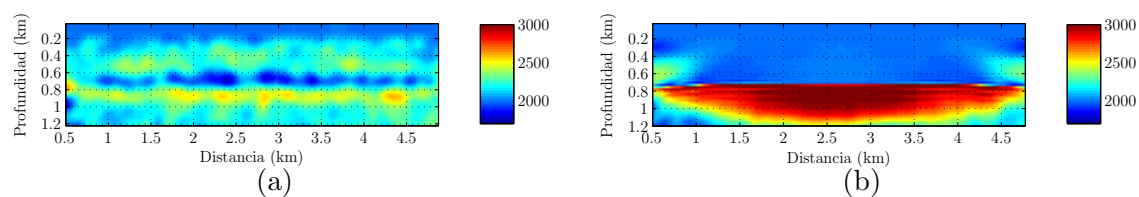


Figura 125: \mathbf{m}_{full} dos capas estimado por $\widehat{C}S_{TC} + \widehat{C}S_{FT} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

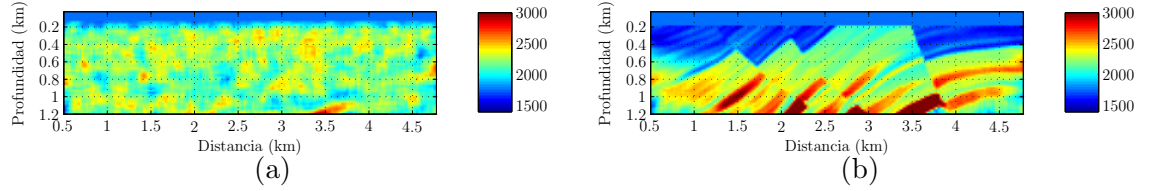


Figura 126: \mathbf{m}_{full} marmousi estimado por \hat{C}_{corr} (a) PSO (b) FWI.

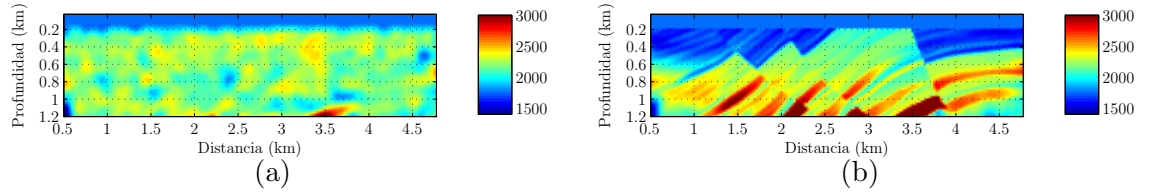


Figura 127: \mathbf{m}_{full} marmousi estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{FT}$ (a) PSO (b) FWI.

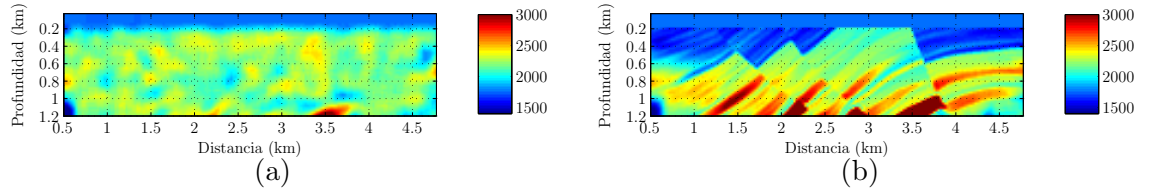


Figura 128: \mathbf{m}_{full} marmousi estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{FT} + \hat{C}\hat{S}_{TC}$ (a) PSO (b) FWI.

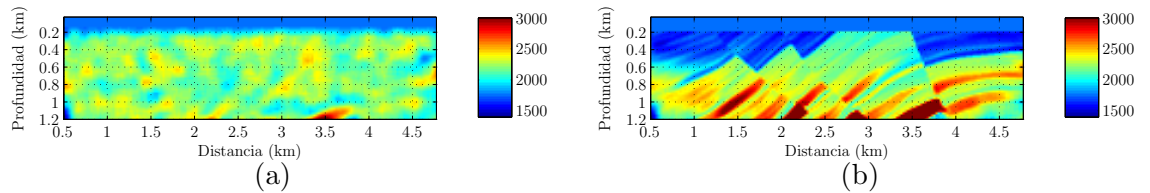


Figura 129: \mathbf{m}_{full} marmousi estimado por $\hat{C}_{corr} + \hat{C}\hat{S}_{TC}$ (a) PSO (b) FWI.

