



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

ANÁLISIS DE SEÑALES TRANSIENTES EN SERIES DE TIEMPO DE GNSS MEDIANTE INFERENCIA BAYESIANA TRANSDIMENSIONAL

Habilitación Profesional

Por: Juan Pablo Merino Merino

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la
Universidad de Concepción para optar al título profesional de Geofísico.

Agosto 2021

Concepción, Chile

**Profesores Guía: Dr. Marcos Moreno S., Dr. Roberto Benavente
B.**

**Comisión: Dr. Marcos Moreno S., Dr. Roberto Benavente B., Dr.
Klaus Bataille**

© 2021, Juan Pablo Merino Merino

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento

A mi abuelo Alberto Merino Villagrán

AGRADECIMIENTOS

En primera instancia quiero agradecer al Dr. Marcos Moreno Switt y al proyecto ACT192169 por darme la oportunidad de trabajar con ellos y financiar este trabajo. También agradecer al Dr. Roberto Benavente Bravo por su gran ayuda y guía durante todo este tiempo de tesis, como también al Dr. Klaus Bataille Bollweg por ser parte de la comisión evaluadora. Agradecer también al Dr. Felipe Donoso por compartir parte de su investigación para el desarrollo de este trabajo, como también a Sebastián Castro por su ayuda en el desarrollo de la metodología.

Quiero agradecer también a mi familia por el apoyo incondicional durante todo este proceso y en los momentos donde mas los necesité, en particular quiero darle infinitas gracias a mi abuelo por estar todo el tiempo motivándome con su característico sentido del humor a seguir adelante y cumplir mis metas.

Agradecer a todos mis amigos, tanto de la vida como de la universidad por todo lo que me han entregado y por ser un soporte esencial en mi vida y por sobre todo durante este proceso. Finalmente agradecer a Catalina Morales Yañez y Luis Bustamante Silva por todas las reuniones y discusiones que sin duda aportaron de manera significativa al desarrollo de este trabajo.

Resumen

Como consecuencia del ciclo sísmico en la zona de subducción chilena, es posible visualizar desplazamientos en la superficie terrestre gracias a series de tiempo de estaciones GNSS, siendo los desplazamientos mas visibles los asociados al periodo co-sísmico, marcados por grandes desplazamientos hacia el este. De todas maneras, en el resto de la serie de tiempo se pueden encontrar múltiples desplazamientos con distintos orígenes, entre los cuales se pueden encontrar los desplazamientos transientes, los cuales corresponderían a una variación súbita de la señal, que decaerá en el tiempo a lo largo de días hasta meses, y que según algunos autores podrían actuar como eventos precursores para grandes terremotos. En particular, para caracterizar estos desplazamientos se utilizó un modelo de trayectoria conocido, en donde ocupan una suma de 3 términos exponenciales para modelar un desplazamiento transiente, razón por la cual en este trabajo se les refiere como desplazamientos multi-transientes.

Para abordar el problema inverso y a diferencia de otras investigaciones, donde se utiliza un enfoque optimal, se recurrió a un enfoque bayesiano transdimensional, esto debido a sus propiedades asociadas a la selección de modelo y al contenido de información que presentan los datos. Tomar un enfoque transdimensional se debe a que en un principio no se conocen la cantidad de multi transientes que se puedan encontrar en la serie de tiempo, por ende tampoco se conoce la dimensión del modelo de trayectoria, luego, la dimensión de este también será una incógnita. Como la inferencia bayesiana esta basada en el teorema de Bayes, se buscará encontrar una función de densidad de probabilidad (PDF) que permita describir que tan probable es un modelo, según los datos. Luego las soluciones del problema inverso vendrán dadas por el muestreo de dicha PDF. Para este muestreo se recurrió a una técnica conocida como *reversible jump Markov Chain Montecarlo* la cual permite abordar el problema transdimensional de forma computacional. Para probar la metodología se hicieron una serie de simulaciones para series de tiempo de distintas características y complejidades obteniendo muy buenos resultados para los casos ideales (i.e con ruido blanco) pudiendo caracterizar de muy buena manera la serie de tiempo. Sin embargo, al momento de probar el método con datos reales, el ajuste realizado puede parecer no del todo correcto, esto asociado principalmente a una incorrecta caracterización del ruido de la

señal. No obstante, se pudieron identificar de todas formas ciertas deformaciones transientes que pudieron ser contrastadas con investigaciones previas y que además podrían ser atribuibles a fenómenos tales como un terremoto lento. Ahora bien, esta metodología aun se encuentra en desarrollo, por lo que los resultados obtenidos en este trabajo sirven como base para el desarrollo de futuras investigaciones utilizando inferencia bayesiana transdimensional.

Índice general

AGRADECIMIENTOS	I
Resumen	II
1. Introducción	1
1.1. Contexto	1
1.1.1. Hipótesis	8
1.1.2. Objetivos generales	8
1.1.3. Objetivos específicos	8
2. Marco Teórico	9
2.1. Modelo de trayectoria	9
2.1.1. Selección de modelo	15
2.2. Inferencia Bayesiana	17
2.2.0.1. Inferencia bayesiana transdimensional	18
2.3. Técnicas de muestreo	20
2.3.1. Markov Chain Monte Carlo (MCMC)	20
2.3.2. Reversible jump Markov Chain Monte carlo (RjMCMC)	21
3. Metodología	24
3.1. Generalidades	24
3.2. Modificación modelo de trayectoria	26
3.3. Elección de PDFs	30
3.3.1. Función de verosimilitud (Likelihood)	30
3.3.2. Prior	32
3.3.3. Proposal	33
3.4. Cálculo de la evidencia	34
4. Resultados	36
4.1. Simulaciones	36
4.1.1. Serie de tiempo sintética para un terremoto lento	41
4.2. Análisis de datos reales	47
4.2.1. Análisis individual	49
4.2.2. Análisis en varias componentes	55
4.3. Comparación GrAtSiD	62

Índice general	v
<hr/>	
5. Análisis y Discusión	65
6. Conclusión	74
Referencias	77
Apéndices	81
A. Test	81
A1. Notación	81
A2. Greedy automatic signal decomposition (GrAtSiD) method	83

Índice de cuadros

3.2.1. Clasificación de multi transientes según su dimensión	29
4.1.1. Características de las particiones encontradas por el rjMCMC en la serie de tiempo sintética. Cabe mencionar que el hecho de que en la primera partición el valor de k sea nulo se debe a que en dicha partición solo existe contribución del término global del modelo de trayectoria	38
4.1.2. Características del modelo utilizado para generar la serie de tiempo sintética.	40
4.1.3. Características del modelo seleccionado por el rjcmc para la serie de tiempo sintética.	41
4.1.4. Características del modelo seleccionado por el rjMCMC para la componente E-O de la estación SLWS	45
4.1.5. Desplazamientos hacia el este según la simulación y el modelo escogido por el rjcmc	46
4.1.6. Desplazamientos hacia el este según la simulación y el modelo escogido por el rjcmc solo considerando los multi transientes en los tiempos 221.946689 y 596.703601	47
4.2.1. Características del modelo seleccionado por el rjMCMC para la serie de tiempo individual de la componente E-O para la estación CGTC.	50
4.2.2. Características del modelo seleccionado por el rjcmc para la serie de tiempo individual de la componente E-O para la estación IQQE.	52
4.2.3. Características del modelo seleccionado por el rjcmc para la serie de tiempo individual de la componente E-O para la estación BN01.	53
4.2.4. Características del modelo seleccionado por el rjcmc para la serie de tiempo en conjunto para la componente E-O y N-S de la estación CGTC.	57
4.2.5. Características del modelo seleccionado por el rjcmc para la serie de tiempo en conjunto para la componente E-O y N-S de la estación IQQE.	59
4.2.6. Características del modelo seleccionado por el rjcmc para la serie de tiempo en conjunto para la componente E-O y N-S de la estación BN01.	62

Índice de figuras

1.1.1. Esquema ideal para el desplazamiento durante el ciclo sísmico con sus tres fases principales, intersísmico, cosísmico y post sísmico (Reddy et al., 2011).	2
2.1.1. Serie de tiempo de la componente vertical de estación GNSS MTCO ubicada en el Amazonas, Brasil. Figura extraída de las bases de datos del laboratorio geodético de Nevada (http://geodesy.unr.edu/NGLStationPages/stations/MTCO.sta)	11
2.1.2. Representación gráfica de tres exponenciales con amplitud aleatoria y constantes de decaimiento de distinto orden. La curva de color negro representa la suma de estas tres exponenciales, la cual vendría a representar la señal multi transiente. (Bedford and Bevis, 2018)	13
2.1.3. Serie de tiempo de tres componentes para la estación SANT ubicada en las cercanías de la comuna de Colina, Región Metropolitana, Chile. Figura extraída de las bases de datos del laboratorio geodético de Nevada (http://geodesy.unr.edu/NGLStationPages/stations/SANT.sta)	14
2.1.4. Regresión de mínimos cuadrados, con distintos modelos polinomiales para un set de datos sintéticos generado a partir de un polinomio de primer grado	16
2.2.1. Ilustración del principio natural de parsimonia, explicado en función del poder predictivo de dos hipótesis \mathcal{H}_1 y \mathcal{H}_2 (Sambridge et al., 2006)	19
2.3.1. Esquema ilustrativo del rjMCMC (Sambridge et al., 2006)	23
3.1.1. Serie de tiempo esquemática para ejemplificar el concepto de partición	25
3.2.1. Contribución de las componentes transiente y Heaviside posterior al tiempo de inicio $t_r = 500$	27
3.2.2. Comparación entre el comportamiento de una función Heavisde y una función exponencial con constante de decaimiento lo suficientemente pequeña	28

4.1.1.Resultados del muestreo hecho por el rjMCMC. La primera figura muestra los datos junto con la predicción hecha con el modelo seleccionado junto con el modelo inicial usado para generar los datos. Los histogramas muestran en porcentaje los criterios de selección de modelo por parte del rjMCMC, como también de manera analítica (calculado de evidencia) respectivamente.	37
4.1.2.Resultado del muestreo realizado por el rjMCMC comparados con el calculo de evidencia. La curva naranja muestra la predicción hecha con el medelo seleccionado por el muestreo. Los dos histogramas del centro muestran la ubicación de una partición según lo encontrado por el rjMCMC (primer histograma) y el calculo analítico de la evidencia (segundo histograma)	38
4.1.3.Serie de tiempo sintética con 5 particiones de multi transientes de distinta dimensión	39
4.1.4.Resultado de inversión para ejemplo sintético de varias particiones, la figura de mas arriba muestra la predicción hecha en base al modelo seleccionado. Por otra parte la figura del medio muestra los histogramas	40
4.1.5.Campo de deslizamiento para el primer terremoto lento simulado, con deslizamiento máximo de 0.05m y 100 días de duración	42
4.1.6.Serie de tiempo sintética para cada componente de la estación SLWS con ruido realista simulada para una red geodética de datos en base a la metodología descrita en Carr Agnew (2013)	43
4.1.7.Serie de tiempo para la componente E-O de la estación SLWS. La curva naranja corresponde a la predicción hecha por el modelo seleccionado por el rjMCMC y las lineas horizontales grises marcan el tiempo de inicio de los multi transientes encontrados. La linea vertical roja indica el inicio de un multi transiente de dimensión 4.	44
4.1.8.Histogramas resultantes del muestreo realizado por el rjMCMC. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo histograma indica la cantidad de veces que se muestreó un número determinado de particiones.	44
4.1.9.Comparación entre desplazamientos hacia el este según lo hecho en la simulación junto con lo mostrado mediante el modelo seleccionado por el rjMCMC.	46
4.1.10Comparación entre desplazamientos hacia el este según lo hecho en la simulación y los dos multi-transientes principales encontrados por el rjMCMC	47
4.2.1.Distribución espacial de estaciones GNSS utilizadas para el análisis de datos reales	48

4.2.2.Serie de tiempo para la componente E-O de la estación CGTC. La curva naranja corresponde a la predicción hecha con el modelo seleccionado por el muestreo del rjMCMC. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un Heaviside puro.	49
4.2.3.Histogramas resultantes del muestreo realizado por el rjMCMC para la componente E-O de la estación IQQE. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones. . .	50
4.2.4.Serie de tiempo para la componente E-O de la estación IQQE. La curva naranja corresponde a la predicción hecha con el modelo seleccionado por el muestreo del rjMCMC. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un Heaviside puro.	51
4.2.5.Histogramas resultantes del muestreo realizado por el rjMCMC para la componente E-O de la estación IQQE. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones. . .	51
4.2.6.Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado por el muestreo del rjMCMC. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un Heaviside puro.	52
4.2.7.Histogramas resultantes del muestreo realizado por el rjMCMC para la componente E-O de la estación BN01. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones. . .	53
4.2.8.Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado, siendo el con menor misfit dentro de todos los modelos de la mayor dimensión muestreada. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un heaviside puro.	55

4.2.9. Histogramas resultantes del muestreo en conjunto realizado por el rjMCMC para la componente E-O y N-S de la estación CGTC. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.	56
4.2.10. Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado, siendo el con menor misfit dentro de todos los modelos de la mayor dimensión muestreada. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un heaviside puro.	58
4.2.11. Histogramas resultantes del muestreo en conjunto realizado por el rjMCMC para la componente E-O y N-S de la estación IQQE. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.	59
4.2.12. Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado, siendo el con menor misfit dentro de todos los modelos de la mayor dimensión muestreada. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un heaviside puro.	60
4.2.13. Histogramas resultantes del muestreo en conjunto realizado por el rjMCMC para la componente E-O y N-S de la estación BN01. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.	61
4.3.1. Comparación entre términos locales y globales de la predicción obtenida mediante el muestreo del rjMCMC con la obtenida mediante GrAtSiD para la componente E-O de la estación CGTC	63
4.3.2. Comparación de la predicción tanto para el modelo seleccionado por el rjMCMC como para la predicción obtenida por GrAtSiD para la componente E-O de la estación CGTC	64
5.0.1. Residuales para componentes E-O y para componente N-S de la estación CGTC a partir de los modelo seleccionados. Las figuras de la segunda columna corresponden a una auto correlación del residual obtenido	66

5.0.2. Residuales para componentes E-O y para componente N-S de la estación IQQE a partir de los modelos seleccionados. Las figuras de la segunda columna corresponden a una auto correlación del residual obtenido	67
5.0.3. Residuales para componentes E-O y para componente N-S de la estación BN01 a partir de los modelos seleccionados. Las figuras de la segunda columna corresponden a una auto correlación del residual obtenido	67
5.0.4. Multi transientes detectados entre enero y febrero del año 2014 para las estaciones CGTC, IQQE y BN01	69
5.0.5. Multi transientes detectados en el mes de marzo/2014 para las estaciones CGTC, IQQE y BN01	70
5.0.6. Multi transientes detectados entre junio y finales de junio de 2013 para las estaciones CGTC, IQQE y BN01	71
5.0.7. Multi transientes detectados al rededor del mes de octubre de 2013 para las componentes N-S de las estaciones IQQE y BN01	72
5.0.8. Multi transientes detectados al rededor del mes de abril de 2013 para las estaciones CGTC, IQQE y BN01	72

Capítulo 1

Introducción

1.1. Contexto

Gran parte del territorio chileno se encuentra ubicado geográficamente frente a la convergencia de dos placas tectónicas, la placa de Nazca y la placa Sudamericana, cuya velocidad de convergencia es de las más altas del planeta, tomando valores entre 60-70 mm/yr (Angermann et al. 1999; Sdrolias and Müller 2006). Por esta razón, Chile un país constantemente afectado por diferentes fenómenos tales como terremotos, tsunamis o erupciones volcánicas.

Estas zonas de convergencia son llamadas zonas de subducción, donde una placa tectónica oceánica se mete por debajo de una placa tectónica continental y son además los lugares donde se concentran los mayores terremotos registrados (Rikitake 1976; Ruff 1996; Bilek and Lay 2018). En el contacto entre ambas placas, ocurre un fenómeno conocido como ciclo sísmico, el cual es un proceso repetitivo durante el cual se acumula energía por décadas a siglos, y se libera repentinamente por un terremoto. Este fenómeno tiene además una contribución constante a la deformación de la corteza terrestre, la cual se traduce en desplazamientos en la superficie.

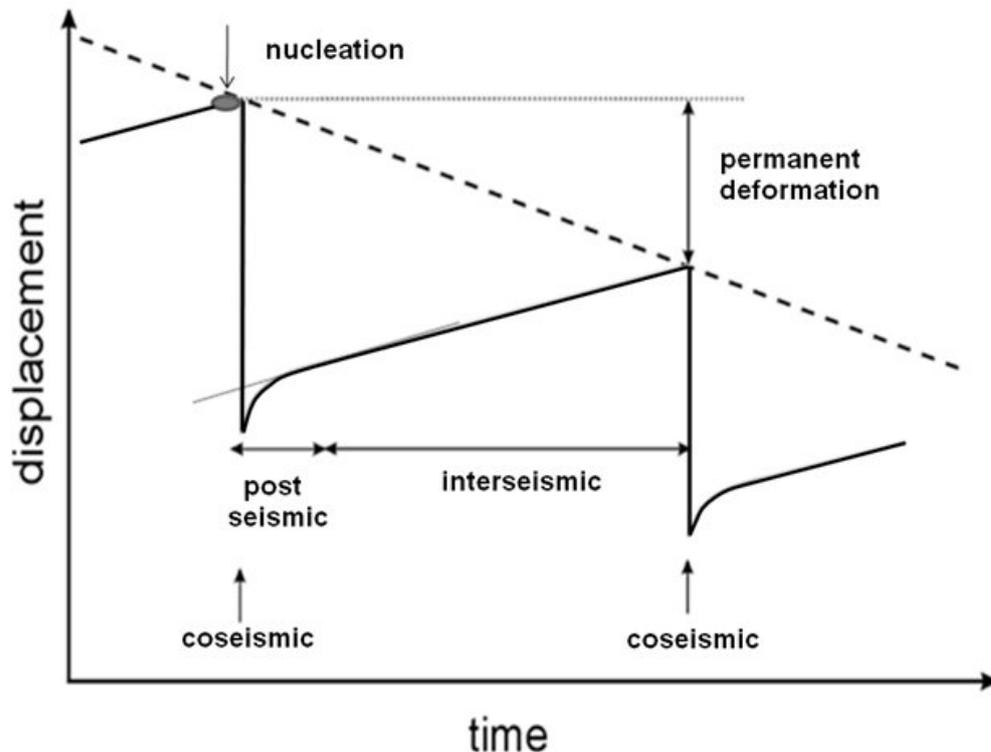


Figura 1.1.1: Esquema ideal para el desplazamiento durante el ciclo sísmico con sus tres fases principales, intersísmico, cosísmico y post sísmico (Reddy et al., 2011).

El ciclo sísmico consta de tres periodos principales (ver fig 1.1.1), el primero es el periodo intersísmico, donde ambas placas tectónicas se encuentran acopladas, acumulando tensión elástica en la zona de contacto, produciendo desplazamientos sostenidos en el tiempo en la superficie terrestre (Govers et al., 2018) y su duración puede abarcar de décadas a siglos. Posterior al periodo intersísmico, se encuentra el periodo cosísmico, donde la sección bloqueada se desacopla (i.e ocurre un terremoto) produciendo un deslizamiento libre, provocando así, una extensión rápida de la placa continental, generando grandes desplazamientos en un periodo corto de tiempo, dicho periodo es el de menor duración de los cuatro, abarcando tiempos del orden de minutos. El último periodo es el postsísmico, el cual puede dividirse entre temprano y tardío. En el postsísmico temprano ocurre un deslizamiento tardío (*afterslip*) (Lin et al., 2013) que sigue al terremoto en el mismo plano de falla, generando pequeños desplazamientos posterior a la ocurrencia del sismo. En este periodo también comienza el re acople entre las placas. Finalmente se encuentra el periodo postsísmico tardío donde predominan

desplazamientos asociados a la respuesta del material viscoso al rededor de la falla como respuesta al cambio de tensión producido por el terremoto (Freed et al. 2007; Pollitz et al. 2000). Con el tiempo, la señal postsísmica decaerá, y las observaciones se aproximarán nuevamente a un patrón intersísmico. La duración de estos periodos es muy relativa dependiendo de la ubicación geográfica y de la magnitud de los eventos previos, sin embargo los tiempos de duración mencionados hacen referencia a lo observado para el caso chileno.

El ciclo sísmico tiene una constante influencia en los desplazamientos observados en la superficie terrestre, ya sea por la ocurrencia de un terremoto, u otros procesos relacionados a este. Sin embargo, estas no son las únicas fuentes que pueden contribuir a la deformación de la corteza terrestre, otros fenómenos que pueden tener algún aporte son por ejemplo, el ajuste isostático glacial (Khan et al. 2010; Jiang et al. 2010), o la carga hidrológica (Heki 2001). Tomando en cuenta todo lo anterior, es interesante preguntarse si es posible caracterizar estos desplazamientos en el tiempo y estudiar los eventos que hacen que la superficie terrestre se mueva de la forma que lo hace. En particular durante el periodo intersísmico, con el objetivo de caracterizar ciertos fenómenos que puedan actuar como señales precursoras a un gran terremoto.

Lo anterior es posible gracias al Sistema Global de Navegación por Satélite (GNSS), el cual nos permite, mediante un instrumento, determinar la posición en cualquier lugar de la Tierra de un punto con cierta exactitud, pudiendo generar así series de tiempo que nos muestren la evolución temporal del desplazamiento en dicho punto de la superficie terrestre. Series de tiempo de GNSS han sido usadas para estudiar algunos de los fenómenos anteriormente mencionados y sus contribuciones a la serie de tiempo (Knowles et al. 2020; Ferreira et al. 2019a; Klein et al. 2019a), pero entre otras contribuciones que podemos encontrar, existe un tipo de señales llamadas transientes. Estas señales corresponden a una variación súbita en la amplitud de la señal, que decae en un periodo relativamente corto de tiempo (decenas de días a meses). Dichas señales pueden ser originadas por distintos fenómenos, tales como los asociados al periodo post sísmico, sin embargo algunas investigaciones sugieren que desplazamientos asociados a transientes pueden aparecer en ventanas de tiempo previas a grandes terremotos (e.g Melbourne and Webb 2002; Kumazawa et al. 2010; Bedford et al. 2020) por lo que resulta interesante poder identificar y caracterizar dichas señales en el tiempo.

La manera de poder hacer esto, es mediante la creación de un modelo de trayectoria. Con modelo de trayectoria (y modelo en general), nos referimos a una ecuación matemática que nos permita hacer una predicción del desplazamiento de un punto a lo largo del tiempo. La creación de estos modelos de trayectoria ha sido uno de los tópicos principales desde que se comenzó a desarrollar la tecnología de GNSS. El primero de estos modelos consistió solo de un modelo de velocidad constante (Michael Bevis, 2014), sin embargo, a medida que nuevas redes de estaciones GNSS se fueron desarrollando, fueron también desarrollándose nuevos modelos que permitían ajustar otros fenómenos que contribuían a la señal, y que no podían ser explicados solo con un modelo de velocidad constante. Un ejemplo de esto, es el modelo de trayectoria propuesto por Bedford and Bevis (2018), el cual tiene la forma:

$$\begin{aligned} \mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}(t - t_r) + \sum_{j=1}^{n_j} b_j \mathbf{H}(t - t_j) + \sum_{k=1}^{n_k} [s_k \sin(\omega_k t) + c_k \cos(\omega_k t)] \\ + \sum_{r=1}^{n_r} \sum_{i=1}^{n_i} [A_i (1 - e^{-(t-t_r)^{1/T_i}})] + \xi(t) \end{aligned} \quad (1.1.1)$$

Donde cada término de la ecuación corresponde a la contribución de los distintos fenómenos previamente mencionados. De forma particular, los dos primeros términos hacen referencia al movimiento secular o de velocidad constante, el segundo se encarga de ajustar saltos repentinos en la señal, el tercero corresponde a movimientos relacionados con la carga estacional y por ultimo se encuentra el termino encargado de describir los desplazamientos asociados a transientes. Cabe mencionar que esta ecuación será ligeramente modificada (ver sección 3.2) y sus distintas componentes también serán revisadas en mayor detalle en secciones posteriores.

