

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE OBRAS CIVILES
VALPARAÍSO-CHILE



APLICACIÓN DEL MÉTODO BAYESIANO Y ABIC PARA
OBTENER INCERTIDUMBRE ASOCIADA A LA INVERSIÓN
DE TSUNAMI EN EL EVENTO DE IQUIQUE 2014

CAMILA PATRICIA CONTRERAS DELGADO

Ingeniero civil

Mayo 18, 2018

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE OBRAS CIVILES
VALPARAÍSO-CHILE



APLICACIÓN DEL MÉTODO BAYESIANO Y ABIC PARA
OBTENER INCERTIDUMBRE ASOCIADA A LA INVERSIÓN
DE TSUNAMI EN EL EVENTO DE IQUIQUE 2014

Memoria de título presentada por

CAMILA PATRICIA CONTRERAS DELGADO

Memoria para optar al título de

Ingeniero civil

Profesor guía

Dr. Patricio Andrés Catalán Mondaca

Mayo 18, 2018

Camila Contreras

*APLICACIÓN DEL MÉTODO BAYESIANO Y ABIC PARA OBTENER INCERTIDUMBRE
ASOCIADA A LA INVERSIÓN DE TSUNAMI EN EL EVENTO DE IQUIQUE 2014*

Memoria de Título, Mayo 18, 2018

Revisores: Patricio Catalán Marcos Valdebenito y Roberto Benavente

Profesor Guía: Patricio Catalán

Av. España 1680

Valparaíso

Resumen

En este trabajo se presenta el método bayesiano y ABIC para calcular incertidumbre en la inversión de tsunamis basado en las observaciones de los mareógrafos y boyas DART. De esta forma se calcula retrospectivamente la deformación inicial de la superficie del mar y su propagación hacia la costa, obteniendo estimaciones de altura de ola y tiempos de arribo, del evento de Iquique 2014 y sus correspondientes incertidumbres. Para este cálculo se aplica el método de Inversión de la forma de la onda de Tsunami [Tsushima et al., 2009] que consiste en una serie de funciones de Green, las cuales, al ser modeladas en el programa de simulación de tsunamis COMCOT, generan variaciones del nivel del mar en los mareógrafos y boya DART que son almacenadas en una base de datos. Al generarse un tsunami estos mareógrafos, que en el caso del evento de Iquique son en especial Arica, Iquique, Patache, Mejillones y la DART 32401, captan las variaciones de nivel, las cuales son el dato de entrada al algoritmo. De las series de tiempo de la base de datos, se plantean combinaciones lineales de parámetros desconocidos, calculándolos mediante mínimos cuadrados y Akaike's Bayesian Information Criterion (ABIC) [Akaike, 1980] para obtener de forma estimada las variaciones que captan los mareógrafos. Este método, además de entregar un resultado promedio de la superficie inicial, también arroja la incertidumbre asociada a cada parámetro de la combinación lineal, la cual disminuye a medida que se obtienen más datos.

Abstract

This study presents the bayesian method and ABIC to calculate uncertainty in tsunami inversion based on tide and DART's data observations. By this way it's calculated sea's initial deformation and its propagation to the coast in real time, obtaining seawave heights and arriving times estimates, applied to Iquique 2014 event. For this calculation it's implemented

the tsunami waveform inversion method [Tsushima et al., 2009] consisting of a set of unit Gaussian sources, modeled by the software of tsunami simulations COMCOT, that produce variations of the ocean's level in the tides and Dart buoy that are stored in a database. When a tsunami occurs, these tides, which in the case of Iquique's event are Arica, Iquique, Patache, Mejillones and DART 32401 buoy, they record these variations that are the data input for the algorithm. From the time series of the database, it is proposed lineal combinations of unknown parameters, calculating them by least squares method and Akaike's Bayesian Information Criterion (ABIC) [Akaike, 1980] to obtain estimates of the variations recorded by the tides. This method, in addition of delivering an average result of the initial surface, it also shows the uncertainty associated to each parameter of the lineal combination, which decline when more information in real time is available.

Agradecimientos

Nada de esto hubiese sido posible sin el apoyo de mi profesor guía, Patricio Catalán, el cual estuvo siempre dispuesto a guiarme, apoyarme y sobre todo a crear esa curiosidad. La curiosidad por conocer, investigar, que hace reflexionar y poner en duda lo conocido. Esa curiosidad es la base del conocimiento del cual me falta tanto por descubrir! y aún así me siento muy contenta del proceso, muchas gracias.

A los que aportaron esa pieza clave cuando estaba horas parada sin poder resolver algo. Entre ellos están Ale y Natalia por toda esa paciencia y ayuda cuando no sabía los comandos o algo me fallaba y que aún ayudan a todos a guiar el trabajo. A Alejandro Urrutia por enseñarme sobre Geoclaw, aunque no lo utilicé en la tesis, aprendí a usarlo y ahora es una herramienta más en mi formación. A Matías Araya que en algún momento del inicio me envió información crucial del monstruo. A Felipe, mi electrónico favorito. Muchas gracias por enseñarme de señales y los filtros. A Iván por acudir siempre a la Iván llamada y asistirnos en cosas de programación. A Jonathan que si bien no participó activamente en algo de mi tesis, siempre estuvo aquí para apoyarme en cualquier cosa urgente igual que todos en esta sala de memoristas. A Álvaro por ayudarme a escribir los agradecimientos y por supuesto a Mili que me ayudó a preparar la presentación con un diseño genial.

Obviamente a mi familia por el apoyo incondicional, por ellos estoy aquí en la universidad e hicieron lo que soy hoy! Mis padres son unos profesores geniales y siento una profunda admiración hacia lo que son y la vocación que tienen de enseñar. Los amo.

Y finalmente, pero de hecho uno de los más importantes, a Roberto Benavente quien me guió por los caminos de la estadística Bayesiana durante toda la segunda parte de esta investigación y que siempre tuvo esa disposición a escuchar, apoyarme y corregirme.

Bueno y a Álvaro por ayudarme a escribir los agradecimientos.

Contenido

1	Introducción	1
1.1	Objetivos Generales	2
1.2	Objetivos Específicos	2
2	Revisión Bibliográfica	4
3	Metodología	11
3.1	Software de propagación de oleaje: COMCOT	12
3.2	Datos disponibles	14
3.3	Método de inversión tFISH	16
3.3.1	Fuentes Unitarias	17
3.3.2	Problema de Inversión	19
3.4	Teorema de Bayes	20
3.4.1	Distribución a priori	21
3.4.2	Likelihood	22
3.4.3	Distribución a posteriori	24
3.4.4	Aplicación del teorema de Bayes	25
3.4.5	Cálculo de Hiperparámetros	26
4	Resultados	30
4.1	Experimento numérico	30
4.1.1	Condiciones del problema	31
4.1.2	Inversión	32
4.2	Aplicación al tsunami de Iquique 2014	36
4.2.1	Resultados de la inversión	37
5	Discusión de los resultados	42
5.1	Análisis de las hipótesis	42

5.1.1	Elección de los hiperparámetros	42
5.1.2	Covarianza relativa de los datos	46
5.1.3	No linealidad	52
5.1.4	Implicación a Futuro	56
6	Conclusión	58
	Referencias	60

Índice de Figuras

3.1	Esquema del proceso de inversión de Tsunami realizado en este trabajo	12
3.2	Propagación del programa COMCOT	14
3.3	Ubicación de mareógrafos y boya DART 32401 utilizadas para la inversión. El área de color blanco es el área abarcada por las fuentes unitarias.	15
3.4	Serie de tiempo del mareógrafo de Arica para el evento de Iquique 2014.	15
3.5	Esquema de el vector $\mathbf{b}$ de datos reales.	16
3.6	Fuentes unitarias utilizadas	17
3.7	Grillas utilizadas en el programa COMCOT	18
3.8	Esquema de la matriz $\mathbf{A}$ de base de datos de las funciones de Green.	19
3.9	Matriz de Covarianza residual no correlacionada, con valores diferentes γ_k para cada punto de pronóstico k . Los puntos aislados fuera de la diagonal representan valores nulos.	23
3.10	Matriz de Covarianza residual correlacionada	23
3.11	Diagrama de flujo del proceso de inversión.	29
4.1	Fuentes utilizadas en el ejemplo sintético	31
4.2	Comparación de valores de los parámetros para $\sigma_1 = 10$	32
4.3	Comparación de la señal invertida, usando 60 minutos de datos, con las originales con ruido y sin ruido en Iquique.	33
4.4	Resultados de la inversión para nivel de ruido $\sigma^2 = 10$	34
4.5	Rangos de valores de skill para periodos de 15 minutos en Iquique.	35
4.6	Comparación entre resultados de inversión para Iquique y datos medidos utilizados en la inversión de 1 hora	37
4.7	Resultados de la inversión con 1 hora de datos	39
4.8	Comparación de soluciones de deformación inicial.	39
4.9	Resultados de la inversión para el evento de Iquique 2014.	41
5.1	Variación de ABIC respecto de hiperparámetros para el experimento sintético	43

5.2	Valores de los parámetros ubicados en la superficie del mar del experimento numérico.	44
5.3	Resultados Inversión Iquique para diferentes valores de σ	45
5.4	Resultados de inversión con una hora de datos para distintos valores de matriz de covarianza	47
5.5	Predicción para el mareógrafo de Iquique	49
5.6	Mapa de calor de valores de Skill en el mareógrafo de Arica para cada tiempo de inversión en [min	50
5.7	Rangos de valores de skill para cada periodo de 15 minutos (s_i , comenzando desde el minuto i) en Arica	52
5.8	Puntos de pronóstico para el ejemplo no lineal, correspondientes a las ubicaciones de los mareógrafos de Arica, Iquique, Patache y Mejillones desplazados en 0.4 grados mar adentro y la boya DART 32401	54
5.9	Resultado superficies iniciales para caso No lineal	55
5.10	Resultados de diferentes tiempos de inversión a 0.4 grados mar adentro del mareógrafo de Iquique.	56
6.1	Resultados de la inversión con covarianza estacionaria correlacionada para 45 minutos de datos.	63
6.2	Resultados de la inversión con matriz no estacionaria y no correlacionada para 45 minutos de datos.	64
6.3	Resultados de la inversión con matriz no estacionaria y correlacionada para 45 minutos de datos	65
6.4	Resultados de la inversión con matriz estacionaria y correlacionada para 60 minutos de datos	66
6.5	Resultados de la inversión con matriz no estacionaria y no correlacionada para 60 minutos de datos	67
6.6	Resultados de la inversión con matriz no estacionaria y correlacionada para 60 minutos de datos.	68
6.7	Resultados de la inversión con matriz estacionaria y correlacionada para 180 minutos de datos.	69
6.8	Resultados de la inversión con matriz no estacionaria y no correlacionada para 180 minutos de datos	70
6.9	Resultados de la inversión con matriz no estacionaria y correlacionada para 180 minutos de datos	71

6.10	Mapa de calor para valores de Skill en el mareógrafo de Iquique	72
6.11	Mapa de calor para valores de Skill en el mareógrafo de Patache	72
6.12	Mapa de calor para valores de Skill en el mareógrafo de Mejillones	73
6.13	Mapa de calor para valores de Skill en el mareógrafo de DART	73
6.14	Rangos de valores de skill para inversión de 45 minutos en Arica.	75
6.15	Rangos de valores de skill para inversión de 180 minutos en Arica.	76
6.16	Rangos de valores de skill para inversión de 45 minutos en Iquique.	77
6.17	Rangos de valores de skill para inversión de 60 minutos en Iquique.	78
6.18	Rangos de valores de skill para inversión de 180 minutos en Iquique.	79
6.19	Rangos de valores de skill para inversión de 45 minutos en Patache.	80
6.20	Rangos de valores de skill para inversión de 60 minutos en Patache.	81
6.21	Rangos de valores de skill para inversión de 180 minutos en Patache.	82
6.22	Rangos de valores de skill para inversión de 45 minutos en Mejillones.	83
6.23	Rangos de valores de skill para inversión de 60 minutos en Mejillones.	84
6.24	Rangos de valores de skill para inversión de 180 minutos en Mejillones.	85
6.25	Rangos de valores de skill para inversión de 45 minutos en DART.	86
6.26	Rangos de valores de skill para inversión de 60 minutos en DART.	87
6.27	Rangos de valores de skill para inversión de 180 minutos en DART.	88

Índice de tablas

4.1	Resumen de valores obtenidos en el experimento numérico.	36
-----	--	----

Introducción

A lo largo de la historia, Chile se caracteriza por ser un país que sufre por diversos desastres naturales como erupciones, aludes, terremotos y en ocasiones también tsunamis. Estos últimos fueron los causantes de un tercio de las víctimas fatales del terremoto del 2010 y de pueblos enteros arrasados por las olas. Frente a esto es importante tener un sistema de alerta, que no sólo permita una evacuación rápida, si no que además entregue información detallada del impacto del tsunami en cada región de Chile.

Desde el 2010, luego de que el sistema de alerta temprana fallara, se comienzan a generar modelos de predicción de tsunamis en Chile y así estar preparados para eventos futuros. Algunos de ellos se basan en modelos existentes en otros países que han demostrado tener una buena capacidad predictiva. El mejor ejemplo es el operativo en Chile con el nombre de SIPAT (Sistema Integrado de Predicción y Alarma de Tsunamis), el cual está basado en el método japonés. Este sistema consta de un catálogo de terremotos con una magnitud y epicentro específico de los cuales se tienen premodelados sus respectivos tsunamis y en consecuencia, el grado de amenaza que podrían causar en cada región. Al generarse un terremoto, se buscan en el catálogo los más similares en magnitud y cercanos al epicentro para así ver la amenaza en la costa. Este modelo entrega una solución en segundos luego de ocurrido el terremoto, lo que permite una evacuación oportuna. Sin embargo, la precisión del modelo está sujeta al uso de instrumentos al medir la magnitud y epicentro del sismo. La complejidad de la ruptura cosísmica hace que escenarios similares pueden ocasionar tsunamis muy distintos entre sí a pesar de su similitud y cercanía [Geist, 2002]. En este contexto es que se transforma en algo fundamental generar otros modelos de la condición inicial, que puedan contrastar la información. Más aun, desarrollos recientes sugieren que es necesario tener una medida de la incertidumbre asociados a ellos [Tatsumi et al., 2014].

Un caso particular de métodos de predicción de tsunami es el método de inversión, desarrollado tanto para tsunamis como para terremotos. El método de inversión de Tsunamis [Satake 1989; Tsushima et al., 2009] intenta reconstruir la deformación inicial en el mar que genera el terremoto a partir de datos de serie de tiempo del nivel del mar en mareógrafos y otros sensores. Para ello requiere de una base de datos de series de tiempo en los mareógrafos

al propagar fuentes unitarias o perturbaciones en el mar. Si bien se tienen soluciones, éstas no son exactas debido a factores como la ubicación espacial de los mareógrafos, las distintas combinaciones lineales posibles entre ellas, variación de la batimetría, etc. Por consiguiente, se hace relevante una solución que considere toda la incertidumbre asociada al método y cómo ésta va cambiando a medida que se tienen más datos.

Esta memoria tiene por objetivo principal calcular la incertidumbre asociada a la inversión de tsunami por medio de la inferencia bayesiana, en específico el llamado Akaike's Bayesian Information Criterion (ABIC), el cual ajusta además ciertos parámetros relacionados con la inversión y la condición de suavidad de la superficie del mar (Hiperparámetros).

Todo este análisis se realiza primeramente utilizando un programa de propagación de olas (COMCOT) para el método de inversión y luego el programa Matlab para el procesamiento de la información y el cálculo de incertidumbre.

1.1 Objetivos Generales

- Crear un algoritmo para cuantificar la incertidumbre de la predicción de Tsunamis a través de un proceso de inversión, usando para ello el método Akaike's Bayesian Information Criterion (ABIC) para, a futuro, poder complementar el sistema de alerta en Chile.

1.2 Objetivos Específicos

- Generar deformaciones iniciales y sus ubicaciones espaciales para la red de funciones de Green del caso en estudio.
- Propagar las fuentes unitarias a lo largo de la costa, obteniendo un catálogo de series de tiempo en 5 boyas.
- Obtener un algoritmo que permita procesar los datos de la inversión.
- Conocer los valores de los hiperparámetros, tanto para el que limita la suavidad de superficie, como para el asociado a los datos.

- Analizar la precisión del método y la información que entrega en ciertos tiempos luego del inicio del evento.

Revisión Bibliográfica

“ *Unless we have some reasonable prediction for the future, we cannot choose appropriate actions for the present.*

— **Hirotsugu Akaike**
(Estadístico Japonés)

Un Tsunami es una serie de olas que se propagan alcanzando alturas que pueden llegar por sobre los 30 metros en la costa. Estas paredes de agua se producen cuando algún fenómeno extraordinario desplaza verticalmente una gran masa de agua, entre los cuales se destacan terremotos, erupciones volcánicas y deslizamientos de tierra. Este trabajo está enfocado específicamente a tsunamis provocados por terremotos debido a que son los más frecuentes en Chile.

Los efectos de los tsunamis en costas pobladas son devastadores, dejando víctimas fatales y pérdidas materiales importantes. Específicamente en Chile, los tsunamis tienen un tiempo de arribo de minutos, debido a la cercanía del foco de los terremotos a la costa, llamados tsunamis de campo cercano. Es por ello que en la predicción de tsunami no sólo es necesario el resultado, al llegar a la costa, sino también el tiempo con que lo arroja y su precisión.

A la fecha se han desarrollado varios métodos de predicción de tsunami, por medio de la proyección de la evolución del tsunami una vez ocurrido el evento sísmico, con el fin de tener un sistema que alerte a la población de forma oportuna y eficiente. Entre ellos cabe destacar el existente en Japón, el cual interpola de la base de datos de escenarios de tsunami precalculados basándose en un análisis de la magnitud y el epicentro de las ondas sísmicas obtenidas por una red de observaciones de la costa [Tatehata, 1997]. Este último es en el que se fundamenta el ya nombrado SIPAT, el cual es el sistema que se encuentra vigente en Chile.

En Estados Unidos el llamado Short-term Inundation Forecasting for Tsunamis (SIFT) es el sistema utilizado por la Administración Nacional Oceánica y Atmosférica (NOAA, por sus siglas en inglés) que modela en tiempo real [Titov et al., 2005]. Este sistema combina datos en tiempo real del evento y modelos numéricos para producir predicciones de tiempos de llegada y amplitudes. Al igual que el de Japón, tiene escenarios precalculados, pero de fuentes unitarias de magnitud de momento 7.5 y un metro de deslizamiento. A continuación éstas son escaladas y combinadas linealmente basados en la magnitud y ubicación del sismo. Una de las características principales de este método es que la predicción es para tsunamis de campo lejano, es decir, generados fuera de la costa de Estados Unidos, por lo que la llegada de las olas de tsunamis son más tardías que para campo cercano. Además, detalles de la ruptura son menos relevantes y la ventana de tiempo para alertar a la población es mayor.

Otro método explorado ha sido el de la utilización de ecuaciones de aguas poco profundas para aplicar un esquema de diferencias, volúmenes o elementos finitos donde la información de entrada sea la batimetría y la deformación inicial de la superficie libre. Sin embargo, por la capacidad computacional requerida para eventos de campo cercano, éstos no son viables para su modelación en tiempo real y alerta temprana. Además se requiere contar con información de gran calidad de la deformación inicial.

Abarcando el problema de otra perspectiva se ha desarrollado un método que reconstruye la deformación inicial de la superficie libre, llamado método de inversión de tsunami que ha sido utilizado tanto para determinar algunas características de los terremotos [Satake, 1989, Tanioka et al., 2006, Fujii, Satake, et al., 2011; Saito et al., 2011], como para la predicción de tsunami en sí [Tsushima et al., 2009]. El modelo se basa en fuentes unitarias, o funciones de Green, que al ser propagadas hacia la costa y los puntos de pronóstico es posible reconstruir la propagación real mediante la combinación lineal de ellas.

Una ventaja al utilizar el modelo de inversión de tsunami es que su propagación puede ser modelada de forma más precisa que la inversión sísmica porque la batimetría es mejor conocida y la velocidad de la superficie del mar sólo depende de ésta [Satake, 1989]. En cambio, en las ondas sísmicas, la estructura de las velocidades sísmicas depende de muchos factores como el patrón de radiación o heterogeneidades laterales del manto [por ejemplo, Woodhouse and Dziewonski, 1984]. Un caso en el que la inversión con ondas de tsunami es más confiable que los datos sísmicos son los llamados "Tsunami Earthquakes", los cuales son

eventos inusuales que estan acompañados de tsunamis de gran amplitud en contraste con ondas sísmicas de amplitud relativamente pequeña [Kanamori, 1972].

En Japón se cuenta con diversos sistemas de medición los cuales han sido utilizados en el método de inversión. Uno de ellos es el "Global Navigation Satellite Systems" (GNSS), sistema de navegación satelital, que provee posicionamiento geo-espacial con cobertura global [Blewitt et al., 2009].

Otro sistema son los OBTMs (Ocean bottom tsunami meters) [Tsushima et al., 2009] sensores de presión, los cuales aportan rapidez al método debido a su ubicación en el océano y proveen información sin las distorsiones causadas por la topografía compleja de la costa [Gonzalez et al., 1991]. Existen 6 redes de estos sensores en Japón, entre las cuales esta la denominada Dense Oceanfloor Network system for Earthquake and Tsunami (DONET).

En el caso de Chile, que no cuenta con gran instrumentación, se tiene fácil acceso a GPS (Global Positioning System) [Kawai et al., 2013; Melgar et al., 2016; Chen et al., 2016], un tipo de GNSS, debido a su bajo costo y su precisión. Además cuenta con 35 mareógrafos en la costa, 4 en la Isla de Pascua, Juan Fernández, San Félix y la Antártica y 5 boyas del NOAA, sistema llamado Deep-ocean Assessment and Reporting of Tsunamis (DART). Estos sistemas estan compuestos por una boya mar adentro y un registrador de presión de fondo, el cual tiene la capacidad de captar las variaciones del nivel del mar producidas por el paso de un tsunami, debido a que puede discriminar los largos periodos de onda de éste. A continuación envía los datos por sistema satelital a la costa donde puedan ser utilizados.

Debido a que el área abarcada por los sistemas de información es la línea de la costa y sólo cinco boyas mar adentro, en Chile se hace más complejo el sistema de inversión, lo que hace necesario la implementación de sistemas que, además de entregar resultados, entreguen su incertidumbre. Otro problema asociado a la instrumentación es el de la no linealidad y necesidad de batimetrías finas alrededor de éstos debido a su ubicación cercana a la costa [Fujii and Satake, 2006]. En consecuencia, variados autores resuelven el problema utilizando sólo los primeros ciclos de los datos de mareógrafos para hacer la inversión. Incluso tomando estas consideraciones, se puede encontrar a menudo en las inversiones una diferencia en la llegada de la onda a los mareógrafos. Esta llegada temprana de los datos sintéticos son generalmente producidas por los mismos mareógrafos o por las batimetrías poco precisas. Por esta razón se han propuesto formas de alinear estos resultados con los datos [Romano et al., 2016] a través de la minimización de una función de costo dependiente de un tiempo

de desplazamiento T , el cual puede ser distinto entre estaciones. Los resultados muestran que los deslizamientos obtenidos con este método son mucho mejores en datos sintéticos y en casos reales.

La estadística Bayesiana especifica un modelo de probabilidad a un parámetro del cual se tiene un conocimiento previo llamado a priori. El parámetro se ajusta a un conjunto de datos observados, obteniendo una distribución de probabilidad a posteriori que resumirá todo el conocimiento del investigador sobre los parámetros de interés dados los datos y sus creencias a priori. Como resultado de la estadística bayesiana se cuantifica la incertidumbre o grado de certeza. En un problema lineal, asumiendo que el prior y los residuales de los datos distribuyen normal, la distribución a posteriori también resulta normal y su máximo puede ser encontrado por métodos de mínimos cuadrados. Relativo a las inversiones, la restricción o información previa más comunmente utilizada es la de suavidad. Sin embargo, la relación de pesos entre el ajuste con los datos y la suavidad de los desplazamientos o variaciones en el nivel del mar no está determinada. Un criterio estadístico utilizado es el método ABIC (Akaike Bayesian information criterion), el cual aplica el teorema de máxima entropía al teorema de Bayes [Akaike, 1980] determinando los pesos, o también llamados hiperparámetros, que compensan el ajuste a los datos y la restricción de la información a priori. Este criterio ha sido utilizado en inversiones sísmicas por diversos autores [Uchida, 1993; Fukahata et al., 2003] como también para datos electromagnéticos [Mitsuhata et al., 2001] entre otras áreas.

Finalmente, se entrega los parámetros que más se acomodan al problema lineal y el nivel de ruido de los datos. Entre algunas de las ventajas del método se destaca que muestra la incertidumbre en los parámetros del modelo y en su predicción, que no depende de la cantidad de datos a utilizar y que puede trabajar con problemas mal condicionados. Esto último es importante para el caso de tsunamis debido a que la variación de alturas en los mareógrafos varía mucho en intervalos de tiempo corto.

En el caso especial de los deslizamientos sísmicos, citados anteriormente [Fukuda and Johnson, 2008], se impone una condición inicial para el algoritmo, la que restringe la negatividad de éstos admitiendo sólo deslizamientos en ciertas direcciones evitando los deslizamientos en direcciones opuestas a la falla lo que requiere una base computacional robusta y mayor tiempo de análisis debido a que la distribución a posteriori deja de ser gaussiana y no puede ser escrita analíticamente. En este contexto es que se utilizan técnicas bayesianas no lineales que incluyen muestreos numéricos, como Markov Chain Monte Carlo

para ser evaluados directamente en la distribución a posteriori. Estas técnicas no requieren una estimación de los pesos de la información a priori, ya que viene dada naturalmente al incluirlas como desconocidas en el problema [Tatsumi et al., 2014]. Sin embargo, es computacionalmente demandante debido a la evaluación de los muestreos de un número muy grande de candidatos. Es por ello que existe una forma de linealizar el problema [Benavente 2016] discretizando la superficie de falla. Una de las desventajas es la utilización de un gran número de parámetros y que se requiere una aproximación de la distribución a posteriori para la obtención de la solución, lo que conlleva errores.

En el caso de los tsunamis, en cambio, las deformaciones o variaciones del nivel del mar pueden ser tanto positivas como negativas, por lo que no es necesario una restricción de positividad, acelerando y simplificando los cálculos.

Entre inversiones de tsunami destaca Dettmer et al. (2016) el cual utiliza un método bayesiano no lineal transdimensional (trans-D), que consiste en grillas irregulares, es decir, discretizaciones desconocidas divididas mediante los coeficientes de Wavelet. La dimensión más pequeña es dividida en grillas de 15 km y con fuentes unitarias con forma de coseno alargado que se traslapan en a lo largo de media grilla. Luego utilizan el muestreo de Monte Carlo en la función a posteriori, moviéndose a través del espacio paramétrico trans-D con los algoritmos Metropolis-Hastings-Green (MHG) o salto reversible. Una de las ventajas es que utiliza la cantidad de parámetros de forma eficiente, es decir, los necesarios para describir la deformación inicial que en algunos lugares es más compleja que en otros.

Otro estudio [Takagawa, 2016] que, a diferencia del caso anterior, utiliza ABIC como método de cálculo para la incertidumbre; método de muestreo Monte Carlo para obtener soluciones de superficie inicial a partir de la función de densidad de probabilidad (escenarios estocásticos para mostrar incertidumbre); y simulación de inundación, muestra resultados con 10% de error incluso usando 10 minutos de datos observados, validándolos igualmente en Nankai para un evento de M9.1. Además utiliza dos informaciones a priori: la suavidad temporal y espacial. Cabe señalar que, en el cálculo, a diferencia de la inversión por desplazamientos, los parámetros no tienen condición de no negatividad, por lo que el análisis es más rápido.

La teoría bayesiana utiliza información a priori, como la suavidad, la cuales influye en la calidad de la inversión y de los resultados. Una de las hipótesis que se deben plantear es la forma de la matriz de covarianza residual de los datos, la cual se toma como una

matriz y su correspondiente hiperparámetro, σ , que cuantifica el peso del ajuste de los datos. Usualmente ésta se ingresa como una matriz identidad, lo que no entrega información, pero en trabajos como Fukahata et al., (2003) se resalta su importancia y se proponen matrices no diagonales, es decir, correlacionada y por tanto el ruido se relaciona temporalmente (cambia con un cierto patrón a medida que transcurre el tiempo). Cuando se tiene información de los datos imprecisa o insuficiente, el modelo da prioridad a la información a priori, en ciertos casos la suavidad. Sin embargo, en otros casos, cuando los datos son demasiados en periodos cortos de inversión, se produce un sobremuestreo en que los resultados del método calzan de forma excesiva con los datos incluyendo el ruido. Por ejemplo, si se toman datos cada 15 segundos como en esta tesis, la cantidad de información observada es mayor que si se toman cada 1 minuto. En un periodo corto de tiempo los datos contendrán poca información sobre la onda de tsunami (por su largo periodo), pero por la cantidad de datos el método supone una mayor información y por tanto le da mayor valor en el resultado de la inversión. Es por ello que, tomando en cuenta la covarianza, se puede reducir la información de sobremuestreo, lo cual evita una inversión sesgada. Una de las formas en que resuelve este problema [Duputel et al., 2012] es agregando una relación fuera de la diagonal debido al sobremuestreo, el cual decae exponencialmente dependiendo del tiempo de muestras y el periodo mínimo en el contenido de la información.

Otra de las aristas de este problema las presenta Yagi y Fukahata (2011), quienes realizan inversión a partir de datos sísmicos, muestran que la matriz de covarianza además del error en la discretización Yagi y Fukahata (2008), debe contener los errores del modelo que se originan a partir de las fuentes unitarias. Esto también es tratado por Minson et al. (2014) el cual genera dos matrices de covarianza, las cuales sumadas resultan en la matriz de covarianza total. Estas matrices tienen origen en los ruidos que contienen los datos y en los errores propios del modelo. Ambas se les consideran independientes y Gaussianas, y se considera que el error del modelo, el cual es un defecto particular de la teoría, es el predominante de la covarianza total. Estos estudios notaron que al considerar sólo una de las matrices resultaba en enormes sobreajustes del modelo a los datos, debido a que los coeficientes multiplicadores en la inversión sísmica son valores altos y al tener errores en el modelo, éstos se multiplicaban por los coeficientes y resultaban en diferencias de varios órdenes de magnitud con respecto a lo real. Si bien, la incertidumbre ha sido tratada con un prior de suavidad como regulador, ésta puede no ser suficiente para evitar la dependencia de la elección de los parámetros. Esto está referido a que la parametrización debería basarse en la resolución de los datos afectando la cantidad de parámetros libres. Una

sobre parametrización puede llevar a altas incertidumbres y por el contrario a predicciones que llevan a interpretaciones erróneas [Malinverno, 2000]. Para ello, es posible regular la cantidad de datos a utilizar en la inversión, o la búsqueda de mayor cantidad de información inicial, como puede ser la suavidad temporal. No obstante, tener una mayor cantidad de información inicial puede traer consecuencias en los tiempos necesarios para la ejecución del método, lo cual debe ser analizado con detalle.