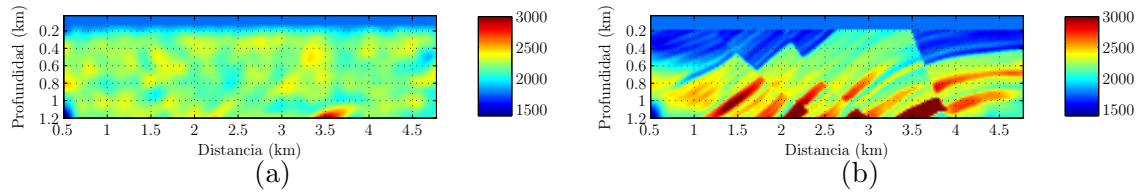


Figura 130: \mathbf{m}_{full} marmousi estimado por $\widehat{C}_{corr} + \widehat{C}S_{TC} + \widehat{C}S_{FT}$ (a) PSO (b) FWI.

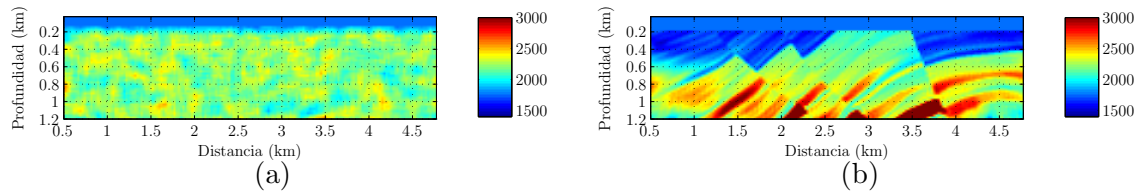


Figura 131: \mathbf{m}_{full} marmousi estimado por $\widehat{C}S_{FT}$ (a) PSO (b) FWI.

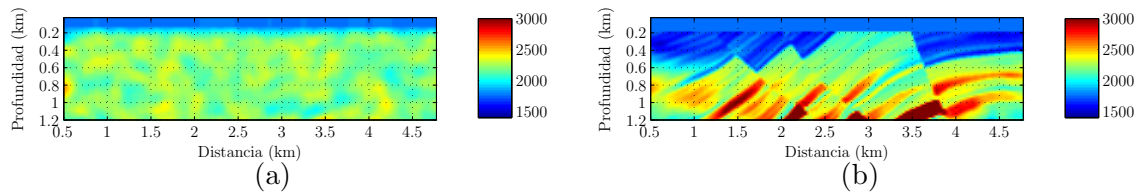


Figura 132: \mathbf{m}_{full} marmousi estimado por $\widehat{C}S_{FT} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

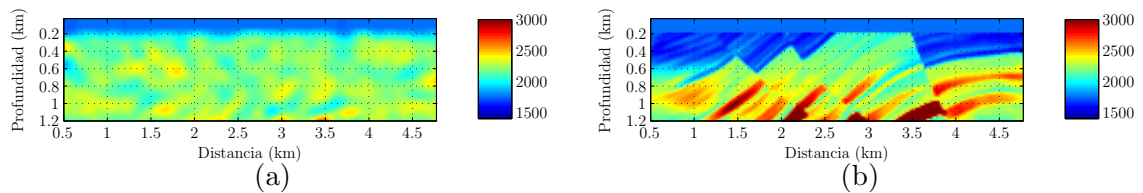


Figura 133: \mathbf{m}_{full} marmousi estimado por $\widehat{C}S_{FT} + \widehat{C}_{corr} + \widehat{C}S_{TC}$ (a) PSO (b) FWI.

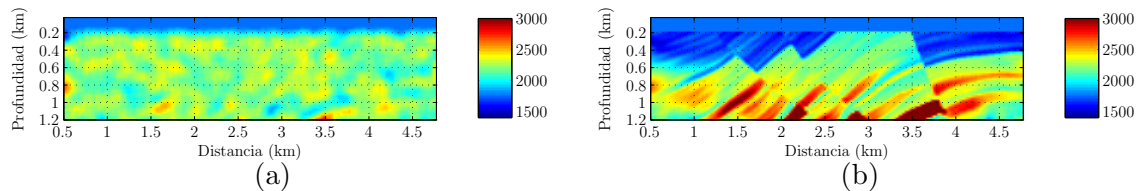


Figura 134: \mathbf{m}_{full} marmousi estimado por $\widehat{C}S_{FT} + \widehat{C}S_{TC}$ (a) PSO (b) FWI.