Es importante notar que la ecuación 2.1.4 permite caracterizar la señal por medio de un conjunto de parámetros asociados a los distintos términos de la ecuación (e.g b_j , s_k , c_k , etc). Típicamente, un sistema físico o modelo, puede describirse en base a un vector \mathbf{m} cuyos valores caracterizan por completo el modelo, luego es posible describir el sistema de la forma:

$$\mathbf{d} = \mathbf{G}(\mathbf{m}), \quad (1.1.2)$$

donde \mathbf{d} corresponderá a las observaciones y $\mathbf{G}(\mathbf{m})$ corresponderá a las predicciones hechas por nuestro modelo. En particular $\mathbf{G}(\mathbf{m})$ representará la teoría física detrás del modelo (Ye et al., 2019) y será dependiente del vector \mathbf{m} . Sin embargo, si nuestro modelo es lineal, podremos escribir entonces:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \quad (1.1.3)$$

donde \mathbf{G} corresponderá a la matriz de diseño y \mathbf{m} será el vector que contenga los parámetros del modelo (b_j, s_k, c_k , etc).

El poder hacer predicciones en base a una ecuación (modelo), se conoce como problema directo, sin embargo, nosotros no conocemos a priori el valor numérico de cada parámetro presente en nuestro modelo, no obstante, sí sabemos cuanto se desplazó un punto determinado gracias a los datos entregados por estaciones GNSS. Esto nos lleva a resolver el problema inverso, que es lo realmente importante, puesto que así podremos determinar dichos parámetros y caracterizar la señales con todas sus contribuciones.

Existen muchas formas de poder resolver el problema inverso, las cuales pueden tener un enfoque bayesiano o uno optimal, es este último el que suele usarse para resolver este tipo de problemas relacionados a modelos de trayectoria (e.g Bedford and Bevis 2018). El objetivo principal de este enfoque, es encontrar un vector \mathbf{m} que contenga los valores de los parámetros del modelo mediante una función que minimice los residuales entre el modelo y las observaciones. El tamaño de los residuales suele medirse normalmente con la norma L2, es decir, mínimos cuadrados.

$$\min(|\mathbf{G}(\mathbf{m}) - \mathbf{d}|_2^2) \quad (1.1.4)$$

Ahora bien, como en geofísica los problemas suelen ser mal condicionados y además de no contar con suficiente información a priori, en este tipo de metodologías se suelen ocupar métodos de regularización que podrían incorporar un sesgo erróneo a los resultados.

Por otro lado, un enfoque bayesiano, aborda el problema inverso en términos del contenido de información que brindan nuestros datos sobre los parámetros. Tal información es codificada por medio de funciones de distribución de probabilidad

(PDF), las que se relacionan por medio del teorema de Bayes.

$$P(\mathbf{m}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{m})P(\mathbf{m})}{P(\mathbf{d})}, \quad (1.1.5)$$

donde cada termino corresponde a una función de distribución de probabilidad (PDF). Luego por consecuencia, toda información debe expresarse en forma de una PDF, tanto para las incógnitas (i.e parámetros del modelo, *prior distribution*) como para los datos, específicamente las discrepancias entre el modelo y las observaciones (*likelihood distribution*) (Bretthorst, 1988). En nuestro caso, además de los parámetros del modelo, la dimensión de este también será una incógnita, puesto que a priori no sabemos cuantos transientes habrán en nuestra serie de tiempo, ni la complejidad de estos, haciendo que el número de parámetros varíe según el caso, esto hace que nuestro problema tome un carácter transdimensional

El objetivo principal entonces, será encontrar una PDF que nos describa la probabilidad de tener un modelo \mathbf{m} dado nuestros datos \mathbf{d} (*posterior PDF*). Lo interesante de abordar el problema desde una perspectiva bayesiana, es que serán los mismos datos los que nos entreguen toda la información que necesitamos sobre el modelo, puesto que no es necesario que la información a priori sea del todo precisa e informativa (Bretthorst, 1988).

Si bien, el objetivo del enfoque bayesiano en nuestro caso es encontrar una PDF que nos describa la probabilidad de que nuestros datos hayan sido generados por un cierto modelo, nuestra solución final no será esta. Luego, la solución del problema inverso estará expresada en función de las muestras de la PDF obtenida. Esto significa entonces, que la solución de un problema inverso no será solo un modelo, sino una colección de ellos que son consistentes con los datos y con la información a priori, a diferencia del enfoque optimal, donde encontramos un solo modelo que minimice los residuales.

Como la solución de nuestro problema vendrá expresada por el muestreo de nuestro *posterior* es necesario recurrir con métodos de muestreo para poder encontrar las soluciones del problema inverso. Entre estas técnicas, podemos encontrar la cadena de Markov Monte Carlo (MCMC), la cual, a grandes rasgos consiste en una cadena de Markov que parte desde un modelo inicial que se compara con uno nuevo propuesto, luego un modelo será seleccionado por sobre otro mediante la

regla de Metropolis-Hasting (Green, 1995). Esta técnica es ampliamente usada en geo ciencias (Sambridge et al., 2006), y es bastante efectiva para problemas donde la dimensión del modelo es conocida, sin embargo, en nuestro caso esto no se cumple. Como alternativa a esto, se recurrirá a una modificación de esta técnica, llamada *reversible jump Markov Chain Monte Carlo* (rjMCMC) (Green, 1995) la cual sigue la misma lógica del MCMC pero modifica la regla de Metropolis-Hasting para la aceptación de un modelo por sobre otro tomando en cuenta el cambio de dimensión.

Ahora bien, debido al gran número de modelos consistentes con nuestros datos resultantes del muestreo del posterior, es sumamente importante contar con herramientas de selección de modelos para poder elegir el mas adecuado dentro de todas las opciones. El rjMCMC cuenta con herramientas inherentes que permiten hacer selección de modelo (siguiendo el principio de parsimonia (Sambridge et al., 2006)), sin embargo, es interesante poder comparar esto con otras alternativas, como lo es el calculo de la evidencia. La evidencia corresponde al denominador de la ecuación(1.1.5) la cual puede interpretarse como cuan bien el modelo propuesto es capaz de describir las observaciones (Sambridge et al., 2006). Consecuentemente se favorecerá entonces el modelo que presente una mayor evidencia.

Como se ha estado mencionado, para este trabajo se utilizará un enfoque bayesiano para resolver el problema inverso, sin embargo, este enfoque tiene algunos detractores, que indican que las propiedades físicas de la Tierra no pueden ser tratadas como variables aleatorias, puesto que existe solo una Tierra. Otro motivo es la gran dependencia del resultado final con la elección de un prior, puesto que pueden entrar en juego información que no tenga que ver con el problema. Sin embargo, tanto el enfoque bayesiano como el optimal suelen entregar resultados similares (Sambridge et al., 2006), pero a diferencial del enfoque optimal, el enfoque bayesiano entrega una gran cantidad de soluciones. De este modo, la caracterización de la incertidumbre es un proceso natural en inferencia bayesiana. Ahora bien, tomando en cuenta lo anterior, resulta interesante comparar los resultados de ambas metodologías frente a un mismo problema, es por esto que en este trabajo también se buscará comparar los resultados del enfoque bayesiano, con una metodología optimal, en particular la usada por Bedford and Bevis 2018.

1.1.1. Hipótesis

- Es posible recuperar el número y la complejidad de los desplazamientos transientes a partir de observaciones GNSS, utilizando un modelo simple de trayectorias y selección bayesiana de modelos.

1.1.2. Objetivos generales

- Aplicar inferencia bayesiana transdimensional en inversiones de series de tiempo GNSS, tanto sintéticas como reales.
- Aplicar métodos de muestreo (rjMCMC) para encontrar soluciones del modelo de trayectoria
- Comparar resultados de rjMCMC con otros métodos optimales.

1.1.3. Objetivos específicos

- Encontrar tiempo de inicio y dimensión de multi transientes en series de tiempo sintéticas y reales de GNSS.
- Realizar análisis en mas de una componente.
- Identificar fenómenos menos visibles tales como terremotos lentos.
- Comparar técnicas de selección de modelo analíticas con resultados de rjMCMC

Capítulo 2

Marco Teórico

2.1. Modelo de trayectoria

El sistema sistema global de navegación por satélite (Global Navigation Satellite System, GNSS) nos permite determinar la posición en cualquier lugar de la Tierra de un objeto con cierta exactitud dependiendo del instrumento. En particular, para el monitoreo de desplazamientos en la superficie terrestre, se suelen utilizar estaciones continuas que entregan constantemente datos de su posición, pudiendo así tener una idea general de como se está moviendo la corteza terrestre. Sin embargo, es necesario tener algún modelo que nos permita poder caracterizar la posición de las estaciones geodésicas en función del tiempo y así proporcionar un medio para predecir la posición de una estación como también para poder analizar las distintas contribuciones que hicieron que la señal tuviera dicha forma. Esto es conocido como modelo de trayectoria.

A lo largo del tiempo y a medida que se han ido incrementando los conocimientos en el área, se han ido implementando distintos modelos de trayectoria yendo desde lo mas simple a modelos un poco mas complejos (ver [Ferreira et al. \(2019b\)](#); [Michael Bevis \(2014\)](#); [Bedford and Bevis \(2018\)](#); [Klein et al. \(2019b\)](#); [Sobrero et al. \(2020\)](#)). El primero de estos modelos, y el más simple para una tierra dinámica, considera un modelo de velocidad constante (CVM, [Michael Bevis \(2014\)](#)) en donde cada coordenada cartesiana geocéntrica para cada estación es expresada como una función lineal en el tiempo según el vector.

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}(t - t_r) \quad (2.1.1)$$

Donde \mathbf{x}_0 corresponde a la posición inicial, $\mathbf{v}(t)$ velocidad de la estación y t_r el tiempo de referencia. Esta velocidad viene del hecho que la estación al estar puesta en superficie, sobre una placa tectónica, ésta se moverá acorde lo haga dicha placa. Este desplazamiento producto de una velocidad constante es también llamado movimiento secular.

Conforme fueron pasando los años, el modelo de velocidad constante fue insuficiente debido a que no se podía acomodar a saltos repentinos o cambios de antenna. Una manera simple de poder modelar esto fue agregando al CVM una función escalón (Heaviside) $\mathbf{H}(t)$, quedando de la siguiente forma:

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}(t - t_r) + \sum_{j=1}^{n_j} b_j \mathbf{H}(t - t_j), \quad (2.1.2)$$

donde b_j caracteriza los n_j saltos que ocurren en el tiempo t_j como un desplazamiento instantáneo del vector $\mathbf{x}(t)$.

A medida que el uso y desarrollo de estaciones continuas fue proliferando, se comenzaron a encontrar oscilaciones con cierta periodicidad, las cuales eran mucho mas visibles en la componente vertical que en las horizontales y que estarían asociadas a respuestas elásticas de la litosfera producto de la carga estacional (Heki, 2001).

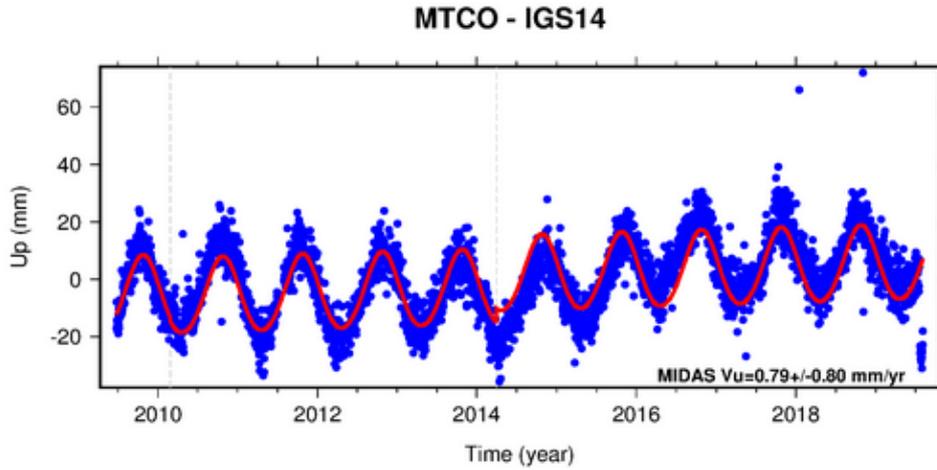


Figura 2.1.1: Serie de tiempo de la componente vertical de estación GNSS MTCO ubicada en el Amazonas, Brasil. Figura extraída de las bases de datos del laboratorio geodético de Nevada (<http://geodesy.unr.edu/NGLStationPages/stations/MTCO.sta>)

Este tipo de señales son mucho mas visibles en lugares donde existen grandes cantidades de precipitaciones, tales como lluvias o nieve. Como muestra la figura 2.1.1 se aprecia una importante variación sinusoidal en la amplitud de la componente vertical con un periodo relativamente constante. En particular para la estación MTCO, este ciclo está regulado por las variaciones estacionales en las masas de agua de la cuenca del Amazonas que es donde se encuentra. Estas oscilaciones pueden ser aproximadas mediante una serie de Fourier que consta de dos periodos anuales y dos semi anuales (Dong et al., 2002) quedando el modelo de la siguiente forma:

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}(t - t_r) + \sum_{j=1}^{n_j} b_j \mathbf{H}(t - t_j) + \sum_{k=1}^{n_k} [s_k \sin(\omega_k t) + c_k \cos(\omega_k t)], \quad (2.1.3)$$

donde n_k corresponde al número de frecuencias usadas en el modelo.

Por último, se ha registrado un gran número de saltos cosísmicos seguidos de una pronunciada señal transiente post sísmica que no puede ser aproximada con los modelos mencionados anteriormente. Las causas de dicho desplazamiento pueden ser *afterslip* en el plano de falla generado algún terremoto, relajación visco elástica de material alrededor y bajo la falla producto de variaciones en las tensiones provocadas por la ruptura (Freed et al. (2007); Pollitz et al. (2000)) o también

por *poroelastic rebound* causado por flujo de fluidos contenidos en poros que fueron liberados por perturbaciones de estrés posterior al terremoto. Estas señales transientes pueden ser modeladas mediante una función logarítmica (Michael Bevis, 2014) como también mediante a la suma de dos o mas funciones de decaimiento exponencial simples (Bedford and Bevis, 2018). En particular para este trabajo se usará el modelo que considera la señal transiente como una suma de exponenciales debido a que la suma de estas funciones alcanza un valores asintóticos lo que permite aproximar movimientos finitos de varias señales transientes (ver figura 2.1.2). Cabe mencionar también que esta suma de multi-transientes permite formar un serie de señales distintas según el signo y magnitud de su amplitud lo que lo convierte en una función muy versátil al momento de aproximar variedad de señales transientes a cualquier escala de tiempo (Bedford and Bevis, 2018). Finalmente tomando en cuenta todo lo anterior el modelo de trayectoria en el que se basará este trabajo vendrá dado por:

$$\begin{aligned} \mathbf{x}(\mathbf{t}) = \mathbf{x}_0 + \mathbf{v}(t - t_r) + \sum_{j=1}^{n_j} b_j \mathbf{H}(t - t_j) + \sum_{k=1}^{n_k} [s_k \sin(\omega_k t) + c_k \cos(\omega_k t)] \\ + \sum_{r=1}^{n_r} \sum_{i=1}^{n_i} [A_i (1 - e^{-(t-t_r)1/T_i})] + \xi(t) \end{aligned} \quad (2.1.4)$$

En donde el último término de la ecuación 2.1.4 corresponde es una combinación lineal de n_j decaimientos exponenciales con un tiempo de inicio idéntico t_r y un valor predeterminado de constantes de decaimiento T_i que para efectos de simulaciones posteriores se considerará una señal transiente como la suma de hasta 3 exponenciales cuyas constantes de decaimiento tendrán ordenes de 10^3 , 10^2 y 10^1 días respectivamente.

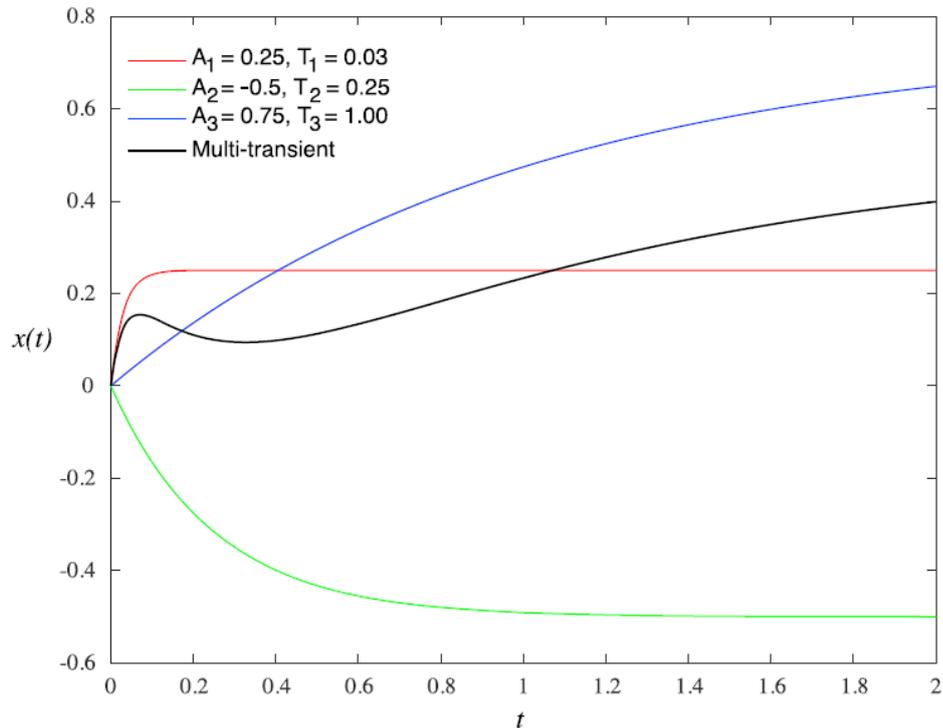
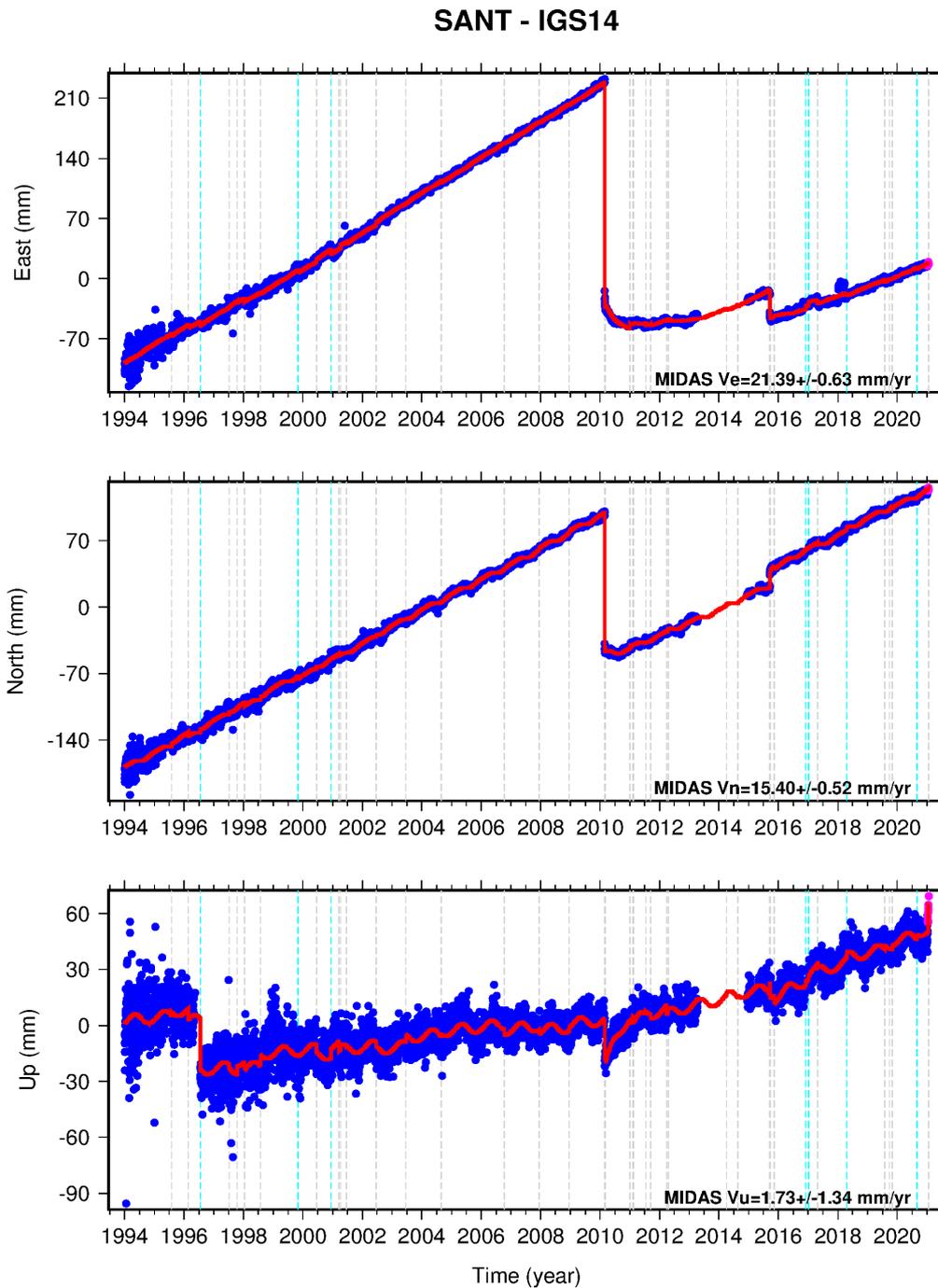


Figura 2.1.2: Representación gráfica de tres exponenciales con amplitud aleatoria y constantes de decaimiento de distinto orden. La curva de color negro representa la suma de estas tres exponenciales, la cual vendría a representar la señal multi transiente. (Bedford and Bevis, 2018)

En particular en el trabajo de Bedford and Bevis (2018) se proponen un número fijo de exponenciales, que como se mencionó anteriormente será de tres, y cada una de ellas con una constante de decaimiento predeterminada, tal como se muestra en la figura 2.1.2 donde cada exponencial tiene una constante de decaimiento de un orden de magnitud diferente esto debido a que ellos consideran que la suma de estar tres exponenciales permite de forma versátil representar un gran número de fenómenos. También al fijar el número de exponenciales con sus respectivas constantes de decaimiento se reducen el número de incógnitas en la ecuación del modelo de trayectoria, lo que es favorable mediante un enfoque optimal.



24 Hour Positions Using Final Orbits (blue) and Rapid Orbits (magenta).
 Processed by the Nevada Geodetic Laboratory.
 Plotted on 2021-Jan-23. Last data on 2021-Jan-21.