El método de inversión utilizando ABIC, que incorpora suavidad, asocia incertidumbre según la calidad de los datos y errores propios del modelo no ha sido implementado en Chile para sistemas de alerta temprana. Esta investigación tiene como fin complementar y encontrar una solución alternativa para mejorar el sistema de alerta en Chile, por medio de la reconstrucción de la deformación inicial en el mar a partir de datos de mareógrafos.

En la siguiente memoria se muestra en primer lugar la metodología aplicada, explicando con detalle el software utilizado, los datos de entrada, el problema de inversión, el teorema de Bayes y finalmente el método ABIC. A continuación en la sección 4 se encuentran los resultados de la aplicación del método, tanto para un experimento numérico simple, hasta para el evento de Iquique 2014. En la sección 5 se discuten los resultados y se analizan distintas aristas, como la elección de los hiperparámetros, la covarianza relativa de los datos y la no linealidad, además de implicaciones a futuro de este trabajo. La sección 6 resume finalmente lo visto en la tesis y destaca retos para futuras investigaciones.

Metodología

Para complementar el sistema de alerta temprana se desarrolla en la presente sección el algoritmo tFISH (tsunami Forecasting based on Inversion for initial sea-Surface Height) [Tsushima et al., 2009] y el método ABIC, mostrando los pasos a seguir y explicando las hipótesis de ambos.

En la primera parte de la metodología se describe el programa COMCOT (Cornell Multigrid Coupled Tsunami model), en su última versión, utilizado para la simulación de la propagación del oleaje.

Luego de esto se describirán dos métodos. El primero es el método de inversión de tsunami tFISH y con él se genera una base de datos de series de tiempo. Estas series de tiempo son también llamadas funciones de Green y se generan como respuesta a un desplazamiento unitario en un punto en la superficie del mar para un cierto lugar de estudio (en este caso mareógrafos o boya DART). Una de sus ventajas es que permite la predicción en tiempo real usando datos provenientes del tsunami. Este algoritmo permite obtener resultados en cuestión de minutos.

A diferencia de lo hecho por Tsushima et al., 2009 el problema lineal, con los datos reales y las funciones de Green, se resolverá con un método que entrega mayor información llamado Akaike's Bayesian Information Criterion, el cual permitirá obtener la incertidumbre asociada a la inversión. El método de inversión al ser un problema lineal, sin restricción de positividad en sus coeficientes, hace que ABIC sea el más adecuado y rápido, ya que entrega la incertidumbre asociada al modelo y el valor promedio de superficie inicial, haciendo posible la generación de escenarios estocásticos.

En la figura 3.1 se muestra el proceso de inversión, donde en azul aparece la creación de la base de datos (tFISH), previo a cualquier evento, en naranja la situación en tiempo real y en verde el método ABIC para resolver el problema lineal. El primer paso para la creación de la base de datos consta en la generación de fuentes unitarias, las cuales se propagan con el programa COMCOT hacia los puntos de pronóstico, que en este caso son ciertos mareógrafos

y la boya DART, y se almacenan las series de tiempo. Cuando la señal en tiempo real llega (naranja) los puntos de pronóstico reciben las series de tiempo y a continuación se ingresa a la ecuación de combinación lineal (en el esquema *Problema lineal*). Para su resolución se utiliza el método ABIC, explicado más adelante, con el cual se tienen múltiples soluciones de diferentes probabilidades, obteniendo una deformación inicial más probable e incertidumbres en cada fuente unitaria.

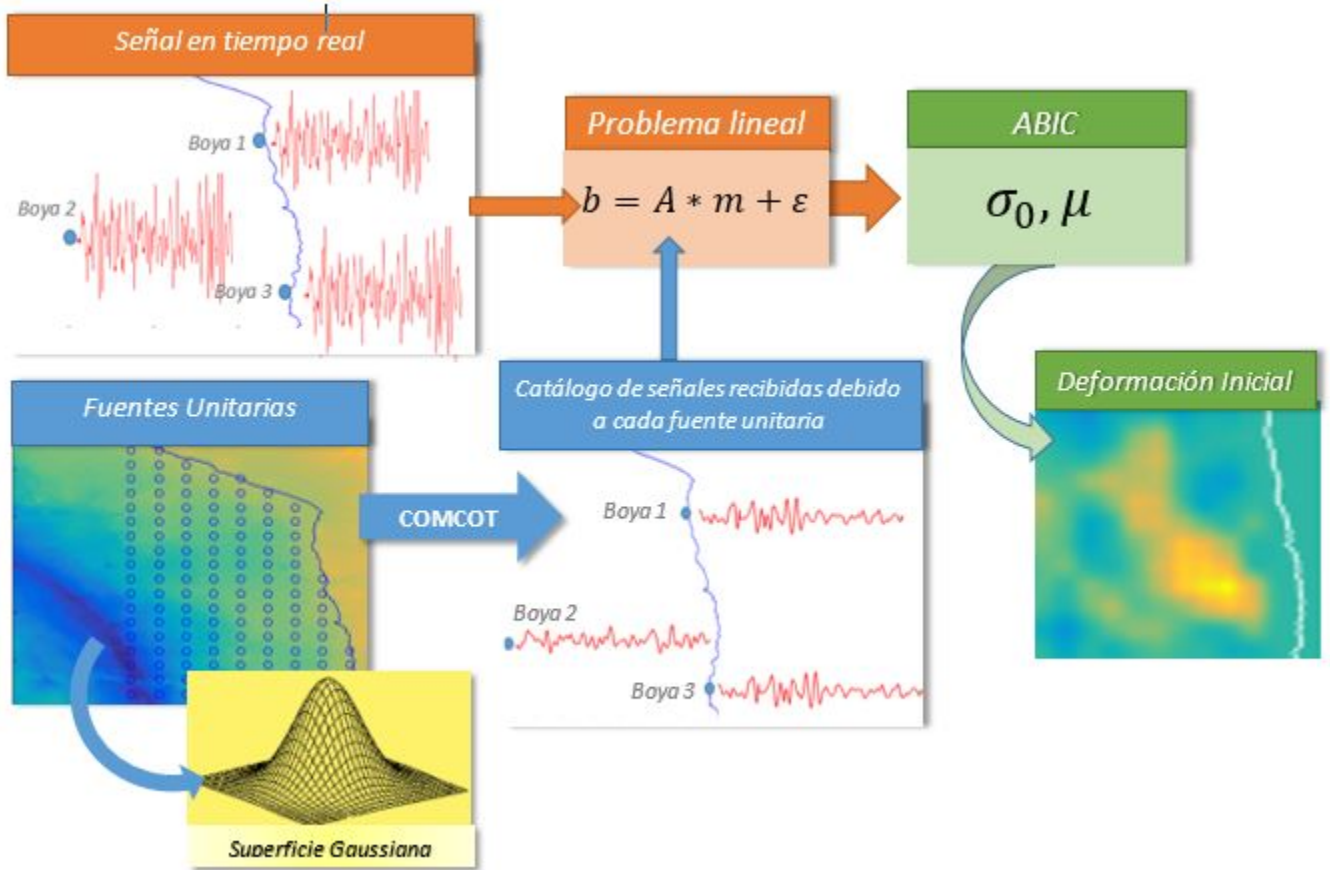


Fig. 3.1: Esquema del proceso de inversión de Tsunami realizado en este trabajo. En el problema lineal b es la señal en tiempo real, A el catálogo de las fuentes unitarias precalculadas, m son los parámetros a calcular y ϵ los errores o ruidos. En el método ABIC se muestran los hiperparámetros σ_0 y μ que se calculan para determinar el peso de los datos y del prior en la solución.

3.1 Software de propagación de oleaje: COMCOT

El programa COMCOT es un software capaz de simular todo el ciclo de vida de un tsunami desde su generación, propagación hasta su llegada a la costa y su área de inundación.

El archivo de entrada puede ser una superficie inicial de deformación del mar, un modelo de falla, deslizamientos de tierra o incluso un perfil personalizado.

En la mayoría de los eventos de tsunami la longitud de onda es mucho mayor que la profundidad del mar, por lo que el efecto de dispersión, es decir efecto nulo del fondo sobre el oleaje, puede ser despreciado. Siendo así las ecuaciones en aguas someras son adecuadas para eventos de tsunami, sin embargo su resolución es compleja ya que tienen derivadas de alto orden asociadas a la no linealidad, debido a la complejidad de las batimetrías que producen distintos efectos sobre las olas. COMCOT desarrolla estas ecuaciones con modelos numéricos, usando un método de diferencias finitas en el que se aplican directamente las ecuaciones sin necesidad de resolverlas explícitamente y así generar sus simulaciones. La forma de calcular la propagación se basa en grillas con distintas resoluciones y un sistema de anidado para hacer más eficiente el cálculo. Los resultados pueden ser entregados en diferentes formas y cantidades, dependiendo de lo requerido por el usuario. Uno de los resultados se muestra en la figura 3.2 el cual es la propagación del evento de 2014 a partir de la deformación inicial. Además, a un costado, se puede apreciar como reciben las variaciones del nivel del mar los mareógrafos, registrándolas como series de tiempo.

El software ha sido usado para investigar diversos eventos de tsunami como el de Chile en 1960, Indonesia en 1992 [P. L.-F. Liu et al., 1995], Algeria en 2003 [Wang and P. Liu, 2005] y en el Océano Indico en 2004 [Wang and P. L.-F. Liu, 2006].

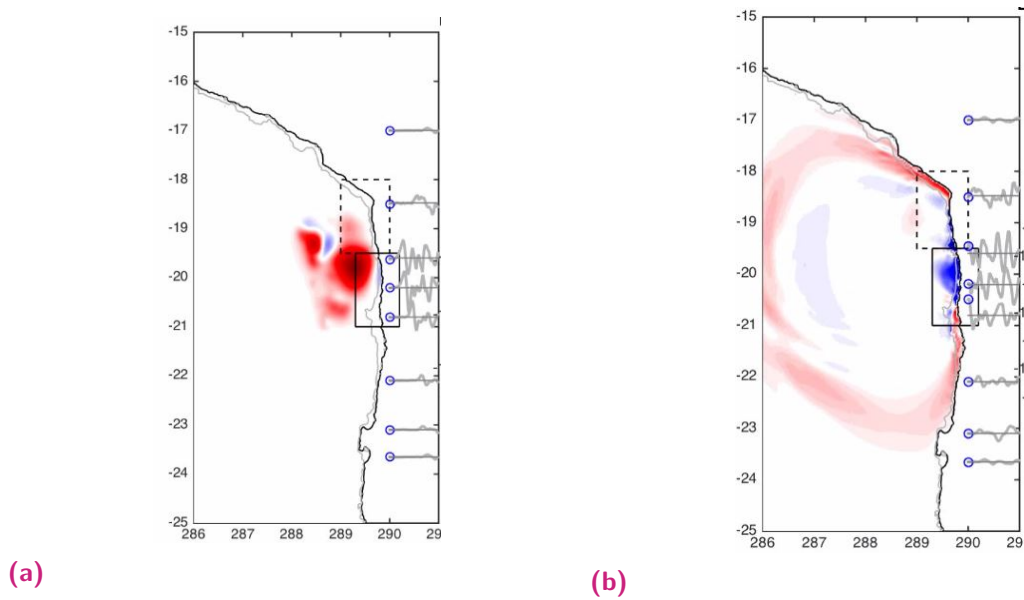


Fig. 3.2: Propagación del programa COMCOT: (a) Deformación inicial de la superficie del mar debido al terremoto de 2014; (b) Propagación de la deformación hacia la costa. Fuente: Patricio Catalán.

3.2 Datos disponibles

Los datos utilizados en este trabajo corresponden a 4 mareógrafos de Chile para el evento de Iquique ($M_w=8.2$) y los de la boya DART 32401 que aparecen en la figura 3.3.

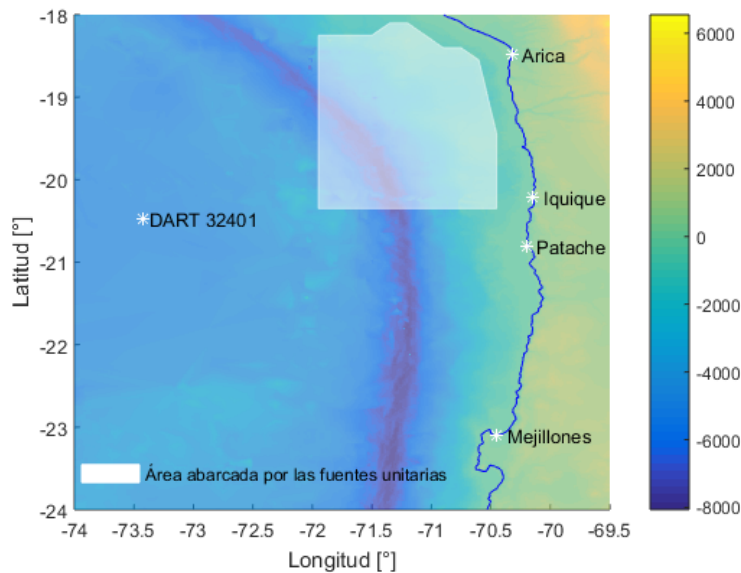


Fig. 3.3: Ubicación de mareógrafos y boya DART 32401 utilizadas para la inversión. El área de color blanco es el área abarcada por las fuentes unitarias.

Estos datos son procesados para eliminar el efecto de las mareas y poder invertirlos. En la figura 3.4 se muestra el registro de Arica desde el tiempo de inicio del terremoto a las 20:46 del día 1 de Abril de 2014. Los registros utilizados son de un largo máximo de cuatro horas y mínimo de 30 minutos después de transcurrido el terremoto, de los cuales para cada mareógrafo se utiliza sólo desde el tiempo en que arriba el tsunami (25, 0, 20, 27 y 18 [min] para Arica, Iquique, Patache y Mejillones respectivamente) para evitar considerar información del movimiento debido al sismo en la inversión. Estos rangos de registros son necesarios para verificar que el modelo es capaz de acercarse a la solución real en distintos tiempos después de ocurrido el tsunami y así conocer su capacidad predictiva. Los registros conforman el vector $\mathbf{b}$ como se muestra en la figura 3.5

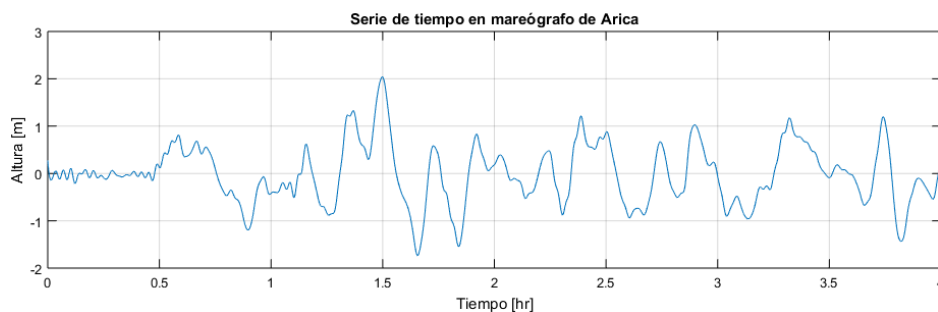


Fig. 3.4: Serie de tiempo del mareógrafo de Arica para el evento de Iquique 2014.

$$\mathbf{b} = \begin{bmatrix} \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \begin{matrix} \left\{ \begin{matrix} b_1(t_1) \\ b_1(t_2) \\ \vdots \end{matrix} \right\} \\ \left\{ \begin{matrix} b_2(t_1) \\ b_2(t_2) \\ \vdots \end{matrix} \right\} \\ \left\{ \begin{matrix} b_3(t_1) \\ b_3(t_2) \\ \vdots \end{matrix} \right\} \end{matrix} \end{bmatrix}$$

Punto de
pronóstico

Fig. 3.5: Esquema de el vector $\mathbf{b}$ de datos reales.

Adicionalmente, para la propagación de las ondas, es necesario contar con una batimetría fina que permita obtener resultados precisos. Para ello se divide el mapa en grillas con diferentes ubicaciones y dimensiones, las cuales puedan abarcar el plano general y además, con resolución más fina, la batimetría compleja de la costa cercana a la boya. Se utiliza, entonces, batimetría de 30 segundos de arco de datos de GEBCO08 para la grilla 1, mientras que para los niveles 2 y 3 se construyeron a partir de batimetría con mayor resolución de los mapas náuticos del SHOA.

3.3 Método de inversión tFISH

Para lograr la inversión del tsunami se desarrolla el algoritmo tFISH, el cual consta básicamente de dos partes. La primera es la creación de las series de tiempo de las fuentes unitarias en cada uno de los puntos de pronóstico mencionados anteriormente. La segunda es el planteamiento del problema lineal, el cual es resuelto en la siguiente sección. Éste es un proceso lento, debido a que cada fuente unitaria debe ser propagada hacia la costa con ciertos niveles de resolución espacial. Para ello es necesario contrastar los valores obtenidos al propagar una situación real donde se tengan las series de tiempo en la costa con una resolución definida que sea suficiente para entregar buenos resultados, sin ser excesivamente fina para que el tiempo de procesamiento sea demasiado. Se utilizó como guía lo planteado

por Tsushima et al., (2009) para las fuentes unitarias en cuanto a dimensiones y distancias entre ellas.

3.3.1 Fuentes Unitarias

Para generar la superficie inicial de tsunami es necesario crear fuentes, cuya máxima altura es unitaria. La forma de las fuentes ha ido variando de estudio en estudio, como cajas sobrepuestas tipo "furgón" [Tsushima et al., 2009] o como pirámide con techo plano [Satake, 1989]. La forma escogida, para este caso, es una gaussiana [Saito et al. 2010] debido a que su combinación puede generar fácilmente superficies iniciales "suaves", característica fundamental de la superficie inicial de tsunami. La ecuación que la caracteriza es:

$$B(x, y) = d_0 \exp \left[-\frac{(x - x_0)^2 + (y - y_0)^2}{(L/2)^2} \right] \quad (3.1)$$

donde $d_0 = 1$ m, $L = 22.26$ km $= 0.2^\circ$ y (x_0, y_0) son las coordenadas del centro de la gaussiana. El largo escogido L se considera de forma tal que, con la separación utilizada por Tsushima et al., (2009) de 0.15° (16.7 [km]) entre los centros de cada fuente unitaria, pueda crear una superficie suave, es decir, que exista superposición entre ellas, como se muestra en la figura 3.6.

El área abarcada por las fuentes unitarias, para efectos de caracterizar el tsunami del 2014 en Iquique, se muestra también en la figura 3.6, considerando resultados de superficies iniciales de otros autores como referencia [An et al. 2014]

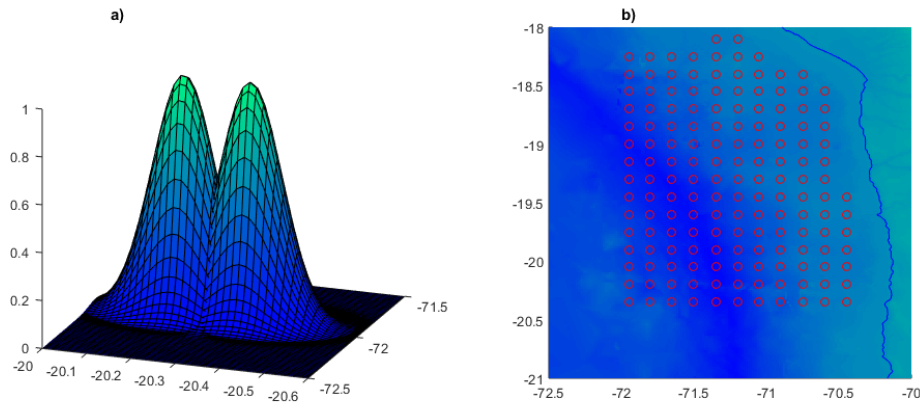


Fig. 3.6: Fuentes unitarias: (a) Esquema con dos fuentes adyacentes separadas 0.15° (b) Ubicación de las fuentes

Estas fuentes unitarias son propagadas individualmente en el programa COMCOT, obteniendo de éste series de tiempo en los puntos de pronóstico. Los puntos de pronóstico, son las boyas que para el evento recogen los datos de variación de alturas para posteriormente ser invertidos. Para el evento de Iquique 2014 se utiliza la boya DART 32401 y los mareógrafos correspondientes a Arica, Iquique, Patache y Mejillones. Estas boyas son escogidas debido a su cercanía y a la cantidad de datos registrados, los cuales debiesen completar la serie de tiempo algunas horas después de generado el terremoto. Las grillas usadas para la propagación en el programa, mostradas en la figura 3.7, son de distintas dimensiones para cada uno de los mareógrafos, dependiendo del requerimiento computacional para encontrar valores de variación de altura del mar de forma precisa y efectiva. En negro aparece la primera grilla con resolución de 120 seg. de arco. Para las grillas intermedias de colores *cián*, *magenta*, *amarillo*, *azul*, resolución de 60 segundos de arco. Dentro de este segundo nivel se encuentran otros aún más finos con tamaño de grilla de 30 seg de arco, los que describen las características de la batimetría cerca de las ubicaciones de los mareógrafos con más detalles.

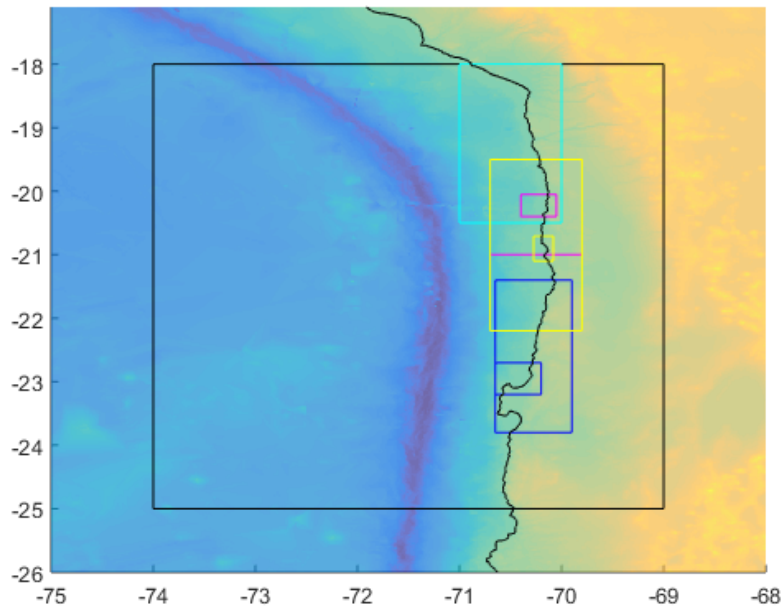


Fig. 3.7: Grillas utilizadas en el programa COMCOT. En *negro* la primera grilla y general. Anidadas a ésta están en color: *cián* para Arica; *magenta* para Iquique; *amarillo* para Patache; *azul* para Mejillones.

Como resultado de este proceso se tendrá el catálogo $\mathbf{A}$ de funciones de Green en los mareógrafos debido a cada una de las fuentes unitarias, el cual se utiliza para el problema lineal y se muestra en la figura 3.8.

$$\mathbf{A} = \begin{array}{c} \text{Función de} \\ \text{Green} \end{array} \begin{array}{c} 1 \quad 2 \quad \dots \quad 156 \end{array} \begin{array}{c} \left[\begin{array}{c} 1 \left\{ \begin{array}{c} A_{11}(t_1) \quad A_{12}(t_1) \quad \dots \quad A_{1 \ 156}(t_1) \\ A_{11}(t_2) \quad A_{12}(t_2) \quad \dots \quad A_{1 \ 156}(t_2) \\ \vdots \quad \vdots \quad \quad \quad \vdots \end{array} \right. \\ 2 \left\{ \begin{array}{c} A_{21}(t_1) \quad A_{22}(t_1) \quad \dots \quad A_{2 \ 156}(t_1) \\ A_{21}(t_2) \quad A_{22}(t_2) \quad \dots \quad A_{2 \ 156}(t_2) \\ \vdots \quad \vdots \quad \quad \quad \vdots \end{array} \right. \\ \vdots \\ 5 \left\{ \begin{array}{c} A_{51}(t_1) \quad A_{52}(t_1) \quad \dots \quad A_{5 \ 156}(t_1) \\ A_{51}(t_2) \quad A_{52}(t_2) \quad \dots \quad A_{5 \ 156}(t_2) \\ \vdots \quad \vdots \quad \quad \quad \vdots \end{array} \right. \end{array} \right] \end{array} \begin{array}{c} \text{Punto de} \\ \text{pronóstico} \end{array} \begin{array}{c} \Sigma_{k=1}^5 t_k \times \Sigma_{k=1}^5 t_k \end{array}$$

Fig. 3.8: Esquema de la matriz $\mathbf{A}$ de base de datos de las funciones de Green.

3.3.2 Problema de Inversión

El método de inversión tiene como hipótesis de trabajo la linealidad de la fuente debido a que las longitudes de onda de tsunami con respecto a la batimetría son muy grandes (ecuaciones de aguas someras) y que las amplitudes de la onda de tsunami son mucho menores a las profundidades mar adentro y por lo tanto las ecuaciones que describen su movimiento son de primer orden, es decir, lineales. Entonces, se resuelve la siguiente combinación lineal:

$$\mathbf{A}_{(tN) \times M} \mathbf{m}_{M \times 1} + \boldsymbol{\varepsilon}_{(tN \times 1)} = \mathbf{b}_{(tN \times 1)}, \quad (3.2)$$

donde $\mathbf{A}$ es el catálogo de la superficie libre del mar en un tiempo t en los puntos N de pronóstico (boyas) de la propagación de las M fuentes unitarias, $\mathbf{m}$ son los coeficientes a calcular para cada fuente unitaria M , $\boldsymbol{\varepsilon}$ es el término de error aleatorio para cada tiempo t y $\mathbf{b}$ es el vector de series de datos en tiempo real t obtenidos de las mismas N boyas. Esta combinación lineal se resuelve para cierto instante, por ejemplo 30 minutos transcurridos desde el terremoto, utilizando todos los datos disponibles hasta ese tiempo.

El error de la ecuación anterior se asume como ruido gaussiano, requerimiento para poder utilizar mínimos cuadrados y el método ABIC en la siguiente etapa. La veracidad de esta hipótesis se escapa al objetivo de esta tesis, sin embargo, puede ser explicado brevemente con el teorema del límite central, debido a que la muestra aleatoria es suficientemente grande. Además, ha sido demostrado que, si la información contiene datos aislados, es posible con esta hipótesis identificarlos y removerlos de los datos en el dominio del tiempo, como lo realizado por Tatsumi et al. (2014). Si bien es posible realizarlo, la incertidumbre debería abarcar los errores de los datos de los mareógrafos, por lo que no se desarrolla en esta tesis.

A diferencia de lo realizado por Tsushima et al., (2009) -que resuelve imponiendo restricciones de suavidad, área de influencia (la cual se ve afectada por los desplazamientos de la superficie libre) y distancia al epicentro-, se resuelve el problema de inversión utilizando la estadística bayesiana y el método ABIC, incorporando la restricción de suavidad como información a priori del modelo, lo cual se trata en la siguiente sección.

3.4 Teorema de Bayes

El teorema de Bayes, basado en la estadística bayesiana, describe la incertidumbre que se tiene sobre un cierto evento basado en un conocimiento previo o creencia del evento (que puede ser mínimo en caso de no poseerlo) y su actualización por medio de la evidencia. Esta basado en una interpretación de la probabilidad como una medida condicional de incertidumbre [Bernardo, 2003]. Una de las diferencias claves de este enfoque con el frecuentista es que los parámetros no son fijos y desconocidos, si no que aleatorios e inciertos. El teorema de Bayes muestra que la probabilidad para ciertos parámetros aleatorios $\mathbf{m}$ dada por los datos $\mathbf{b}$ es

$$p(\mathbf{m}|\mathbf{b}) = \frac{p(\mathbf{b}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{b})}, \quad (3.3)$$

la cual es llamada densidad de probabilidad a posteriori. El numerador $p(\mathbf{b}|\mathbf{m})$ es la función de verosimilitud o likelihood del modelo, $p(\mathbf{m})$ es la densidad de probabilidad de a priori y el denominador $p(\mathbf{b})$ es la función de densidad de probabilidad de los datos, la cual es una constante de normalización. $p(\mathbf{m}|\mathbf{b})$ es la función de densidad de probabilidad a posteriori, la cual muestra la probabilidad del modelo elegido dados los datos. En los siguientes párrafos se explican con más detalle las distribuciones de probabilidad del teorema y se determinan para el problema de inversión.

3.4.1 Distribución a priori

La distribución a priori en la teoría bayesiana se caracteriza por introducir información inicial del evento, que depende de las creencias o juicio. Esta distribución es de vital importancia para el método bayesiano, ya que las conclusiones obtenidas dependen de éste y por tanto su elección debe ser rigurosa. Usualmente para el caso de inversión se utiliza la condición de suavidad de la superficie [Yabuki and Matsu'Ura, 1992; Fukuda and Johnson, 2008; Satake, 1989;]. En otras palabras, cada parámetro debe ser similar a los parámetros de las fuentes unitarias aledañas, que viene de igualar el campo del Laplaciano a cero. La ecuación general que la describe es de la forma

$$-4m_{i+(j-1)r_j} + m_{i-1+(j-1)r_j} + m_{i+1+(j-1)r_j} + m_{i+(j-2)r_j} + m_{i+jr_j} = 0 \quad (3.4)$$

donde m es el vector de coeficientes en la posición espacial de las fuentes unitarias i, j , donde $i = 1, \dots, r_j$ se numera de oeste a este con un r_j variable por la forma de la costa según j (Ver figura 3.6) y $j = 1, 2, \dots, 16$ de sur a norte. En otras palabras, se realiza un promedio entre las cuatro fuentes más cercanas. De forma matricial queda como:

$$S = \begin{bmatrix} S_1 & I & 0 & \dots & 0 \\ I & S_2 & I & & \vdots \\ 0 & I & \ddots & \ddots & 0 \\ \vdots & & \ddots & S_{n-1} & I \\ 0 & \dots & 0 & I & S_n \end{bmatrix}_{M \times M} \quad (3.5)$$

donde la matriz

$$S_k = \begin{bmatrix} -4 & & & & \\ 1 & -4 & & & \\ 0 & 1 & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \dots & 0 & 1 & -4 \end{bmatrix}_{r_j \times r_j} \quad (3.6)$$

es simétrica.