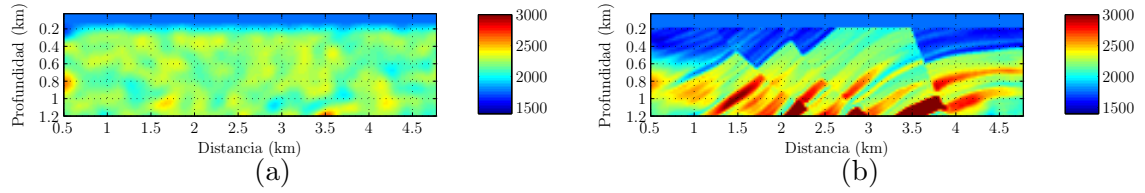


Figura 135: \mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{FT} + \widehat{CS}_{TC} + \widehat{C}_{corr}$ (a) PSO (b) FWI.

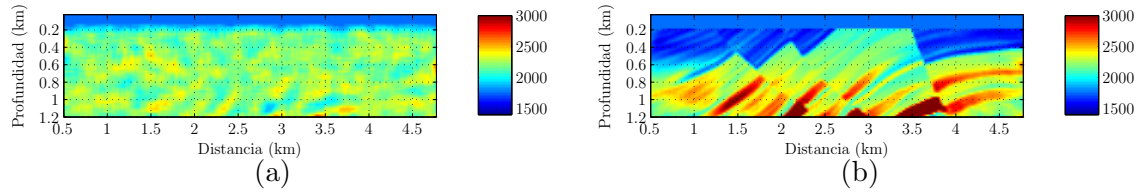


Figura 136: \mathbf{m}_{full} marmousi estimado por \widehat{CS}_{TC} (a) PSO (b) FWI.

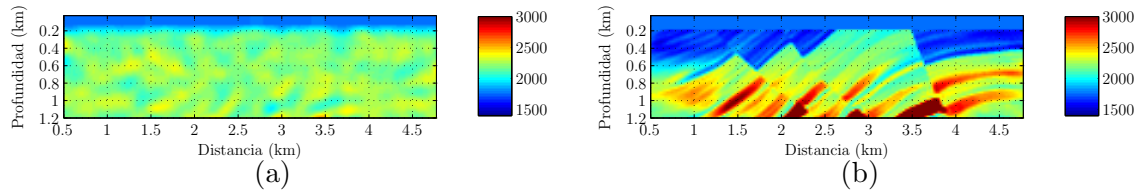


Figura 137: \mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr}$ a) PSO b) FWI.

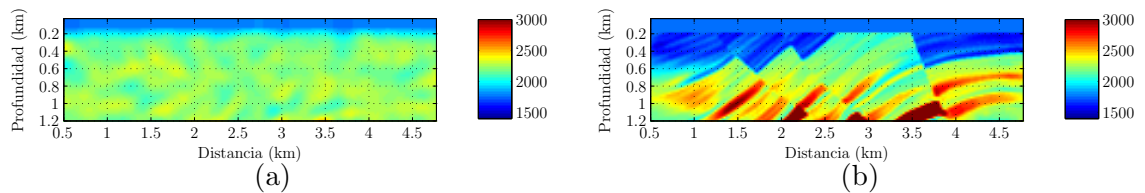


Figura 138: \mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{C}_{corr} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

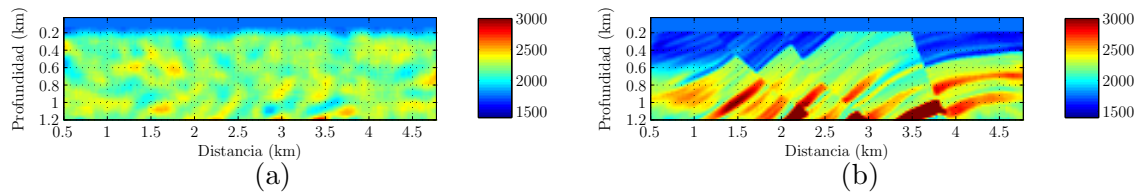


Figura 139: \mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT}$ (a) PSO (b) FWI.

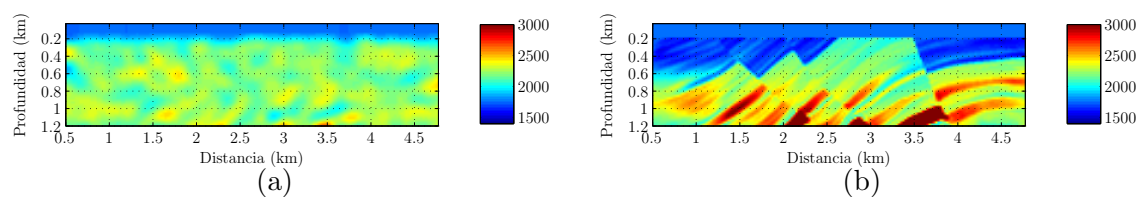


Figura 140: \mathbf{m}_{full} marmousi estimado por $\widehat{CS}_{TC} + \widehat{CS}_{FT} + \widehat{C}_{corr}$ (a) PSO (b) FWI.