Figura 2.1.3: Serie de tiempo de tres componentes para la estación SNT ubicada en las cercanías de la comuna de Colina, Región Metropolitana, Chile. Figura extraída de las bases de datos del laboratorio geodético de Nevada (<http://geodesy.unr.edu/NGLStationPages/stations/SNT.sta>)

Cada componente anterior mencionada es posible verla explícitamente en la figura 2.1.3. En particular podemos apreciar la contribución del CVM en los primeros años desde 1996 hasta 2010 donde el desplazamiento asociado a la estación SANT consistió solo en un movimiento secular, sin embargo, a inicios del año 2010 se puede apreciar el efecto del terremoto del Maule el 27 de febrero del año 2010 el cual generó una gran deformación hacia el oeste y en menor medida (pero no menos importante) hacia el sur. Posterior a este evento y de forma casi inmediata podemos observar una deformación con cierto comportamiento exponencial asociado al desplazamiento post sísmico del terremoto del Maule, dicho desplazamiento vendría a ser representado por el multi-transiente de nuestro modelo de trayectoria. De forma adicional podemos ver también la influencia (en menor medida) del terremoto de Illapel, el 15 de septiembre de 2015 donde también se pueden ver los fenómenos mencionados. Por último, entre los años 2017 y 2018 se pueden ver pequeños saltos en los datos que no vienen seguidos de alguna deformación multi transiente, por lo que pueden estar asociados a cambios en la ubicación de la estación y pueden ser modelados con una función Heaviside.

Ahora bien, tener un modelo de trayectoria es solo el primer paso, lo realmente importante es poder a partir de las observaciones poder recuperar los parámetros de nuestro modelo para sí poder caracterizar la señal y estudiar los distintos fenómenos que hicieron que la señal tuviera dicha forma. En particular para este trabajo nos interesa encontrar y caracterizar desplazamientos transitorias a partir de observaciones GNSS tanto para datos sintéticos como para datos reales. La forma de hacer es resolviendo el problema inverso a partir de las observaciones mediante un enfoque bayesiano.

2.1.1. Selección de modelo

Como se mencionó en secciones anteriores, se entiende por modelo a una ecuación matemática que nos permite hacer una predicción en base a alguna variable, como por ejemplo el tiempo (e.g (2.1.4)). Por otra parte, la dimensión del modelo hará referencia a la cantidad de parámetros que tendrá dicho modelo.

Para ejemplificar lo anterior, supongamos un modelo polinomial de primer grado, de la forma:

$$y(x) = Ax + B \tag{2.1.5}$$

Luego, el modelo (2.1.5) tendrá dimensión dos, debido a que tiene dos parámetros (A y B). Si en base a este modelo generamos una serie de predicciones con valores aleatorios para sus parámetros y posteriormente le agregamos ruido blanco, tendremos un conjunto de puntos los cuales nos servirán como datos sintéticos (ver figura 2.1.4). Realizando una regresión de mínimos cuadrados a nuestros datos sintéticos, estimando los parámetros de modelos polinomiales de distinto grado de la forma:

$$\mathbf{m} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d}, \quad (2.1.6)$$

donde \mathbf{m} será el vector que contenga los valores de los parámetros estimados, \mathbf{G} corresponderá a la matriz de diseño del modelo y \mathbf{d} serán los datos sintéticos.

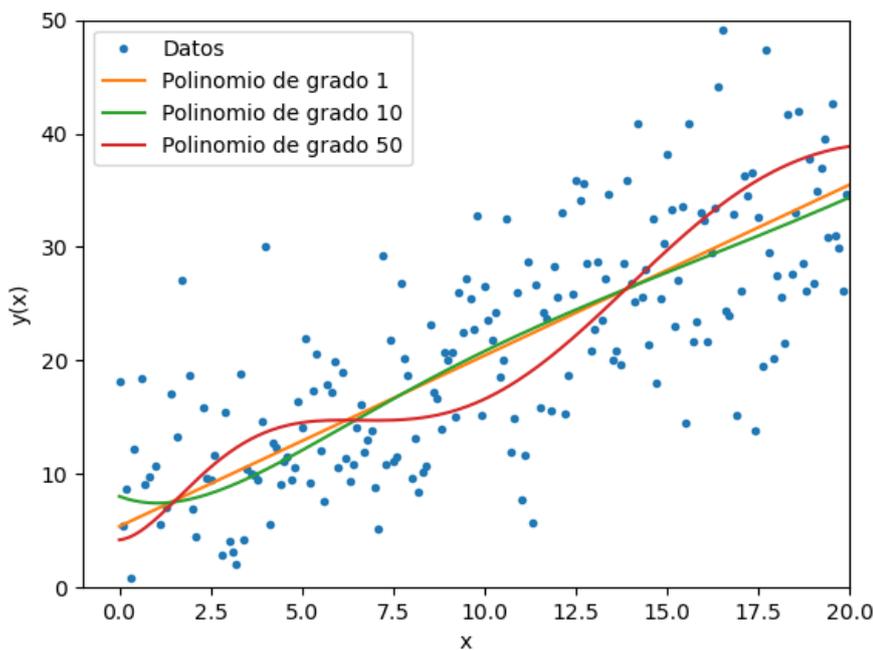


Figura 2.1.4: Regresión de mínimos cuadrados, con distintos modelos polinomiales para un set de datos sintéticos generado a partir de un polinomio de primer grado

Como los datos sintéticos fueron generados a partir de un polinomio de primer grado, resulta lógico pensar que el polinomio que mejor se ajuste a dichos datos corresponderá a un polinomio de primer grado. Sin embargo, en la figura (fig 2.1.4) podemos ver que mientras mas complejo el modelo, este aparentemente

se ajustará de mejor manera a los datos, puesto que la curva ajustada tratará de pasar por una mayor cantidad de puntos. El problema de esto, es que para modelos demasiado complejos, la curva ajustada tenderá a ajustar el ruido, luego el resultado final terminará incorporando información al modelo que nunca estuvo. Considerando lo anterior, resulta sumamente importante contar con técnicas de selección de modelos, con el objetivo de encontrar el modelo que mejor soporte a los datos, sin necesidad de sobre ajustar el ruido. Para este caso problema, la inferencia bayesiana cuenta con herramientas inherentes que permiten seleccionar un modelo, sin embargo, para ciertos casos específicos, también existen expresiones analíticas que permiten hacerlo, tal como el calculo de la evidencia.

2.2. Inferencia Bayesiana

La teoría física nos permite hacer predicciones, es decir, dado un cierto modelo nosotros podemos tener una predicción para alguna cierta medida (e.g tiempo, temperatura, velocidad, etc.). Haciendo la analogía a nuestro caso, podemos decir que dado un cierto modelo de trayectoria, podemos predecir la posición de una estación GNSS en función del tiempo. Esto es llamado un problema directo, sin embargo, lo que nosotros buscamos es que a partir de una predicción, por ejemplo desplazamiento, podamos inferir los valores de los parámetros de nuestro modelo y caracterizarlo, es decir, resolver el problema inverso. Si consideramos un enfoque bayesiano para resolver el problema inverso, tenemos que el estado del conocimiento sobre un grupo de incógnitas (i.e, parámetros de un modelo) es representado por funciones de probabilidad, es decir, toda información conocida sobre el modelo a priori (*prior distribution*) y la que nos entregan las observaciones (*likelihood distribution*) se expresan por medio de funciones de distribución de probabilidad (probability density function, PDF). Como su nombre indica, la inferencia bayesiana se centra en el teorema de Bayes, el cual establece que:

$$P(\mathbf{m}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{m})P(\mathbf{m})}{P(\mathbf{d})}, \quad (2.2.1)$$

donde $P(\mathbf{m}|\mathbf{d})$ es la PDF que buscamos encontrar (*posterior*), la cual nos describirá la probabilidad de tener un modelo \mathbf{m} , dado un set de datos \mathbf{d} .

Analizando el resto de la ecuación, tenemos que el término $P(\mathbf{d}|\mathbf{m})$ corresponde a la función de verosimilitud o *likelihood* la cual, por otro lado, describe la

probabilidad de que nuestros datos \mathbf{d} hayan sido originados por un modelo caracterizado por un vector de parámetros \mathbf{m} . En la práctica, la función de verosimilitud mide la probabilidad de que las discrepancias entre las observaciones y las predicciones del modelo sean producto solo de ruido aleatorio y pueden ser descritas por una PDF conocida (Sambridge et al., 2013).

El siguiente término $P(\mathbf{m})$ corresponde a la función de probabilidad *prior*, la cual se puede interpretar como el estado de conocimiento previo acerca del modelo que estamos buscando, es decir, nos indica la probabilidad que tiene cada parametrización del modelo descrita por nuestro vector \mathbf{m} previo a cualquier análisis. La elección del prior es una de las partes más controversiales, puesto que todos los resultados de la inversión dependen de dicho prior y deben ser interpretados acorde a su definición.

Finalmente, el término $P(\mathbf{d})$ de la ecuación 2.2.1 se conoce como evidencia, pero en muchos casos es sumamente complicado de calcular, por lo que suele ignorarse simplemente considerándolo como una constante de normalización. Sin embargo este término puede cobrar relevancia en selección de modelos (Sambridge et al., 2006).

2.2.0.1. Inferencia bayesiana transdimensional

Ahora bien, si dos modelos igualmente competentes van a ser testeados frente a un set de datos es necesario compararlos. Llamaremos una hipótesis \mathcal{H} a cualquier alternativa de modelo que vaya a ser ajustado a los datos (por ejemplo, dos modelos de distinta dimensión). Entonces, tomando en cuenta una hipótesis cualquiera, el teorema de Bayes tendrá la forma:

$$p(\mathbf{m}|\mathbf{d}, \mathcal{H}) = \frac{p(\mathbf{d}|\mathbf{m}, \mathcal{H})p(\mathbf{m}|\mathcal{H})}{p(\mathbf{d}|\mathcal{H})}. \quad (2.2.2)$$

Es en esta parte donde toma relevancia la evidencia en el problema de selección de modelos, puesto que frente a dos hipótesis, igual de competentes, es necesario escoger una por sobre la otra. En esta línea, la inferencia bayesiana transdimensional sigue naturalmente el principio de parsimonia (Sambridge et al., 2006), es decir, la hipótesis más simple será favorecida si es que se tiene el mismo nivel de ajuste entre dos teorías. Como las PDF se encuentran normalizadas, tendremos que el área bajo la curva entre dos hipótesis es la misma, variando

solo el área de cobertura de cada una. Si consideramos dos hipótesis \mathcal{H}_1 y \mathcal{H}_2 con \mathcal{H}_2 mas compleja que \mathcal{H}_1 , tendremos según la figura 2.2.1 que \mathcal{H}_2 tendrá un poder predictivo para una mayor cantidad de datos, sin embargo, si evaluamos ambas hipótesis en un intervalo mas acotado, donde ambas teorías tengan efecto, tendremos que la hipótesis \mathcal{H}_1 presentará una mayor evidencia que \mathcal{H}_2 y por ende, sera favorecida por sobre esta

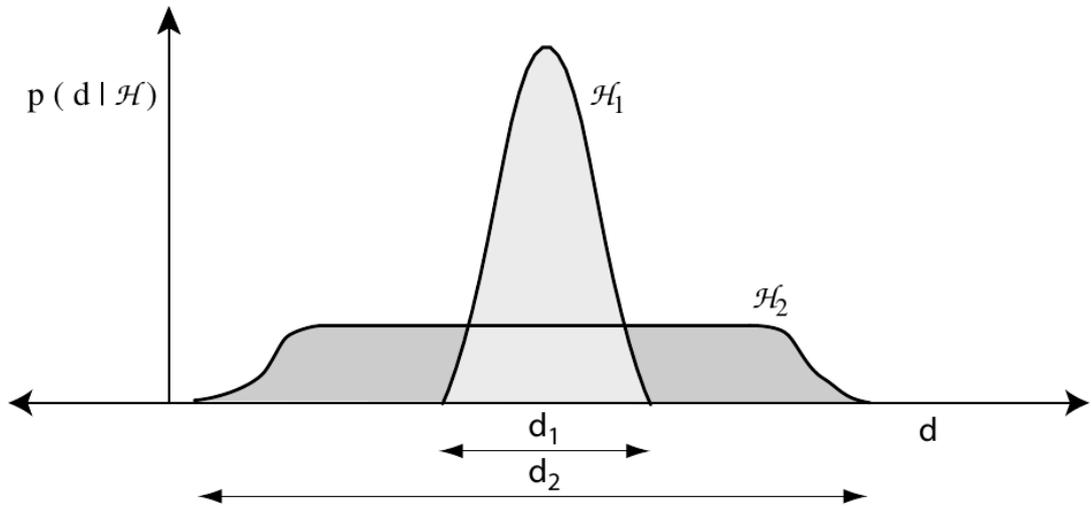


Figura 2.2.1: Ilustración del principio natural de parsimonia, explicado en función del poder predictivo de dos hipótesis \mathcal{H}_1 y \mathcal{H}_2 (Sambridge et al., 2006)

De forma particular la evidencia, según 2.2.2 vendrá dada por:

$$p(\mathbf{d}|\mathcal{H}) = \int p(\mathbf{d}|\mathbf{m}, \mathcal{H})p(\mathbf{m}|\mathcal{H})d\mathbf{m}. \quad (2.2.3)$$

Un problema transdimensional será entonces donde la dimensión k del vector \mathbf{m} que describe los parámetros de un modelo también es una incógnita. Un ejemplo de este tipo de problemas vendría ser el que estamos abordando en este trabajo, ya que al no conocer a priori la cantidad de saltos o multi transientes que pueda llegar a tener nuestra señal no podemos establecer el tamaño de nuestro vector \mathbf{m} , el cual dependerá de la cantidad de parámetros presentes en nuestro modelo. Reescribiendo la ecuación 2.2.2 considerando que la hipótesis es que nuestro modelo tiene una dimensión k , tendremos entonces:

$$p(\mathbf{m}|\mathbf{d}, k) = \frac{p(\mathbf{d}|\mathbf{m}, k)p(\mathbf{m}|k)}{p(\mathbf{d}|k)}, \quad (2.2.4)$$

Este enfoque transdimensional presenta una serie de ventajas, entre las más importantes se destaca que sigue el principio de parsimonia como se mencionó anteriormente, donde las soluciones más simples serán favorecidas. Otra propiedad importante es que no es necesario contar con alguna técnica de suavización o regularización para encontrar soluciones, ya que las PDF utilizadas serán escogidas según el contenido de información de los datos. También es común que en problemas mal condicionados no se cuente con información suficiente a prior del modelo, lo que en este enfoque no resulta un problema importante, debido que se busca tener un *prior* PDF poco informativa con el objetivo de disminuir sesgos.

Ahora bien, para resolver nuestro problema no basta solo con obtener la PDF *a posteriori* tanto para problemas de dimensión fija como para problemas de dimensión variable, sino que es necesario aplicar técnicas de muestreo a nuestra PDF para poder obtener los parámetros de nuestro modelo y poder finalmente caracterizarlo, lo cual en muchos casos es bastante complejo, sin embargo, existen formas hacerlo, entre ellas el Markov chain Monte Carlo (MCMC) para problemas de dimensión fija y reversible jump Markov chain Monte Carlo (rjMCMC) para problemas transdimensionales.

2.3. Técnicas de muestreo

2.3.1. Markov Chain Monte Carlo (MCMC)

La cadena de Markov Monte Carlo (MCMC) corresponde a la técnica insignia de muestreo para una PDF arbitraria de dimensión fija y ha sido usado ampliamente en aproximaciones bayesianas en problemas inversos en geofísica ([Sambridge et al., 2006](#))

Esta técnica consiste básicamente en una cadena de Markov donde se parte de un modelo propuesto \mathbf{m} de dimensión fija k y se compara con un nuevo modelo \mathbf{m}' de igual dimensión creado a partir del modelo \mathbf{m} siguiendo una distribución de probabilidad $q(\mathbf{m}'|\mathbf{m})$ (proposal), luego el modelo \mathbf{m}' será aceptado con una probabilidad α dada por

$$\alpha = \text{Min} \left[1, \frac{p(\mathbf{m}'|k, \mathbf{d}) \cdot q(\mathbf{m}|\mathbf{m}')}{p(\mathbf{m}|k, \mathbf{d}) \cdot q(\mathbf{m}'|\mathbf{m})} \right]. \quad (2.3.1)$$

Donde $p(\mathbf{m}'|k, \mathbf{d})$ y $p(\mathbf{m}|k, \mathbf{d})$ corresponderán al *posterior* de nuestro modelo propuesto y el modelo inicial respectivamente. Cabe mencionar que en este caso \mathbf{m}' y \mathbf{m} tienen la misma dimensión, puesto que k es igual para todos los modelos propuestos. Luego, de forma mas general aplicando el teorema de Bayes a las PDF de los *posterior*, con el objetivo de dejar la ecuación en función del *likelihood*, *prior* y *proposal* nos queda:

$$\alpha = \text{Min} \left[1, \frac{l(\mathbf{d}|\mathbf{m}', k) \cdot p_r(\mathbf{m}'|k) \cdot q(\mathbf{m}|\mathbf{m}')}{l(\mathbf{d}|\mathbf{m}, k) \cdot p_r(\mathbf{m}|k) \cdot q(\mathbf{m}'|\mathbf{m})} \right]. \quad (2.3.2)$$

Donde $l(\mathbf{d}|\mathbf{m}, k)$ corresponderá al *likelihood* y $p_r(\mathbf{m}|k)$ al prior de nuestro modelo \mathbf{m} de dimensión k . De forma sencilla y fácil de recordar, la proporción entre los modelos comparados se puede representar como

$$\text{likelihood ratio} \times \text{prior ratio} \times \text{proposal ratio} \quad (2.3.3)$$

La ecuación (2.3.2) es conocida como la regla Metropolis-Hasting ([Metropolis et al. \(1953\)](#); [Hastings \(1970\)](#)). Esto nos dice básicamente que si la razón entre el modelo propuesto \mathbf{m}' y el modelo inicial \mathbf{m} es mayor a 1, el modelo propuesto se aceptara con una probabilidad de 1 (100%), es decir, siempre se ira aceptando el mejor modelo, sin embargo, si esto empeora obteniendo por ejemplo un valor de 0.5 en la razón de los modelos, significará que nuestro modelo propuesto será aceptado con una probabilidad del 50%.

En la práctica, una vez encontrado el valor de α , se generará un número aleatorio u entre cero y uno, luego, si u es menor a α el modelo \mathbf{m}' será aceptado, en caso contrario, la cadena se mantendrá en el modelo \mathbf{m} . Los modelos \mathbf{m}^k aceptados de esta forma convergerán a muestrear el posterior $p(\mathbf{m}, k|\mathbf{d})$, esto como consecuencia de la regla Metropolis-Hasting [Green \(2003\)](#), en donde los modelos muestreados aceptados terminarán convergiendo asintóticamente al posterior de dimensión fija.

2.3.2. Reversible jump Markov Chain Monte carlo (RjMCMC)

Como vimos anteriormente el MCMC es muy buena herramienta para poder muestrear PDFs de dimensión fija, sin embargo para problemas transdimensionales como el abordado en este trabajo, el MCMC no es útil, debido a que como

indica en su nombre, la dimensión k no es fija por lo que se debe recurrir a una extensión de este método, llamado Reversible Jump Markov Chain Monte Carlo (rjMCMC) (Green, 1995). A grandes rasgos sigue la misma lógica del MCMC, sin embargo al cambiar de dimensión se requiere cambiar la regla de aceptación de modelos usada en el MCMC regular donde toma la forma

$$\alpha = \text{Min} \left[1, \frac{p(\mathbf{m}', k' | \mathbf{d}) \cdot j_m(k', \mathbf{m}') \cdot g_{k'}(\mathbf{u}_{k'})}{p(\mathbf{m}, k | \mathbf{d}) \cdot j_m(k, \mathbf{m}) \cdot g_k(\mathbf{u}_k)} \cdot |J| \right]. \quad (2.3.4)$$

Donde $j_m(k, \mathbf{m})$, $g_k(\mathbf{u}_k)$ son los términos asociados al *proposal* de modelos de distinta dimensión y $|J|$ es un termino asociado al cambio de dimensión entre modelos propuestos. Expandiendo la ecuación (2.3.4) de la misma forma que (2.3.2) nos queda:

$$\alpha = \text{Min} \left[1, \frac{l(\mathbf{d} | k', \mathbf{m}') \cdot p_r(k', \mathbf{m}') \cdot j_m(k', \mathbf{m}') \cdot g_{k'}(\mathbf{u}_{k'})}{l(\mathbf{d} | k, \mathbf{m}) \cdot p_r(k, \mathbf{m}) \cdot j_m(k, \mathbf{m}) \cdot g_k(\mathbf{u}_k)} \cdot |J| \right]. \quad (2.3.5)$$

De igual manera que 2.3.3 podemos escribir de forma intuitiva la proporción entre los modelos comparados como Green (1995):

$$\text{likelihood ratio} \times \text{prior ratio} \times \text{proposal ratio} \times \text{Jacobian} \quad (2.3.6)$$

En este caso, podemos ver que el *proposal ratio* tiene dos términos en vez de uno, como ocurre en el MCMC de dimensión fija, estos términos son $\frac{j_m(k', \mathbf{m}')}{j_m(k, \mathbf{m})}$ que corresponde a la proporción del modelo propuesto sobre el modelo inicial y $\frac{g_{k'}(\mathbf{u}_{k'})}{g_k(\mathbf{u}_k)}$ que corresponde a la proporción entre las PDF asociadas a las variables auxiliares presentes en el cambio de dimensión. Es en ese cambio de dimensión que aparece un Jacobiano, ya que si suponemos que el cambio de (\mathbf{m}, k) a (\mathbf{m}', k') es un difeomorfismo (la transformación y su inversa son diferenciables) es necesario entonces que el Jacobiano sea distinto de cero (Green, 2003). En la práctica dicho Jacobiano funciona como un factor de escala al momento de pasar de una dimensión a otra.

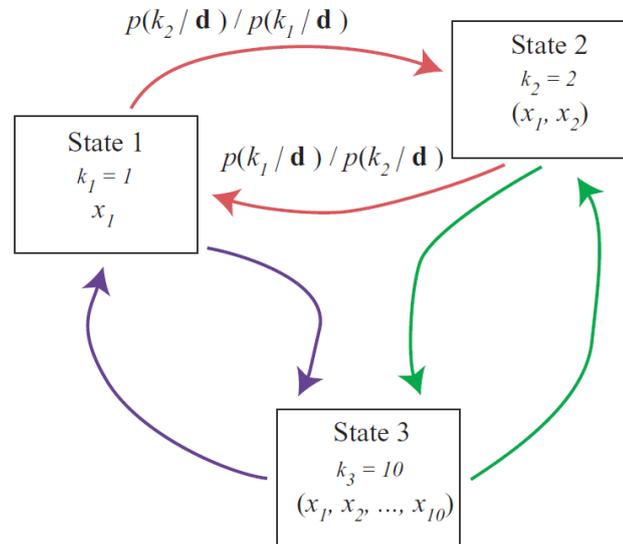


Figura 2.3.1: Esquema ilustrativo del rjMCMC (Sambridge et al., 2006)

La figura (2.3.1) ilustra de forma esquemática como funciona el rjMCMC, en donde se representan distintos estados o hipótesis con distinto número de variables cada uno. Análogo a nuestro caso, cada hipótesis representaría un modelo caracterizado por un vector de parámetros \mathbf{m} de dimensión k (es decir, contiene k parámetros). Tendremos entonces que el método irá “saltando” entre los distintos estados, muestreando distintos \mathbf{m}_k y comparándolos con el anterior, determinando si son aceptados o no siguiendo la ecuación (2.3.5). Cabe mencionar que el rjMCMC por sí mismo es un buen método de selección de modelos ya que las dimensiones más visitadas serán los modelos con mayor evidencia (Sambridge et al., 2006)

Capítulo 3

Metodología

3.1. Generalidades

Para poder abordar este problema de forma computacional, se utilizó un módulo de python proporcionado por el grupo *iEarth*¹, con algunas modificaciones para poder aplicarlo a nuestro problema. Este módulo que entrega las herramientas necesarias para aplicar el rjMCMC en el muestreo del *posterior* que buscamos caracterizar.