La relevancia de la suavidad se ajusta con el hiperparámetro α , el cual, a mayor valor, permite mayores variaciones en la superficie y, al ser menor, una mayor suavidad. Al igual que los errores antes mencionados, se trabaja con la hipótesis que las probabilidades asociadas

a los parámetros se comportan como gaussianas multivariadas a medida que se obtienen mayor cantidad de datos. La probabilidad a priori queda entonces:

$$p(\mathbf{m}) = (2\pi\alpha^2)^{-\frac{M}{2}} |\mathbf{\Lambda}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\alpha^2} \mathbf{m}^T \mathbf{S}^T \mathbf{S} \mathbf{m}\right], \quad (3.7)$$

la cual corresponde directamente a la probabilidad de los parámetros dada una información inicial $\mathbf{S}$, con una media igual a cero y una covarianza de $\alpha^2 \mathbf{\Lambda}$, siendo $\mathbf{\Lambda}$ el determinante $\mathbf{S}^T \mathbf{S}$ [Fukahata et al., 2003]. Cabe mencionar que el cálculo de $\mathbf{\Lambda}$ no es necesario finalmente para la obtención de los coeficientes ni para la incertidumbre debido al desarrollo matemático que se explica en las siguientes secciones.

3.4.2 Likelihood

El likelihood o verosimilitud es el que introduce información del evento (datos $\mathbf{b}$) como forma de actualizar la incertidumbre del parámetro aleatorio.

Se determina la verosimilitud dados los datos con la forma de gaussiana multivariada, la que se muestra continuación:

$$p(\mathbf{b}|\mathbf{m}, \sigma_0) = (2\pi\sigma_0^2)^{-\frac{N}{2}} |\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma_0^2} (\mathbf{b} - \mathbf{A}\mathbf{m})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{A}\mathbf{m})\right] \quad (3.8)$$

En la función $\mathbf{A}\mathbf{m}$ es el valor medio, donde $\mathbf{A}$ es la matriz de la base de datos y $\mathbf{m}$ son los coeficientes de la ecuación 3.2, $|\cdot|$ es el determinante y ^{T} es la traspuesta de la matriz. En consecuencia, si $\mathbf{A}\mathbf{m}$ es la estimación de la señal mediante la inversión y $\mathbf{b}$ son los datos, su diferencia es el residual entre ambos y por tanto representa su distribución de probabilidad, asumido como una gaussiana multivariada, siendo $\sigma_0^2 \mathbf{\Sigma}$ la matriz de covarianza, donde σ_0 es un factor de corrección que compensa la estimación de la covarianza relativa explicado con mayor detalle en 3.4.5.

Como se dijo anteriormente la covarianza relativa es estimada, ya que no se tienen valores de los parámetros aún para conocer la covarianza del residual. Este es un factor de discusión entre diversos autores [Dettmer et al., 2007, Duputel et al., 2012], variando desde una matriz identidad a una con correlaciones (elementos fuera de la diagonal), asumiendo entonces una dependencia entre los valores en el tiempo. En el primer caso con una matriz identidad, $\sigma^2 \mathbf{I}$, lo que supone es nula correlación entre el primer dato recogido y el segundo y que incluso tienen el mismo valor entre un punto de pronóstico y otro igual a σ^2 . En el

segundo caso, figura 3.9, los elementos no poseen correlación entre ellos al igual que en el primero (matriz diagonal), pero si existe diferencia entre el valor de un punto de pronóstico respecto al otro.

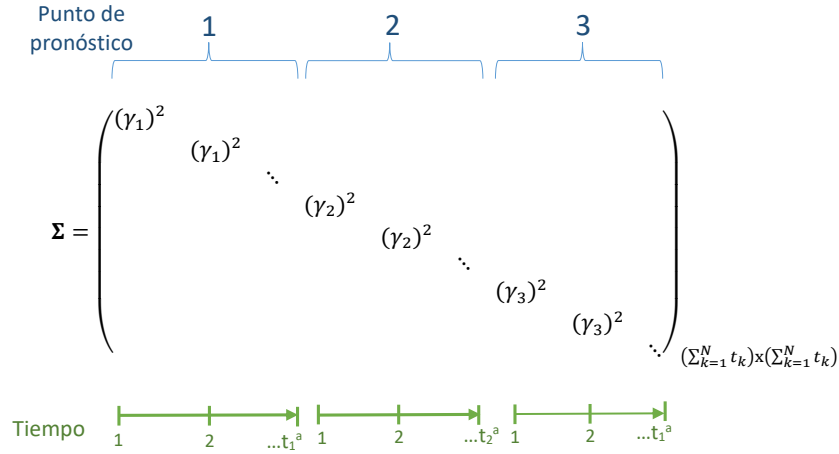


Fig. 3.9: Matriz de Covarianza residual no correlacionada, con valores diferentes γ_k para cada punto de pronóstico k . Los puntos aislados fuera de la diagonal representan valores nulos.

En el tercer caso, figura 3.10 se utilizan correlaciones o valores fuera de la diagonal. Estas correlaciones son del tipo temporal, siendo los elementos de cada mareógrafo correlacionados entre sí, pero los elementos entre mareógrafos, espaciales, nulos.

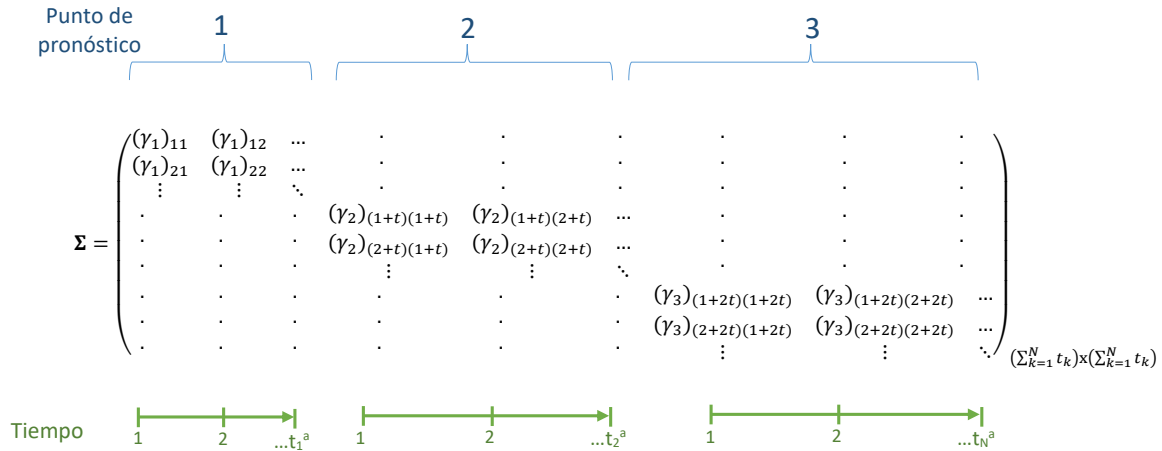


Fig. 3.10: Matriz de Covarianza residual correlacionada, con valores diferentes γ_k para cada punto de pronóstico k y correlaciones temporales. Los puntos aislados fuera de la diagonal representan valores nulos.

La estimación de Σ utilizada en el primer caso de esta tesis, recoge la varianza de los datos de cada mareógrafo desde que se inicia el terremoto hasta que llegan las primeras ondas. Para cada uno de los 5 conjuntos de datos (4 mareógrafos y la boya DART) se aplica la ecuación

$$(\gamma_k)^2 = \frac{\sum_{u=1}^{t_k^a} [(b_k)_u - \bar{b}_k]^2}{t_k^a} \quad (3.9)$$

en la cual $(b_k)_u$ representa al dato u del subconjunto de datos de $\mathbf{b}$ de cantidad t_k^a que es el tiempo de arribo del tsunami al mareógrafo k . Luego se construye la matriz de dimensiones $\sum_{k=1}^N t_k \times \sum_{k=1}^N t_k$ como la de la figura 3.10 con cinco puntos de pronóstico ($N = 5$) y con una cantidad de datos t_k dependiente de la cantidad de datos disponibles para la inversión.

Las correlaciones utilizadas para este caso han sido las de decaimiento exponencial [Duputel et al., 2012] del tipo

$$(\gamma_k)_{ij} = (\gamma_k)^2 \exp(-|\Delta t_{ij}|/t_0), \quad (3.10)$$

la cual es una matriz temporal con Δt_{ij} como la diferencia de tiempo entre las componentes i, j y t_0 como el periodo más corto contenido en los datos, que en este caso se escogió como media igual a 2 [min]. Además es importante destacar que estas correlaciones se cumplen sólo para datos i, j captados a partir del mismo punto de pronóstico k .

3.4.3 Distribución a posteriori

Una vez descritas las funciones de densidad de probabilidad (PDF) es posible utilizar el teorema de Bayes (Ver ecuación 3.12) y para simplificar el cálculo matemático, el hiperparámetro α [Mitsuhata et al., 2001, Yadav et al., 2013] se reemplaza por,

$$\alpha^2 = \frac{\sigma_0^2}{\mu}, \quad (3.11)$$

donde μ es una variable auxiliar. De acuerdo a todo lo anterior, la ecuación 3.3 queda como

$$p(\mathbf{m}|\mathbf{b}) = c(2\pi)^{-\frac{N}{2}} \left(\frac{\mu}{2\pi\sigma_0^2} \right)^{\frac{M}{2}} |\sigma_0^2 \Sigma|^{-\frac{1}{2}} |\mathbf{S}^T \mathbf{S}|^{-\frac{1}{2}} \exp \left[-\frac{U(\mathbf{m})}{2\sigma_0^2} \right], \quad (3.12)$$

donde

$$U(\mathbf{m}) = (\mathbf{b} - \mathbf{A}\mathbf{m})^T \Sigma^{-1} (\mathbf{b} - \mathbf{A}\mathbf{m}) + \mu \mathbf{m}^T \mathbf{S}^T \mathbf{S} \mathbf{m}. \quad (3.13)$$

y

$$c = \frac{1}{p(\mathbf{b})} = \frac{1}{\int p(\mathbf{b}|\mathbf{m}; \sigma_0^2) p(\mathbf{m}; \mu) d\mathbf{m}} \quad (3.14)$$

3.4.4 Aplicación del teorema de Bayes

El teorema de Bayes, explicado al inicio de ésta sección (sección 3.4) tiene como datos de entrada las funciones de densidad de probabilidad a priori ($p(\mathbf{m})$) y de verosimilitud ($L(\mathbf{b}|\mathbf{m}, \sigma_0)$). En la primera se tiene la hipótesis que la superficie del mar se deforma de forma suave y que la distribución tiene la forma de una gaussiana multivariada. La segunda, que comparte la misma forma de gaussiana multivariada, ingresa al algoritmo las series de tiempo recibidas por los mareógrafos en tiempo real.

Con este método se obtiene una función de distribución de probabilidad de los coeficientes, tipo gaussiana multivariada, que contiene una media de los parámetros, o los valores más probables, y su varianza.

De la teoría bayesiana es posible afirmar que los parámetros del modelo más probable se encuentran en el máximo de la distribución a posteriori (MAP) y, en consecuencia, el mínimo $U(\mathbf{m})$. Lo que define los valores esperados de $\mathbf{m}$.

Dado que la matriz de covarianza es una matriz simétrica definida positiva, es aceptable descomponerla con la factorización de Cholesky a $\Sigma = \mathbf{V}\mathbf{V}^T$, donde $\mathbf{V}$ es una matriz triangular inferior. Luego, la ecuación (3.14) puede ser escrita como:

$$U(\mathbf{m}) = \|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2, \quad (3.15)$$

donde,

$$\mathbf{d} = \begin{bmatrix} \mathbf{V}^{-1}\mathbf{b} \\ 0 \end{bmatrix} \quad (3.16)$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{V}^{-1}\mathbf{A} \\ \sqrt{\mu}\mathbf{S} \end{bmatrix} \quad (3.17)$$

el cual es un problema de mínimos cuadrados, en el que la incógnita son los coeficientes $\mathbf{m}$. Derivando (3.15) con respecto a $\mathbf{m}$ e igualando a cero se obtiene:

$$(\mathbf{G}^T \mathbf{G})\mathbf{m} = \mathbf{G}^T \mathbf{d}. \quad (3.18)$$

Para obtener los parámetros más probables, $\mathbf{m}^*$, dado el prior y los datos, se debe resolver la ecuación anterior 3.18. En cuanto a su covarianza, ésta está dada por

$$\Sigma_{\mathbf{m}^*} = \sigma^2 \mathbf{A}^T \Sigma^{-1} \mathbf{A} + \mu \mathbf{S}^T \mathbf{S} = \sigma^2 (\mathbf{G}^T \mathbf{G})^{-1}. \quad (3.19)$$

Finalmente, la ecuación de densidad de probabilidad de $\mathbf{m}$ o a posteriori queda de la forma:

$$p(\mathbf{m}|\mathbf{b}) = c(2\pi)^{-\frac{M}{2}} |\Sigma_{\mathbf{m}^*}|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{m} - \mathbf{m}^*)^T \Sigma_{\mathbf{m}^*}^{-1} (\mathbf{m} - \mathbf{m}^*) \right], \quad (3.20)$$

donde $\mathbf{m}$ es un conjunto de coeficientes multiplicadores de las fuentes unitarias con cierta probabilidad, $\mathbf{m}^*$ es el valor más probable y que, al ser $p(\mathbf{m}|\mathbf{b})$ una función gaussiana, también corresponde al promedio, c es la constante de valor $\frac{1}{p(\mathbf{b}|\sigma^2, \mu)}$, ya mencionada, que normaliza la función y $\Sigma_{\mathbf{m}^*}$ es la matriz de covarianza en la cual, por propiedad de las distribuciones normales multivariadas, cada elemento de la diagonal es la varianza de la distribución marginal $p(\mathbf{m}_i|\mathbf{b})$ de cada parámetro de media $\mathbf{m}_i^*$ donde i es la posición espacial del parámetro. Esto explicado en otras palabras, significa que para cada uno de los parámetros existe una distribución unidimensional gaussiana que las describe (distribución marginal), independiente de los demás parámetros, que está definida sin la necesidad de hacer un cálculo más complejo.

3.4.5 Cálculo de Hiperparámetros

Si bien se ha explicado en la sección anterior cómo obtener el modelo más probable, dentro de este análisis se tienen dos hiperparámetros desconocidos: σ_0 y μ . Estos hiperparámetros regulan la información a priori y la función de verosimilitud, siendo relevantes para encontrar el estimador $\mathbf{m}^*$ y su covarianza.

Para su cálculo se utiliza el método ABIC, propuesto por Akaike (1980), cuya ecuación representa la cantidad de información perdida como consecuencia de la regularización (distribución de pesos entre prior y verosimilitud) que, en este caso particular, ocurre en el proceso de inversión. La ecuación queda dada por

$$ABIC = -2\ln(L(\sigma_0, \mu)) + 2N_H \quad (3.21)$$

siendo N_H el número de hiperparámetros y $L(\sigma_0, \mu)$ la evidencia dados los hiperparámetros $(1/c)$, la cual se escribe también como

$$L(\sigma_0^2, \mu) = p(\mathbf{b}|\sigma_0^2, \mu) = \int p(\mathbf{b}|\sigma_0, \mathbf{m})p(\mathbf{m}|\sigma_0, \mu)d\mathbf{m}. \quad (3.22)$$

La cual corresponde a la integral sobre m , que son los coeficientes a obtener a través del ABIC, del numerador de la ecuación a posteriori, implicando que la probabilidad de ocurrencia de los datos b es una función de los hiperparámetros μ y σ_0 . Los hiperparámetros se eligen de forma tal que de la ecuación (3.21) se tenga el valor mínimo. Reemplazando las ecuaciones (3.7) y (3.3) en (3.22) y resolviendo se obtiene

$$ABIC = N\ln(2\pi\sigma_0^2) - 2\ln(\det(\mathbf{W})) - M\ln(\mu) + \ln(\det(\mathbf{G}^T\mathbf{G})) + \frac{U(\mu)_{min}}{\sigma_0^2} + 4, \quad (3.23)$$

donde N es el número de puntos de pronóstico. Para encontrar el mínimo en la ecuación se recurre a

$$\frac{\partial L(\sigma^2, \mu)}{\partial \sigma^2} = 0, \quad (3.24)$$

dando como resultado

$$\sigma^2 = \frac{U_{min}(\mu)}{N}. \quad (3.25)$$

Luego se sustituye (3.25) en la ecuación (3.21). Entonces, se puede reescribir ABIC en la forma

$$ABIC = N\ln\left(2\pi\frac{U(\mu)_{min}}{N}\right) - M\ln(\mu) - 2\ln(\det(\mathbf{W})) + \ln(\det(\mathbf{G}^T\mathbf{G})) + N + 4. \quad (3.26)$$

La búsqueda del valor de μ se realiza mediante iteraciones, específicamente se utiliza "Golden section search" [Press et al., 1992]. Este método consiste buscar el extremo de la función entre dos rangos que sucesivamente se hacen más pequeños, acortándose a razón del número aureo,

$$\varphi = \frac{1 + \sqrt{5}}{2}. \quad (3.27)$$

Finalmente, teniendo los coeficientes o parámetros $\mathbf{m}^*$ y sus covarianzas es posible obtener la superficie inicial más probable y sus rangos de valores al multiplicar cada fuente unitaria por su coeficiente.

La figura 3.11 [Mitsuhata et al., 2001] describe el proceso para encontrar el mínimo ABIC y por ende, los resultados de la inversión. El problema radica en la cantidad de datos (tiempo transcurrido desde el terremoto) necesaria para lograr una inversión que tenga una varianza reducida o, en otras palabras, que no exista una incertidumbre demasiado grande en la predicción. Es por esto que en la sección siguiente se analiza la calidad de la inversión al realizar predicciones para distintos mareógrafos con datos sintéticos y luego para un caso real.

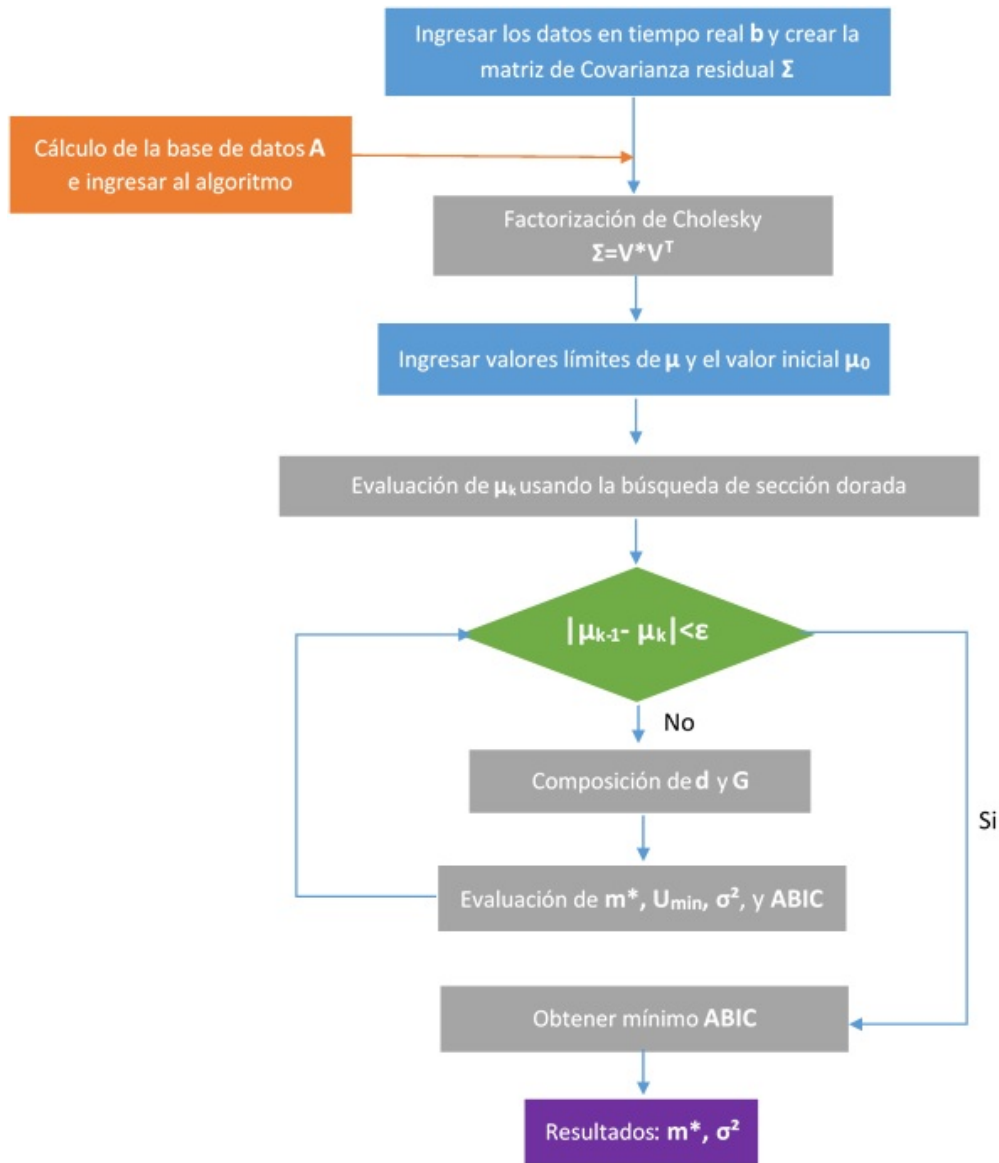


Fig. 3.11: Diagrama de flujo del proceso de inversión. En *azul* se destacan los valores a ingresar. En *naranja* la matriz con las series de tiempo precalculadas. En *gris* la aplicación del algoritmo explicado en las secciones 3.4.4 y 3.4.5. En *verde* se muestra el criterio para iterar. Finalmente en *morado* se obtienen los resultados. Fuente: Modificación de Mitsuata et al., 2001

Resultados

En esta sección se muestran los resultados de la inversión y se verifica la capacidad del método.

En primer lugar, se realiza un experimento numérico para verificar la capacidad predictiva del método según datos sintéticos y ruidos conocidos. En una segunda parte, se aplica el método al evento de Iquique 2014, arrojando resultados más probables de deformaciones y las incertidumbres asociadas, además de predicciones en los mareógrafos. Además, se compara con soluciones de dos autores diferentes y con métodos de inversión distintos.

El principal objetivo de estos análisis es poder validar el método y conocer sus alcances dependiendo de los tiempos de inversión. Esto con base en la capacidad predictiva tanto para el modelo más probable como para la incertidumbre asociada a los resultados.

4.1 Experimento numérico

Para validar el modelo se realiza un experimento numérico consistente en 12 fuentes unitarias y en datos sintéticos con error aleatorio conocido. Cada una de ellas tiene sus correspondientes series de tiempo en cada punto de pronóstico almacenadas en la base de datos **A**. Los puntos de pronóstico están ubicados en los mareógrafos de Arica, Iquique, Patache, Mejillones y la boya DART.

El experimento numérico se realiza de la siguiente forma. (1) Se crea una señal que es combinación lineal de las series de tiempo de la base de datos, creando una superficie suave. (2) Se añade un ruido de promedio nulo y varianza proporcional de las alturas absolutas de la señal y de un multiplicador. (3) Se invierte la señal dada una matriz de suavidad y se reconstruye la deformación inicial. El método se valida cuando se compara el valor de σ^2 (ver sección 3.25), al del multiplicador de la varianza del ruido conocido.

4.1.1 Condiciones del problema

Se consideran 12 de las fuentes unitarias creadas para la inversión del evento de 2014, las cuales se muestran en 4.1. A continuación, se diseña la base de datos de series de tiempo en los puntos de pronóstico, a partir de la propagación de estas 12 fuentes unitarias ($\mathbf{A}$). Para generar datos sintéticos que puedan ser generados por una combinación lineal de éstas, se multiplica la base de datos por coeficientes, $\mathbf{Ax}$. Estos coeficientes $\mathbf{x}$ se eligen de forma tal, que generen una condición suave, es decir, que la deformación del mar no tenga cambios abruptos de alturas en el espacio.

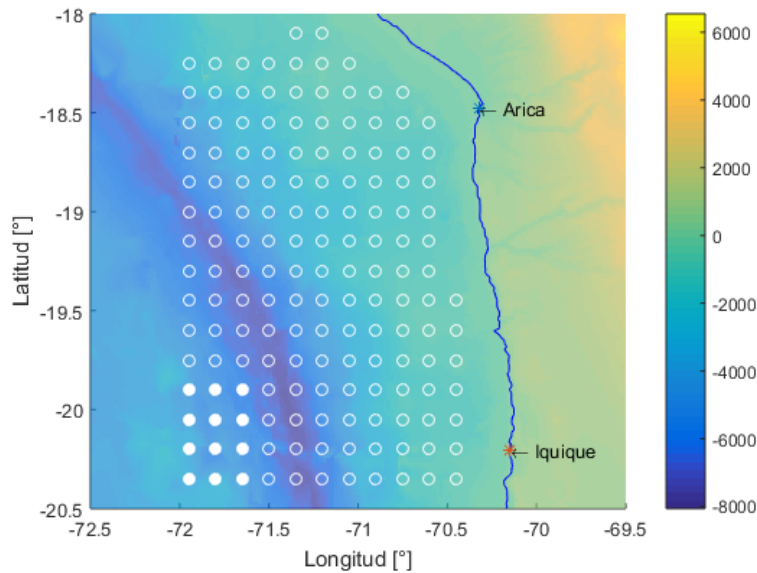


Fig. 4.1: Fuentes unitarias creadas para el evento de Iquique 2014. Los círculos achurados son las que se utilizan en este ejemplo. Los colores representan la batimetría y altimetría, siendo la línea azul la costa.

Luego se incorpora a la señal un ruido representando los errores aleatorios en las mediciones. Se asume que estos errores se distribuyen de forma normal, debido a que la muestra es suficientemente grande (cantidad de datos mayor a 30) como se explica en el capítulo 3.3.2. Este ruido es caracterizado por una media igual a cero y por una covarianza proporcional a las alturas promedio de la señal en cada punto de pronóstico, $\varepsilon \sim N(0, \sigma_1^2 \Sigma)$. A σ_1^2 se le asigna un valor de 10, 0.5 y 0.05 en la primera prueba.

Se crea la matriz de suavidad simétrica mostrada en la ecuación 3.5 y covarianza en la figura 3.9 de la sección anterior, donde β_i es el promedio de las amplitudes absolutas de

las olas ($|\eta|$) del modelo para cada punto de pronóstico i . Los puntos aislados en la matriz simbolizan valores nulos. En cuanto a las dimensiones de las matrices las fuentes unitarias, M , son en este caso 12, el tiempo, t , considerado para la inversión es de 60 minutos, de 15 segundos por dato, generando un total de 240 datos sintéticos y los puntos de pronóstico, N , en total 5.

4.1.2 Inversión

Con las 12 fuentes unitarias se tienen 12 parámetros a calcular, además de los hiperparámetros σ y μ con los que se controla la estructura del modelo paramétrico. Para ello se busca numéricamente μ que minimiza el ABIC definido en la ecuación 3.26. El valor de μ para un nivel de ruido de 10, da un valor de 7.6418, correspondiente a un ABIC de 8366. Con el valor óptimo de μ el mejor estimador de los parámetros $\mathbf{m}^*$ y el hiperparámetro σ se obtienen con las ecuaciones 3.18 y 3.25. En la figura 4.2 se puede ver que la deformación inicial es casi completamente reconstruida en el caso del mínimo ABIC.

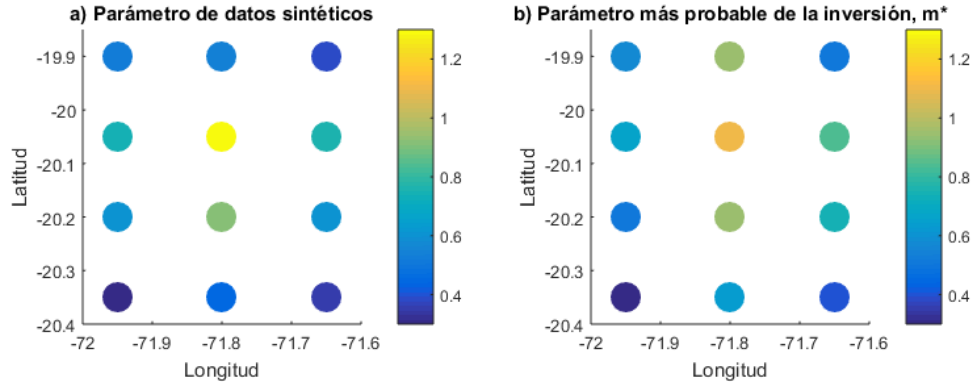


Fig. 4.2: Comparación de valores de los parámetros para $\sigma_1 = 10$. a) Valores de los parámetros de los datos sintéticos. b) Valores de los parámetros del modelo más probable, $\mathbf{m}^*$, obtenidos tras la inversión de 60 minutos.

Entonces, para verificar la consistencia del método, se compara el hiperparámetro σ (10.8729) con el valor del nivel de ruido de los datos sintéticos ($\sigma_1^2 = 10$). Estos valores son comparables a causa de la definición de σ^2

$$C_{od} = \sigma^2 \Sigma \quad (4.1)$$

donde C_{od} es la covarianza de los residuales y Σ es la covarianza estimada al comenzar la iteración. Por consiguiente, al acercarse la estimación de la inversión a los datos sintéticos sin el ruido, se espera que la covarianza sea muy similar a la del error aleatorio, impuesto en este caso. Para los niveles de ruido inferiores, ya sea con 0.5 o 0.05, el valor óptimo de coeficientes, es decir, la deformación más probable se comporta incluso mejor que lo mostrado en la figura 4.2, ya que el nivel de ruido es menor.