En esta sección y en el resto del documento se utilizará cierta terminología para referirse a la caracterización e interpretación de las series de tiempo. En primer lugar, se utilizará el termino “partición” para referirse al periodo de tiempo posterior a la ocurrencia de un multi transiente y anterior al inicio de otro, luego una serie de tiempo tendrá $m_t + 1$ número de particiones, donde m_t es el número de multi transientes que se encuentren en la serie. Cabe mencionar que el periodo de tiempo desde el inicio de la serie, hasta el comienzo del primer multi transiente también se considerará una partición

¹<http://www.earth.org.au/codes/rj-MCMC/>

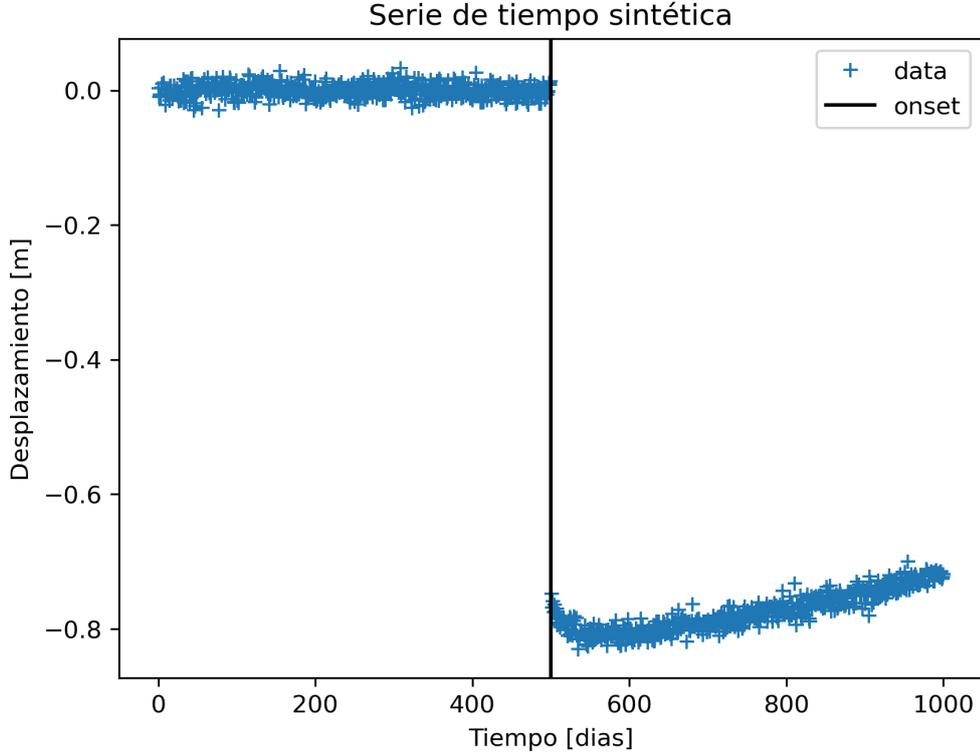


Figura 3.1.1: Serie de tiempo esquemática para ejemplificar el concepto de partición

Para ejemplificar lo anterior tomemos en cuenta la figura (3.1.1), la cual cuenta con un desplazamiento multi transiente que comienza al rededor de los 500 días desde el inicio de la serie. Luego, la serie de tiempo contará con dos particiones, la primera que va desde el día 0 hasta el tiempo de inicio del multi transiente y la segunda desde ese tiempo hasta el final de la serie.

En la misma linea de las particiones, llamaremos “parámetros locales” a los parámetros de los términos del modelo que sólo tienen una contribución desde un tiempo arbitrario. Luego tomando en cuenta la ecuación 2.1.4 tendremos que el término “local” de nuestro modelo de trayectoria se referirá al termino multi transiente, como también a los saltos que puedan haber en la serie, luego:

$$\mathbf{x}(t)_{\text{local}} = \sum_{j=1}^{n_j} b_j \mathbf{H}(t - t_j) + \sum_{r=1}^{n_r} \sum_{i=1}^{n_i} [A_i (1 - e^{-(t-t_i)1/T_i})] \quad (3.1.1)$$

Volviendo a la figura 3.1.1 tendremos que solo en la segunda partición contaremos con la presencia de parámetros locales. Por otra parte, llamaremos “parámetros

globales” a los parámetros de las componentes que tienen una contribución durante toda la serie tiempo, es decir, los términos asociados al movimiento secular y carga estacional, luego:

$$\mathbf{x}(t)_{\text{global}} = \mathbf{x}_0 + \mathbf{v}(t - t_r) + \sum_{k=1}^{n_F} [s_k \sin(\omega_k t) + c_k \cos(\omega_k t)] \quad (3.1.2)$$

3.2. Modificación modelo de trayectoria

Como idea central para nuestro modelo de trayectoria usaremos el propuesto por (Bedford and Bevis, 2018) y descrito en 2.1.4, donde el desplazamiento estará compuesto por cuatro términos principales, el movimiento secular, la carga estacional, saltos y multi transientes. Estos dos últimos términos, como se mencionó, tendrán un carácter local, es decir, tienen valor cero para todo tiempo menor a su tiempo de inicio y tienen una contribución permanente para todo tiempo posterior. De 2.1.4 tenemos que los saltos estarán representados por una función Heaviside y las deformaciones transientes por una suma de tres funciones exponenciales con constantes de decaimiento de distinto orden.

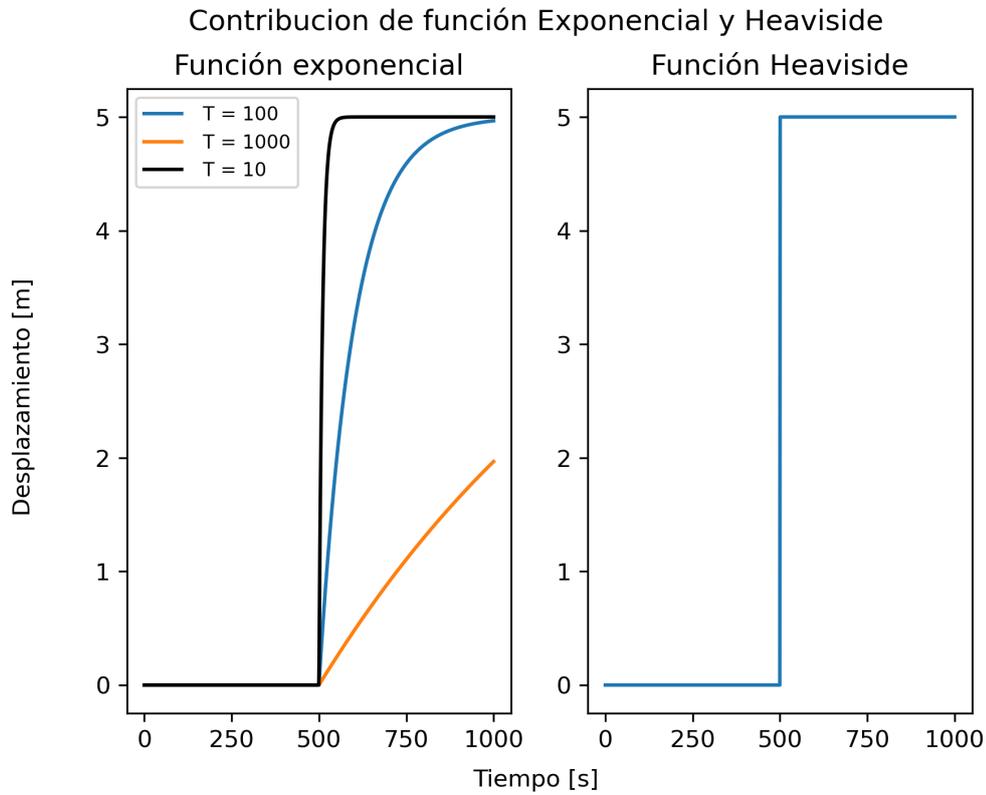


Figura 3.2.1: Contribución de las componentes transiente y Heaviside posterior al tiempo de inicio $t_r = 500$

Si suponemos un desplazamiento, nulo tal y como muestra la figura (3.2.1), veremos que previo a los 500 segundos no existe alguna contribución de los términos locales. Sin embargo, para un tiempo de referencia ubicado alrededor de los 500 segundos, veremos que la amplitud del desplazamiento comienza a aumentar gradualmente para el caso del termino multi transiente y de forma inmediata para el salto representado por una función Heaviside. Notar que las funciones exponenciales tenderán a converger gradualmente a un valor constante, sin embargo, a medida que la constante de decaimiento disminuye, se alcanzará dicho valor de manera mas rápida.

De la ecuación 2.1.4 tenemos que el desplazamiento producido por solo un exponencial vendrá dado por:

$$\mathbf{x}(t) = A \left(1 - e^{-(t-t_r)1/T} \right), \quad (3.2.1)$$

sin embargo, consideremos el hecho de que

$$\lim_{T \rightarrow 0} e^{-(t-t_r)1/T} = 0, \quad (3.2.2)$$

donde t_r corresponde al tiempo de inicio de un transiente y T la constante de decaimiento. Aplicando 3.2.2 en 3.2.1 tendremos que:

$$\lim_{T \rightarrow 0} A (1 - e^{-(t-t_r)1/T}) = A. \quad (3.2.3)$$

Luego, podemos establecer que para una constante de decaimiento lo suficientemente pequeña, tendremos que el término transiente tomará un valor constante en todo el tiempo desde su inicio, aproximándose al comportamiento de una función Heaviside.

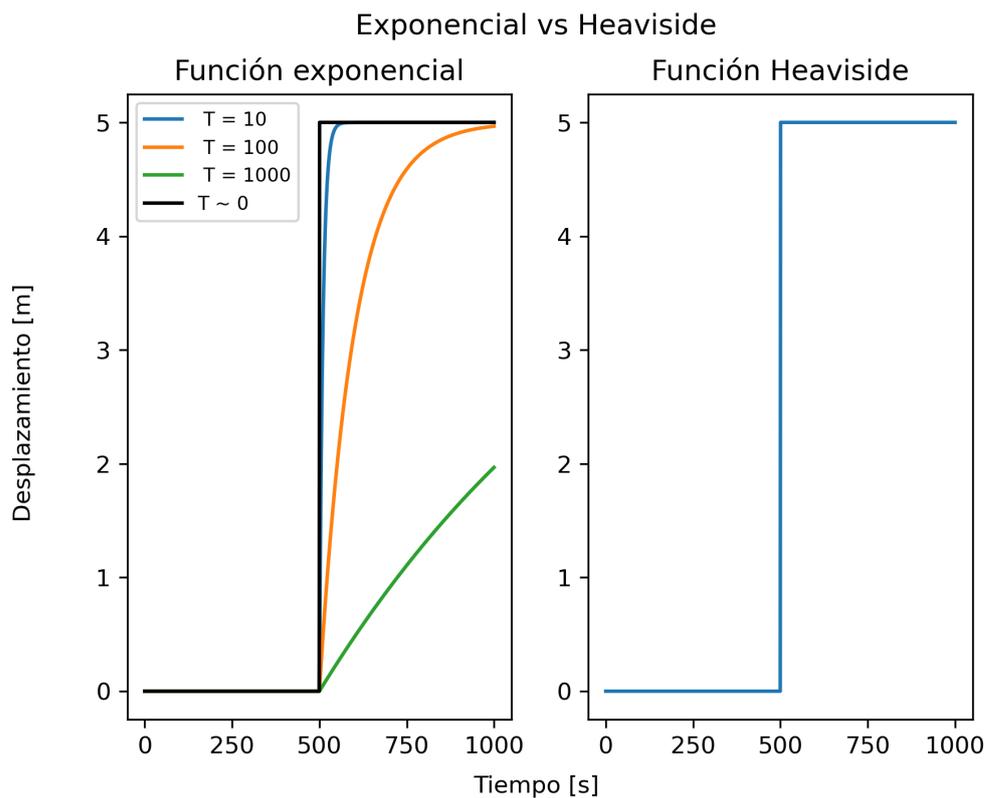


Figura 3.2.2: Comparación entre el comportamiento de una función Heaviside y una función exponencial con constante de decaimiento lo suficientemente pequeña

En la figura 3.2.2 podemos ver de forma gráfica el comportamiento anterior mencionado, donde al haber una constante de decaimiento cercana a 0, el término que contiene una función exponencial terminará comportándose como una función

Heaviside. Finalmente, incorporando los saltos en la componente multi transiente, nuestro modelo de trayectoria quedará descrito por:

$$\begin{aligned} \mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}(t - t_r) + \sum_{k=1}^{n_F} [s_k \sin(\omega_k t) + c_k \cos(\omega_k t)] \\ + \sum_{r=1}^{n_r} \sum_{i=1}^{n_i} [A_i (1 - e^{-(t-t_r)^{1/T_i}})] + \xi(t) \end{aligned} \quad (3.2.4)$$

Donde el término multi transiente, en vez de ser la suma de tres exponenciales con constantes de decaimiento de distinto orden de magnitud, en nuestro caso será una suma de hasta cuatro exponenciales donde una ellas tendrá una constante de decaimiento cercana a cero. Aunque a primera vista no parezca una modificación importante, esto nos permitirá no tratar como fijo el número de transientes que compone un multi transiente, a diferencia del trabajo de [Bedford and Bevis \(2018\)](#).

Con nuestro nuevo modelo de trayectoria, tendremos entonces que la contribución local vendrá dada por sólo por el término

$$\mathbf{x}(t)_{local} = \sum_{r=1}^{n_r} \sum_{i=1}^{n_i} [A_i (1 - e^{-(t-t_r)^{1/T_i}})] + \xi(t). \quad (3.2.5)$$

Dicho término local tendrá una dimensión variable $k \leq n_i$ la cual nos indicará el número de exponenciales que contribuirán a la señal local. Luego llamaremos multi transiente a todo término local que corresponda a la suma de 2 o más exponenciales.

k	N° de exponenciales	Ctes. Decaimiento T [días]	Tipo
1	1	≈ 0 ,	Heaviside puro
2	2	≈ 0 , 1000	Multi transiente
3	3	≈ 0 , 1000, 100	Multi transiente
4	4	≈ 0 , 1000, 100, 10	Multi transiente

Cuadro 3.2.1: Clasificación de multi transientes según su dimensión

Esto convierte nuestro problema en uno jerárquico, es decir, que todo término local de dimensión k contendrá a los términos asociados a dimensiones menores. Es decir, un multi transiente de dimensión 3 será la suma de un multi transiente de dimensión 2 con una exponencial con constante de decaimiento $T = 100$. Esto nos dice que desplazamientos mas largos corresponderán a términos locales mas

simples y a medida que el desplazamiento presente una menor duración se irán agregando términos exponenciales.

3.3. Elección de PDFs

Como se ha mencionado en secciones anteriores, la regla de aceptación de un modelo por sobre otro vendrá dado por la regla de Metropolis-Hasting descrita en 2.3.5 donde se comparan las probabilidades de cada modelo \mathbf{m}_k respecto a uno propuesto $\mathbf{m}_{k'}$. Luego cada expresión matemática mostrada a continuación será equivalente tanto para un modelo \mathbf{m}_k como para $\mathbf{m}_{k'}$.

3.3.1. Función de verosimilitud (Likelihood)

Por lo general los valores predichos no son idénticos a las observaciones, esto debido a la incertidumbre de las mediciones y las imperfecciones del modelo (e.g limitaciones en el modelo de trayectoria), por esta razón es de suma importancia tener un cuidadoso análisis de las incertidumbres del modelo. En nuestro caso dichas consideraciones son abordadas al momento de escoger una PDF adecuada como likelihood.

Para construir nuestra función de verosimilitud debemos tomar la diferencia entre nuestra predicción hecha con el modelo propuesto y los datos. Esto se traduce en encontrar una función de densidad de probabilidad que describa la forma en que se distribuyen los residuales de nuestros datos [Bretthorst \(1988\)](#). Si asumimos que para nuestro caso el los residuales se distribuyen de forma gaussiana, la elección lógica para nuestra función de verosimilitud vendría a ser una distribución normal multivariada de la forma:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\mathbf{C}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})}. \quad (3.3.1)$$

Teóricamente \mathbf{x} corresponde a un vector de dimensión N , normalmente distribuido con promedio $\boldsymbol{\mu}$ y con matriz de covarianza \mathbf{C} . En nuestro caso, el término $(\mathbf{x} - \boldsymbol{\mu})$ corresponderá al residual entre nuestro modelo propuesto y las observaciones, luego podemos reescribir 3.3.1 como:

$$l(\mathbf{d}|\mathbf{m}, k) = \frac{1}{(2\pi)^{N/2} |\mathbf{C}_d|^{1/2}} e^{-\frac{1}{2}(\mathbf{G}(\mathbf{m})-\mathbf{d})^T \mathbf{C}_d^{-1}(\mathbf{G}(\mathbf{m})-\mathbf{d})}, \quad (3.3.2)$$

con \mathbf{C}_d matriz de covarianza de los residuales y k el tamaño del vector \mathbf{m} . Debido a las grandes dimensiones que pueden llegar a tener los datos y su respectiva matriz de covarianza existen complicaciones numéricas y computacionales que impiden utilizar 3.3.2, por lo que es común utilizar el logaritmo natural de dicha ecuación para facilitar los cálculos, caso que aplica a nuestro problema. Finalmente nuestro likelihood quedará expresado por:

$$\text{loglikelihood} = -\frac{1}{2} (N \ln(2\pi) + \ln(|\mathbf{C}_d|)) + \frac{1}{2} (\mathbf{G}(\mathbf{m}) - \mathbf{d})^T \mathbf{C}_d^{-1} (\mathbf{G}(\mathbf{m}) - \mathbf{d}) \quad (3.3.3)$$

El motivo de la elección de esta PDF para nuestro likelihood fue siguiendo la línea de investigaciones similares, donde se estudiaron deformaciones de la corteza terrestre por medio de datos geodéticos (e.g., Fukuda and Johnson, 2010; Yabuki and Matsu'ura, 1992; Fukuda and Johnson, 2008) mediante métodos bayesianos. La elección de una PDF gaussiana se basa principalmente en el teorema del límite central, el cual establece a grandes rasgos que la suma de un gran número (> 30) de variables aleatorias independientes de distinta fuente (es decir, con distintas distribuciones aleatorias) tenderán a converger a una distribución normal (Schatte, 1988), lo que de cierta podría describir como de distribuye el ruido de nuestros datos, asumiendo que éste es originado por una multitud de distintas fuentes independientes.

El hecho de que escojamos una PDF normal multi variada restringe la forma de nuestra matriz de covarianza, puesto que al establecer que el ruido se origina de forma aleatoria e independiente nos dice que no deberían existir correlaciones en este. Si bien, una PDF normal multivariada admite correlaciones en la matriz de covarianza, por simplicidad se asumirá que dichas correlaciones en los residuales son despreciables. La pertinencia de esta simplificación se discutirá en secciones posteriores. Por consecuencia de lo anterior, nuestra matriz de covarianza seña proporcional a una matriz identidad, de la forma:

$$\mathbf{C}_d = C_\sigma \cdot \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & . & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}_{n \times n}, \quad (3.3.4)$$

donde C_σ corresponde a un hiper parámetro de escala para la varianza, esto con el objetivo de lograr un mejor ajuste a los datos. Notar que en casos donde no se conozca la desviación estándar, esta formulación permitirá estimarla mediante los mismos datos. La forma de hacerlo, es tomando un intervalo de tiempo de los datos (cientos de días) al cual se le elimina una tendencia lineal y el promedio. Finalmente se ocupará la desviación estándar correspondiente a dicho intervalo de tiempo, de modo que el parámetro C_σ se encargue de ajustar un mejor en caso de ser necesario.