Cuando se contrastan los resultados de series de tiempo en los mareógrafos con los datos sintéticos sin ruido (ver figura 4.3) se puede percibir que el método logra obtener los datos con mucha precisión para σ_1^2 de 0.05. Sin embargo, al aumentar el nivel de ruido a 0.5

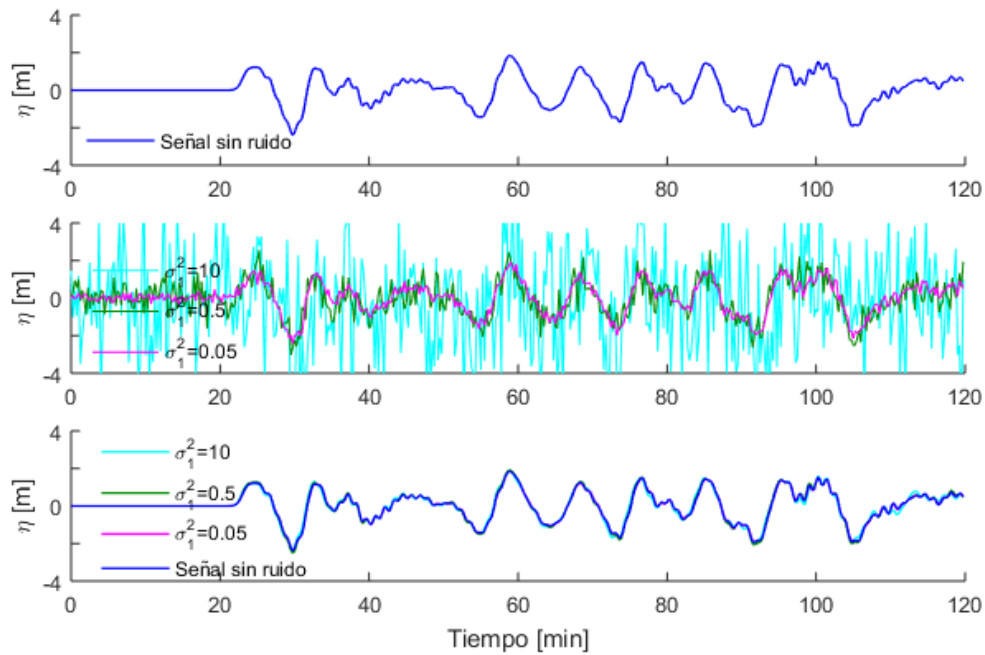


Fig. 4.3: Comparación de la señal invertida, usando 60 minutos de datos, con las originales con ruido y sin ruido en Iquique. *a)* Señal de datos sintéticos sin ruido. *b)* Señal de datos sintéticos con ruido; *cián* para $\sigma^2 = 10$; *verde* para $\sigma^2 = 0.5$; *magenta* para $\sigma^2 = 0.05$. *c)* Inversión para los tres niveles de ruido y señal sin ruido en color azul.

la incertidumbre aumenta, reflejándose en una línea más gruesa en el gráfico de inversión en 4.4.b). Esto es debido a que el objetivo de la incertidumbre es abarcar estos puntos donde los valores más probables de los resultados de la inversión no son coincidentes con los de los datos sintéticos, incorporándolos en la región y complementando los resultados. Se aprecia claramente cuando el nivel de ruido es mayor, con valor de 10 en la figura 4.4, donde en

negro se muestra el valor más probable de la solución, en azul el valor de los datos sintéticos sin ruido y la franja lila es la incertidumbre asociada a la inversión.

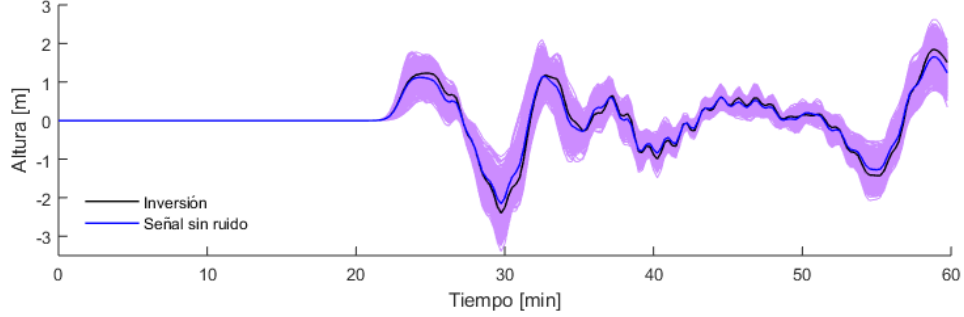


Fig. 4.4: Resultados de la inversión para nivel de ruido $\sigma^2 = 10$. En *negro* se muestra el valor medio de los parámetros obtenidos en la inversión, en *azul* la señal sin ruido y en *lila* la incertidumbre asociada.

Una forma de reconocer la diferencia entre las soluciones de la inversión y los datos sintéticos es el índice de skill. El índice de skill es también llamado la puntuación de Willmott [Willmott, 1981] y está basado en el promedio del error cuadrático entre lo observado y lo predicho por la inversión. La fórmula del skill se muestra a continuación,

$$sk = 1 - \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2}{\frac{1}{n} \sum_{i=1}^n O_i^2}} \quad (4.2)$$

donde n es la cantidad de datos (cada 30 minutos), P_i es la serie predicha en cierto tiempo y O_i la serie observada. Este índice toma valores cercanos a 1, cuando la diferencia entre lo observado y lo predicho es cercano a cero. Al acercarse a 0 muestra que la diferencia aumenta, comparándose con el valor del denominador. Finalmente, cuando toma valores negativos, da a conocer que el término de la derecha está por sobre la unidad y por tanto el error cuadrático tiene un mayor valor, lo que sugiere un muy mal ajuste.

Se muestran a continuación (Figura 4.5) las gráficas de skill para la ciudad de Iquique de $\sigma^2 = 10, 0.5$ y 0.05 . En ellas se toman muestras de 15 minutos ($n = 60$ datos, de intervalos de 15 [s]), partiendo del tiempo 1 (s_1) hasta los últimos 15 minutos (s_180) evaluando un periodo total de 60 minutos, correspondiente al eje de las abscisas. Además, esto se reproduce con diferentes soluciones obtenidas a partir de la función de densidad de probabilidad a posteriori

por lo que para cada s_i existen diversos valores de skill o de desempeño según los valores sintéticos (líneas verticales).

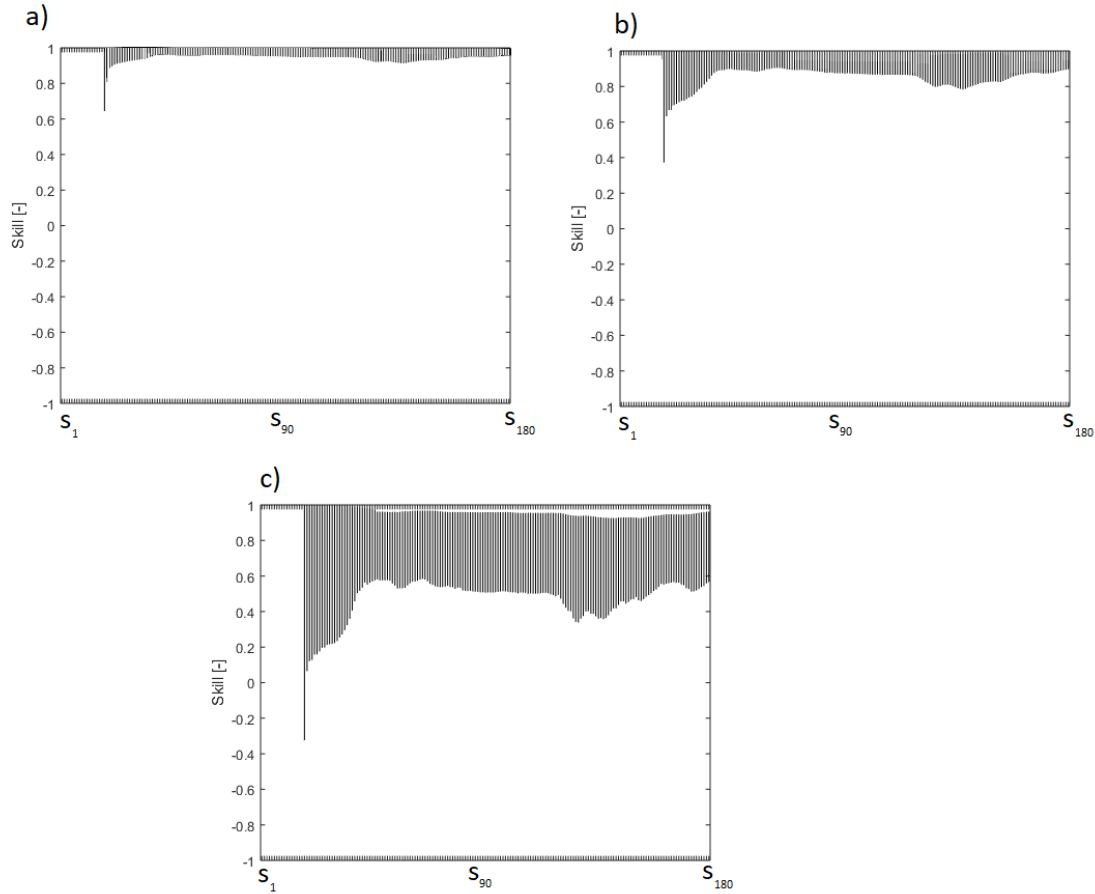


Fig. 4.5: Rangos de valores de skill para periodos de 15 minutos en Iquique. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos o líneas verticales corresponden a diferentes escenarios estocásticos de los resultados de inversión para 60 minutos de datos. *a)* Gráfico para $\sigma_1 = 0.05$. *b)* Gráfico para $\sigma_1 = 0.5$. *c)* Gráfico para $\sigma_1 = 10$.

Se presenta en la tabla 4.1 los valores del hiperparámetro σ esperado y obtenido y los errores de estimación entre ellos. Es de esperar que a mayor nivel de ruido el error porcentual aumente, ya que la señal está más contaminada. Además se destaca que los errores son menores al 10% en todos los casos.

Tab. 4.1: Resumen de valores obtenidos en el experimento numérico.

σ del ruido	Hiperparámetro σ	Error [%]
0.05	0.0513	2.69
0.5	0.5354	7.07
10	10.8729	8.73

Con este ejemplo se da por validada la aplicación del método, sin embargo se debe destacar que este experimento está limitado a ciertas condiciones. La matriz utilizada en la covarianza es conocida, y obtenida a partir de los datos sintéticos para tener un σ comparable. Sin embargo, en casos reales los ruidos no siempre tendrán un nivel constante de varianza en el tiempo y por ende la matriz en la mayoría de los casos no es diagonal (Ver sección 5.1.2). Además en situaciones reales no es conocida la cantidad de fuentes que deberán ser utilizadas para la inversión por efectos de variaciones en ciertos lugares de la superficie del mar. Por lo tanto, se requerirá utilizar una mayor cantidad de éstas, lo que puede originar errores mayores.

4.2 Aplicación al tsunami de Iquique 2014

Como ejemplo de aplicación de este método se considera el tsunami de Iquique (Mw=8.2), el cual, a pesar de no ser tan destructivo como eventos pasados, se tiene información suficiente para poder realizar la inversión. El epicentro del terremoto está ubicado en 19.642°S y 70.817°W [Barrientos, 2014] entre las costas de Iquique y Pisagua y ocurrió el 1 de Abril de 2014. En esta sección se analizan los resultados comparándolos con soluciones de otras inversiones [An et al., 2014; Hayes et al., 2014]. El primero, [An et al., 2014], utiliza datos de elevación del agua captadas por tres boyas DART (32401, 32402 y 32412) y utiliza un método de inversión para estimar la distribución de los deslizamientos en el plano de falla. Luego calcula el desplazamiento del fondo marino y la correspondiente elevación de la superficie del agua con el modelo de Okada [Okada, 1985]. El segundo [Hayes et al., 2014] utiliza datos sísmicos para la inversión, llamadas inversiones rápidas de falla finita, y obtención de la distribución de deslizamientos de la ruptura del terremoto. Estos deslizamientos son procesados también con el modelo de Okada para conseguir la deformación inicial del mar.

Cabe mencionar que los resultados a mostrar en esta sección están influenciados por la hipótesis aplicada en esta tesis de linealidad, la cual se cumple para mar abierto a grandes

profundidades, pero no cercana a la costa. La discusión sobre esta hipótesis se comenta en el siguiente capítulo.

4.2.1 Resultados de la inversión

La inmediata verificación del método se realiza comparando los datos, b , con los obtenidos por la inversión de 60 minutos totales (desde el inicio del terremoto), Am^* , mostrados en la figura 4.6(a).

En ella se puede notar que los datos medidos y los de inversión son muy similares en la forma con picos coincidentes entre ambos, dando valores máximos en Iquique de 1.9 [m] y 1.8 [m] y mínimos de 2.9 [m] y 2.6 [m], para la inversión y los datos reales respectivamente. Estas diferencias pueden ser causadas por los pesos relativos dados por el algoritmo al prior o condición de suavidad, ya que esta condición promedia los coeficientes, reduciendo la energía y aminorando los valores mínimos y máximos y al traslape de las fuentes unitarias, ya que también generan condiciones de suavidad.

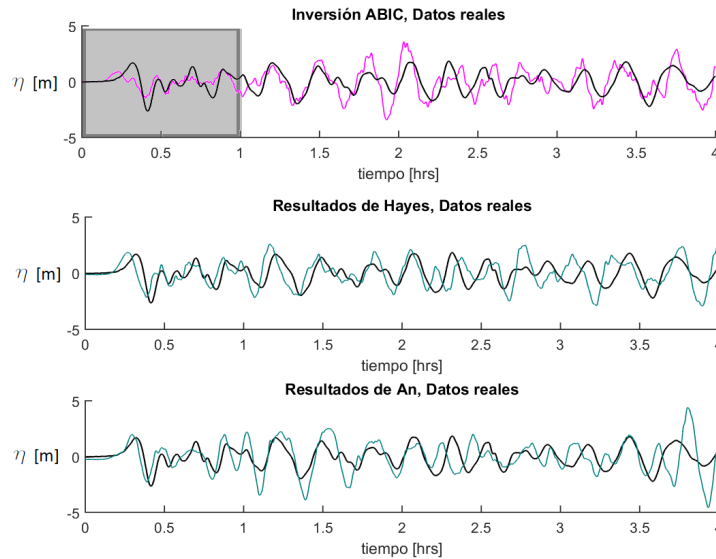


Fig. 4.6: Comparación entre resultados de inversión (línea magenta) de 1 hora de datos (área gris) y datos medidos para Iquique (línea negra). a) Inversión con método ABIC (verde). b) Resultados de Hayes et al., 2014. c) Resultados de An et al., 2014 (verde).

En otras palabras, los datos son relevantes, pero también lo es la suavidad y dependerá de la calidad de los datos, es decir, que los datos estén bien medidos y no contengan errores

aleatorios, cuál es el más importante. Una de las formas de controlar esto es la correcta calibración de los mareógrafos, una ubicación que evite efectos locales por la batimetría compleja y mantención periódica de estos. Además, se puede percibir un buen ajuste para las primeras horas, sin embargo, entre 2.5 y 4 horas las curvas se alejan de los datos reales. La razón principal puede deberse a situaciones de no linealidad o falta de datos, lo cual se comentará en el siguiente capítulo.

Las soluciones de otros autores [An et al., 2014; Hayes et al., 2014] se muestran en la figura 4.8 las cuales fueron propagadas con la misma resolución que los demás análisis realizados en este trabajo, haciendo posible una comparación. La propagación de la solución de Hayes et al., (2014) se aproxima bastante en las primeras 2 horas a los datos reales, donde los máximos y mínimos son muy similares en cuanto a tiempos de ocurrencia y mediciones de alturas. Para la solución de An et al., (2014) se puede notar una gran similitud en la primera hora, sin embargo, los máximos y mínimos tienen valores absolutos un tanto mayores que los reales.

Los valores de m^* y su ubicación espacial en las fuentes unitarias correspondientes se muestran en la figura 4.7. La imagen superior representa los valores de los parámetros para la deformación inicial de superficie más probable, donde la máxima es de aproximadamente un metro. La imagen (b), que muestra la desviación estándar, permite reconocer que la mayor incertidumbre se tiene en la parte nor-oeste del gráfico.

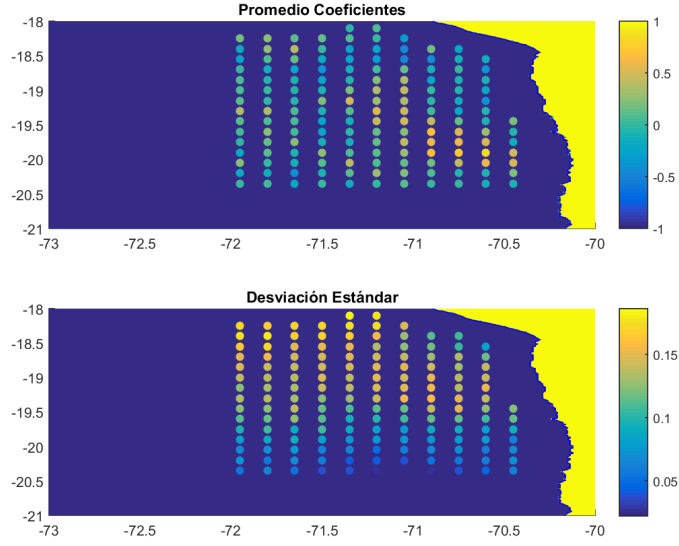


Fig. 4.7: Resultados de la inversión con 1 hora de datos (a) Valor de los parámetros (m^*) en la posición de las fuentes unitarias (b) Desviación estándar asociada a los marginales de m^* .

Al multiplicar los m^* , ya mencionados, por las fuentes unitarias gaussianas se tiene la figura 4.8 (a). Comparando con la deformación inicial encontrada por Hayes et al., 2014 es posible notar que las alturas máximas alcanzadas y su ubicación espacial son similares, con valores de 1.12 y 1.26 metros para la solución con ABIC y la de Hayes et al., 2014 respectivamente. En cambio, la de An et al., 2014, tiene valores máximos de 1.4 metros de altura. Esto puede ser explicado por el área abarcada de deformación, siendo mayor en la solución con ABIC y menor en (b) y (c), lo que fuerza la concentración de deformación.

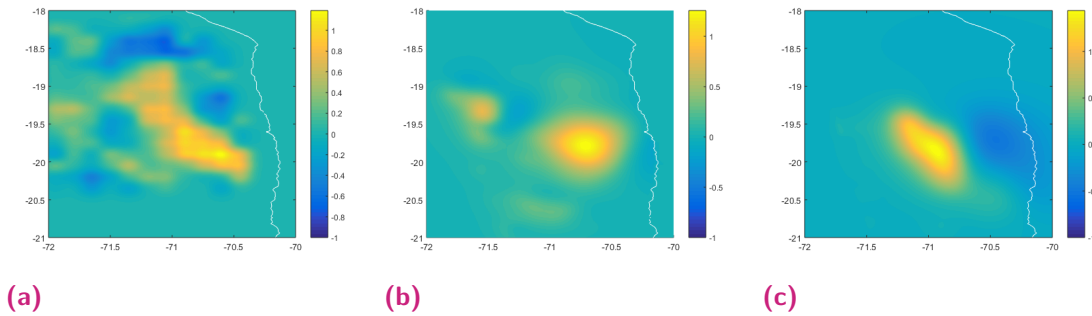


Fig. 4.8: Soluciones de deformación inicial del nivel del mar producto del tsunami 2014 propuestas por: (a) Este trabajo (b) Hayes et al., 2014 (c) An et al., 2014

Las diferencias entre estas soluciones pueden ser ocasionadas a las formas de obtención de las deformaciones, telesísmica para el caso de Hayes et al., 2014, inversión de tsunami para el desarrollo de este trabajo y una combinación de ambas en lo desarrollado por An et al., 2014. Otro punto importante es destacar que, si bien la solución mostrada en 4.8 (a) es obtenida con los parámetros m más probables, existe más de una deformación inicial que se ajusta a los datos y a la suavidad, sin embargo no varía en gran medida, ya que las desviaciones estándar de cada parámetro (Ver figura 4.7) son bajas, no alcanzando a ser mayores a 0.18.

Finalmente en fig. 4.9 se muestra un resumen de los resultados para la inversión con 1 hora de datos. Las curvas lilas en los gráficos son 1000 muestras aleatorias de la función de densidad de probabilidad a posteriori, es decir, soluciones de la inversión obtenidas gracias al muestreo con un software matemático. Para Iquique, que es el lugar más cercano al epicentro, se puede notar que la curva de datos reales (azul) se encuentra dentro de la región dada por la incertidumbre (lila), modelando de buena manera el comportamiento de la variación de altura de oleaje. Sin embargo, en ciertos puntos la incertidumbre tiene rangos muy amplios, donde las máximas superficies libres son metros más grandes que las reales, lo que puede generar una alerta para la población demasiado conservadora en etapas posteriores (al propagar estas soluciones en la costa). Cabe mencionar que para este mareógrafo el tiempo de arribo es corto, de alrededor de 10 minutos, por lo que los datos de éste que contienen las variaciones producidas por el terremoto utilizados en la inversión son mayores que de otros mareógrafos, como el de Arica y Patache, en que el tiempo de llegada es cercano a 20 minutos. Es por ello que para Iquique las alturas invertidas son mucho más cercanas a los datos que para los otros mareógrafos.

En los mareógrafos se marcan claramente los datos que se han invertido, ya que hasta los 60 minutos la inversión es casi igual a los datos, lo que se refleja también en la incertidumbre, la cual es bastante baja. A partir de los 60 minutos los rangos de incertidumbre aumentan y la modelación, aunque similar, no es tan precisa como para los primeros minutos o como en Iquique. Para Mejillones, la boya más distante al epicentro, los resultados de la inversión hasta los 150 minutos son muy precisos. Luego la incertidumbre aumenta, y algunos escenarios estocásticos dan valores muy altos en comparación al valor real. Por último, la boya DART 32401 en el tiempo de arribo, en la amplitud y la duración de sus máximos y mínimos son reproducidas casi por completo, lo cual es de esperar gracias a la

ubicación fuera de la costa, que evita efectos propios de la costa y la no linealidad. Para ver los resultados de otros tiempos de inversión, revisar el Anexo A.

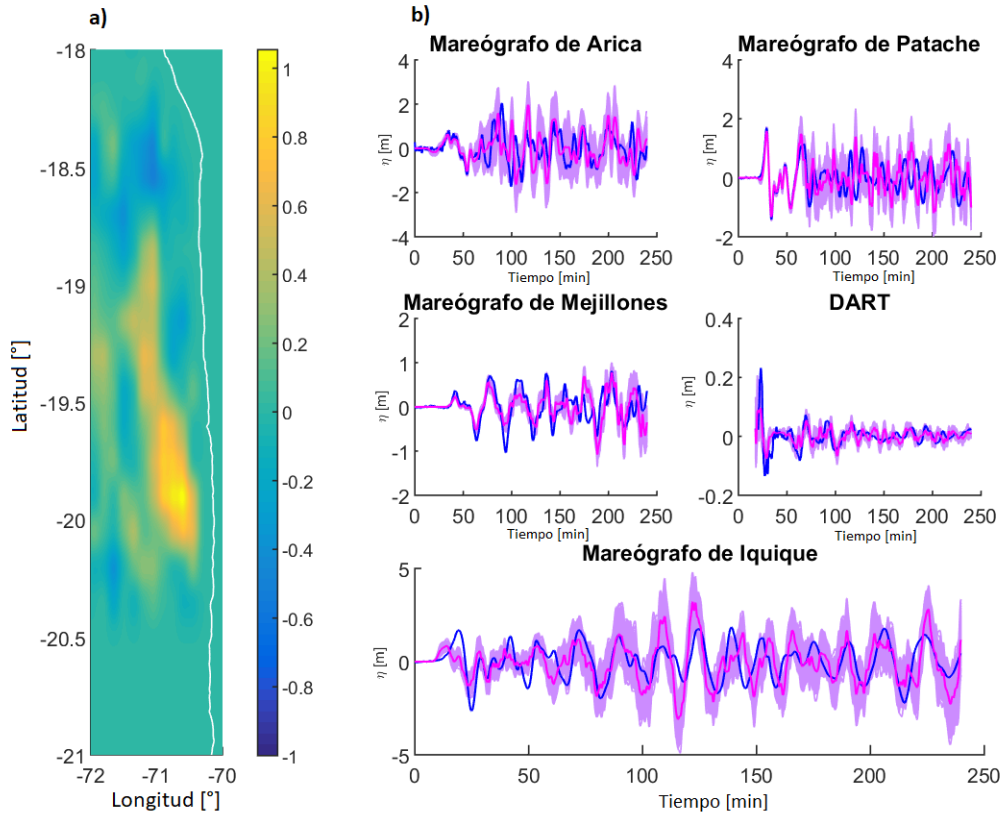


Fig. 4.9: Resultados de la inversión. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

Discusión de los resultados

En la sección anterior se mostraron resultados del experimento numérico para validación y de la inversión del tsunami de Iquique 2014. A partir de éstos se realiza un análisis crítico, revisando las diferencias en los datos y los resultados, comparando los valores en los distintos tiempos de inversión y examinando con mayor detalle las hipótesis planteadas en este método.

5.1 Análisis de las hipótesis

De las hipótesis realizadas para la implementación del método se discutirán dos esenciales, las cuales pueden cambiar drásticamente los resultados. La primera hipótesis es sobre la minimización del ABIC y los hiperparámetros obtenidos. Pese a que la derivación de la función ABIC $L(\sigma, \mu)$ sobre σ o μ da un mínimo, ésta puede no ser tan sensible a la elección de ese hiperparámetro y por lo tanto se puede obtener un rango de valores los cuales amplían el valor de la incertidumbre. La segunda hipótesis trata acerca de la covarianza residual de los datos, la cual dependiendo de los valores que adopte, influirán en las estimaciones de incertidumbre como muestra la ecuación 3.19.

5.1.1 Elección de los hiperparámetros

El método desarrollado se fundamenta en la base de la máxima entropía y por ende encontrar los hiperparámetros que minimicen ABIC (Ecuación 3.21). A pesar de ello, el ABIC no siempre es sensible a estos hiperparámetros y por tanto pueden existir diversos valores de ellos que resulten en valores de ABIC similares. Es por esto que a continuación se muestra el experimento numérico realizado en la sección anterior que ejemplifica la sensibilidad de los hiperparámetros y sus resultados.

Para comenzar se muestran en la figura 5.1 las variaciones del ABIC según cada hiperparámetro. Es posible observar que si μ varía significativamente, el valor del ABIC

se mantiene casi constante. En cambio si σ^2 varía, el valor de ABIC cambia de forma considerable.

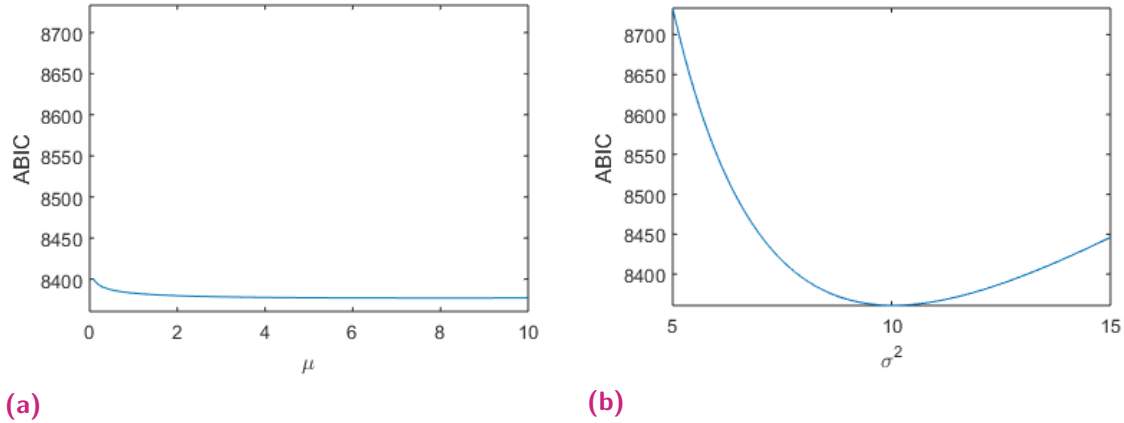


Fig. 5.1: Variación de ABIC respecto de hiperparámetros para el experimento sintético: (a) Variación de μ para σ^2 óptimo obtenido del mínimo ABIC. (b) Variación de σ^2 para μ óptimo obtenido del mínimo ABIC.

De estos gráficos (figura 5.1) se escogen valores a la izquierda o a la derecha del mínimo ABIC, ($\mu = 7.6$), del hiperparámetro μ obteniendo la figura 5.2 donde se aprecian los valores de los parámetros de los datos sintéticos, los obtenidos de la inversión a partir del más probable y las deformaciones para elecciones inapropiadas del valor. En la figura 5.2, en la parte c), se muestra el caso cuando el valor de μ está por debajo del valor óptimo. A menor valor de μ , más débil es la suavidad dada por el prior. Los valores de los parámetros de la inversión se vuelven más irregulares entre sí, con valores más altos que para el de la inversión 5.2 b) y éstos similares a los datos sintéticos, pero también con valores más bajos que en a) o b) en los extremos. Por el contrario, cuando el valor de μ es mucho mayor al valor óptimo, la inversión de la deformación es más suave, como se ve en la figura 5.2 d). En este caso el pico de 1.3 [m] de superficie libre del mar baja considerablemente debido a que el prior es muy importante homogeneizando la superficie. Sin embargo, cuando los valores de los ruidos son muy grandes, en otras palabras, los datos no se ajustan al modelo (por ejemplo, cuando hay fenómenos no lineales), se obtienen resultados de deformaciones con demasiada suavidad, aunque el ABIC sea el menor. El problema radica en saber a priori si se tienen ruidos grandes con datos reales.

No obstante, si en la elección de hiperparámetros se observa la figura 5.1 b), se puede concluir que el ABIC es muy sensible a σ^2 , en este caso, lo que repercute directamente en

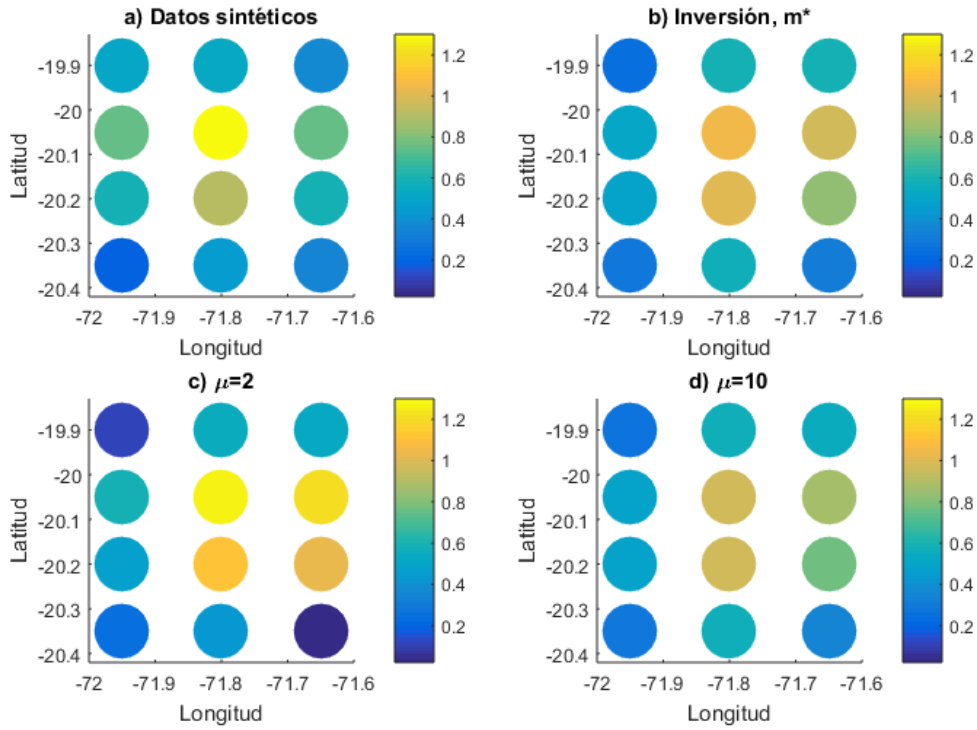


Fig. 5.2: Valores de los parámetros ubicados en la superficie del mar del experimento numérico. *a)* Parámetros de datos sintéticos propuesta. *b)* Parámetros de la inversión para el valor óptimo de μ y más probables, m^* . *c)* Parámetros para μ menor al valor óptimo ($\mu = 2$). *d)* Parámetros para μ mayor al valor óptimo ($\mu = 10$).

los rangos de incertidumbre asociados a los resultados, mostrados en los gráficos a) y b) de la figura 5.3 con σ^2 de 5 y 15 respectivamente obtenidos de escenarios estocásticos de la distribución a posteriori manteniendo μ constante e igual al óptimo de la inversión.