3.3.2. Prior

La elección de un prior suele ser uno de los temas mas controversiales cuando se habla de métodos bayesianos, esto debido a que todos los resultados dependen directamente de esta elección. Recordemos que el prior representa nuestro estado del conocimiento acerca del modelo, sin considerar la información provista por los datos. Es por esto, que nuestro prior no debe ser muy informativo, debido a que el algoritmo podría detectar cosas en el modelo que no están ahí (Bretthorst, 1988), introduciendo un sesgo en los resultados. Tendremos entonces que nuestro prior, para un caso transdimensional vendría a ser definido por una probabilidad P de la forma:

$$P(k, \mathbf{m}) = p(k) \cdot p_r(\mathbf{m}|k). \quad (3.3.5)$$

Donde $p(k)$ corresponderá a la probabilidad que tenga cada modelo de dimensión k , sin embargo, se asumirá que todos los modelos tendrán la misma probabilidad, luego:

$$p(k) = \frac{1}{k_{\max}}, \quad (3.3.6)$$

donde k_{\max} corresponde a la dimensión máxima establecida para el análisis. Por otra parte $p_r(\mathbf{m}|k)$ será la PDF que describirá la probabilidad de nuestras incógnitas (i.e parámetros del modelo) dada una dimensión k . Para esta PDF se utilizará una distribución uniforme, esto quiere decir que antes de hacer alguna evaluación o comparación entre modelos, todos tendrán la misma probabilidad, esto siguiendo la línea de la poca información del prior. Dicha distribución tendrá la forma:

$$p_r(\mathbf{m}|k) = \prod_{i=1}^{n_g+n_k} p(\mathbf{m}^i|k), \quad (3.3.7)$$

donde n_g , n_k corresponden al número de parámetros locales y globales, por otra parte \mathbf{m}^i , hará referencia al i -ésimo parámetro de nuestro vector \mathbf{m} de dimensión k representativo del modelo. Como se trata de una distribución uniforme, es necesario también definir un valor mínimo (U) y máximo (L) en el que se muevan los parámetros de modo que la distribución, esta probabilidad puede ser calculada como:

$$p(\mathbf{m}^i|k) = \frac{1}{\mathbf{m}^{i,U} - \mathbf{m}^{i,L}}. \quad (3.3.8)$$

3.3.3. Proposal

El proposal corresponderá a la PDF que describirá la probabilidad del nuevo modelo propuesto a partir de un modelo anterior. Supongamos que nos encontramos en un estado (k, \mathbf{m}_k) y se propone cambiar a un modelo de dimensión k' , para esto se considerará la siguiente transformación, con n_k el número de parámetros que presenta el modelo \mathbf{m}_k

$$\mathbf{m}_{k'} = \begin{cases} \boldsymbol{\mu}_{k'} + \mathbf{B}_{k'} [\mathbf{B}_k^{-1}(\mathbf{m}_k - \boldsymbol{\mu}_k)], & n_{k'} < n_k \\ \boldsymbol{\mu}_{k'} + \mathbf{B}_{k'} \mathbf{B}_k^{-1}(\mathbf{m}_k - \boldsymbol{\mu}_k), & n_{k'} = n_k \\ \boldsymbol{\mu}_{k'} + \mathbf{B}_{k'} \begin{bmatrix} \mathbf{B}_k^{-1}(\mathbf{m}_k - \boldsymbol{\mu}_k) \\ \mathbf{u}_k \end{bmatrix}, & n_{k'} > n_k \end{cases}$$

donde $\boldsymbol{\mu}_k$ es el promedio de los parámetros \mathbf{m}_k , \mathbf{B}_k es una matriz tal que $\mathbf{B}_k \mathbf{B}_k^T$ es la matriz de covarianza de los parámetros \mathbf{m}_k y \mathbf{u}_k son las variables auxiliares para el cambio de dimensión (Green, 2003). De 2.3.5 tenemos que el proposal está compuesto por dos términos, $j_m(k, \mathbf{m}_k)$ que corresponde a la probabilidad del modelo que se va a utilizar y $g_k(\Gamma_k)$ que nos describe como distribuyen las variables auxiliares en el cambio de dimensión (Green, 2003). Como asumimos

que cada modelo de dimensión k tienen la misma importancia, tendremos que:

$$j_m(k, \mathbf{m}_k) = \frac{1}{k_{max}}. \quad (3.3.10)$$

Por otra parte, para $g_k(\mathbf{u}_k)$, supondremos que distribuye de forma normal para que sea consistente con el resto de las PDF, luego:

$$g_k(\mathbf{u}_k) = (2\pi)^{-(n_{k'}-n_k)/2} e^{\frac{1}{2}\mathbf{u}_k^T \mathbf{u}_k} \quad (3.3.11)$$

Si consideramos:

$$\mathbf{u}_k = [\mathbf{B}_k^{-1}(\mathbf{m}_{k'} - \boldsymbol{\mu}_{k'})] \quad (3.3.12)$$

Finalmente

$$g_k(\mathbf{u}_k) = \frac{(2\pi)^{-n_{k'}/2} \exp \left[-\frac{1}{2}(\mathbf{m}_{k'} - \boldsymbol{\mu}_{k'})^T \mathbf{C}_{k'}^{-1}(\mathbf{m}_{k'} - \boldsymbol{\mu}_{k'}) \right]}{(2\pi)^{-n_k/2} \exp \left[-\frac{1}{2}(\mathbf{m}_k - \boldsymbol{\mu}_k)^T \mathbf{C}_k^{-1}(\mathbf{m}_k - \boldsymbol{\mu}_k) \right]} \quad (3.3.13)$$

Por otra parte el término $|J|$ vendrá dado por:

$$|J| = \frac{|\mathbf{B}_{k'}|}{|\mathbf{B}_k|} = \frac{|\mathbf{C}_{k'}|^{\frac{1}{2}}}{|\mathbf{C}_k|^{\frac{1}{2}}} \quad (3.3.14)$$

Con \mathbf{C}_k la matriz de covarianza del modelo de dimensión k

3.4. Cálculo de la evidencia

La evidencia corresponde al denominador de la ecuación (2.2.4), y en el caso transdimensional, hará referencia a cuan bien una hipótesis (i.e dimensión del modelo) explicará los datos. Este término toma mucha importancia al momento de seleccionar modelos, puesto que si tenemos dos hipótesis (i.e dos modelos con distinta dimensión), la comparación entre estas cantidades entonces favorecerá una hipótesis por sobre otra, en específico, la que presente una mayor evidencia. Ahora bien, se mencionó en secciones previas que el rjMCMC cuenta con herramientas inherentes para hacer selección de modelo en función de la cantidad de veces que

se muestreó una hipótesis. Sin embargo, investigaciones previas establecen que en porcentaje, la cantidad de veces que se visita una hipótesis por medio del rjMCMC es equivalente a la evidencia para una cierta hipótesis sobre el total de hipótesis propuestas (Sambridge et al., 2006). El problema en muchos casos, es que encontrar expresiones analíticas para la evidencia es sumamente complicado, sobre todo para problemas no lineales, haciendo el calculo de esta cantidad muy complicado, razón por la cual se suele omitir y considerada solo como una constante de normalización Sambridge et al. (2006).

No obstante, se tiene que para problemas lineales con *likelihood* gaussiano, la evidencia puede ser obtenida analíticamente de la forma:

$$p(\mathbf{d}) = e^{-\psi(\hat{\mathbf{m}})/2} \sqrt{\frac{|\mathbf{C}_{\mathbf{m}}^{\text{post}}|}{(2\pi)^N |\mathbf{C}_{\mathbf{d}}| |\mathbf{C}_{\mathbf{m}}|}} \quad (3.4.1)$$

Con $\psi(\hat{\mathbf{m}})$ la función de costo, evaluada en el modelo máximo a posteriori (MAP, $\hat{\mathbf{m}}$) (Mitsuhata et al., 2001):

$$\psi(\mathbf{m}) = (\mathbf{G}\mathbf{m} - \mathbf{d})^T \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{G}\mathbf{m} - \mathbf{d}) + (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_{\mathbf{m}}^{-1} (\mathbf{m} - \mathbf{m}_0) \quad (3.4.2)$$

Notar que los valores de la evidencia para problemas muy grandes, tienden a estar fuera de los límites numéricos computacionales. Es por eso que normalmente se trabaja con en logaritmo de ella. En el contexto de este trabajo, estas expresiones nos permitirán comparar las selecciones de modelo hechas tanto por el rjMCMC como por la evidencia en simulaciones sencillas con características similares a series de tiempo de GNSS, de modo que nos permita validar la selección de modelo hecha por el rjMCMC.

Capítulo 4

Resultados

Como hemos estado mencionando, una de las incógnitas de nuestro problema, además de los parámetros de nuestro modelo, será también la dimensión de este. Luego siguiendo la lógica de las particiones, la dimensión de nuestro modelo (i.e. cantidad de parámetros) será directamente proporcional a la cantidad de particiones y además a la dimensión k del término local correspondiente a cada partición. Dichas particiones tendrán entonces una dimensión k desconocida, que tendrá las características descritas en la tabla [3.2.1](#)

4.1. Simulaciones

En primera instancia se generaron una serie de ejemplos sintéticos de distintas características y complejidades con el objetivo de probar las diversas propiedades del rjMCMC, tanto su selección de modelo natural como su equivalencia con expresiones analíticas. Todos estos ejemplos se harán considerando un ruido gaussiano estacionario, con el objetivo de ser coherentes con las PDF escogidas para nuestra formulación. Como primer ejemplo se generó una serie de tiempo con solo una partición, con un multi transiente de dimensión 4, es decir, un Heaviside puro mas 3 exponenciales.

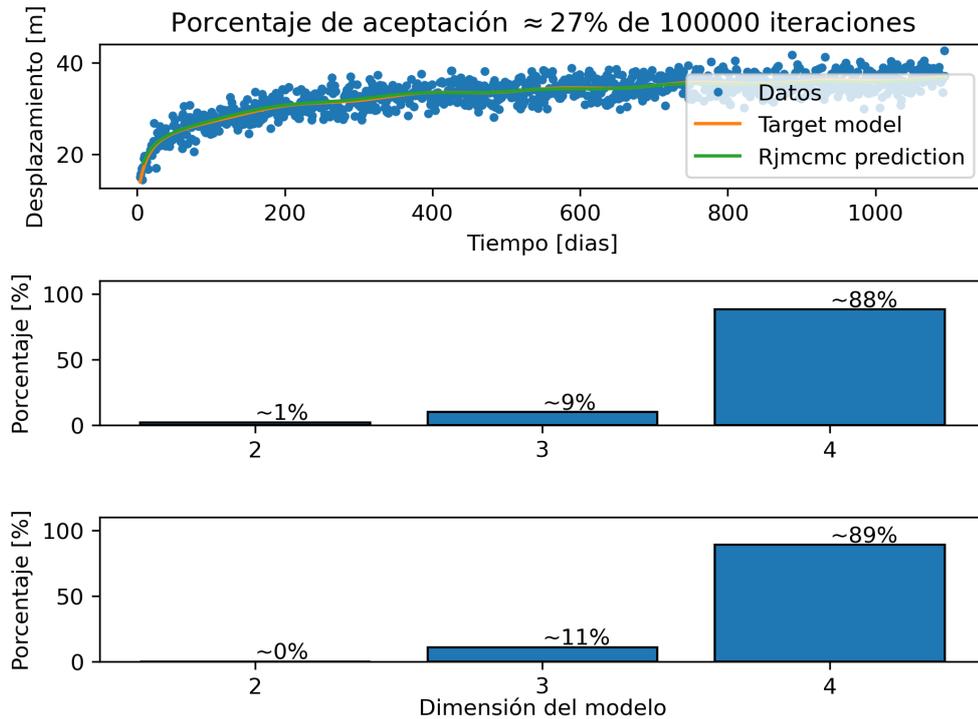


Figura 4.1.1: Resultados del muestreo hecho por el rjMCMC. La primera figura muestra los datos junto con la predicción hecha con el modelo seleccionado junto con el modelo inicial usado para generar los datos. Los histogramas muestran en porcentaje los criterios de selección de modelo por parte del rjMCMC, como también de manera analítica (cálculo de evidencia) respectivamente.

El porcentaje de aceptación hará referencia a la cantidad de modelos que fueron aceptados sobre el total muestreado. Analizando el resto de la figura 4.1.1, tenemos que el primer histograma muestra en porcentaje, la cantidad de veces que el rjMCMC muestreó modelos de una dimensión determinada sobre el total de iteraciones, en este caso cada iteración corresponde a un muestreo de modelo. Por otra parte el segundo histograma muestra en porcentaje cuanto aporta la evidencia de cada parametrización (i.e dimensión del multi transiente) respecto al total de las evidencias. Como podemos ver, los porcentajes son bastante parecidos, lo que nos muestra que tanto técnicas de muestreo transdimensionales como analíticas son equivalentes bajo ciertas circunstancias. (Sambridge et al., 2006).

Ahora bien, el cálculo de evidencia resulta práctico solo para ejemplos simples tales como el anterior mostrado, principalmente para comparar alguna parametrización de una partición en particular. Sin embargo, puesto que para un tiempo de origen para un multi transiente dado, el problema es lineal, es posible utilizarla también

para un ejemplo sencillo con dos particiones, en este caso, se hace una búsqueda en la grilla, evaluando la evidencia en pequeños intervalos de tiempo, luego la evidencia para una parametrización de dos particiones será máxima en el tiempo de inicio del multi transiente. Luego para una serie de tiempo sintética de dos particiones, con un multi transiente de dimensión $k = 4$, ubicado justo a la mitad del dominio del tiempo, tanto el análisis por medio del rjMCMC como de la evidencia deberían ser similares.

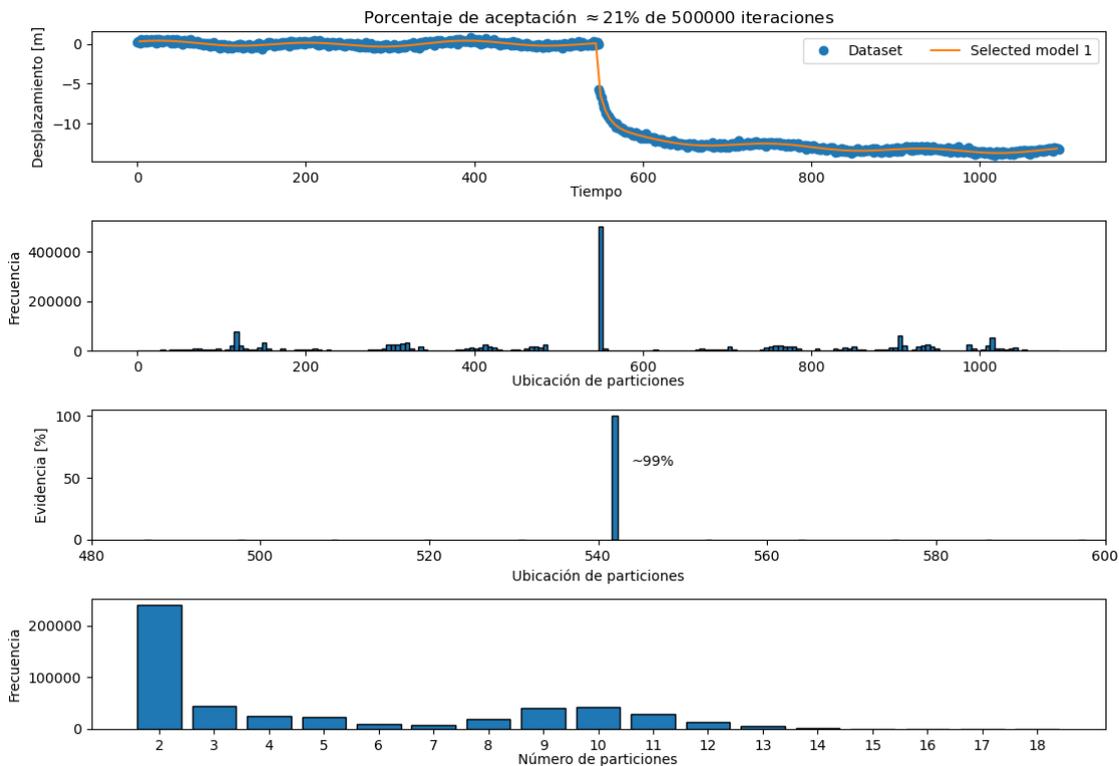


Figura 4.1.2: Resultado del muestreo realizado por el rjMCMC comparados con el calculo de evidencia. La curva naranja muestra la predicción hecha con el medelo seleccionado por el muestreo. Los dos histogramas del centro muestran la ubicación de una partición según lo encontrado por el rjMCMC (primer histograma) y el calculo analítico de la evidencia (segundo histograma)

Tiempo [días]	Número	Dimensión	Tipo
0.000000	1	0	Nulo
547.276	2	4	Mult. trans.

Cuadro 4.1.1: Características de las particiones encontradas por el rjMCMC en la serie de tiempo sintética. Cabe mencionar que el hecho de que en la primera partición el valor de k sea nulo se debe a que en dicha partición solo existe contribución del término global del modelo de trayectoria .

Si analizamos la figura 4.1.2 tenemos que el primer histograma muestra la cantidad de veces que se muestreó un modelo con una partición ubicada en el tiempo indicado. Por otro lado, si vemos el tercer gráfico de la figura 4.1.2), podemos ver que tanto la evidencia como el muestreo hecho por el rjMCMC entregan resultados similares, identificando la partición con muchísima precisión en ambos casos, reafirmando el postulado de Sambridge et al. (2006) sobre la equivalencia de estas aproximaciones. El cuadro ?? muestra en detalle los resultados arrojados por el muestreo del rjMCMC.

Ahora, como último ejemplo sintético con ruido gaussiano, un poco mas complejo, se creó una serie de tiempo compuesta por varios términos locales de distinta dimensión, entre Heaviside puros y multi transientes con el objetivo de verificar si el método es capaz de identificar el número de particiones presentes en la serie como también el tiempo de inicio y dimensión de cada término local.

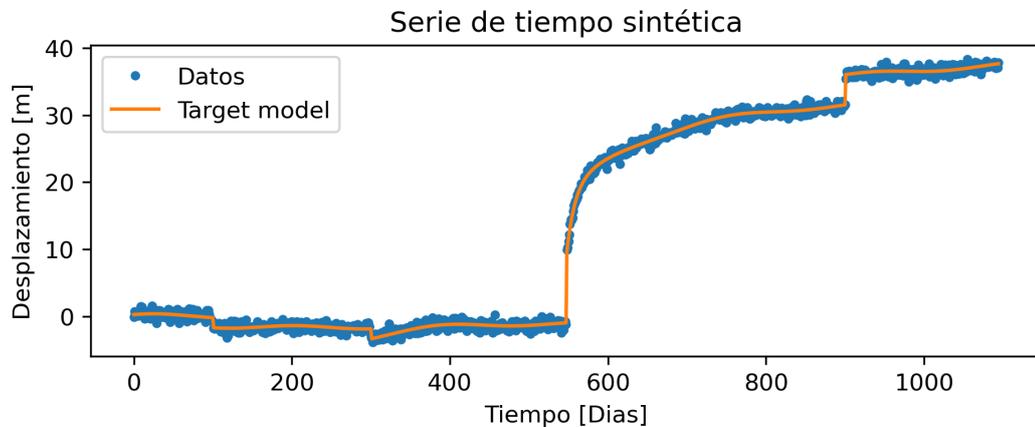


Figura 4.1.3: Serie de tiempo sintética con 5 particiones de multi transientes de distinta dimensión

En particular, para la serie de tiempo sintética mostrada en la figura 4.1.4 se consideraron 5 particiones (i.e 4 términos locales), cada una de distinta dimensión. Notar que en este caso, el uso de la expresión analítica para la evidencia se vuelve poco práctico ya que involucraría una búsqueda en grilla de 4 onset time, por lo que no se utilizará en la comparación. En mayor detalle las características de la señal vendrán descritas a continuación

Partición	k	Tiempo de inicio [día]	Tipo
1	-	0	-
2	1	100	Heaviside puro
3	3	300	Multi transiente
4	4	547.5	Multi transiente
5	1	900	Heaviside puro

Cuadro 4.1.2: Características del modelo utilizado para generar la serie de tiempo sintética.

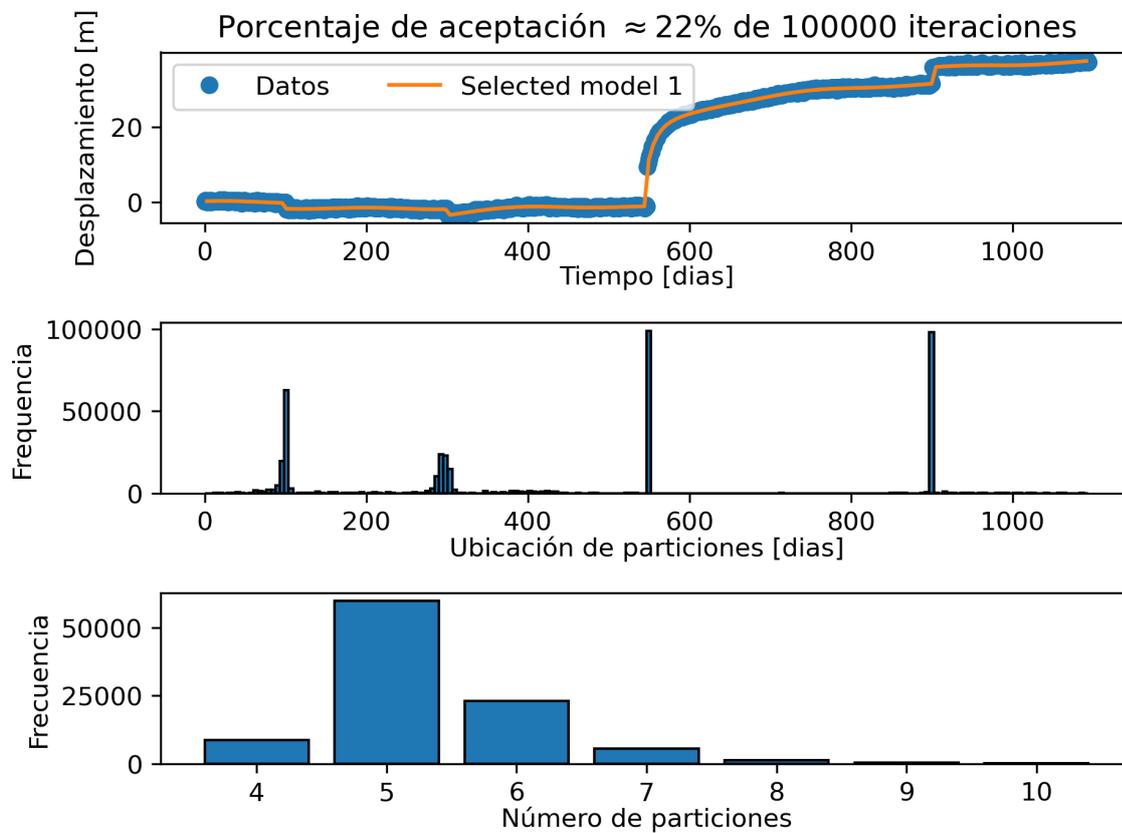


Figura 4.1.4: Resultado de inversión para ejemplo sintético de varias particiones, la figura de mas arriba muestra la predicción hecha en base al modelo seleccionado. Por otra parte la figura del medio muestra los histogramas

Partición	k	Tiempo de inicio [día]	Tipo
1	-	0	-
2	1	100.73	Heaviside puro
3	3	303.48	Multi transiente
4	4	547.96	Multi transiente
5	1	900.22	Heaviside puro

Cuadro 4.1.3: Características del modelo seleccionado por el rjmcmc para la serie de tiempo sintética.

Podemos ver que para este ejemplo mas complejo, el método recupera de manera muy precisa los tiempos de inicio de cada multi transiente, como también la dimensión de este. Además de los histogramas de la figura 4.1.4 podemos observar que pese a muestrear modelos mas complejos (mayor número de particiones), el método se decanta por modelos mas simples.

4.1.1. Serie de tiempo sintética para un terremoto lento

En la siguiente sección, y con el objetivo de probar fenómenos un poco mas parecidos a la realidad, se generaron series de tiempo para una red GNSS distribuida al azar por el territorio chileno. La forma de generar dichas series de tiempo se en el trabajo de Carr Agnew (2013), en donde se muestra una forma realista de generar simulaciones para una red geodética, como también su respuesta a distintos fenómenos, entre ellos, un terremoto. Para generar el desplazamiento producido por el terremoto lento se utilizaron funciones de Green que describieran el comportamiento del medio frente a un impulso (i.e terremoto) y que fueran coherentes con el margen chileno de placas. En nuestro caso, se generaron dos terremotos lentos con igual epicentro, el primero de ellos tuvo un deslizamiento máximo de 0.05 [m] a lo largo de 100 días y el segundo tuvo un deslizamiento máximo de 0.03[m] durante 50 días, ambos frente a las costas del norte de Chile.

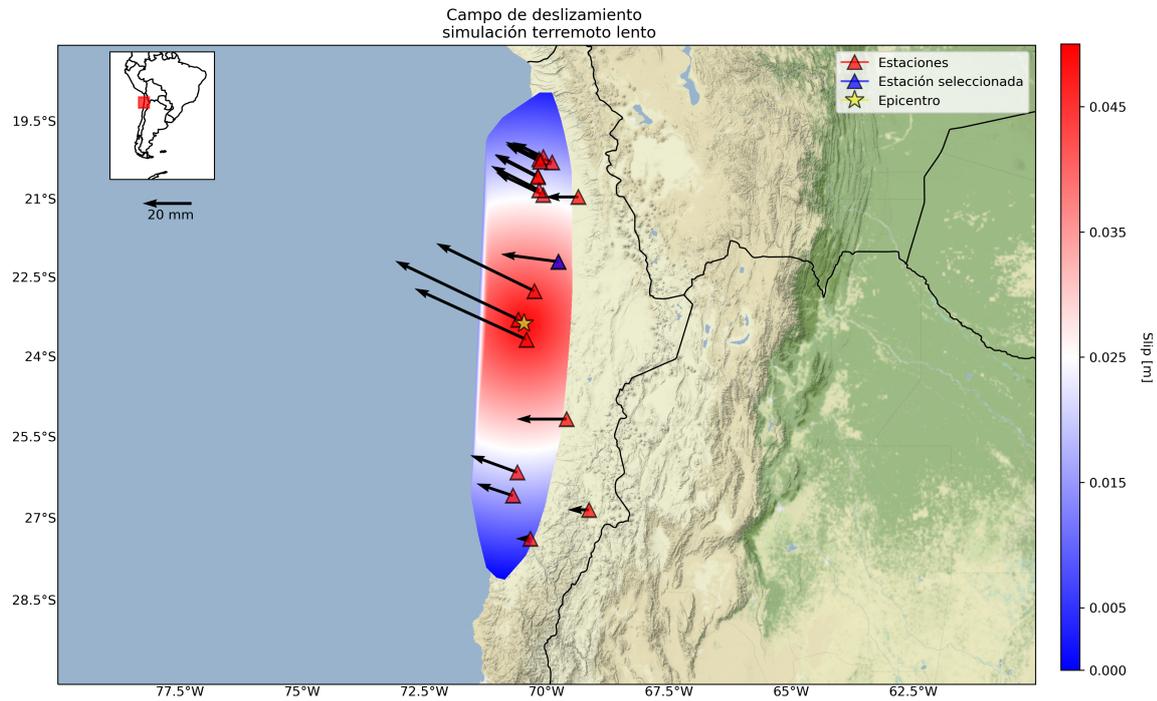


Figura 4.1.5: Campo de deslizamiento para el primer terremoto lento simulado, con deslizamiento máximo de 0.05m y 100 días de duración

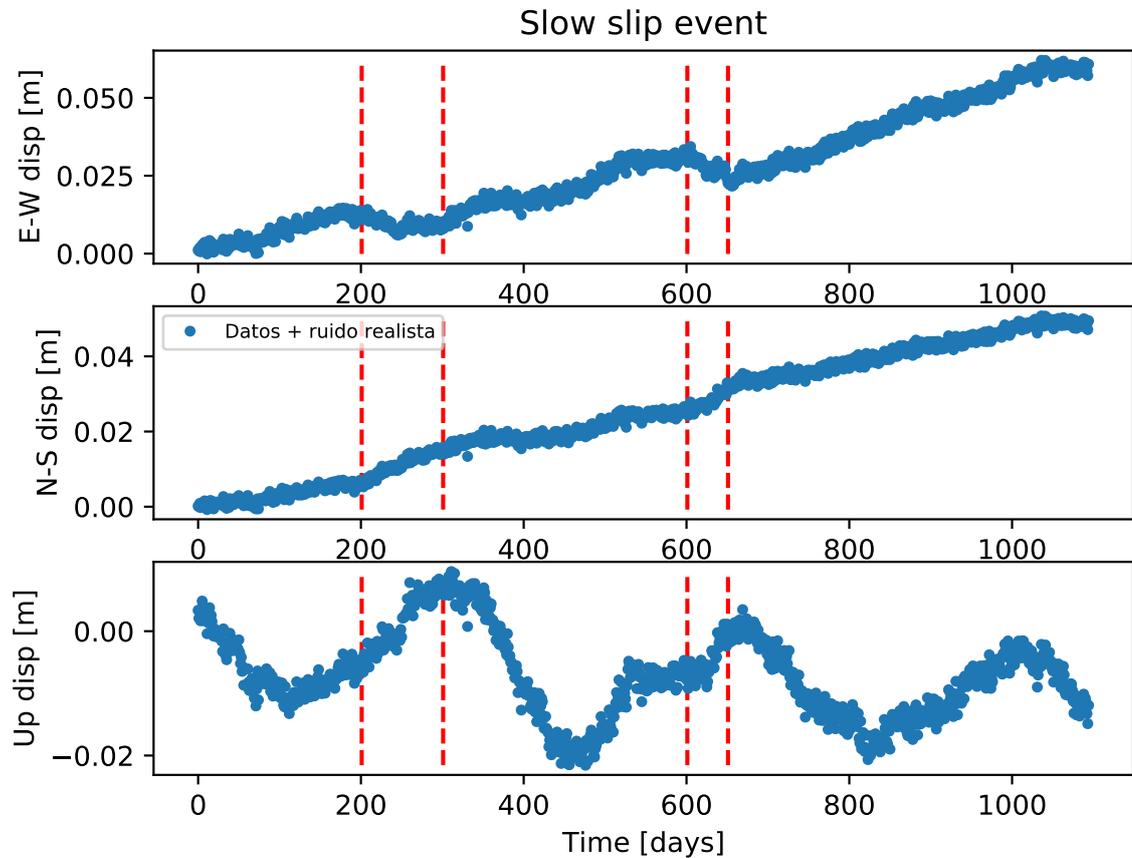


Figura 4.1.6: Serie de tiempo sintética para cada componente de la estación SLWS con ruido realista simulada para una red geodética de datos en base a la metodología descrita en Carr Agnew (2013)

Para efectos del análisis de las series de tiempo, tomaremos las componentes E-O de una estación cercana a la zona de mayor deformación producida por los terremotos lentos, que llamaremos SLWS, (ver figura 4.1.5). Como se puede apreciar en la figura 4.1.6 tendremos la serie tiempo para cada componente de la estación SLWS mostrando en ella los lapsos de tiempo de duración de ambos terremotos lentos, de forma detallada, el primero de ellos comienza el día 201 y el segundo el día 601.

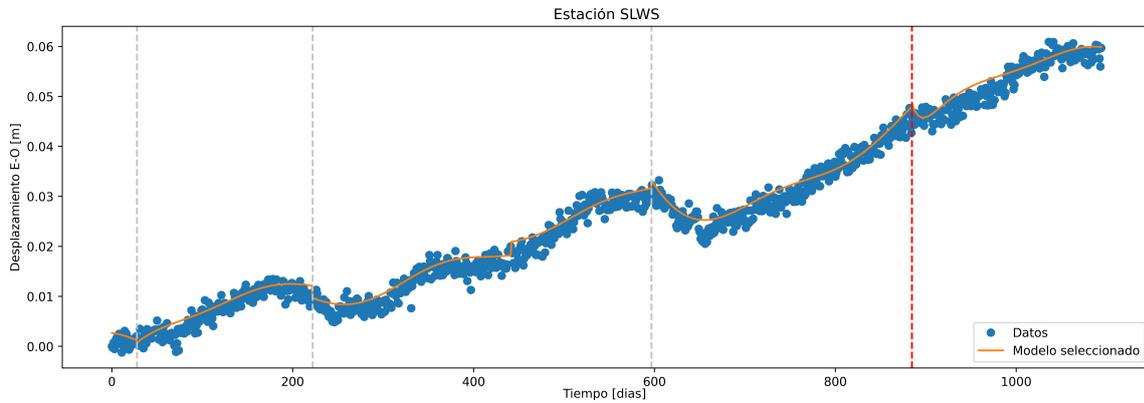


Figura 4.1.7: Serie de tiempo para la componente E-O de la estación SLWS. La curva naranja corresponde a la predicción hecha por el modelo seleccionado por el rjMCMC y las líneas horizontales grises marcan el tiempo de inicio de los multi transientes encontrados. La línea vertical roja indica el inicio de un multi transiente de dimensión 4.

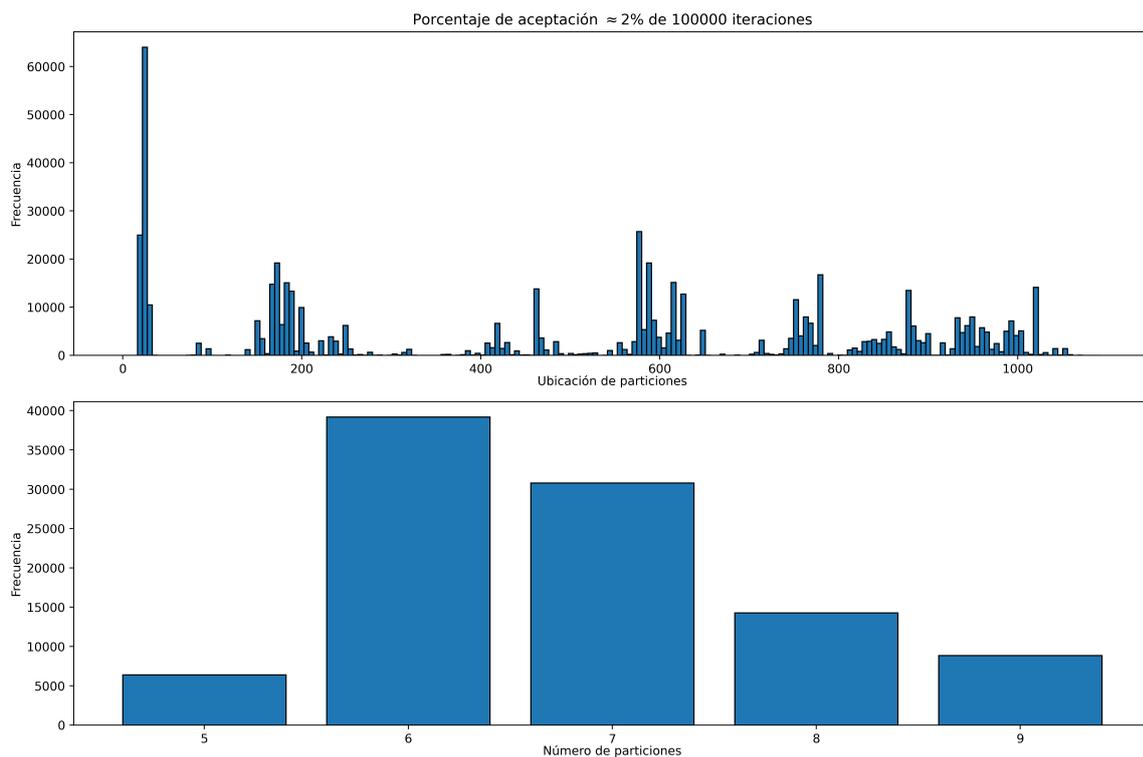


Figura 4.1.8: Histogramas resultantes del muestreo realizado por el rjMCMC. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo histograma indica la cantidad de veces que se muestreó un número determinado de particiones.

Día	Partición	Dimensión	Tipo	C_σ
0.000000	1	0	Nulo	4.32
27.755553	2	3	Mult. trans.	
221.946689	3	3	Mult. trans.	
441.468818	4	1	Heaviside	
596.703601	5	3	Mult. trans.	
884.635128	6	4	Mult. trans.	

Cuadro 4.1.4: Características del modelo seleccionado por el rjMCMC para la componente E-O de la estación SLWS

Este resultado es interesante, puesto que pese a que la señal tuviese un ruido de características desconocidas (y similares a las de estaciones reales), el método pudo ajustar de manera un tanto precisa nuestro modelo de trayectoria a la señal del terremoto lento sintético. En particular las series de tiempo generadas consideran contribuciones de ruido asociadas principalmente al modo común de error y al ruido aleatorio inherente de cada estación. En este caso, para la componente E-O, hubo una variación de pocos días entre el tiempo de inicio original y el encontrado (ver 4.1.4). También es interesante puesto que reconoce la deformación multi transiente asociada a los terremotos lentos como un sub modelo de dimensión 3, lo que de alguna forma nos podría dar un sustento para encontrar este tipo de deformaciones en otras series de tiempo y que no necesariamente consista en un multi transiente de dimensión 4 (i.e suma de 3 exponenciales) como se plantea en (Bedford and Bevis, 2018). Sin embargo, es llamativo también que el método encuentra otras particiones (i.e deformación multi transiente) en lugares donde originalmente no había nada, por lo que afirmar que efectivamente se trata de una deformación real no es del todo correcto. Pese a lo anterior, resulta interesante también analizar la cantidad cuanta distancia se desplazó la estación SLWS tanto para la simulación como para lo encontrado por el rjMCMC.

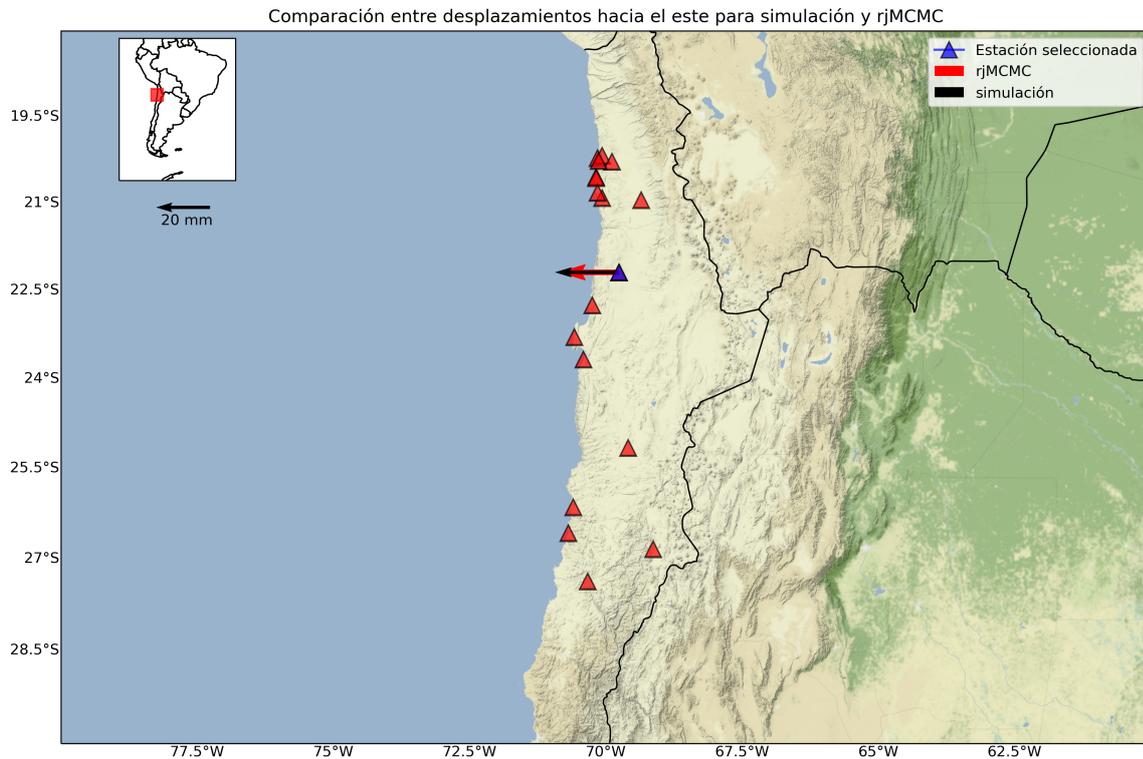


Figura 4.1.9: Comparación entre desplazamientos hacia el este según lo hecho en la simulación junto con lo mostrado mediante el modelo seleccionado por el rjMCMC.

	Simulación	rjMCMC
Desplazamiento [mm]	23.54040839	23.71635457

Cuadro 4.1.5: Desplazamientos hacia el este según la simulación y el modelo escogido por el rjmcmc

Si tomamos la predicción seleccionada por el rjMCMC y consideramos el desplazamiento hacia el este hecho solo por los multi transientes tendremos que es bastante similar al desplazamiento hacia el este producido por los dos terremotos lentos considerados en la simulación (tabla (4.1.5)), sin embargo, hay que tener en consideración que el modelo consideró mas multi transientes de los que originalmente se consideraron en la simulación, por lo que parte de ese desplazamiento puede estar asociado a desplazamientos incorrectos. Por otra parte si consideramos para nuestro resultado, solo el desplazamiento hacia el este producido por los dos multi transientes principales tendremos una diferencia mas notoria entre ambos desplazamientos (ver fig. 4.1.10 y tabla 4.1.6).

	Simulación	rjMCMC
Desplazamiento [mm]	23.54040839	15.6951378489

Cuadro 4.1.6: Desplazamientos hacia el este según la simulación y el modelo escogido por el rjmc solo considerando los multi transientes en los tiempos 221.946689 y 596.703601

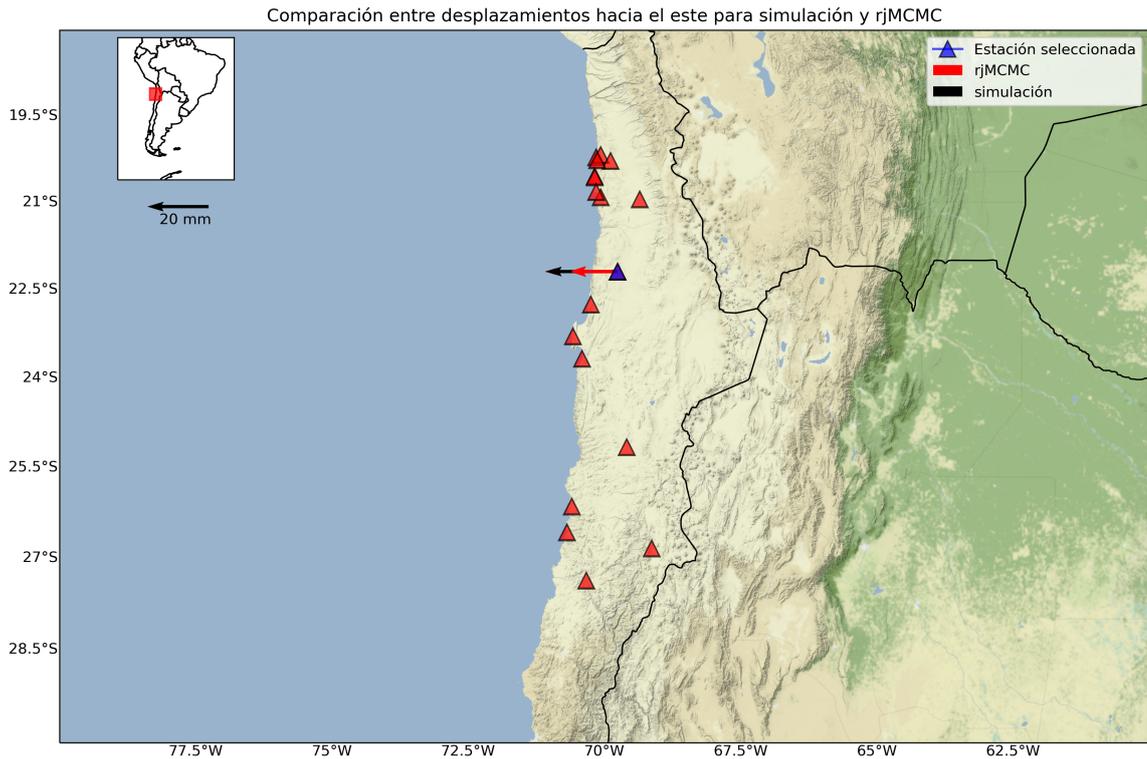


Figura 4.1.10: Comparación entre desplazamientos hacia el este según lo hecho en la simulación y los dos multi-transientes principales encontrados por el rjMCMC

4.2. Análisis de datos reales

Hasta ahora se ha mostrado que para casos sintéticos un tanto ideales (i.e ruido blanco), el método funciona de muy buena manera, pudiendo identificar de manera muy precisa la cantidad de particiones presentes en la serie de tiempo como también la dimensión k de cada una de ellas. Sin embargo, en muchos de los casos sintéticos anteriormente mencionados se conocía la naturaleza del ruido de la señal, lo que no se cumple en su totalidad al momento de considerar series de tiempo reales, de las cuales *a priori* no conocemos en su totalidad como se comporta el ruido, haciendo que resulte interesante probar la formulación que se ha estado planteando para una serie de tiempo obtenida de estaciones GNSS reales.

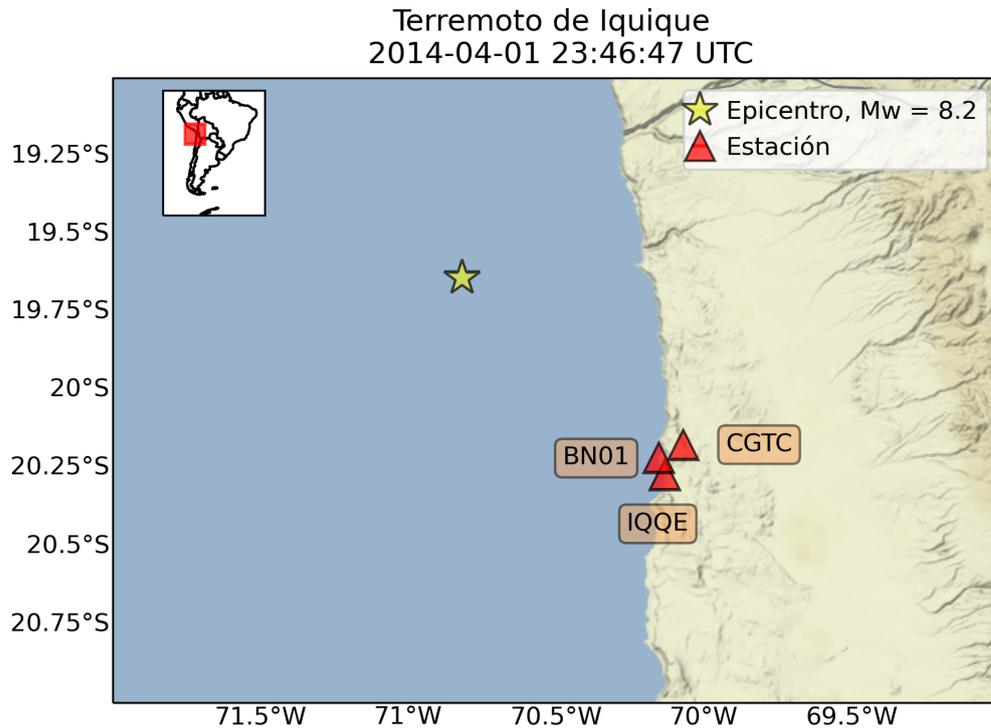


Figura 4.2.1: Distribución espacial de estaciones GNSS utilizadas para el análisis de datos reales

Un caso interesante para poder estudiar, es el del terremoto ocurrido el 01 de abril de 2014 frente a las costas de Iquique, cuya magnitud de momento fue de $M_w = 8.1$. Este terremoto llama la atención puesto que investigaciones previas demuestran que hubo evidencia de actividad precursora a dicho terremoto (Ruiz et al. (2014), Schurr et al. (2014), y en particular se destaca la ocurrencia de terremotos lentos meses antes de la ruptura inicial (Ruiz et al. (2014), Socquet et al. (2017)). Para poder estudiar este terremoto, se ocuparon datos de 3 estaciones GNSS cercanas al epicentro del terremoto, dichos datos fueron obtenidos del laboratorio geodético de Nevada¹ y corresponden a las estaciones CGTC, IQQE y BN04 (ver figura (4.2.1)). En particular, para el análisis de este terremoto se considero una ventana de tiempo entre el 01 de enero de 2013 y el 01 de enero de 2015 con el objetivo de analizar principalmente el periodo previo a la ruptura principal.

¹<http://geodesy.unr.edu/NGLStationPages/gpsnetmap/GPSNetMap.html>

4.2.1. Análisis individual

En primera instancia se analizaron los desplazamientos para la componente E-O de cada estación, ya que es comúnmente donde son mayormente visibles las deformaciones producidas por fenómenos tectónicos. En particular, para las series de tiempo que se expondrán a continuación, se mostrarán solo los tiempos de inicio encontrados para multi transientes, puesto que son dichas deformaciones las que buscamos encontrar, obviando saltos (i.e parámetros locales de dimensión 1) que puedan ser encontrados.