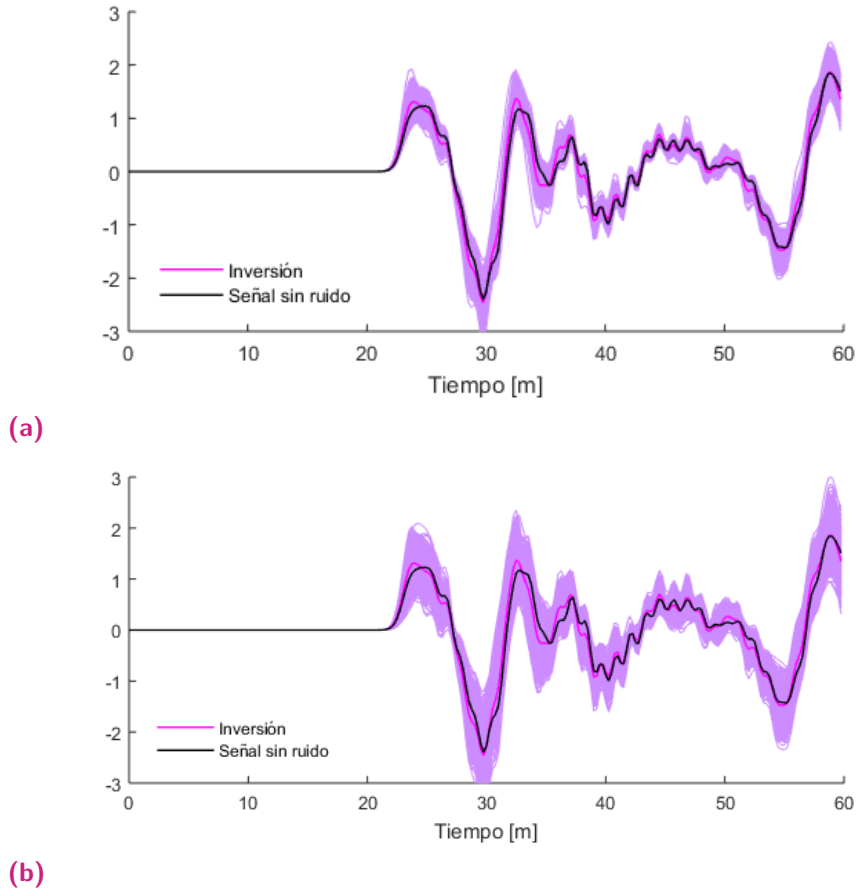


Fig. 5.3: *a)* Resultado de la inversión para Iquique con $\sigma^2 = 5$. *b)* Resultado de la inversión para Iquique con $\sigma^2 = 15$.

Finalmente se observa que la sensibilidad de los hiperparámetros es relevante a la hora de implementar el método y predecir, dado el cambio de incertidumbre con una elección u otra o de los valores medios de la probabilidad a posteriori de los parámetros. Su relevancia radica en la diferencia en los resultados con μ que dan valores similares de ABIC, dándoles mayor o menor suavidad a los resultados, influyendo en los máximos y mínimos obtenidos y que no obstante dan valores similares de ABIC. Sin embargo, no es analizado con más detalles en el presente trabajo, dando prioridad a la segunda hipótesis que se muestra a continuación.

5.1.2 Covarianza relativa de los datos

La covarianza relativa de los datos es una medida del error o ruido presente en los datos. El caso ideal, para el que tenemos soluciones analíticas, es ingresar una matriz en el algoritmo que sea proporcional a la covarianza del ruido de las señales, que al multiplicarla por el hiperparámetro σ^2 pueda reconstruir la matriz de covarianza real. A pesar de ello, al ser una información requerida antes de conocer el modelo, es necesario realizar aproximaciones e hipótesis para generarla.

Una opción es utilizar una matriz identidad, $\Sigma = \mathbf{I}$, lo cual omite cualquier relación fuera de la diagonal de la matriz [Duputel et al., 2012], pese a que es difícil pensar en problemas reales que posean solo ruido sin correlaciones entre los datos [Langbein and Johnson 1997]. Es más, usualmente los errores tienen ciertos periodos característicos de decaimiento o escalas de tiempo. Entonces, se deben considerar componentes fuera de la diagonal en la matriz covarianza si se muestrean los datos con un intervalo más corto que el largo característico y/o la escala de tiempo del ruido. Por ello algunos autores [e. g Duputel et al., 2012] han propuesto matrices de covarianzas de decaimiento exponencial en el tiempo como la utilizada anteriormente [ecuación 3.10]. Otro tipo de matrices correlacionadas son las de inversiones con datos InSAR (Synthetic Aperture Radar Interferometry), que tienen ruidos principalmente originados por la variación del vapor de agua en la atmósfera, por lo que las correlaciones entre los errores observados son principalmente de origen espacial [Fukahata et al., 2003]. En el caso de los mareógrafos, el ruido no es constante temporalmente, lo que puede ser descrito por matrices que cambien su diagonal en el tiempo, llamadas matrices no estacionarias. Una forma de evitar la incorporación de estas matrices complejas es captar la información sólo de los primeros ciclos de inversión, las cuales estarán afectas a menos cambios temporales de ruidos y efectos no lineales (que se comentan en la sección 5.1.3) utilizadas por autores como Satake y Kenji (1989).

Esto muestra que para los distintos sistemas de medición es necesario contar con matrices de covarianzas específicas, que puedan describir de mejor manera los procesos físicos que ocurren, que en el caso de los mareógrafos utilizados en esta tesis es un tema de discusión y análisis para estudios siguientes.

Para analizar el comportamiento de la inversión se utilizan diferentes matrices de covarianza para el tsunami de Iquique 2014, con el objeto de evaluar si mejora la predicción. El primer caso, es el que se ha utilizado hasta ahora en este trabajo, consistente en una

matriz que recoge la varianza de los datos en cada mareógrafo entre el inicio del terremoto y comenzado el tsunami. Esta matriz se conforma de bloques, incorporando la expresión 3.10 donde γ_k es la varianza antes mencionada y t_0 que toma un valor de 10 [min], como aproximación del periodo de las ondas. En un caso real, este valor debe ser estudiado, ya que en ciertos tiempos no es posible conocer aún los periodos de las ondas y es necesario un análisis de sensibilidad en distintos eventos. El segundo caso incorpora una actualización de la matriz sin correlaciones, la cual utiliza como primera matriz la misma del caso 1 (inversión de 30 minutos) y luego para las inversiones posteriores (45 min, 60 min, etc.) la varianza de la diferencia entre los datos recogidos y lo modelado anteriormente, es decir, una matriz que depende de la inversión inmediatamente anterior. Expresado matemáticamente sería

$$\mathbf{r} = \mathbf{A}(\mathbf{m}^*)_{30min} - \mathbf{b} \quad (5.1)$$

y reemplazando en la ecuación 3.9,

$$(\gamma_k)^2 = \frac{\sum_{u=1}^{t_k^a} [(r_k)_u - \bar{r}_k]^2}{t_k^a} \quad (5.2)$$

en la cual $(r_k)_u$ representa al dato u del subconjunto de datos de $\mathbf{r}$ de cantidad t_k^a , que depende de los datos a invertir, del mareógrafo k . El último caso incorpora las correlaciones temporales del caso 1, en otras palabras, aplicando la ecuación 3.10, y además actualizándose en cada inversión utilizando el método recursivo del caso 2.

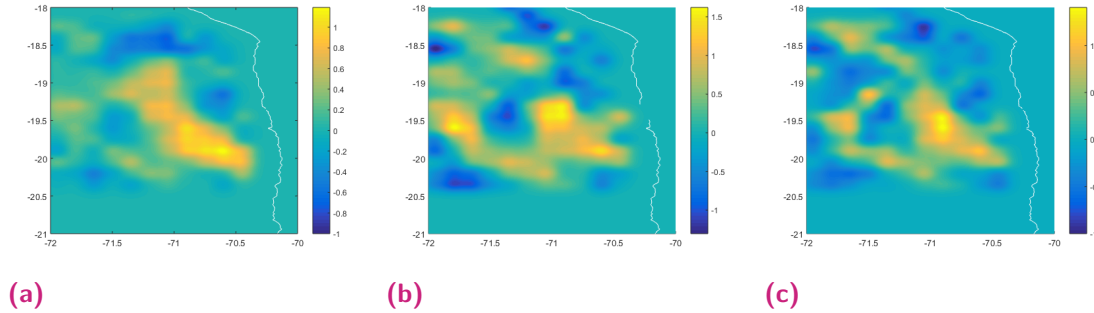


Fig. 5.4: Resultados de inversión con una hora de datos para distintos valores de matriz de covarianza. *a)* Caso 1. Matriz no actualizada en el tiempo y correlacionada. *b)* Caso 2. Matriz que se actualiza en el tiempo y diagonal. *c)* Caso 3. Matriz que se actualiza en el tiempo y correlacionada.

Al observar los resultados de las deformaciones iniciales 5.4 en el caso 1 está mucho más concentrada y con valores mayoritariamente positivos o nulos. Esta deformación es muy

similar al resultado encontrado por An et al., 2014 en ubicación y área abarcada (Ver fig 4.8). En el caso b) por el contrario, se tiene una deformación poco homogénea, con niveles de descenso cercanos a un metro bajo el nivel del mar y deformaciones positivas de 1.4 metros (cercano a lo encontrado por Hayes et al., 2014). Para el caso c) la solución nuevamente es más concentrada, con valores máximos de 1.2 y mínimos de -0.5 metros y su ubicación se asimila a la de An et al., 2014. Sin embargo, se puede apreciar un valor de deformación en el noroeste de la gráfica, el que se aparta de la deformación principal. De los tres casos, los que se asemejan más a soluciones de otros autores son los casos 1 y 3, de los cuales el que pierde menor energía es el caso 3, con picos más altos de superficie libre.

Valores para el mareógrafo de Iquique de cada caso se muestran en la figura 5.5. En el caso a) en que la covarianza se obtiene del ruido anterior al tsunami se tienen curvas muy cercanas a los datos reales e incertidumbres adecuadas para el nivel diferencia entre ambas curvas. Es probable que el nivel de ruido no haya cambiado demasiado desde el inicio del tsunami y por tanto es una buena modelación para este periodo de datos. Por el contrario, en el caso b), se observa poca precisión entre el modelo más probable (magenta) y los datos (azul). Esto puede deberse en parte a un sobremuestreo (cada 15 segundos) lo cual no es amortiguado por la matriz de covarianza correlacionada como el caso anterior. Para ello son necesarios estudios con un muestreo con intervalo distinto de la variación de la elevación de la superficie y analizar cómo se comporta la matriz. En caso de mejorar, se deben hacer cambios en la matriz de covarianza para poder abarcar el ruido producido por el sobremuestreo, utilizando una correlación distinta, y así utilizar la cantidad de datos disponibles. En el último caso c) el desempeño de la predicción se asemeja al del caso 1 con matriz de covarianza estática, sin embargo, se aprecia una mejora respecto al caso 1 alrededor de los 50 minutos y de los 75 minutos. Al realizar la inversión en los 60 minutos se puede notar que el nivel de ruido hasta este tiempo no ha sufrido grandes variaciones respecto al inicio del terremoto, por lo que la matriz de covarianza residual aún puede generar una solución semejante a la real.

La matriz de covarianza, por lo visto en esta sección, resulta de gran relevancia en la inversión. Lo expuesto en esta tesis tiene ciertas limitaciones, ya que no se han utilizado matrices no estacionarias, las cuales varían la diagonal respecto al tiempo (γ_k) y además los datos utilizados para la primera inversión han sido datos desde el inicio del terremoto, hasta la llegada de la primera ola, por lo cual puede contener ruidos producidos por el terremoto, los que después de ocurrido desaparecen. De estos resultados se puede concluir

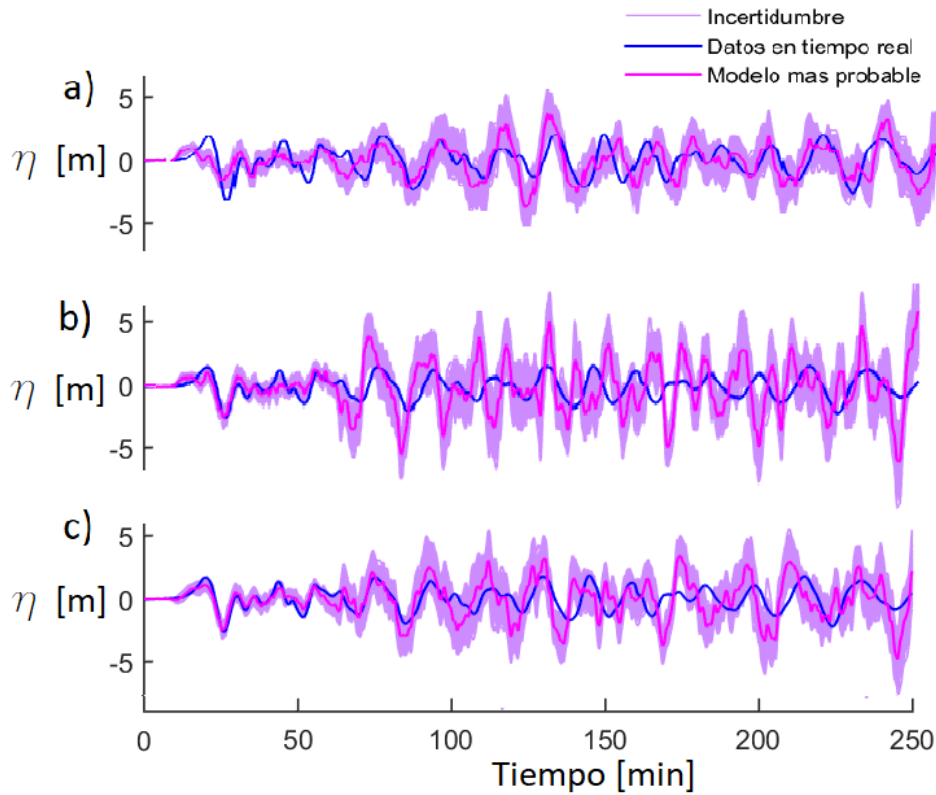


Fig. 5.5: Predicción para el mareógrafo de Iquique. *a)* Caso 1. Matriz no actualizada en el tiempo y correlacionada. *b)* Caso 2. Matriz que se actualiza en el tiempo y diagonal. *c)* Caso 3. Matriz que se actualiza en el tiempo y correlacionada.

que la covarianza es un elemento clave en el desarrollo del método y por tanto es importante poder medir el desempeño de cada uno de los casos.

Comparación de alternativas

Hasta el momento se han mostrado resultados sólo con dos intervalos de datos (30 y 60 min.), sin embargo la inversión debe hacerse periódicamente y conocer su comportamiento a medida que se tienen más datos. La lógica indica que a mayor cantidad de datos, mejor es el desempeño de la inversión y menor es la incertidumbre asociada. La realidad, en cambio, no es exactamente así ya que existen factores locales en los mareógrafos, sobremuestreo o cambios en el origen del ruido a medida que pasa el tiempo.

Es por ello que la medición de la calidad de cada una de las alternativas de generación de matrices de covarianza no es trivial y se recurre al índice de skill. Para distintos tiempos de inversión (30, 45, 60, ..., 240 minutos) se obtienen skills en cada uno de los mareógrafos, utilizando para ello 15 minutos de datos muestreados cada 15 segundos, es decir, $n = 60$ en todos los tiempos de inversión (Ecuación 4.2). Resultados para el mareógrafo de Arica se muestran en la figura 5.6. Ya que los demás mapas muestran cambios similares no se analizan en detalle en esta sección, pero se muestran adjuntos en el Anexo B.

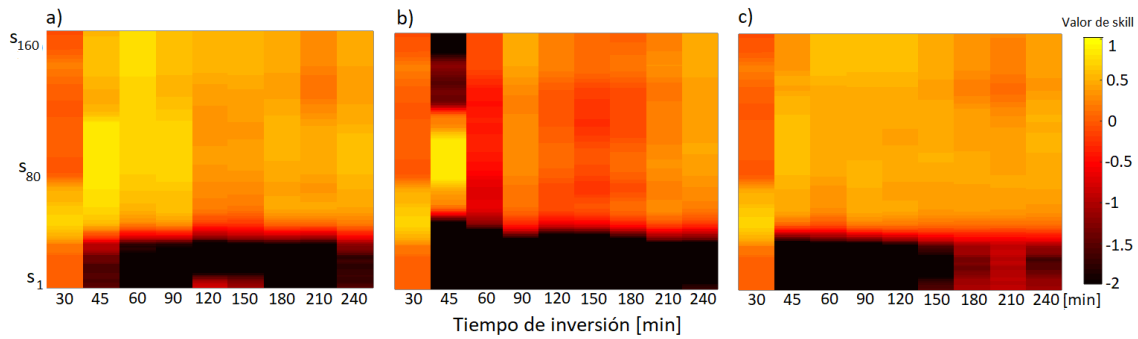


Fig. 5.6: .

]Mapa de calor de valores de Skill en el mareógrafo de Arica para cada tiempo de inversión en [min]. Los valores s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i

El mapa de calor de la figura 5.6 a), con matriz correlacionada muestra valores de skill similares para los primeros 30 minutos. Luego, para las siguientes inversiones, los valores de skill para los primeros minutos son negativos, pero hay un cambio drástico alrededor de los 45 minutos en que la predicción mejora y por lo tanto los valores de skill se acercan a 1. Pasados los 90 minutos de predicción los valores de skill son ligeramente menores. Esto muestra que hasta ese minuto, a pesar de tener una mayor cantidad de datos reales, la predicción no mejora en este caso.

Para la figura b), caso 2, a los 30 minutos es idéntico al caso anterior, pero una vez que se genera la covarianza actualizada a los 45 minutos se tiene una predicción con índices de skill menores que en el caso anterior. En el último caso de análisis, figura c), los índices de skill son similares al caso a), con una leve variación a los 45 minutos, con menores skills, y a los 210 minutos, con mayores skills.

Si bien los resultados analizados sólo corresponden al modelo más probable, es decir no hay información acerca de su incertidumbre, se pueden sacar algunas conclusiones respecto a

su comportamiento. Se destaca en los gráficos que el caso 2 predice de forma menos exacta que 1 y 3, lo que muestra la importancia de las correlaciones en la matriz de covarianza. También se puede inferir que la matriz estacionaria funciona mejor para los primeros ciclos de inversión que para las inversiones posteriores, probablemente porque el ruido va cambiando y por lo tanto la varianza también debe cambiar, además de otros factores como el de la sección 5.1.3. Es por ello que la matriz no estacionaria mejora la predicción para inversiones con más datos, aunque de forma leve.

Complementando el análisis, en la figura 5.7 se muestran los rangos de valores de skill para cada uno de los intervalos de tiempo de 15 minutos de los escenarios estocásticos. En a) (caso 1) se tienen valores para los primeros 40 periodos entre -4 y -1, los cuales como se vió en la gráfica anterior muestran que para los primeros minutos la capacidad de predicción no es tan precisa. Luego, para los siguientes minutos la predicción se acerca a los datos reales, con valores de skill cercanos a 0.7 y rangos de variaciones menores. Para el caso b) se tiene un rango de valores mucho más amplio para los primeros minutos, entre -6 y -12. Al igual que en a) los valores de skill mejoran luego del periodo 40, pero aún menores que en el caso 1. Finalmente el caso 3, aunque con rangos menores que el caso 2, son mayores que en a). Esto muestra que el caso 1 es el que mejor predice a los 60 minutos para el mareógrafo de Arica. No obstante, esto varía dependiendo de la ubicación del mareógrafo y de los datos invertidos, con una ligera mejoría en el caso 3. Para más información sobre gráficos de otros tiempos de inversión y distintos puntos de pronóstico ver Anexo C.

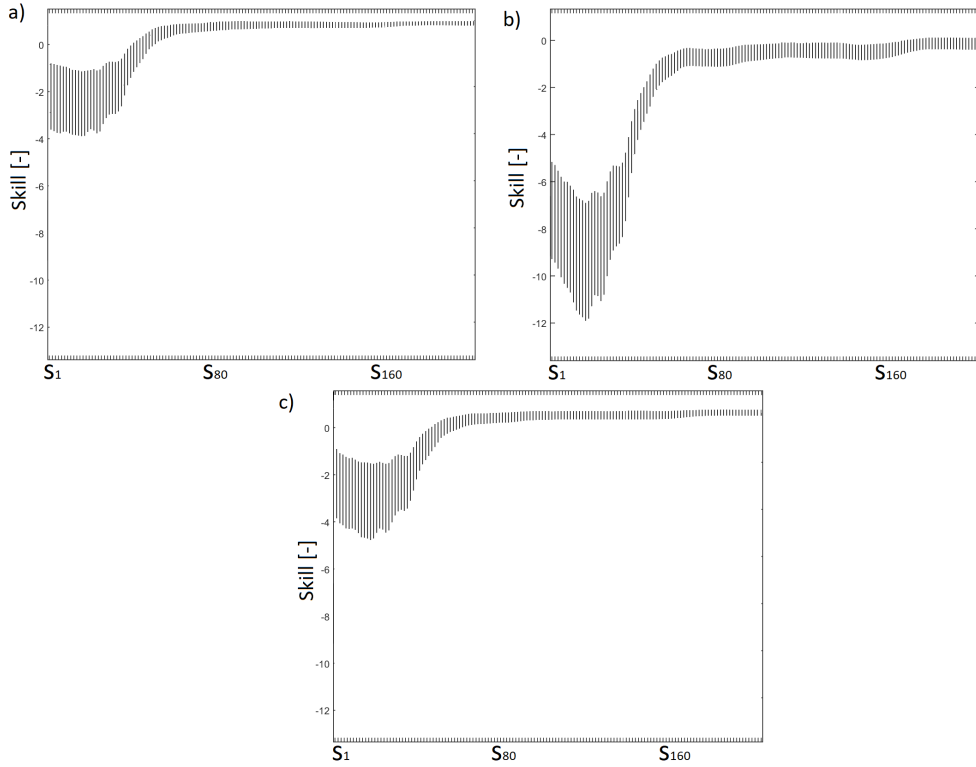


Fig. 5.7: Rangos de valores de skill para cada periodo de 15 minutos (s_i , comenzando desde el minuto i) en Arica. Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 60 minutos tras iniciado el terremoto.

5.1.3 No linealidad

El comportamiento de las olas usualmente, en mar abierto y a grandes profundidades, se asume lineal donde el principio de superposición es admisible. Incluso cercano a la costa, se suele utilizar este principio para que el costo computacional sea más menor. Sin embargo, en estos lugares, procesos no lineales tienen mayor relevancia y se deben considerar al momento de utilizar los datos. Una de las ecuaciones que muestran estos procesos no lineales, por ejemplo, es la ecuación de Euler de momentum:

$$\frac{\partial \vec{u}}{\partial t} + \frac{A}{h} \vec{u} \cdot \nabla \vec{u} + \nabla \eta = O(\mu^2) \quad (5.3)$$

donde $\vec{u}$ es el vector de velocidad, A es la amplitud de la onda, h es la batimetría en ese punto respecto a la superficie del mar, η es la altura de la onda respecto a la superficie del

mar y a la derecha los términos de segundo orden que dependen de $\mu = u = h/L$. El término de advección en tanto, que es el cambio de velocidad debido al movimiento de un fluido, se desprecia sólo cuando la amplitud de la onda es mucho menor a la profundidad del mar, que en caso de zonas cercanas a la costa deja de ser un supuesto válido.

Es por ello que al realizar inversión con datos de mareógrafos cercanos a la costa con procesos no lineales actuando puede generar predicciones sesgadas a mayor tiempo transcurrido desde su llegada a la costa. Otro factor asociado es que las batimetrías utilizadas no son lo suficientemente finas para reproducirlos y se tienen efectos de retraso o adelanto de las ondas, como también señales adicionales no generadas directamente por el tsunami después de los primeros ciclos de llegada. Es por ello que Fujii y Satake, 2006 utilizan sólo el primer ciclo de cada uno de los mareógrafos, explicando que los datos de batimetría pueden no ser suficientes para modelar fases posteriores, como las ondas reflejadas. Una de las hipótesis de las imprecisiones del método es que a través de estos efectos la inversión no logra tener resultados tan certeros, sobre todo luego de los primeros 60 minutos de datos. Para lograr verificar estas hipótesis es que se realiza un segundo ejemplo sintético. Se propaga la solución de Hayes et al., 2014 de deformación inicial y en puntos de pronósticos mar adentro, 0.4 grados al oeste de cada uno (figura 5.8), se obtienen sus series de tiempo para ser utilizados como datos para la inversión (sin ruido). El objetivo del ejemplo será el de simular puntos de pronósticos fuera de la costa y así obviar los procesos no lineales costeros. El problema se abarca tal como en el caso 2 de la sección 5.1.2, donde la matriz de covarianza es no estacionaria y correlacionada.

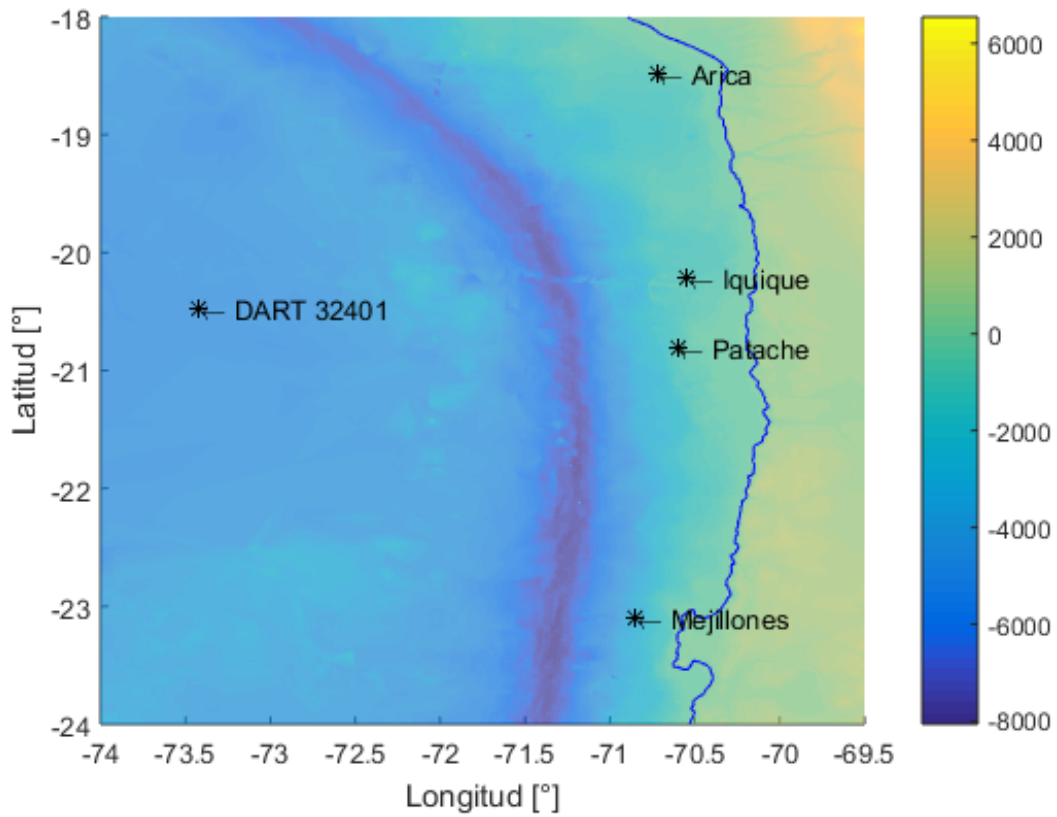


Fig. 5.8: Puntos de pronóstico para el ejemplo no lineal, correspondientes a las ubicaciones de los mareógrafos de Arica, Iquique, Patache y Mejillones desplazados en 0.4 grados mar adentro y la boya DART 32401

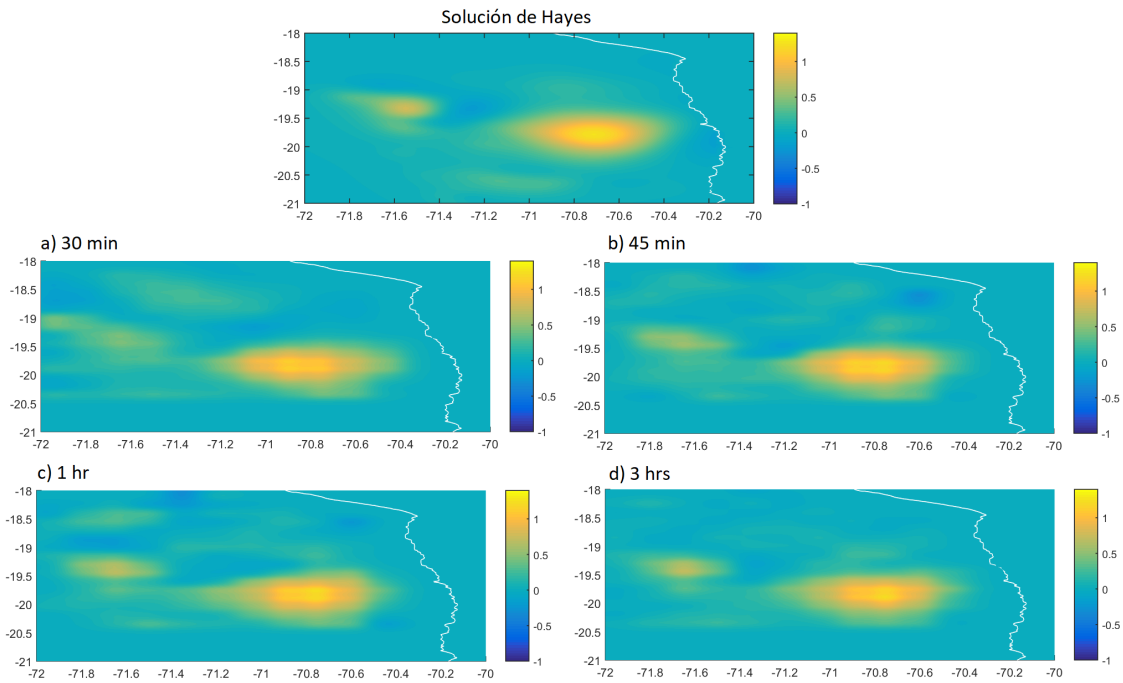


Fig. 5.9: Resultado de superficies iniciales del mar a partir de la solución de Hayes et al., 2014 para distintas cantidades de datos de entrada al método.

De la figura 5.9 se puede comprobar que, para datos no afectados a eventos producidos en la costa, el resultado más probable de la inversión da realmente cercano a la solución, incluso para tiempos cortos de 30 minutos. Y al contrario del ejemplo real, las inversiones mejoran a medida que se tienen más datos.

En cuanto a la incertidumbre asociada a la inversión en la figura 5.10 ubicada 0.4 grados mar adentro del mareógrafo de Iquique, es posible notar que al invertir, la solución promedio se acerca a la solución de Hayes a medida que pasa el tiempo y la incertidumbre disminuye. No obstante, en la inversión de 3 horas de datos la solución promedio no mejora desde la última inversión con una hora de datos y la incertidumbre disminuye nuevamente de forma tal que las demás soluciones obtenidas a través de la ecuación de densidad de probabilidad no generan tampoco una curva exactamente igual a la solución (no se encuentra en los rangos de incertidumbre de la sección achurada lila).

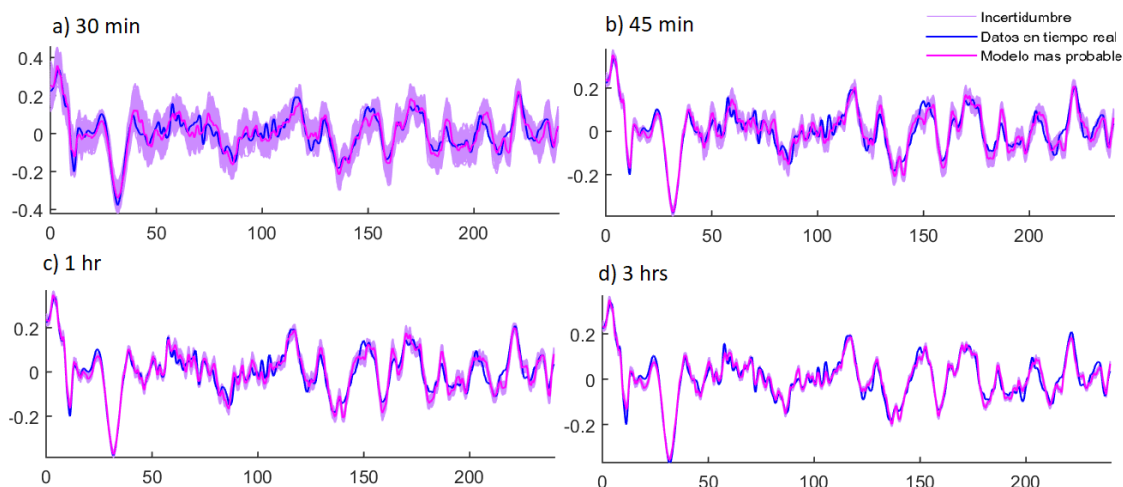


Fig. 5.10: Resultados de diferentes tiempos de inversión a 0.4 grados mar adentro del mareógrafo de Iquique.

De este ejemplo sintético se puede concluir que, para inversiones con datos no afectados a procesos no lineales, los resultados mejoran sustancialmente. Incluso, a mayor cantidad de datos utilizados, la solución promedio mejora, a diferencia del caso anterior. En cuanto a la incertidumbre, ésta disminuye a medida que se tienen más datos y es por ello que alrededor de las 3 horas la sección achurada lila se separa ligeramente de la solución de Hayes en ciertos puntos y esto puede generar en escalas mayores de separación una subestimación de la elevación de la superficie libre.

Es de destacar que si bien, hay una mejoría en la solución con respecto a los mareógrafos de la costa, este ejercicio no tiene iguales condiciones a las del evento real, ya que se invierte con datos de otra solución y sin ruido. Es por ello que se consideran los efectos no lineales como algo relevante, pero aún es necesario estudiarlos con más profundidad, como también la forma de incluirlos en la inversión, no como medida de reducción de incertidumbre, pero sí para que los datos que son sometidos a estos efectos puedan aportar información a la inversión.

5.1.4 Implicación a Futuro

En este trabajo se aplica el método ABIC a procesos de inversión de onda de tsunami. Si bien los resultados de la inversión con tiempos muy grandes de datos (4 horas) no dan tan precisos, es posible trabajar con los resultados hasta los 60 minutos, los cuales son cercanos

a las soluciones encontradas hasta ahora y tienen una buena capacidad de predicción, pero ¿Qué hacer con estas soluciones?

Uno de los objetivos de este trabajo es poder complementar el sistema de alerta en Chile. Entonces, para poder conocer el nivel de amenaza al que se expone la población, es necesario relacionar la deformación inicial del tsunami con el nivel de inundación que generarán al propagarse a la costa. No obstante, al aplicar el método ABIC se tienen múltiples escenarios, al ser una función de densidad de probabilidad, por lo que la solución no es algo trivial. Una de las propuestas de este trabajo es realizar un submuestreo entre las soluciones y así entre los niveles de inundación elegir los peores escenarios. Luego de ello, elegir un nivel de confianza que logre que la alerta y evacuación sea apropiada en cada situación, realizando un análisis de sensibilidad. Ahora bien, esto dependerá de los niveles de incertidumbre asociados a cada solución como también de la cantidad de datos de inversión, ya que las soluciones varían radicalmente en la primera hora de datos.

Por otra parte el método debe trabajar de forma rápida si se quiere alertar a la población. No obstante este sistema depende de los datos recogidos por los mareógrafos casi en su totalidad, en los cuales el oleaje puede llegar luego de 20 minutos. Además puede que las primeras olas sean ya devastadoras, las cuales no se pueden predecir con el modelo. Por otra parte, en la sección anterior se vieron las consecuencias de la ubicación de los mareógrafos por efectos no lineales. Así pues, es necesario mejorar el sistema de sensores para una predicción adecuada y rápida, con sensores que se encuentren mar adentro.

Otro de los desarrollos a futuro implica la mejora de este método. Se destaca el desarrollo de matrices de covarianza que generen mejores resultados posteriores a los 60 minutos. Esto puede ser mediante mejores correlaciones entre los errores, ajustando el t_0 de las correlaciones de manera precisa para cada set de datos (y no uno para todas las series de tiempo), o generando matrices no estacionarias que se ajusten a la variación en los errores en el tiempo y caracterizando de mejor manera las bahías. Sin embargo, los procesos no lineales son importantes y la linealidad es una de las bases del método, por lo que parece difícil poder generar una mejora sustancial con una cantidad muy grande de datos (mayor a los 60 minutos) y por ende esto tampoco ayudaría a disminuir la incertidumbre.

Conclusión

El objetivo principal de este trabajo ha sido la creación de un algoritmo que permita incorporar incertidumbre al método de inversión y así obtener las posibles deformaciones iniciales del mar que origina el tsunami. Este algoritmo, creado por Akaike 1980, mediante la incorporación de información a priori (en este trabajo la suavidad) y en combinación con datos en tiempo real, forma la distribución de probabilidad a posteriori del modelo. De modo de verificación se utiliza el método en un ejemplo sintético y en el evento de Iquique 2014. Para la aplicación del algoritmo se plantean diversas hipótesis como la linealidad, que es base para la superposición de las fuentes unitarias, y los valores y forma de la matriz de covarianza residual de los datos, que caracteriza el error entre lo medido y el resultado de la inversión.

Se verifica el método con datos sintéticos, a los cuales se les aplica un ruido. El método logra predecir la deformación inicial incluso con niveles de ruido relativamente altos. A continuación se aplica el método al evento de Iquique 2014. Desarrollado el método de inversión y analizando los resultados es posible notar que el método, aunque puede predecir de forma aproximada los resultados, a partir de los 60 minutos de datos, comienza a decaer en precisión al igual que su incertidumbre, la cual no logra abarcar la solución real. Sin embargo entre los 30 y los 60 minutos logra buenos resultados, cercano a otras soluciones dadas por Hayes et al., 2014 y An et al., 2014, confirmando la capacidad del método a través de índices de skill.

En este trabajo una parte relevante fue el análisis de la matriz de covarianza, la cual se utiliza como correlacionada y/o iterativa, es decir que se actualiza en cada inversión. Es posible notar que las soluciones varían sustancialmente en uno u otro caso. No obstante, las matrices utilizadas no caracterizan por completo los errores, pudiendo ser no estacionarias y variar los valores de la diagonal respecto al tiempo.

Otro de los retos que enfrenta el método es el uso de los mareógrafos, los cuales debido a su ubicación en la costa, están expuestos a efectos no lineales, los que se recogen luego de los primeros ciclos después de la llegada del tsunami, ya que no se cumplen las hipótesis de

superposición lineal. Esto puede ser evitado utilizando una cantidad de datos que no esté expuesta a estos fenómenos. Además debido a su ubicación, no es posible generar resultados antes de que lleguen las primeras olas a la costa, las que pueden ser devastadoras. Es posible mejorar los tiempos y evitar además la no linealidad utilizando una mayor cantidad de boyas DART, que no están afectas a los efectos no lineales por su ubicación mar adentro, pero que en Chile son escasas.

Referencias

- Akaike, Hirotugu (1980). «Likelihood and the Bayes procedure». In: *Trabajos de Estadística Y de Investigación Operativa* 31.1, pp. 143–166.
- An, Chao, Ignacio Sepúlveda, and Philip L-F Liu (2014). «Tsunami source and its validation of the 2014 Iquique, Chile, earthquake». In: *Geophysical Research Letters* 41.11, pp. 3988–3994.
- Barrientos, Sergio (2014). «Informe técnico terremoto de Iquique, Mw= 8.2, 1 de Abril de 2014». In: *Centro Sismológico Nacional. Universidad de Chile*.
- Bernardo, José M (2003). *Bayesian Statistics. Encyclopedia of Life Support Systems (EOLSS). Probability and Statistics*.
- Blewitt, Geoffrey et al. (2009). «GPS for real-time earthquake source determination and tsunami warning systems». In: *Journal of Geodesy* 83.3, pp. 335–343.
- Chen, Kejie et al. (2016). «Comparing source inversion techniques for GPS-based local tsunami forecasting: A case study for the April 2014 M8.1 Iquique, Chile, earthquake». In: *Geophysical Research Letters* 43.7. 2016GL068042, pp. 3186–3192.
- Dettmer, Jan, Stan E Dosso, and Charles W Holland (2007). «Uncertainty estimation in seismo-acoustic reflection travel time inversion». In: *The Journal of the Acoustical Society of America* 122.1, pp. 161–176.
- Duputel, Zacharie et al. (2012). «Uncertainty estimations for seismic source inversions». In: *Geophysical Journal International* 190.2, pp. 1243–1256.
- Fujii, Yushiro and Kenji Satake (2006). «Source of the July 2006 West Java tsunami estimated from tide gauge records». In: *Geophysical Research Letters* 33.24.
- Fujii, Yushiro, Kenji Satake, et al. (2011). «Tsunami source of the 2011 off the Pacific coast of Tohoku Earthquake». In: *Earth, Planets and Space* 63.7, p. 55.
- Fukahata, Yukitoshi, Yuji Yagi, and Mitsuhiro Matsu'ura (2003). «Waveform inversion for seismic source processes using ABIC with two sorts of prior constraints: Comparison between proper and improper formulations». In: *Geophysical Research Letters* 30.6. 1305, n/a–n/a.

- Fukuda, Jun'ichi and Kaj M. Johnson (2008). «A Fully Bayesian Inversion for Spatial Distribution of Fault Slip with Objective Smoothing». In: *Bulletin of the Seismological Society of America* 98.3, p. 1128. eprint: /gsw/content_public/journal/bssa/98/3/10.1785_0120070194/4/1128.pdf.
- Geist, Eric L (2002). «Complex earthquake rupture and local tsunamis». In: *Journal of Geophysical Research: Solid Earth* 107.B5.
- Gonzalez, F. I. et al. (1991). «The 1987–88 Alaskan Bight tsunamis: Deep ocean data and model comparisons». In: *Natural Hazards* 4.2, pp. 119–139.
- Hayes, Gavin P et al. (2014). «Continuing megathrust earthquake potential in Chile after the 2014 Iquique earthquake». In: *Nature* 512.7514, pp. 295–298.
- Kanamori, Hiroo (1972). «Mechanism of tsunami earthquakes». In: *Physics of the Earth and Planetary Interiors* 6.5, pp. 346–359.
- Kawai, Hiroyasu et al. (2013). «Characteristics of the 2011 Tohoku tsunami waveform acquired around Japan by nowphas equipment». In: *Coastal Engineering Journal* 55.03, p. 1350008.
- Liu, Philip L-F et al. (1995). «Numerical simulations of the 1960 Chilean tsunami propagation and inundation at Hilo, Hawaii». In: *Tsunami: progress in prediction, disaster prevention and warning*. Springer, pp. 99–115.
- Malinverno, Alberto (2000). «A Bayesian criterion for simplicity in inverse problem parametrization». In: *Geophysical Journal International* 140.2, pp. 267–285.
- Melgar, Diego et al. (2016). «Slip segmentation and slow rupture to the trench during the 2015, Mw8.3 Illapel, Chile earthquake». In: *Geophysical Research Letters* 43.3. 2015GL067369, pp. 961–966.
- Mitsuhata, Yuji et al. (2001). «Electromagnetic heterogeneity of the seismogenic region of 1962 M6.5 Northern Miyagi Earthquake, northeastern Japan». In: *Geophysical Research Letters* 28.23, pp. 4371–4374.
- Press, William H et al. (1992). «Numerical recipes in C++». In: *The art of scientific computing*.
- Romano, F et al. (2016). «Optimal time alignment of tide-gauge tsunami waveforms in nonlinear inversions: Application to the 2015 Illapel (Chile) earthquake». In: *Geophysical Research Letters* 43.21.
- Saito, Tatsuhiko et al. (2011). «Tsunami source of the 2011 Tohoku-Oki earthquake, Japan: Inversion analysis based on dispersive tsunami simulations». In: *Geophysical Research Letters* 38.7. L00G19, n/a–n/a.
- Satake, Kenji (1989). «Inversion of tsunami waveforms for the estimation of heterogeneous fault motion of large submarine earthquakes: The 1968 Tokachi-oki and 1983 Japan Sea earthquakes». In: *Journal of Geophysical Research: Solid Earth* 94.B5, pp. 5627–5636.
- Takagawa, Tomohiro (2016). «Observation-based Quantitative Uncertainty Estimation for Realtime Tsunami Inundation Forecast using ABIC and Ensemble Simulation». In: *AGU Fall Meeting Abstracts*.

- Tanioka, Yuichiro et al. (2006). «Rupture process of the 2004 great Sumatra-Andaman earthquake estimated from tsunami waveforms». In: *Earth, Planets and Space* 58.2, pp. 203–209.
- Tatehata, H. (1997). «The New Tsunami Warning System of the Japan Meteorological Agency». In: *Advances in natural and technological hazards research* 9, pp. 175/188.
- Titov, Vasily V et al. (2005). «Real-time tsunami forecasting: Challenges and solutions». In: *Developing tsunami-resilient communities*. Springer, pp. 41–58.
- Tsushima, Hiroaki et al. (2009). «Near-field tsunami forecasting from cabled ocean bottom pressure data». In: *Journal of Geophysical Research: Solid Earth* 114.B6. B06309, n/a–n/a.
- Uchida, T (1993). «Smooth 2-D inversion for magnetotelluric data based on statistical criterion ABIC». In: *Journal of geomagnetism and geoelectricity* 45.9, pp. 841–858.
- Wang, Xiaoming and Philip L-F Liu (2006). «An analysis of 2004 Sumatra earthquake fault plane mechanisms and Indian Ocean tsunami». In: *Journal of Hydraulic Research* 44.2, pp. 147–154.
- Wang, Xiaoming and PL Liu (2005). «A numerical investigation of Boumerdes-Zemmouri (Algeria) earthquake and tsunami». In: *Computer Modeling in Engineering and Sciences* 10.2, p. 171.
- Willmott, Cort J (1981). «On the validation of models». In: *Physical geography* 2.2, pp. 184–194.
- Woodhouse, John H. and Adam M. Dziewonski (1984). «Mapping the upper mantle: Three-dimensional modeling of earth structure by inversion of seismic waveforms». In: *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5953–5986.
- Yabuki, T and M Matsu'Ura (1992). «Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip». In: *Geophysical Journal International* 109.2, pp. 363–375.
- Yadav, Rajeev Kumar et al. (2013). «Coseismic offsets due to the 11 April 2012 Indian Ocean earthquakes (Mw 8.6 and 8.2) derived from GPS measurements». In: *Geophysical Research Letters* 40.13, pp. 3389–3393.

Anexo A

Se muestran a continuación los resultados de las inversiones del evento de Iquique 2014 para los diferentes casos tratados en la tesis. Para cada tiempo de inversión se muestra la deformación inicial más probable y las series de tiempo para cada punto de pronóstico.

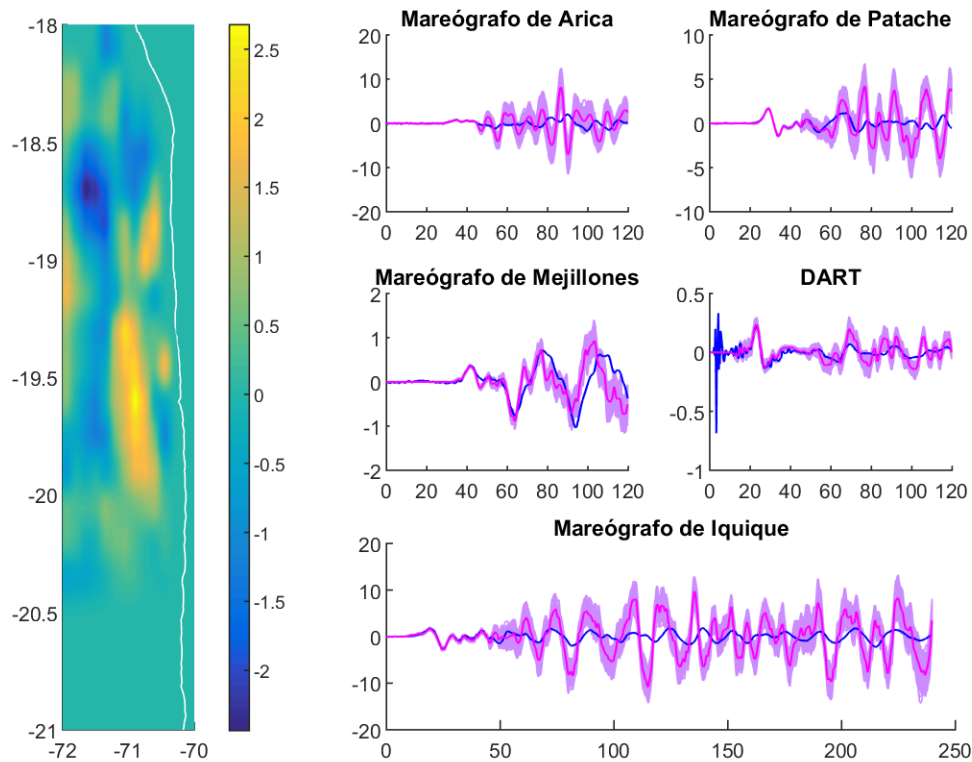


Fig. 6.1: Resultados de la inversión con covarianza estacionaria correlacionada para 45 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

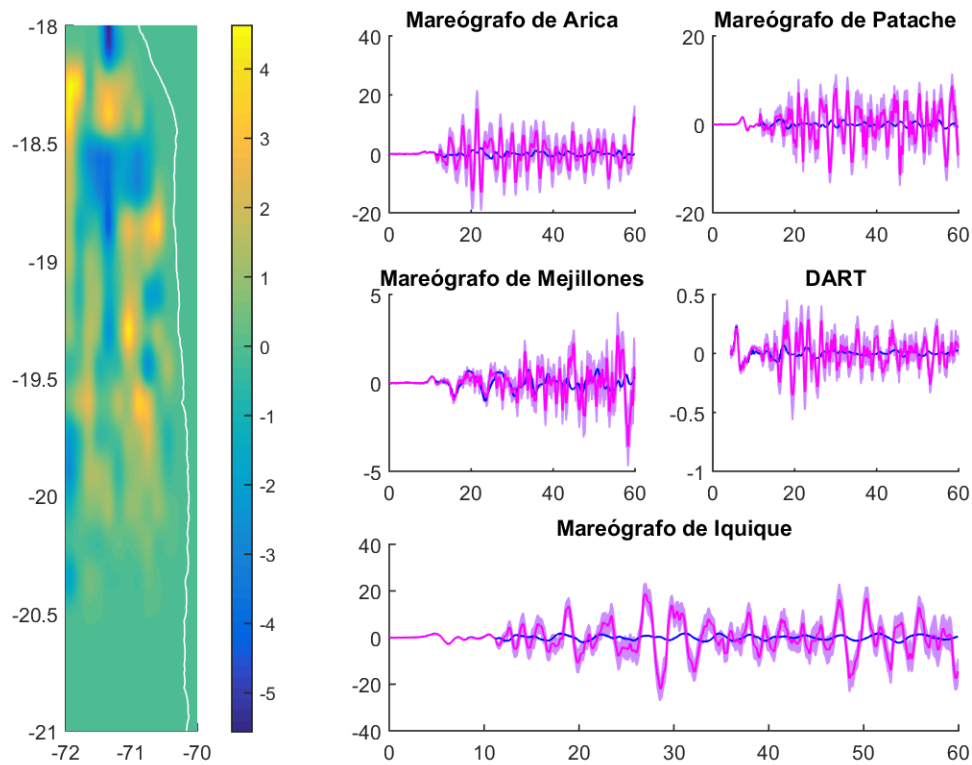


Fig. 6.2: Resultados de la inversión con matriz no estacionaria y no correlacionada para 45 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

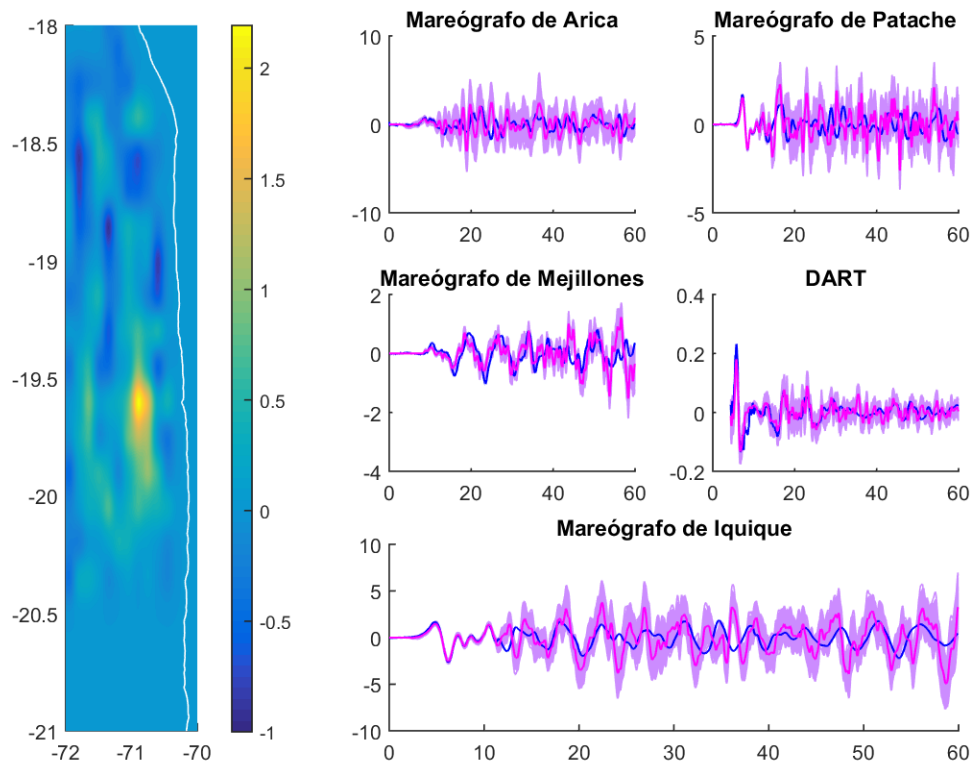


Fig. 6.3: Resultados de la inversión con matriz no estacionaria y correlacionada para 45 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

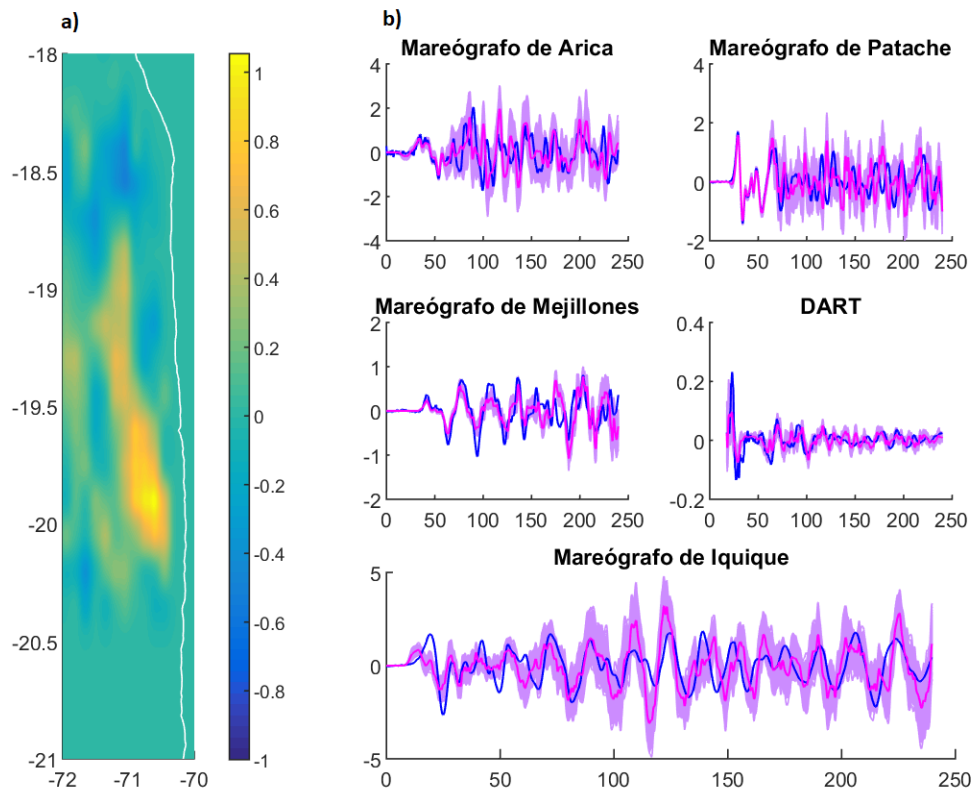


Fig. 6.4: Resultados de la inversión con matriz estacionaria y correlacionada para 60 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

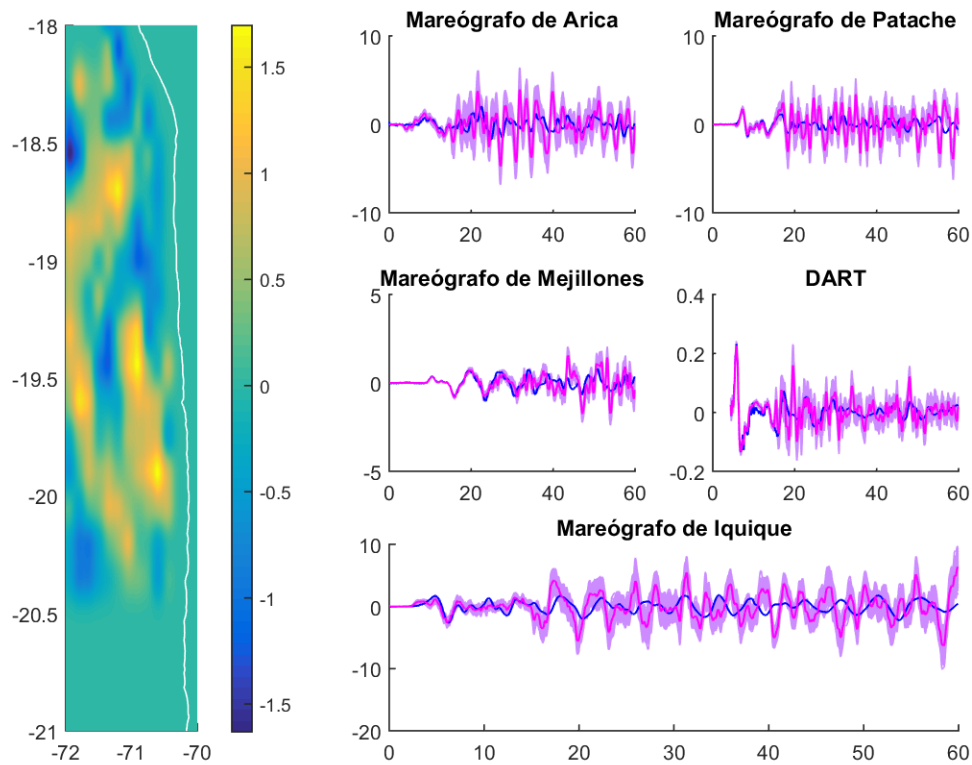


Fig. 6.5: Resultados de la inversión con matriz no estacionaria y no correlacionada para 60 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