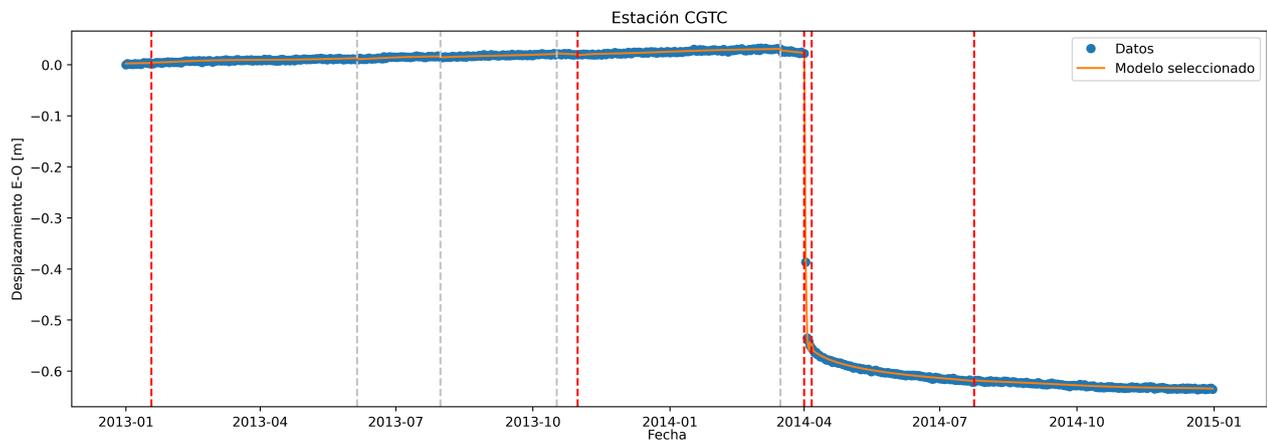


Figura 4.2.2: Serie de tiempo para la componente E-O de la estación CGTC. La curva naranja corresponde a la predicción hecha con el modelo seleccionado por el muestreo del rjMCMC. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un Heaviside puro.

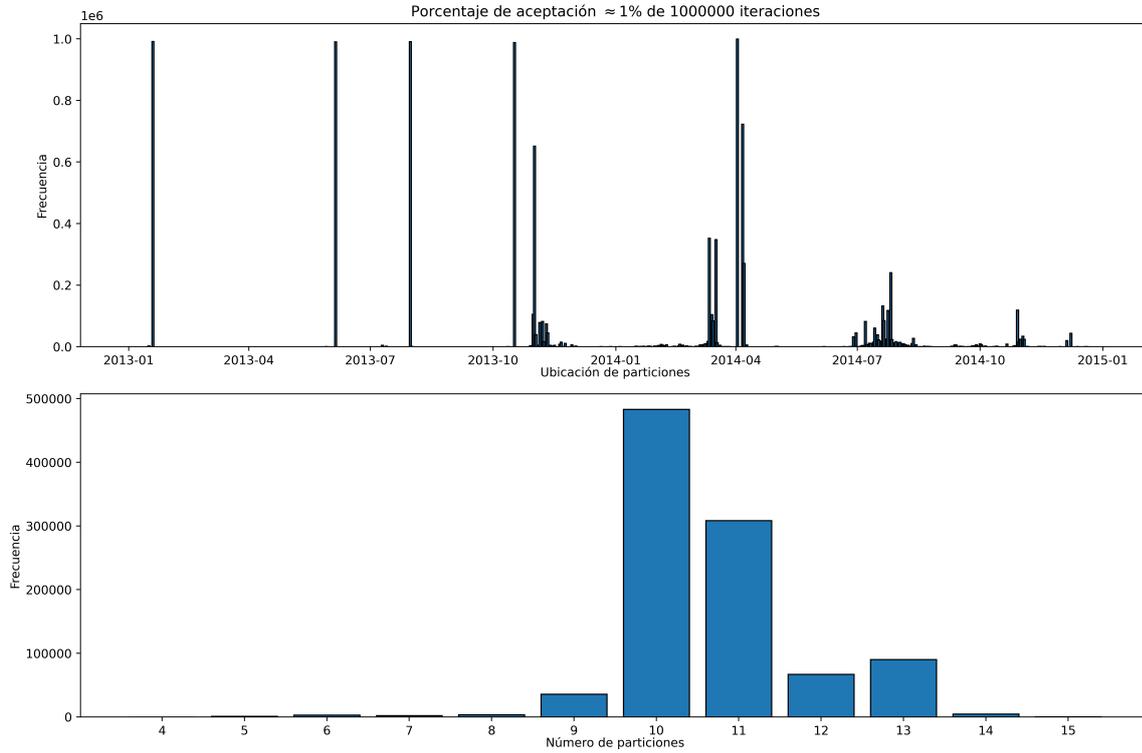


Figura 4.2.3: Histogramas resultantes del muestreo realizado por el rjMCMC para la componente E-O de la estación IQQE. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.

Fecha	Número	Dimensión	Tipo	C_σ
2013-01-01	1	1	Heaviside	5.83
2013-01-19	2	4	Mult. trans.	
2013-06-05	3	2	Mult. trans.	
2013-07-31	4	2	Mult. trans.	
2013-10-17	5	2	Mult. trans.	
2013-11-01	6	4	Mult. trans.	
2014-03-16	7	2	Mult. trans.	
2014-04-02	8	4	Mult. trans.	
2014-04-07	9	4	Mult. trans.	
2014-07-24	10	4	Mult. trans.	

Cuadro 4.2.1: Características del modelo seleccionado por el rjMCMC para la serie de tiempo individual de la componente E-O para la estación CGTC.

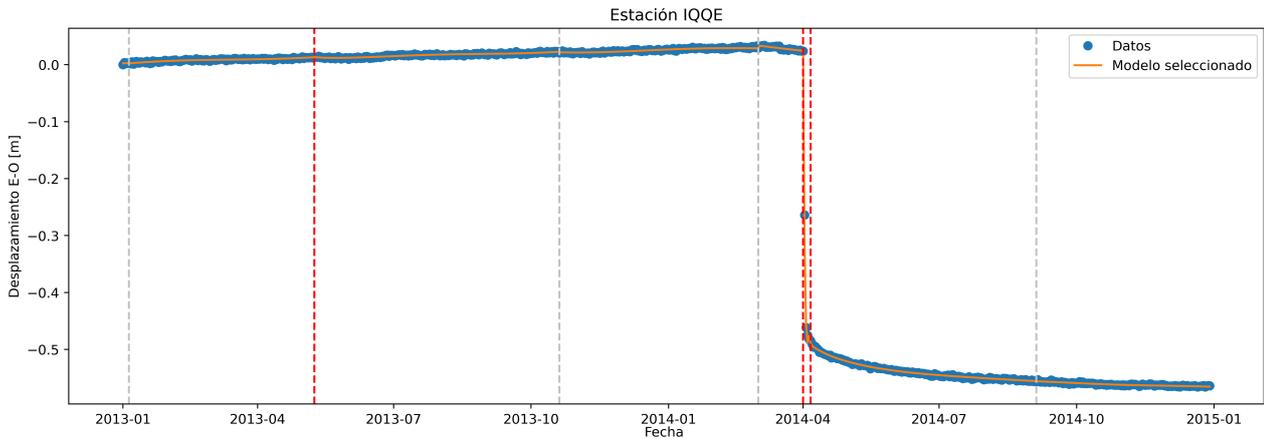


Figura 4.2.4: Serie de tiempo para la componente E-O de la estación IQQE. La curva naranja corresponde a la predicción hecha con el modelo seleccionado por el muestreo del rjMCMC. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un Heaviside puro.

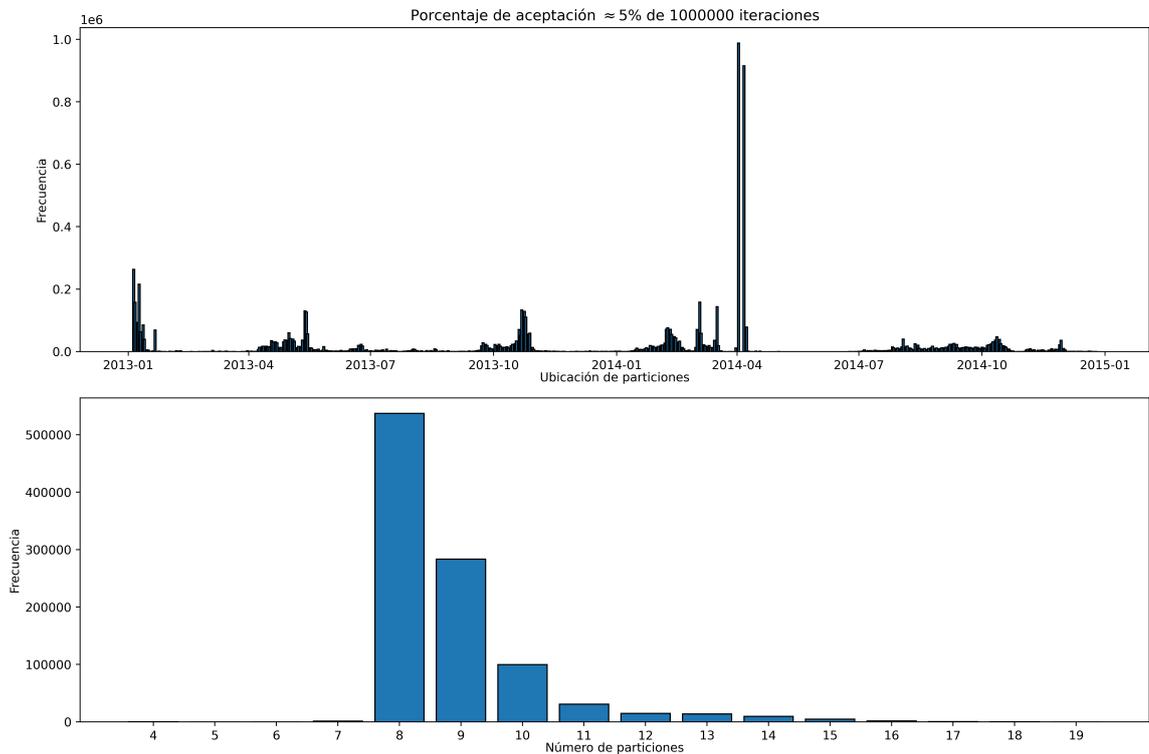


Figura 4.2.5: Histogramas resultantes del muestreo realizado por el rjMCMC para la componente E-O de la estación IQQE. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.

Fecha	Número	Dimensión	Tipo	C_σ
2013-01-01	1	1	Heaviside	5.87
2013-01-06	2	2	Mult. trans.	
2013-05-10	3	4	Mult. trans.	
2013-10-20	4	3	Mult. trans.	
2014-03-03	5	2	Mult. trans.	
2014-04-02	6	4	Mult. trans.	
2014-04-06	7	4	Mult. trans.	
2014-09-05	8	3	Mult. trans.	

Cuadro 4.2.2: Características del modelo seleccionado por el rjmcnc para la serie de tiempo individual de la componente E-O para la estación IQQE.

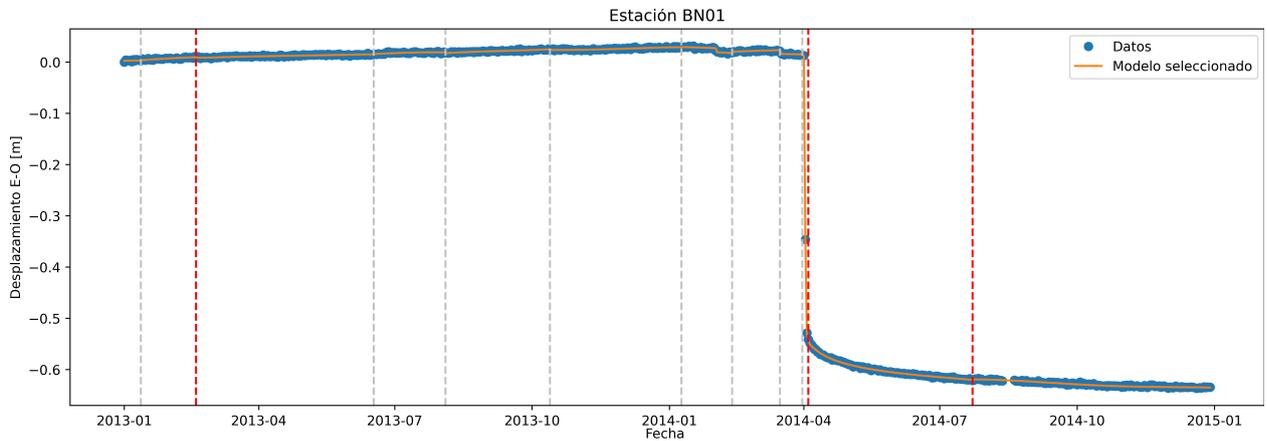


Figura 4.2.6: Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado por el muestreo del rjMCMC. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un Heaviside puro.

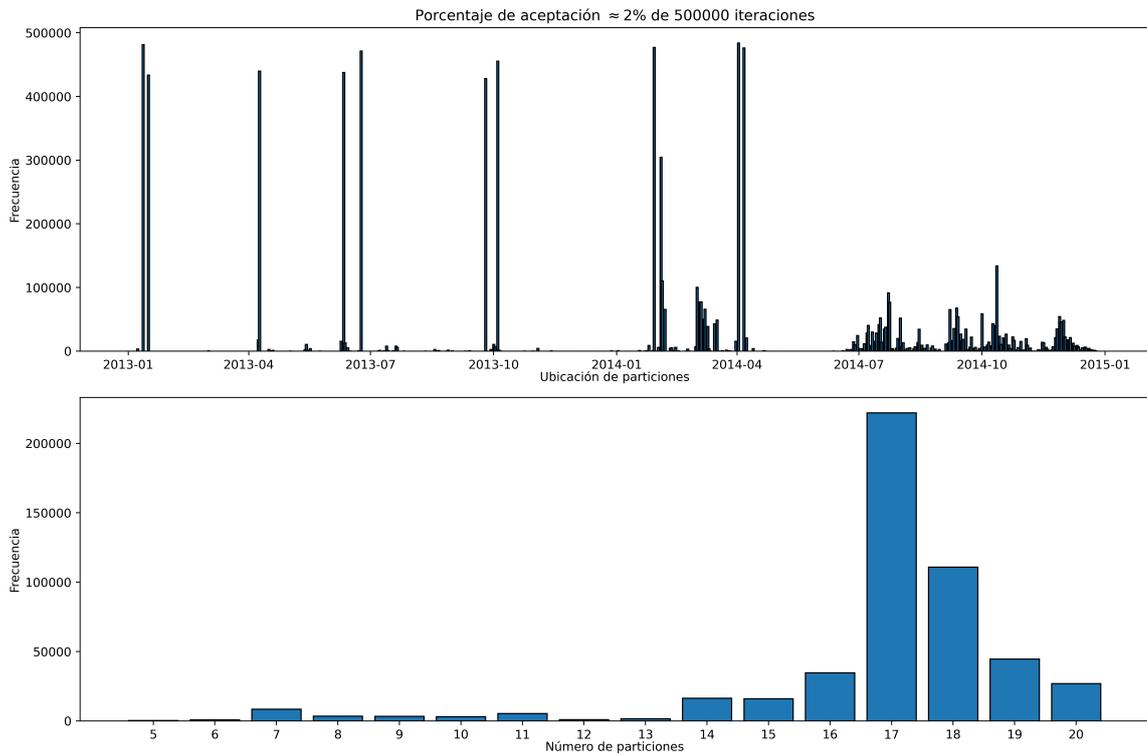


Figura 4.2.7: Histogramas resultantes del muestreo realizado por el rjMCMC para la componente E-O de la estación BN01. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.

Fecha	Número	Dimensión	Tipo	C_σ
2013-01-01	1	1	Heaviside	5.87
2013-01-13	2	2	Mult. trans.	
2013-02-19	3	4	Mult. trans.	
2013-06-18	4	2	Mult. trans.	
2013-08-04	5	2	Mult. trans.	
2013-10-14	6	2	Mult. trans.	
2014-01-09	7	2	Mult. trans.	
2014-02-01	8	1	Heaviside	
2014-02-12	9	2	Mult. trans.	
2014-03-17	10	2	Mult. trans.	
2014-04-01	11	3	Mult. trans.	
2014-04-04	12	4	Mult. trans.	
2014-07-23	13	4	Mult. trans.	

Cuadro 4.2.3: Características del modelo seleccionado por el rjmcnc para la serie de tiempo individual de la componente E-O para la estación BN01.

Podemos ver a simple vista, que en algunas estaciones se encontraron multi transientes en tiempos similares, principalmente desde enero de 2014 hasta el 1 de abril de 2014 que es cuando ocurrió el terremoto Mw 8.1. Destacando entre ellas el terremoto mismo, y un multi transiente aproximadamente dos semanas antes. Sin embargo, fuera de este intervalo de tiempo, no se encuentran eventos compartidos por las 3 estaciones. Pese a lo anterior, si se realiza un análisis en conjunto de ambas componentes horizontales de cada estación podremos encontrar un mayor número de eventos, como también mayor similitud entre las series de tiempo de las 3 estaciones, lo que nos permitiría comparar nuestras estaciones.

4.2.2. Análisis en varias componentes

Una de las aplicaciones de nuestra implementación para el rjMCMC, permite realizar la inversión para múltiples series de tiempo de similares características simultáneamente, de forma que se puedan encontrar particiones en común y por ende, seleccionar un modelo con igual número de multi transientes para ambos sets de datos (la dimensión de cada multi transiente puede variar según la componente). Siguiendo esta idea, es que se realizaron inversiones en conjunto para las componentes horizontales (E-O Y N-S) de cada estación seleccionada. Recordar que solo se mostrarán en las series de tiempo los tiempos de inicio de multi transientes, de modo que si se muestra una partición en una componente que no se señala en la otra, es debido a esto.

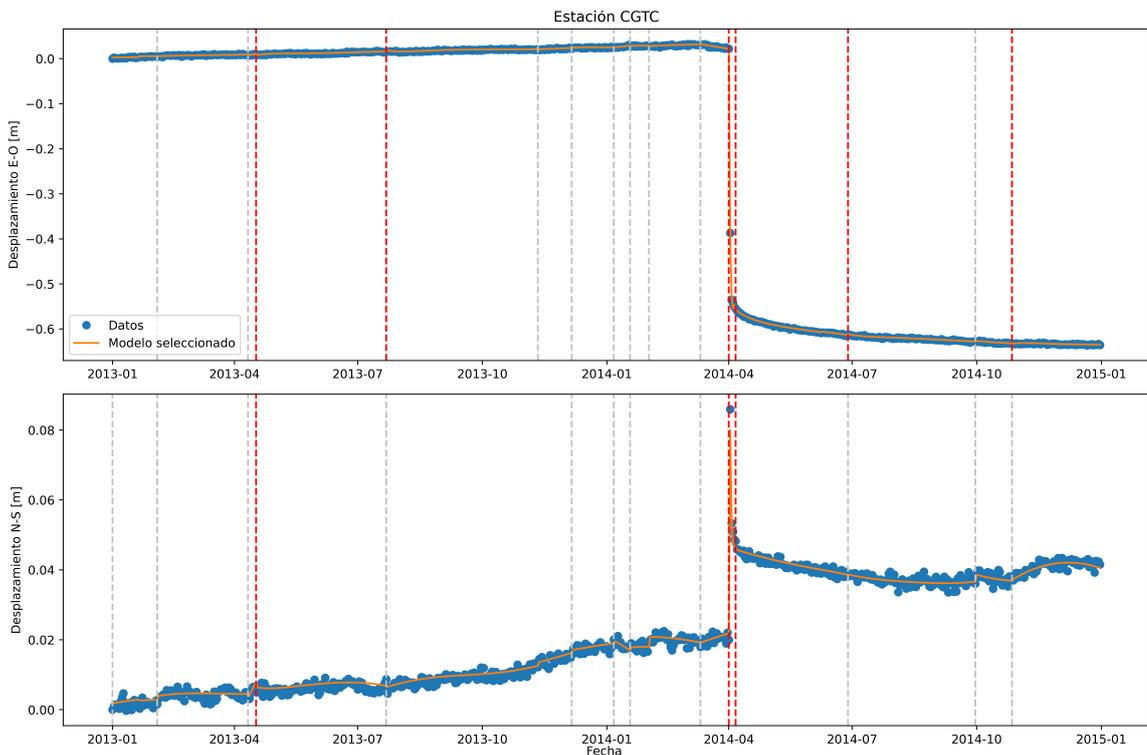


Figura 4.2.8: Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado, siendo el con menor misfit dentro de todos los modelos de la mayor dimensión muestreada. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales mas un heaviside puro.

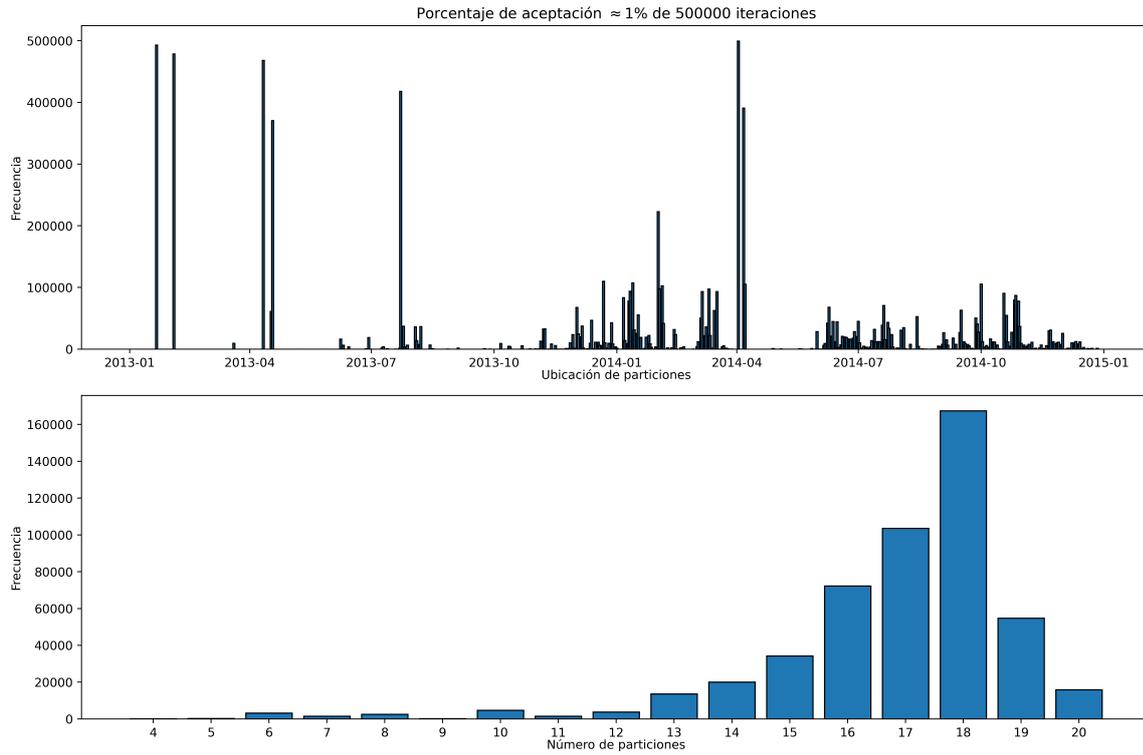


Figura 4.2.9: Histogramas resultantes del muestreo en conjunto realizado por el rjMCMC para la componente E-O y N-S de la estación CGTC. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.