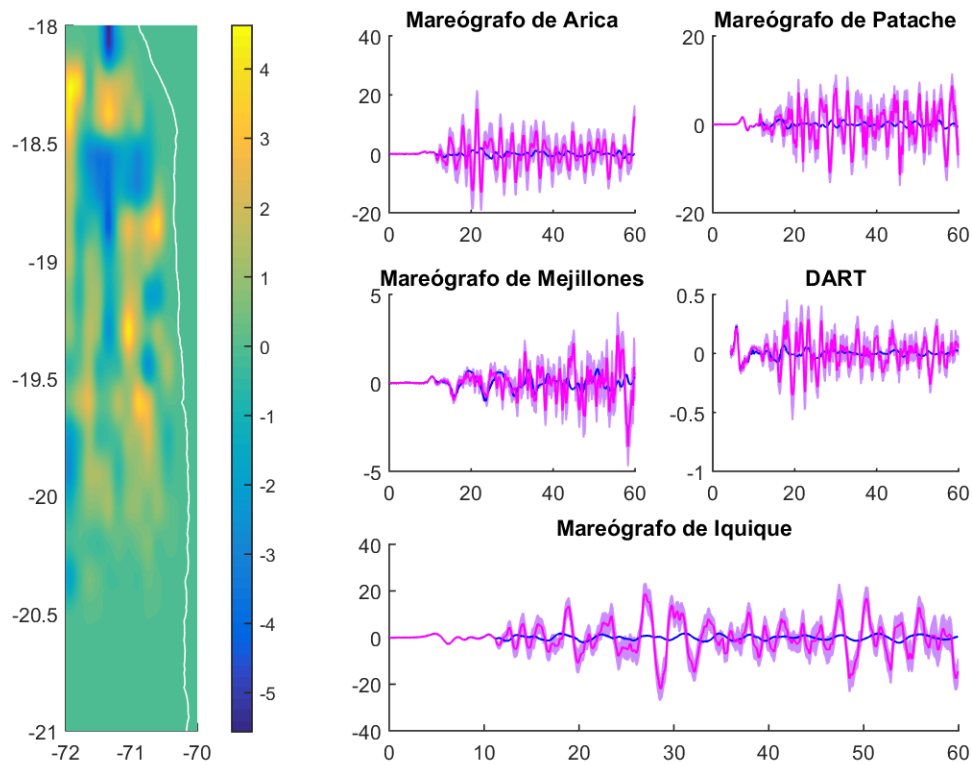


Fig. 6.6: Resultados de la inversión con matriz no estacionaria y correlacionada para 60 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

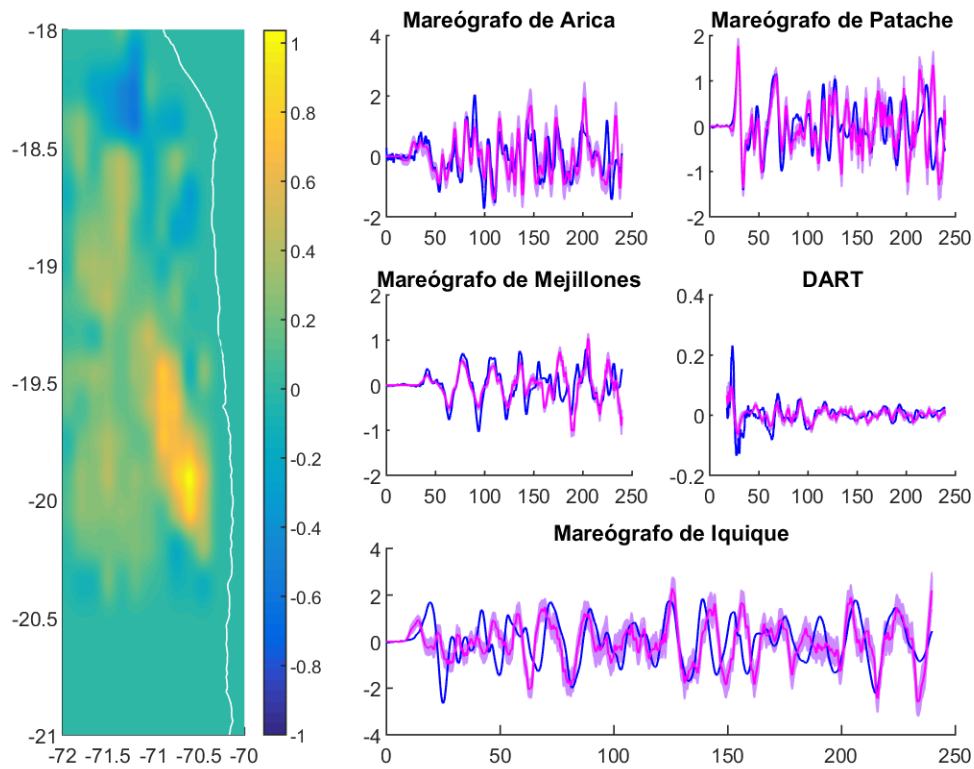


Fig. 6.7: Resultados de la inversión con matriz estacionaria y correlacionada para 180 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

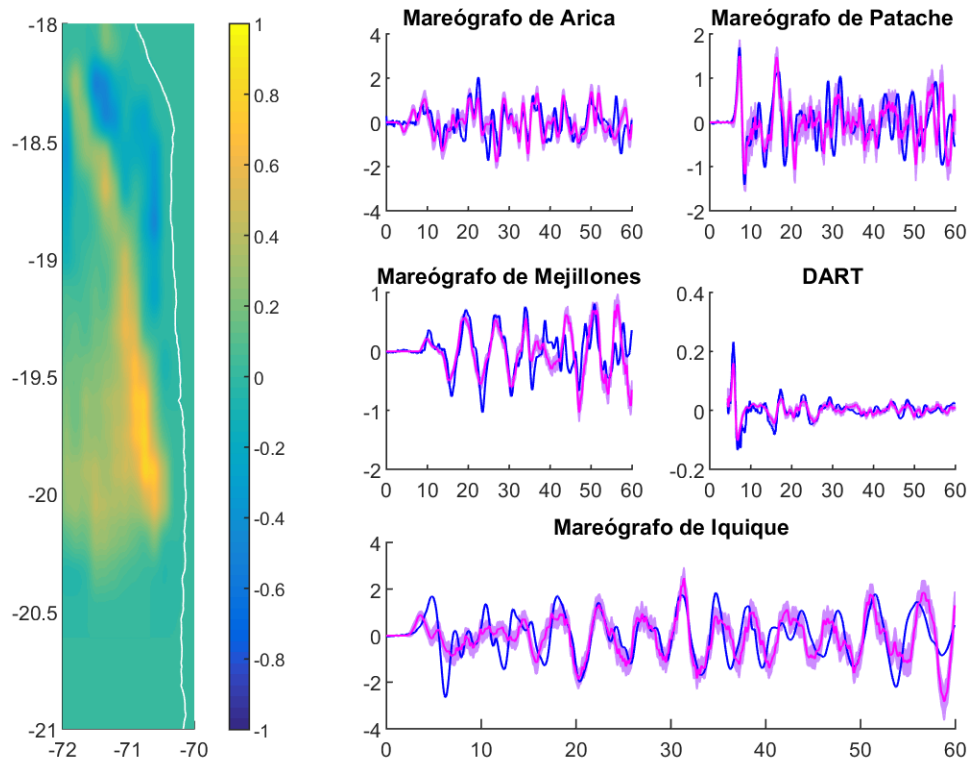


Fig. 6.8: Resultados de la inversión con matriz no estacionaria y no correlacionada para 180 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

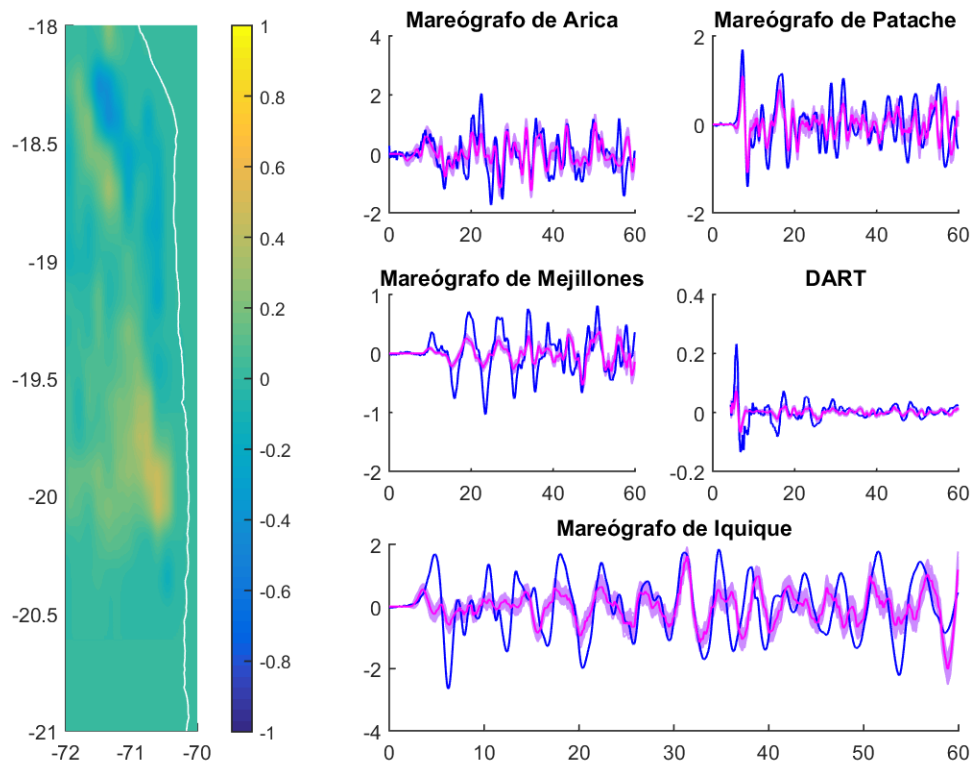


Fig. 6.9: Resultados de la inversión con matriz no estacionaria y correlacionada para 180 minutos de datos. (a) Deformación inicial más probable obtenida de la inversión. (b) Series de tiempo de los mareógrafos. En azul se muestra el valor real, en magenta la predicción para la deformación más probable y en lila el rango de valores posibles dada la incertidumbre.

Anexo B

En este anexo se muestran mapas de calor del índice de skill para el modelo más probable a posteriori para los distintos casos de matrices de covarianza residual.

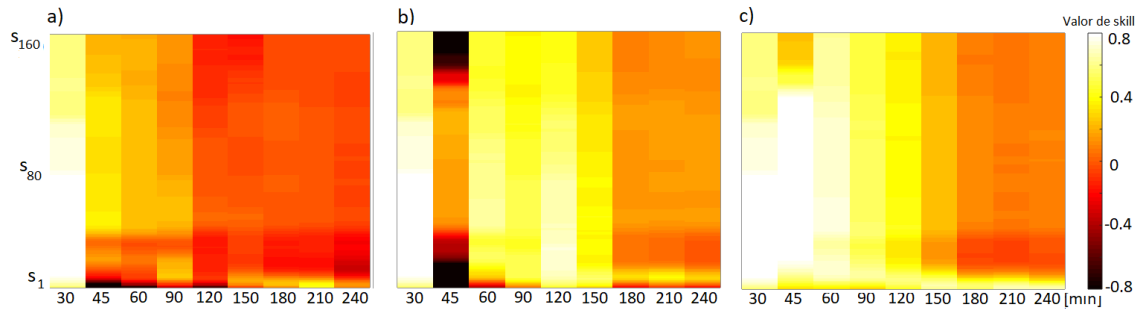


Fig. 6.10: Mapa de calor para valores de Skill en el mareógrafo de Iquique. En el eje y los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i y en el eje de las x se muestran los distintos periodos en que se realiza la inversión.

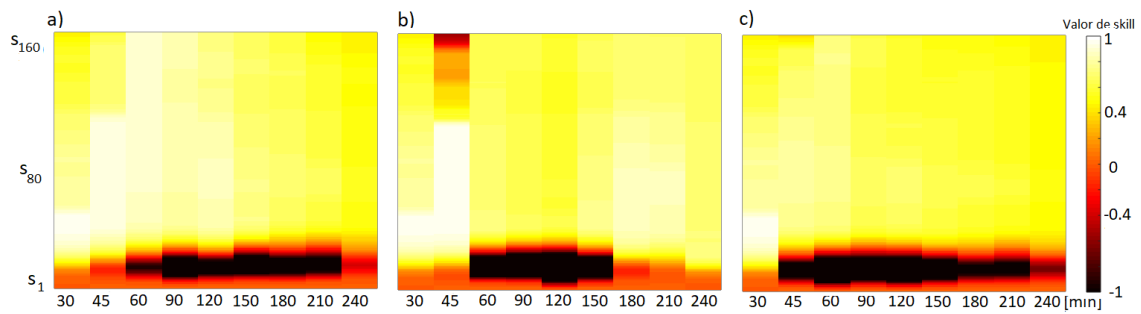


Fig. 6.11: Mapa de calor para valores de Skill en el mareógrafo de Patache. En el eje y los valores s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i y en el eje de las x se muestran los distintos periodos donde se realiza la inversión.

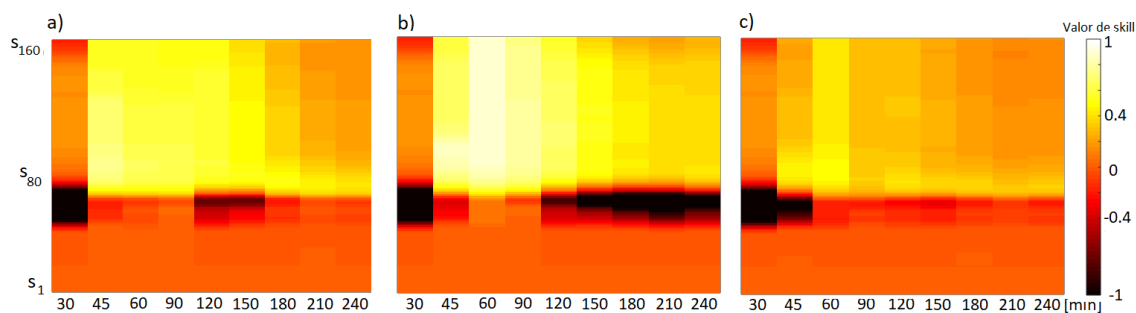


Fig. 6.12: Mapa de calor para valores de Skill en el mareógrafo de Mejillones. En el eje y los valores s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i y en el eje de las x se muestran los distintos periodos donde se realiza la inversión.

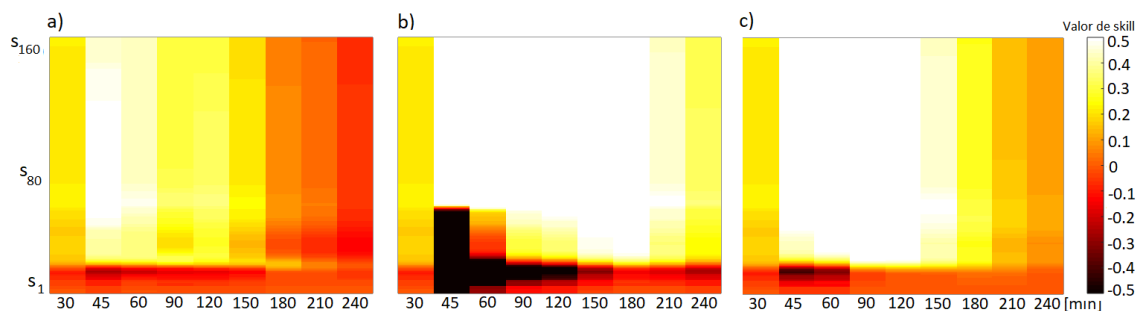


Fig. 6.13: Mapa de calor para valores de Skill en el mareógrafo de DART. En el eje y los valores s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i y en el eje de las x se muestran los distintos periodos donde se realiza la inversión.

Anexo C

En este anexo se muestran los valores de skill para los casos de matriz de covarianza mencionados y para los diferentes escenarios estocásticos.

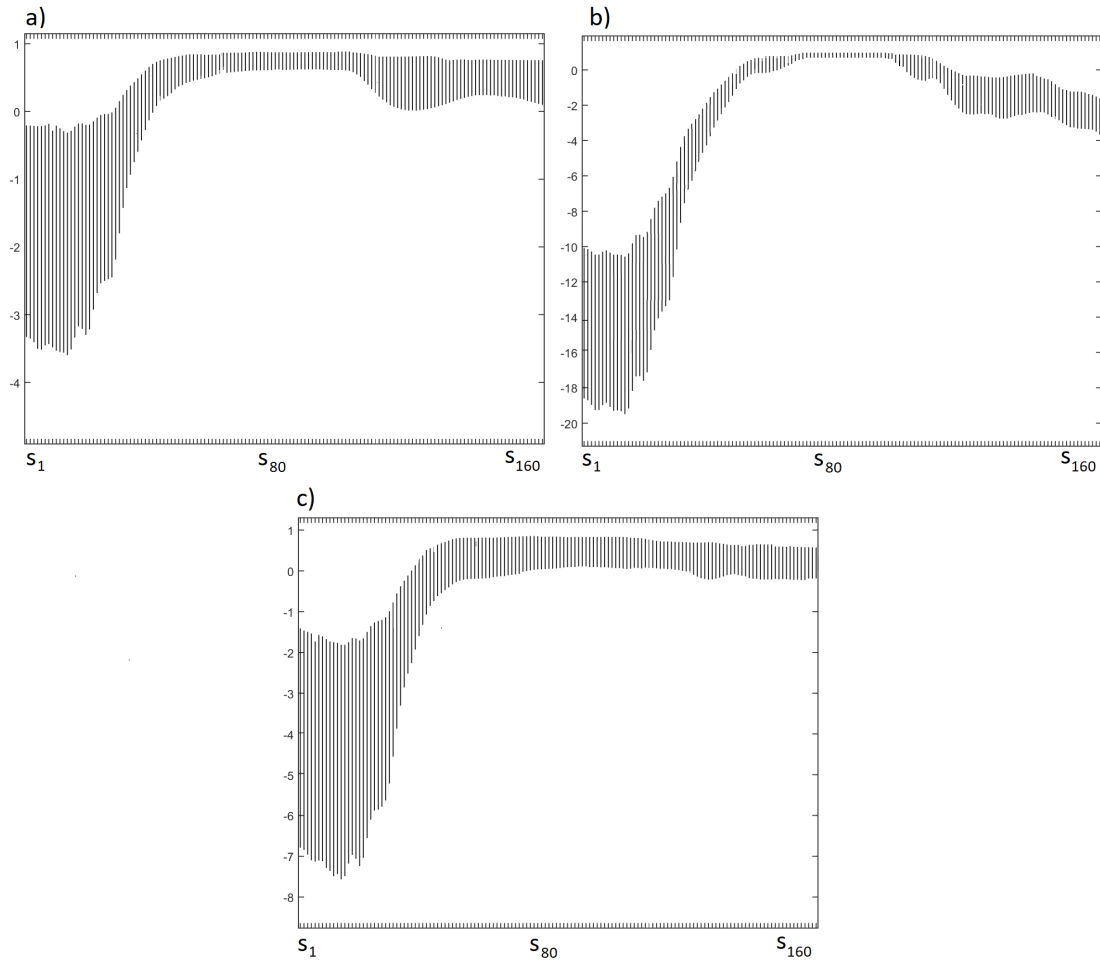


Fig. 6.14: Rangos de valores de skill para periodos de 15 minutos en Arica. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 45 minutos tras iniciado el terremoto.

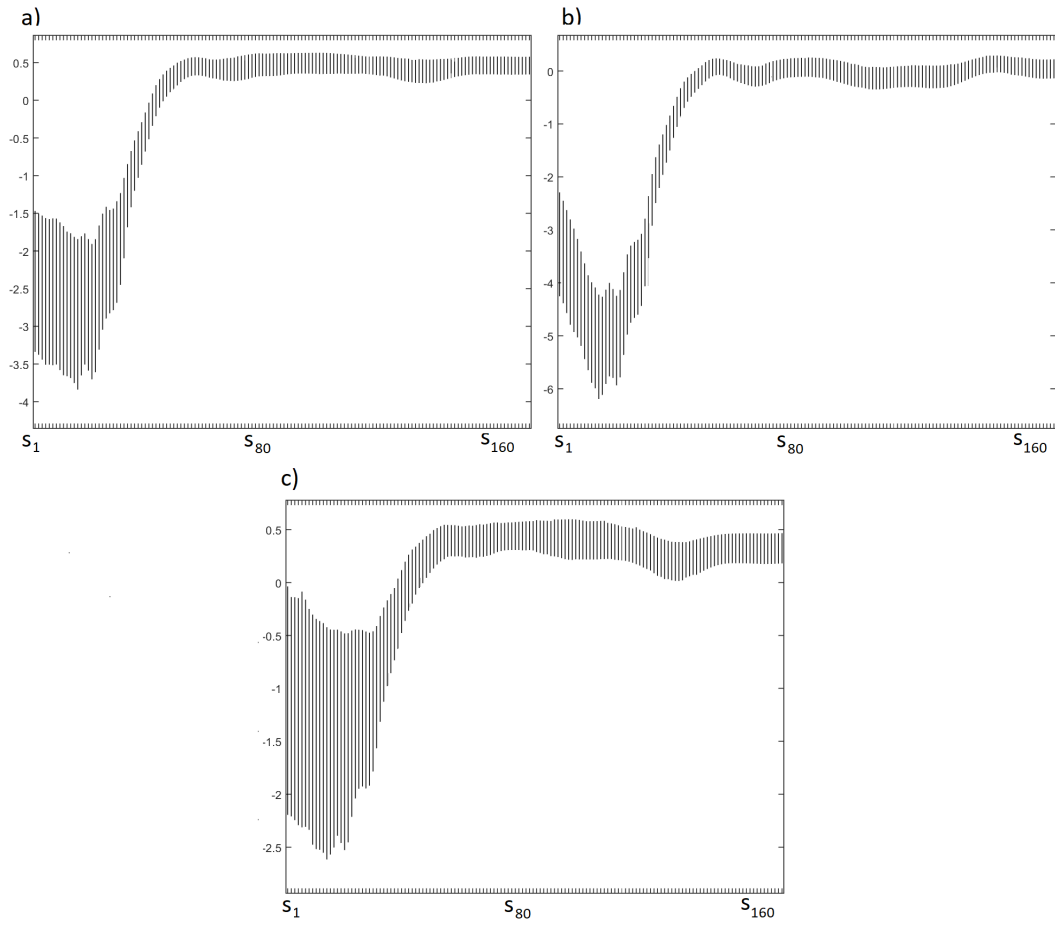


Fig. 6.15: Rangos de valores de skill para periodos de 15 minutos en Arica. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 180 minutos tras iniciado el terremoto.

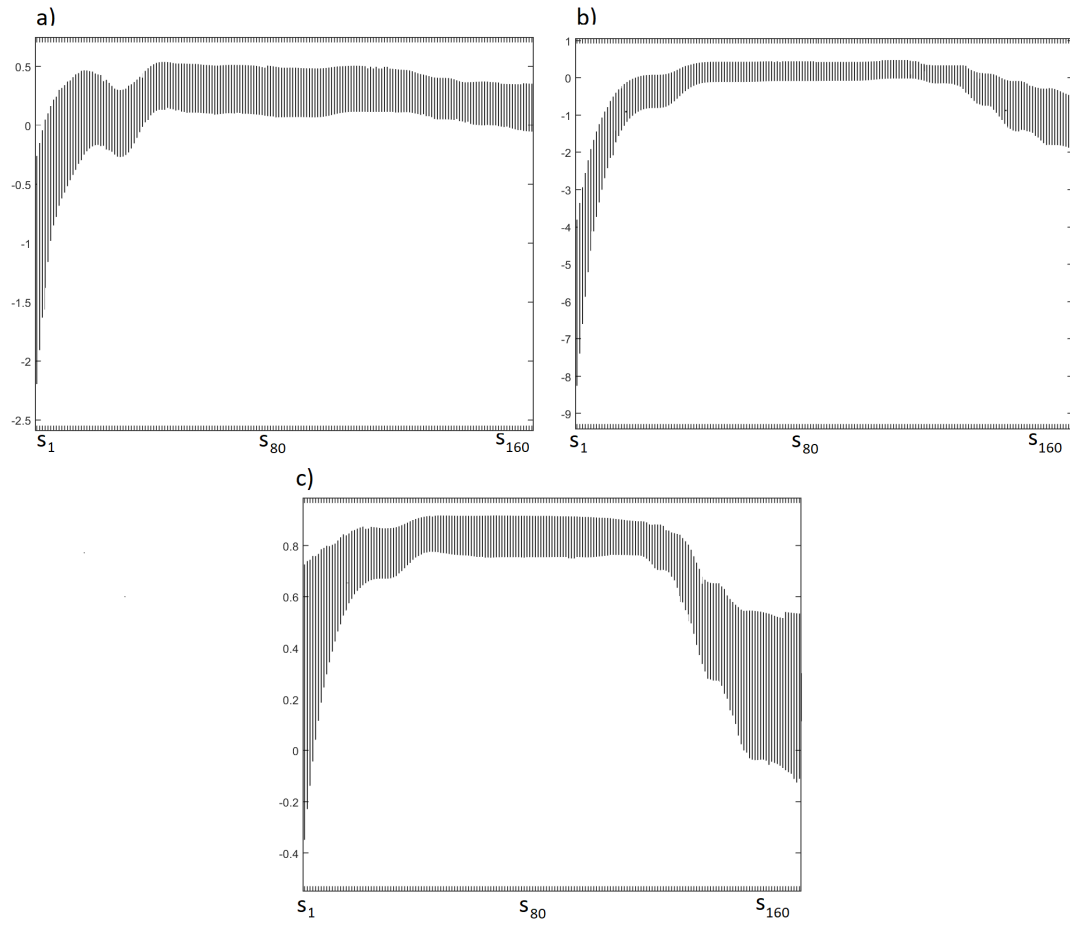


Fig. 6.16: Rangos de valores de skill para periodos de 15 minutos en Iquique. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 45 minutos tras iniciado el terremoto.

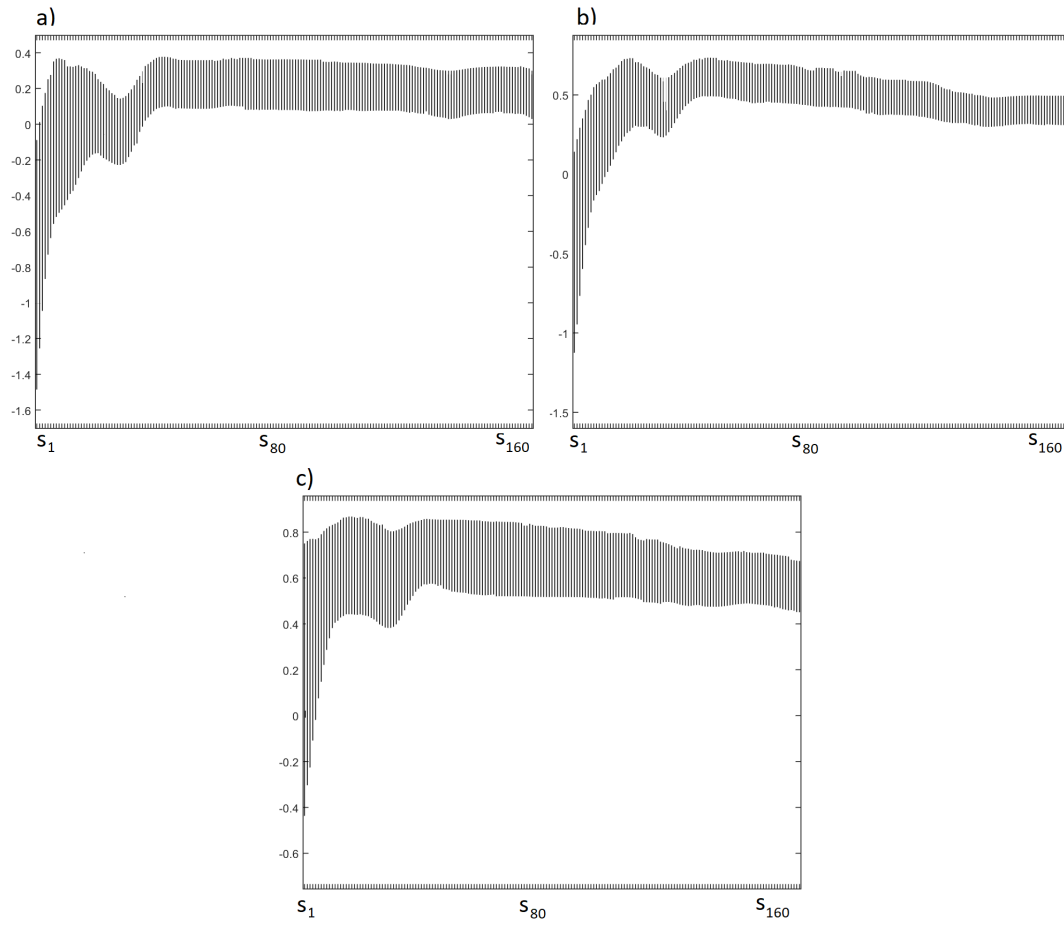


Fig. 6.17: Rangos de valores de skill para periodos de 15 minutos en Iquique. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 60 minutos tras iniciado el terremoto.

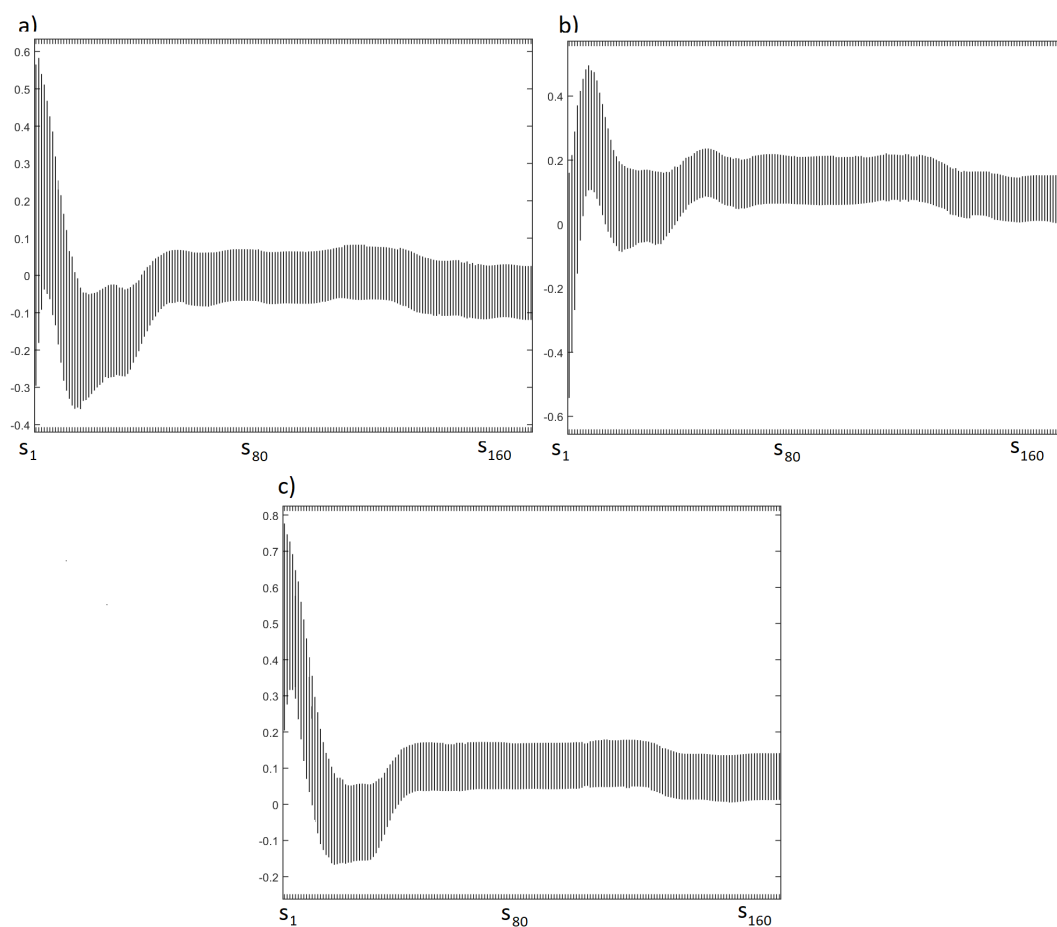


Fig. 6.18: Rangos de valores de skill para periodos de 15 minutos en Iquique. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 180 minutos tras iniciado el terremoto.

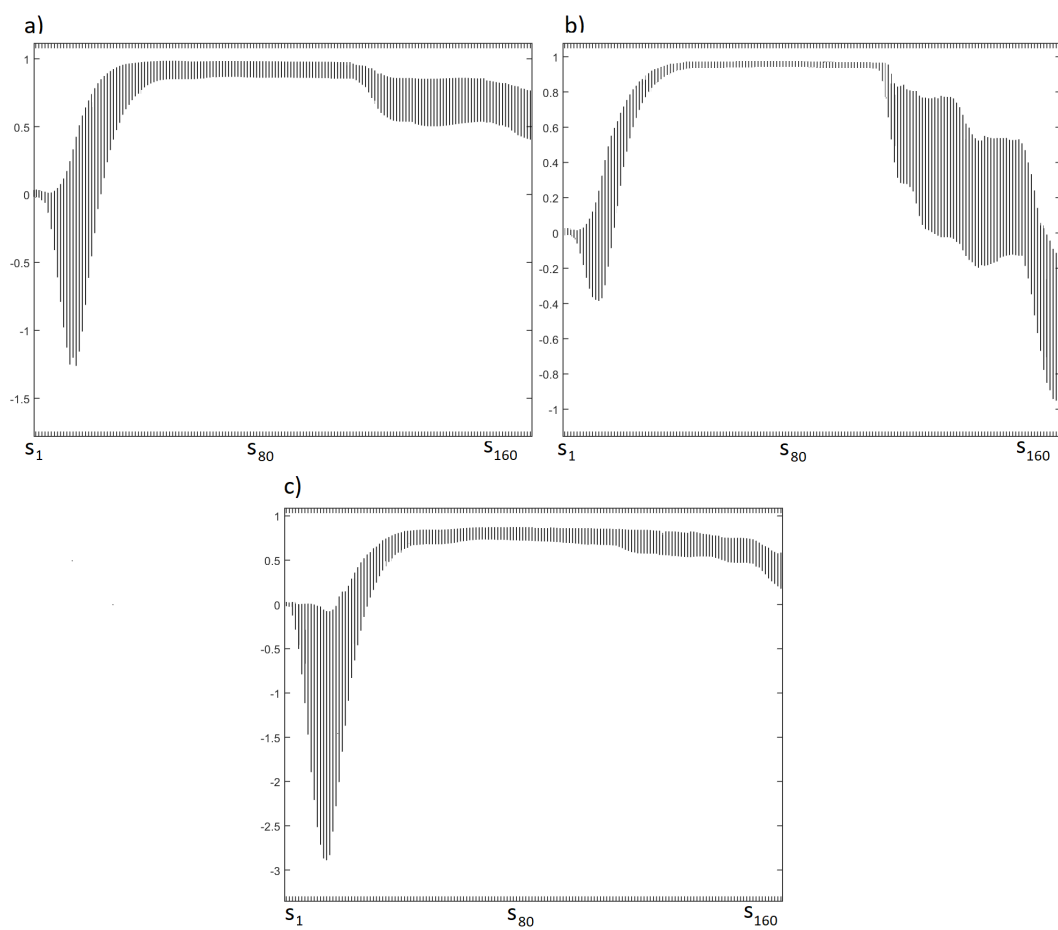


Fig. 6.19: Rangos de valores de skill para periodos de 15 minutos en Patache. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 45 minutos tras iniciado el terremoto.

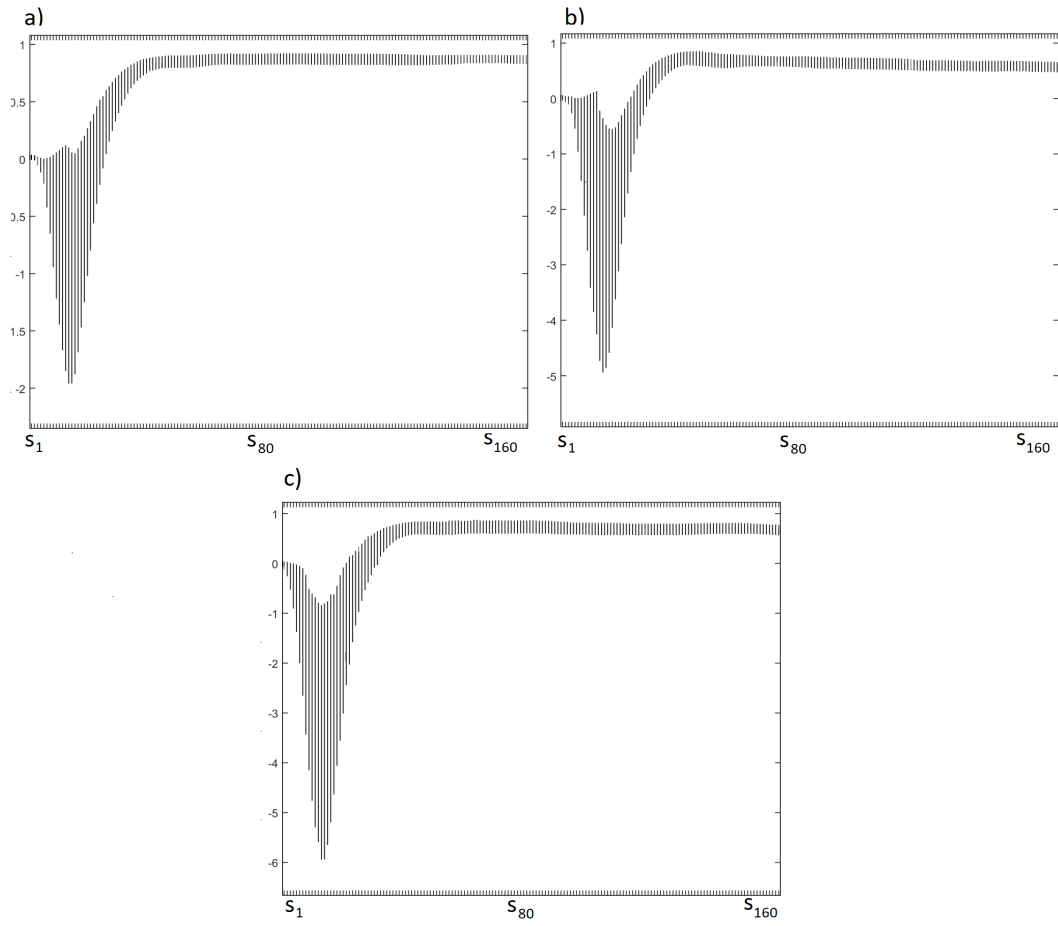


Fig. 6.20: Rangos de valores de skill para periodos de 15 minutos en Patache. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 60 minutos tras iniciado el terremoto.

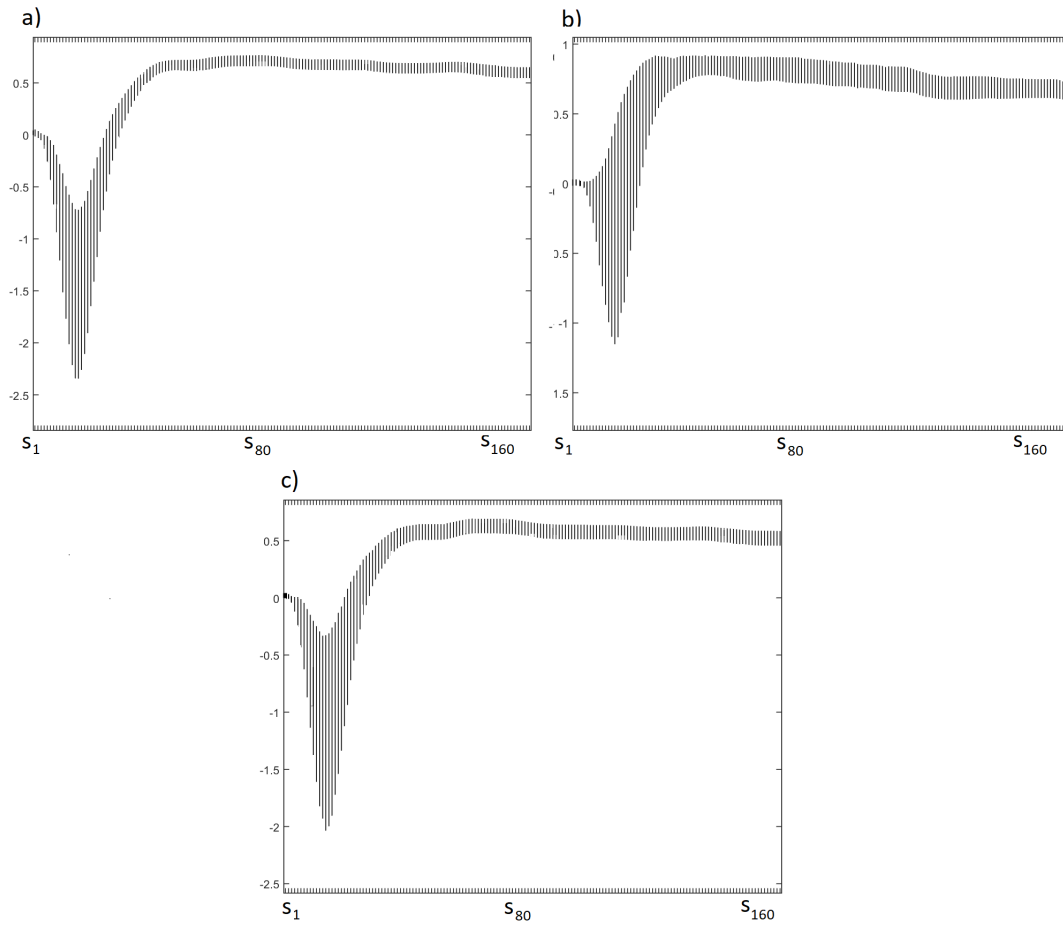


Fig. 6.21: Rangos de valores de skill para periodos de 15 minutos en Patache. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 180 minutos tras iniciado el terremoto.

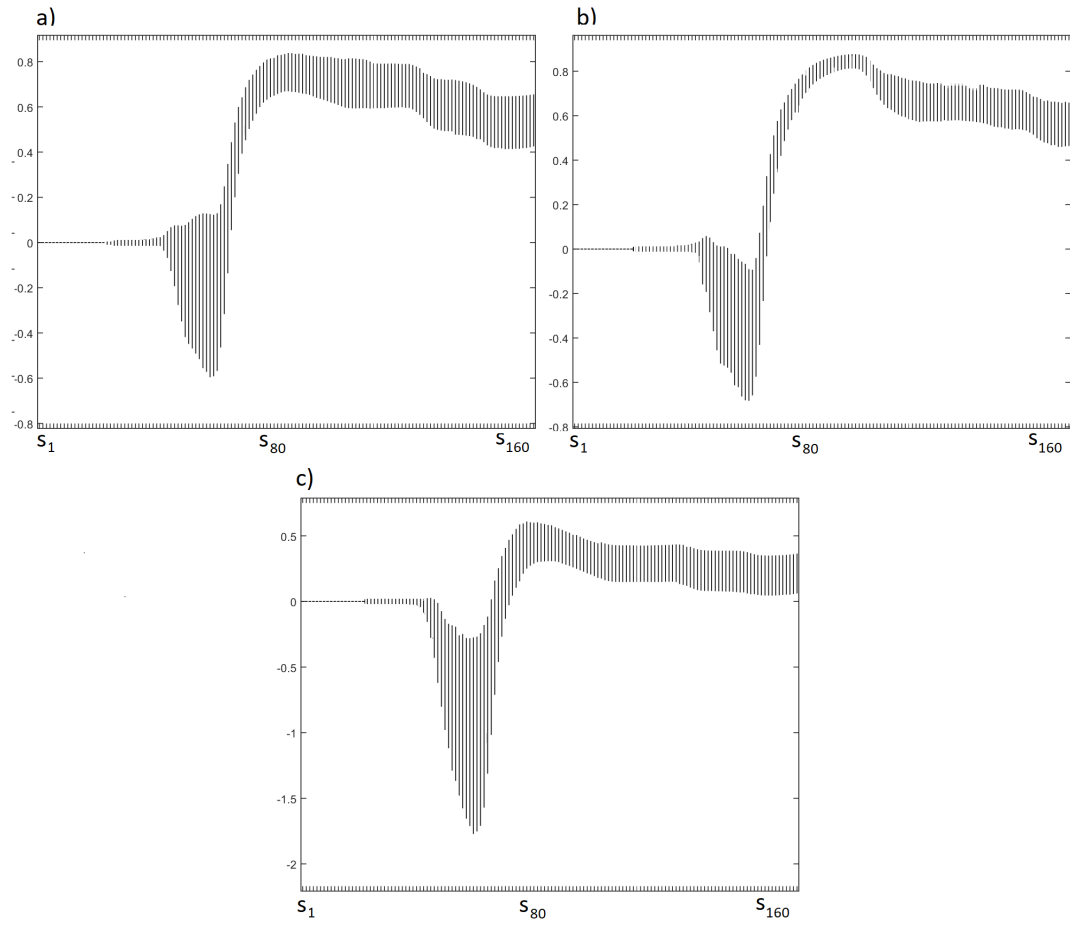


Fig. 6.22: Rangos de valores de skill para periodos de 15 minutos en Mejillones. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 45 minutos tras iniciado el terremoto.

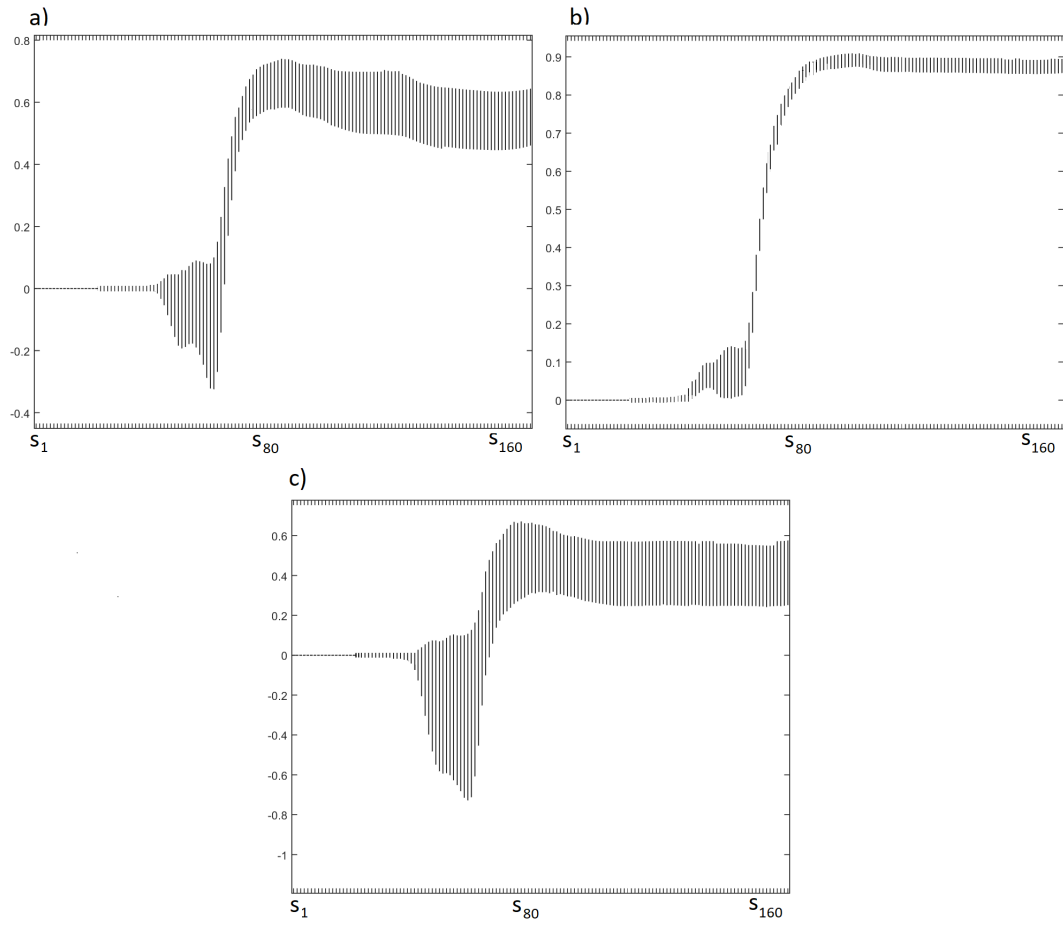


Fig. 6.23: Rangos de valores de skill para periodos de 15 minutos en Mejillones. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 60 minutos tras iniciado el terremoto.

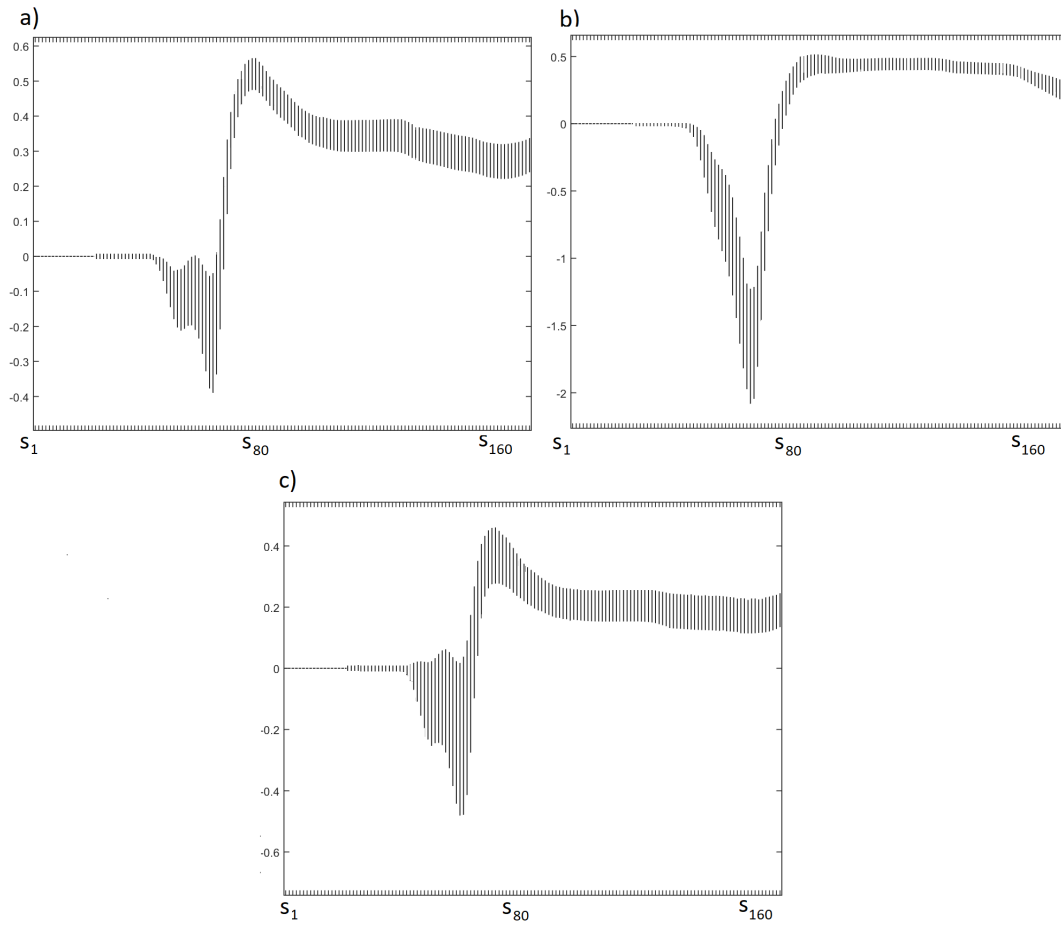


Fig. 6.24: Rangos de valores de skill para periodos de 15 minutos en Mejillones. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 180 minutos tras iniciado el terremoto.

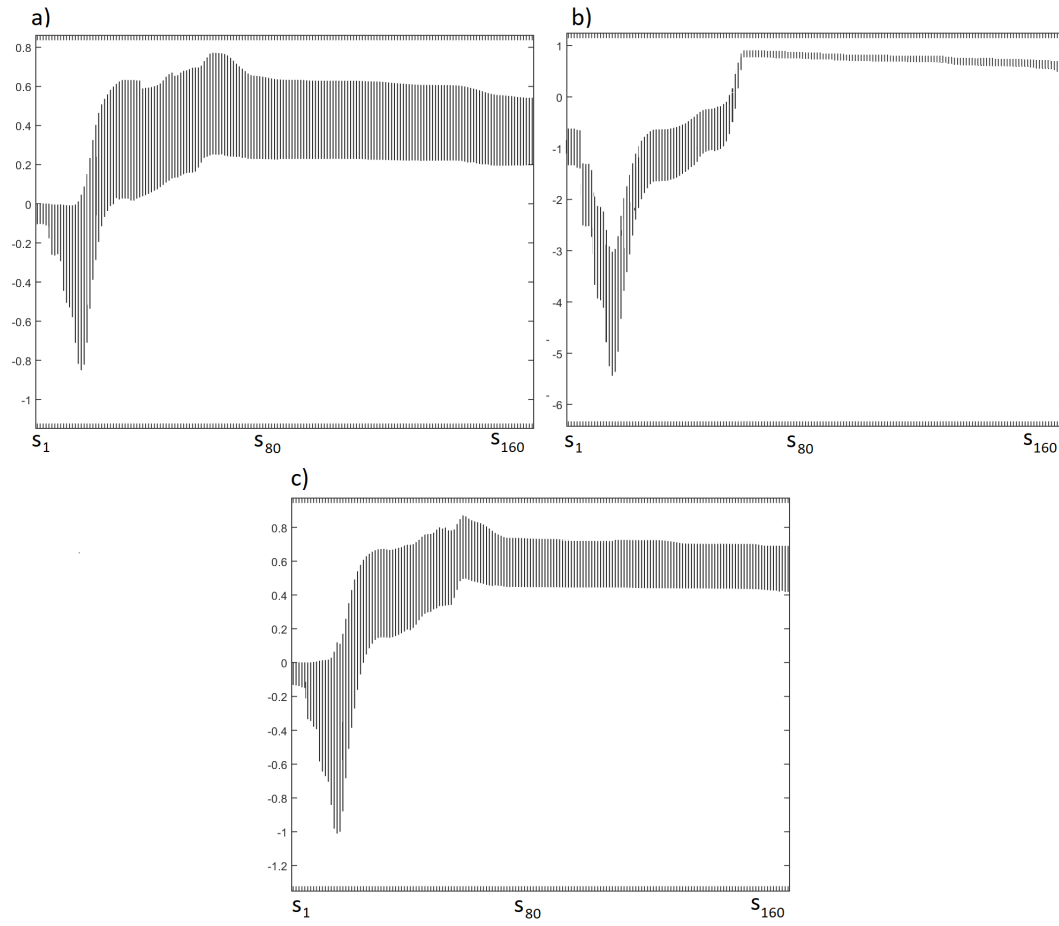


Fig. 6.25: Rangos de valores de skill para periodos de 15 minutos en DART. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 45 minutos tras iniciado el terremoto.

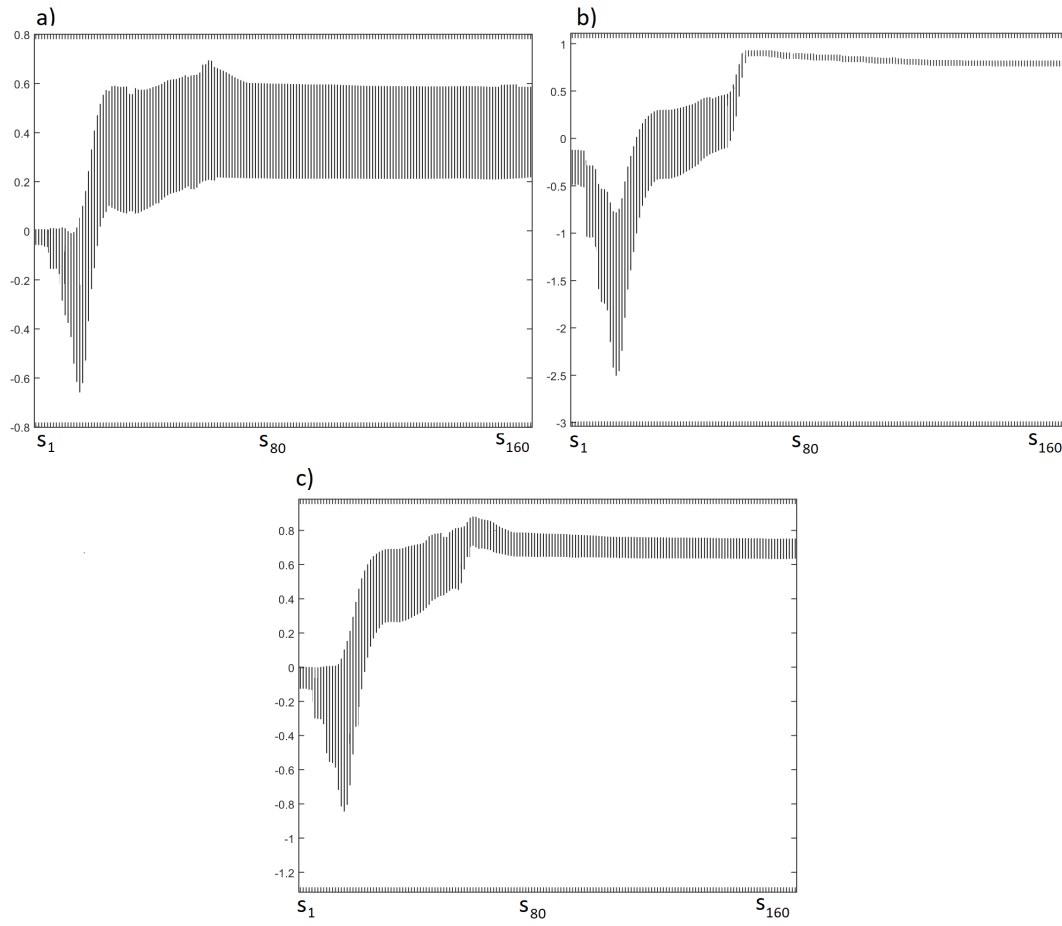


Fig. 6.26: Rangos de valores de skill para periodos de 15 minutos en DART. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 60 minutos tras iniciado el terremoto.

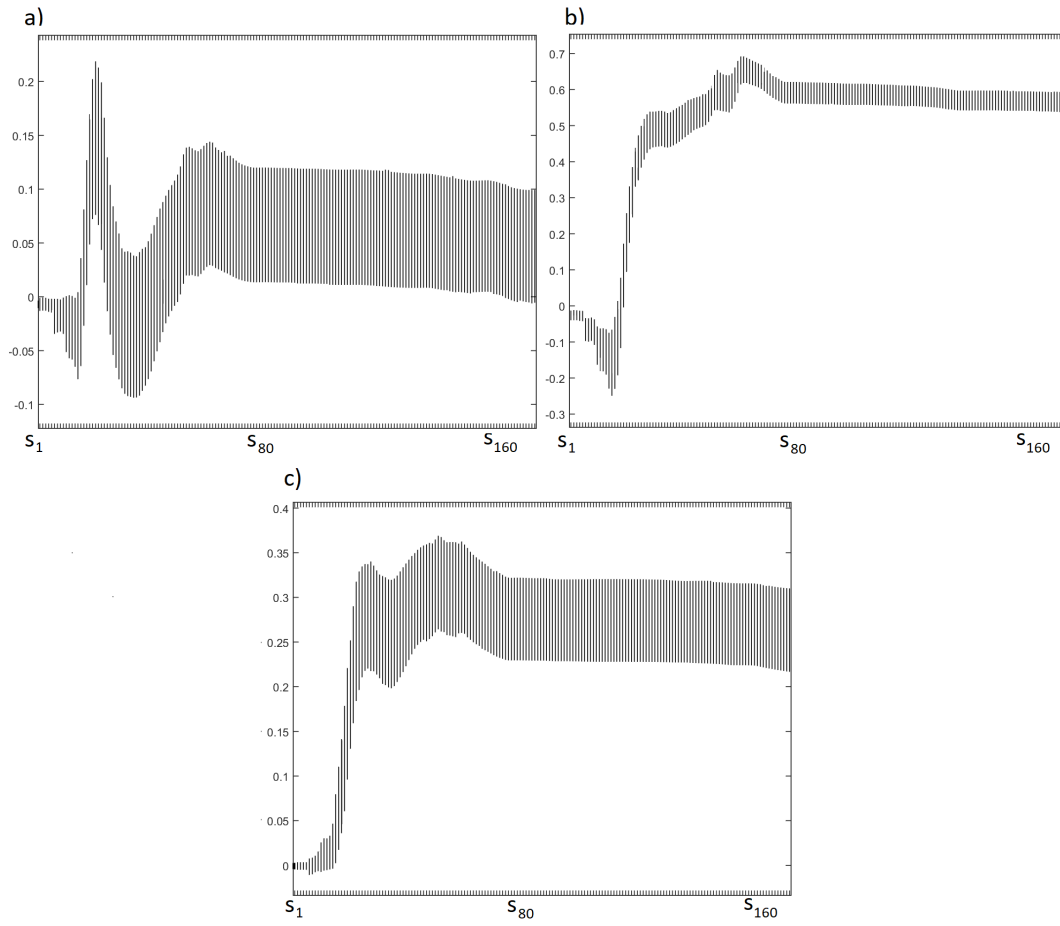


Fig. 6.27: Rangos de valores de skill para periodos de 15 minutos en Iquique. En el eje y aparecen los valores de skill y en el eje x los s_i son los valores de skill para cada periodo de 15 minutos comenzando desde el minuto i . Los rangos corresponden a diferentes escenarios estocásticos de los resultados de inversión para datos en 180 minutos tras iniciado el terremoto.