Fecha	N°	Dim. E-O	Tipo E-O	C_σ	Dim. N-S	Tipo N-S	C_σ
2013-01-01	1	1	Heaviside	5.87	2	Mult. trans.	8.38
2013-01-21	2	1	Heaviside		1	Heaviside	
2013-02-03	3	2	Mult. trans.		2	Mult. trans.	
2013-04-11	4	2	Mult. trans.		2	Mult. trans.	
2013-04-18	5	4	Mult. trans.		4	Mult. trans.	
2013-07-23	6	4	Mult. trans.		3	Mult. trans.	
2013-11-11	7	2	Mult. trans.		1	Heaviside	
2013-12-07	8	2	Mult. trans.		2	Mult. trans.	
2014-01-07	9	2	Mult. trans.		2	Mult. trans.	
2014-01-18	10	2	Mult. trans.		2	Mult. trans.	
2014-02-01	11	2	Mult. trans.		1	Heaviside	
2014-03-11	12	2	Mult. trans.		2	Mult. trans.	
2014-04-02	13	4	Mult. trans.		4	Mult. trans.	
2014-04-07	14	4	Mult. trans.		4	Mult. trans.	
2014-06-09	15	1	Heaviside		1	Heaviside	
2014-06-29	16	4	Mult. trans.		3	Mult. trans.	
2014-10-01	17	2	Mult. trans.		2	Mult. trans.	
2014-10-28	18	4	Mult. trans.		3	Mult. trans.	

Cuadro 4.2.4: Características del modelo seleccionado por el rjmc para la serie de tiempo en conjunto para la componente E-O y N-S de la estación CGTC.

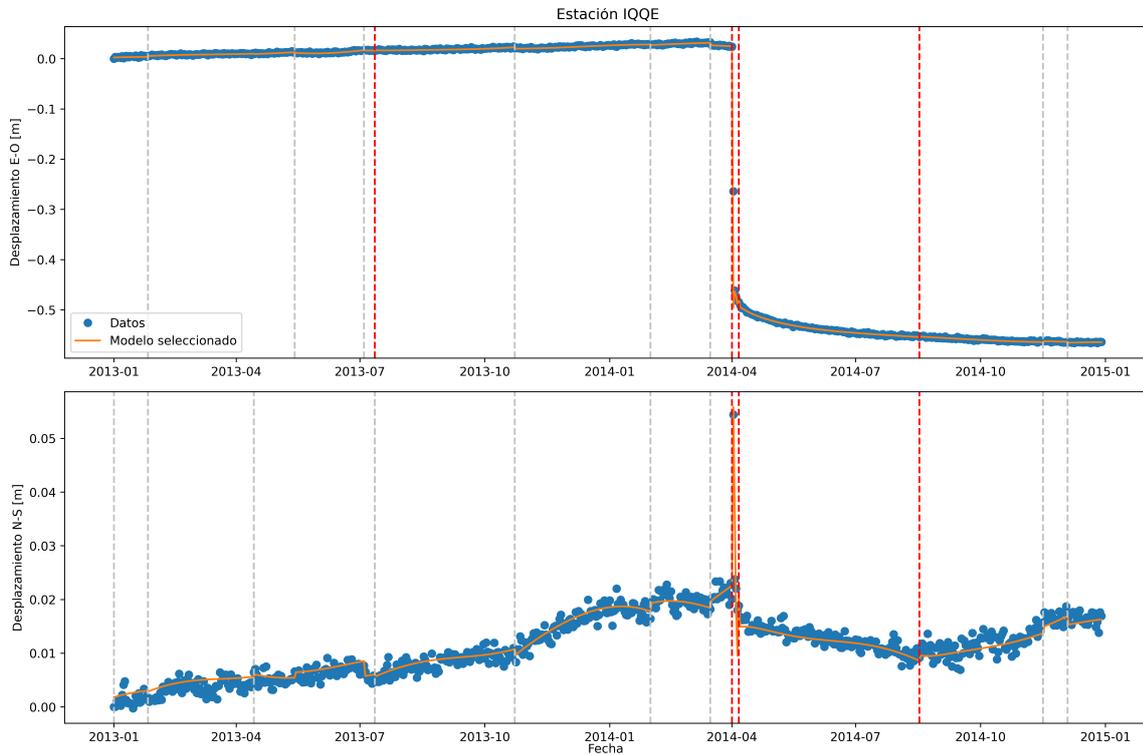


Figura 4.2.10: Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado, siendo el con menor misfit dentro de todos los modelos de la mayor dimensión muestreada. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales más un heaviside puro.

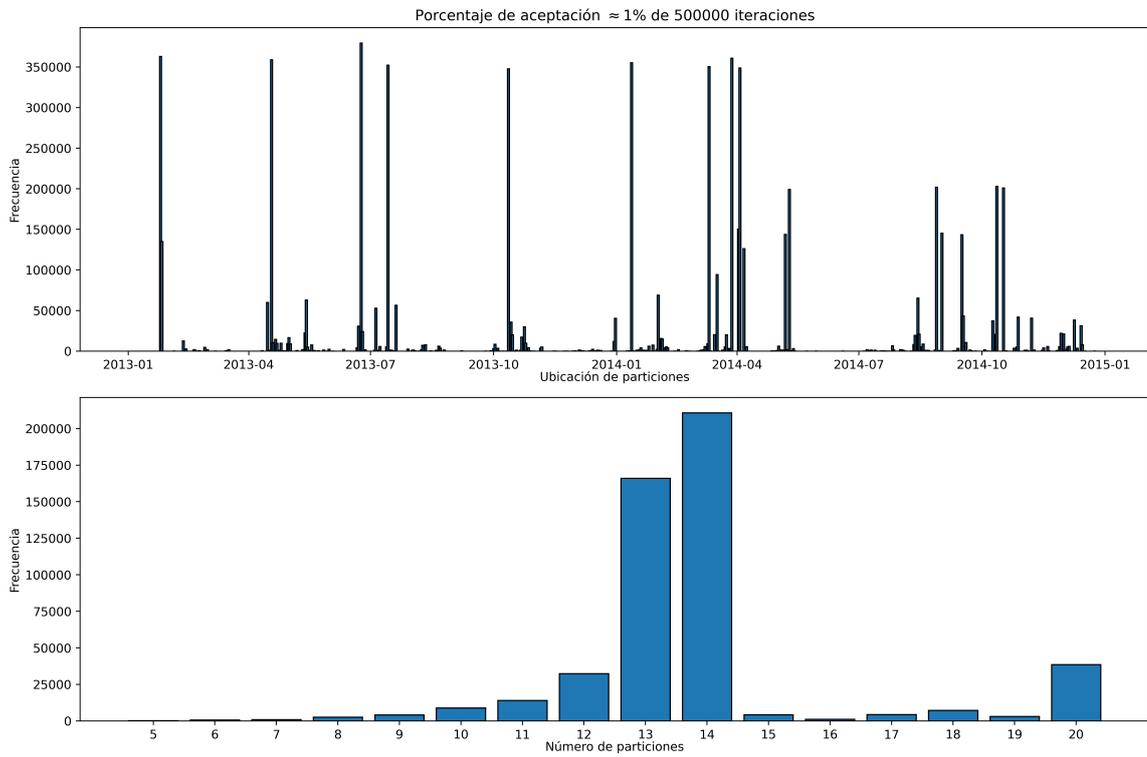


Figura 4.2.11: Histogramas resultantes del muestreo en conjunto realizado por el rjMCMC para la componente E-O y N-S de la estación IQQE. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.

Fecha	N°	Dim. E-O	Tipo E-O	C_σ	Dim. N-S	Tipo N-S	C_σ
2013-01-01	1	1	Heaviside	5.87	2	Mult. trans.	8.38
2013-01-26	2	2	Mult. trans.		2	Mult. trans.	
2013-04-15	3	1	Heaviside		2	Mult. trans.	
2013-05-15	4	3	Mult. trans.		1	Heaviside	
2013-07-05	5	2	Mult. trans.		1	Heaviside	
2013-07-13	6	4	Mult. trans.		2	Mult. trans.	
2013-10-23	7	2	Mult. trans.		3	Mult. trans.	
2014-02-01	8	2	Mult. trans.		2	Mult. trans.	
2014-03-16	9	2	Mult. trans.		2	Mult. trans.	
2014-04-02	10	4	Mult. trans.		4	Mult. trans.	
2014-04-07	11	4	Mult. trans.		4	Mult. trans.	
2014-08-18	12	4	Mult. trans.		4	Mult. trans.	
2014-11-17	13	2	Mult. trans.		2	Mult. trans.	
2014-12-05	14	2	Mult. trans.		2	Mult. trans.	

Cuadro 4.2.5: Características del modelo seleccionado por el rjmcnc para la serie de tiempo en conjunto para la componente E-O y N-S de la estación IQQE.

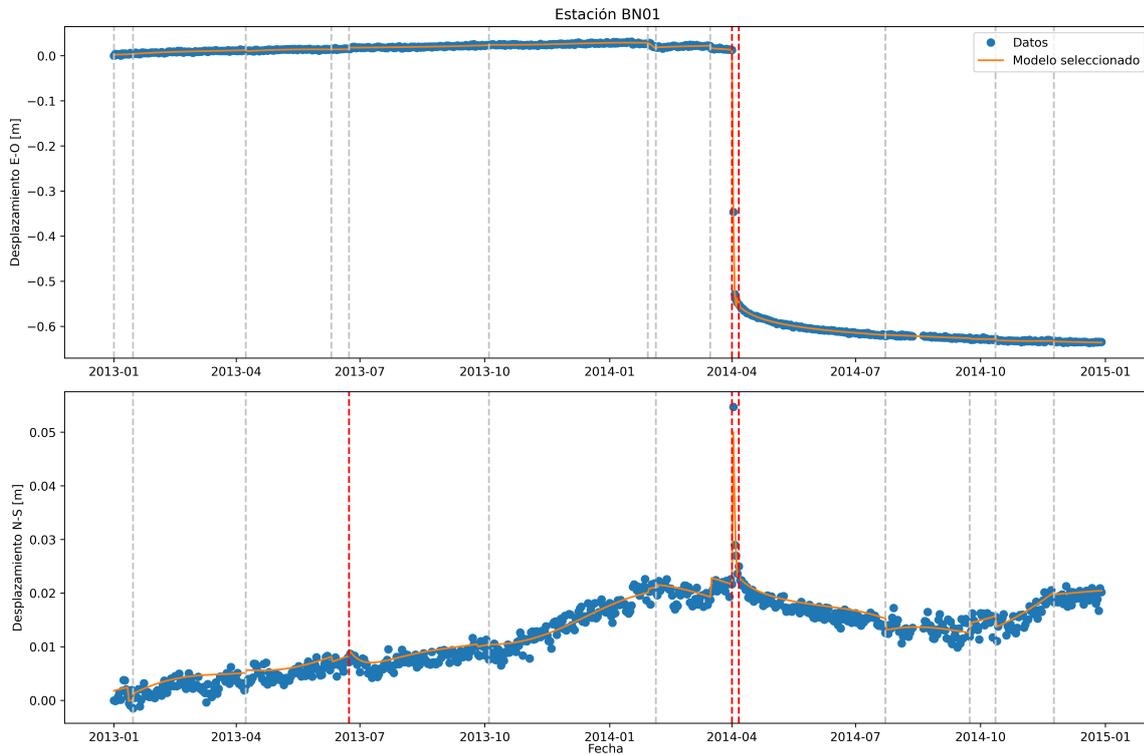


Figura 4.2.12: Serie de tiempo para la componente E-O de la estación BN01. La curva naranja corresponde a la predicción hecha con el modelo seleccionado, siendo el con menor misfit dentro de todos los modelos de la mayor dimensión muestreada. Las líneas punteadas verticales corresponden a los tiempos de inicio para cada multi transiente encontrado en la serie, en particular las líneas rojas corresponderán a los multi transientes de dimensión $k = 4$ es decir, una suma de 3 exponenciales más un heaviside puro.

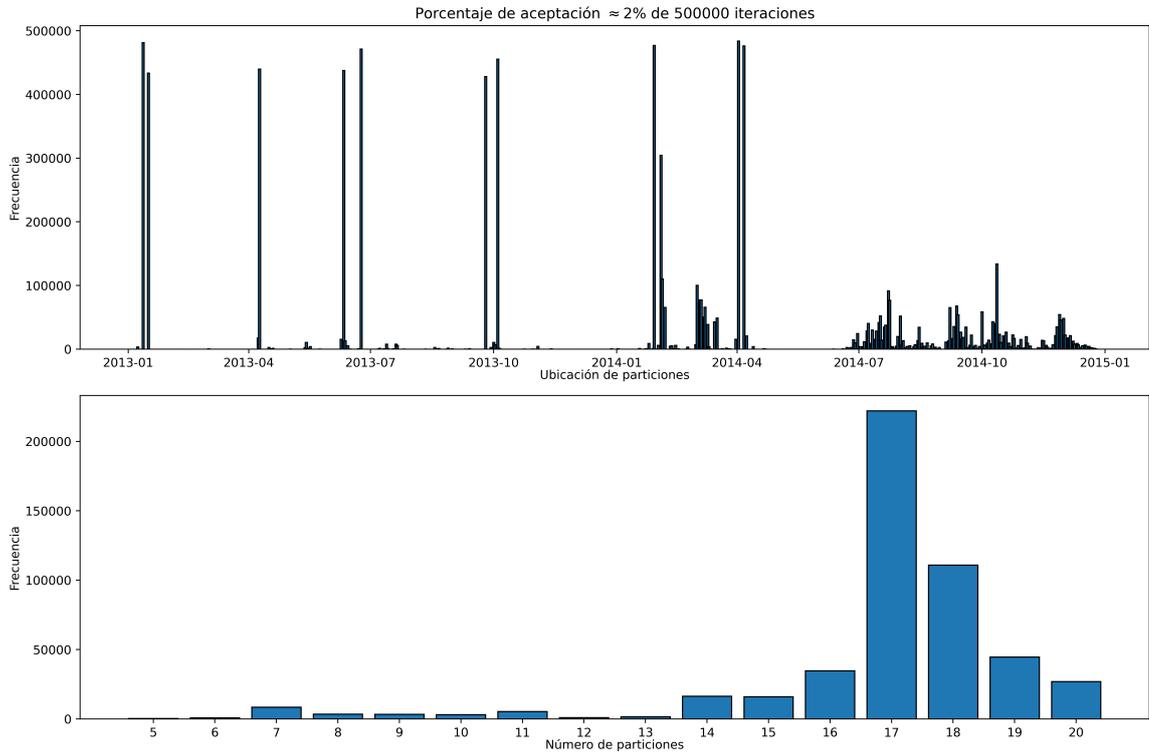


Figura 4.2.13: Histogramas resultantes del muestreo en conjunto realizado por el rjMCMC para la componente E-O y N-S de la estación BN01. El primer histograma muestra cuantas veces se muestrearon modelos con una partición ubicada en el tiempo indicado. El segundo indica la cantidad de veces que se muestreó un número determinado de particiones.

Fecha	Número	Dim. E-O	Tipo E-O	C_σ	Dim. N-S	Tipo N-S	C_σ
2013-01-01	1	2	Mult. trans.	5.87	1	Heaviside	8.39
2013-01-11	2	1	Heaviside		1	Heaviside	
2013-01-16	3	2	Mult. trans.		2	Mult. trans.	
2013-04-09	4	3	Mult. trans.		2	Mult. trans.	
2013-06-10	5	2	Mult. trans.		1	Heaviside	
2013-06-24	6	3	Mult. trans.		4	Mult. trans.	
2013-09-25	7	1	Heaviside		1	Heaviside	
2013-10-04	8	3	Mult. trans.		2	Mult. trans.	
2014-01-29	9	2	Mult. trans.		1	Heaviside	
2014-02-05	10	2	Mult. trans.		2	Mult. trans.	
2014-03-17	11	2	Mult. trans.		1	Heaviside	
2014-04-02	12	4	Mult. trans.		4	Mult. trans.	
2014-04-07	13	4	Mult. trans.		4	Mult. trans.	
2014-07-23	14	2	Mult. trans.		2	Mult. trans.	
2014-09-24	15	1	Heaviside		2	Mult. trans.	
2014-10-12	16	2	Mult. trans.		2	Mult. trans.	
2014-11-25	17	2	Mult. trans.		2	Mult. trans.	

Cuadro 4.2.6: Características del modelo seleccionado por el rjmc para la serie de tiempo en conjunto para la componente E-O y N-S de la estación BN01.

4.3. Comparación GrAtSiD

En [Sambridge et al. \(2006\)](#) se menciona que pese a los detractores que puede llegar a tener la inferencia bayesiana, los resultados obtenidos por este método suelen ser bastante similares a los obtenidos por métodos optimales. Siguiendo esta línea y aprovechando los resultados recién mostrados, los compararemos con resultados obtenidos por un método optimal, en particular el propuesto en la investigación de [Bedford and Bevis \(2018\)](#)(ver Anexo A2). Esta comparación resulta interesante puesto que para este trabajo se ocupó como base el modelo de trayectoria propuesto en dicha investigación. Para hacer la comparación, tomaremos como referencia la estación CGTC, y compararemos los términos globales (movimiento secular mas estacional) y locales (multi transiente mas saltos) de cada modelo.

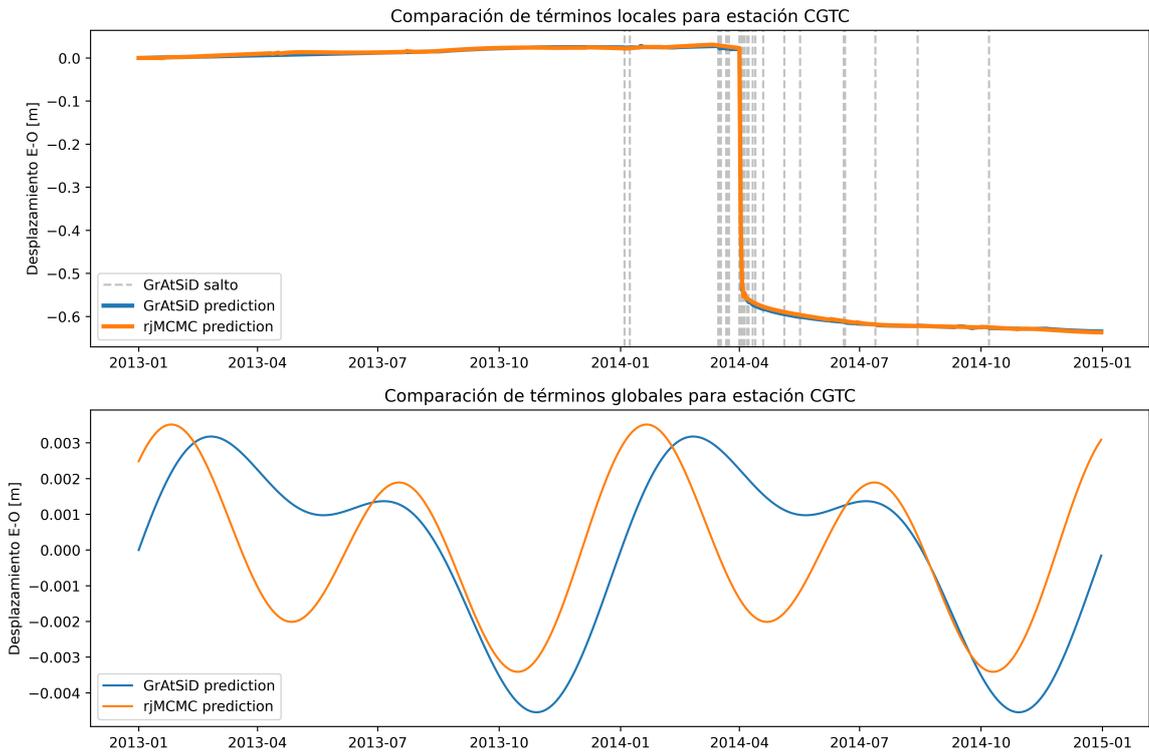


Figura 4.3.1: Comparación entre términos locales y globales de la predicción obtenida mediante el muestreo del rjMCMC con la obtenida mediante GrAtSiD para la componente E-O de la estación CGTC

A simple vista, de la figura 4.3.1 tenemos que para los términos locales no se aprecian grandes diferencias, sin embargo, si se mira con detalle, se ven discrepancias entre ambas predicciones desde del mes de marzo aproximadamente, la razón de esto se puede deber netamente a la cantidad de saltos pre establecidos al momento de hacer la inversión mediante GrAtSiD. Por otra parte, en la componente estacional, se puede ver que coinciden claramente los ciclos anuales y semi anuales dentro de un rango de amplitudes mas o menos similar, sin embargo se muestra un diferencia en las fases entre ambos ciclos.

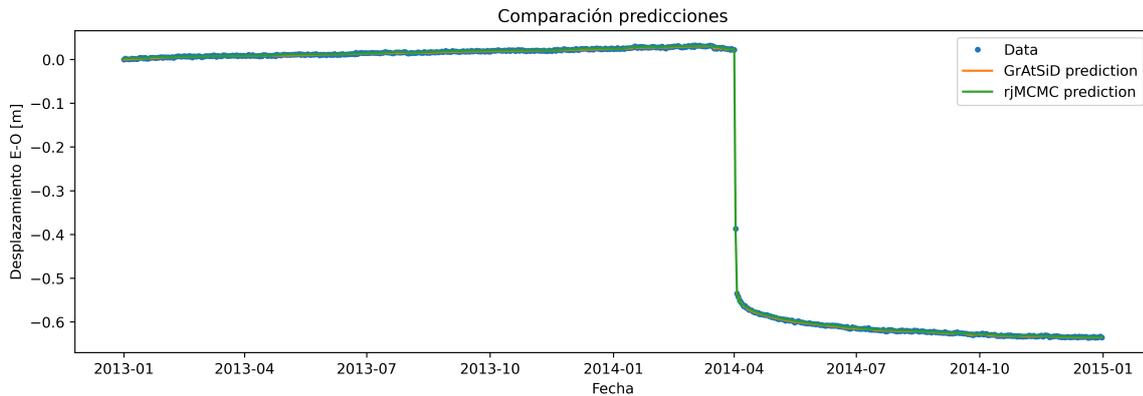


Figura 4.3.2: Comparación de la predicción tanto para el modelo seleccionado por el rjMCMC como para la predicción obtenida por GrAtSiD para la componente E-O de la estación CGTC

De todas formas, al momento de considerar la predicción total (suma de término local y global) encontramos que a simple vista no existen diferencias notorias entre ambas predicciones (ver fig. ??). Esto es bastante interesante, puesto que si bien, ambas predicciones están basadas en un modelo muy similar, tenemos que en nuestra formulación el número de transientes que componen un multi transiente no es fijo, a diferencia del GrAtSiD. Lo anterior, podría decirnos intuitivamente que es en las predicciones de los multi transientes donde existirían mayores diferencias, sin embargo, se mostró que no resultó ser así, mostrando que una formulación bayesiana jerárquica puede ajustar de manera similar que una aproximación optimal.

Capítulo 5

Análisis y Discusión

En nuestro problema, el principal foco para obtener buenos resultados, o que al menos sean coherentes teóricamente, es teniendo un buen conocimiento del ruido que presentan nuestros datos, esto debido a que dicha información nos permitirá escoger una PDF para el *likelihood* que describa de buena manera cómo se distribuye el ruido y por ende, poder hacer una buena estimación del modelo. Como se mostró en la sección de resultados, cuando se tenían casos ideales (i.e ruido distribuido normal) el método lograba encontrar de manera precisa las cantidades de interés, tales como dimensión del modelo, tiempos de inicio de multi transientes o la dimensión de estos. Sin embargo, esto no resulta del todo cierto al momento de realizar la inversión con datos sintéticos que se asemejen a observaciones reales, como con datos obtenidos de estaciones GNSS.

Lo anterior también podemos verlo si comparamos los porcentajes de aceptación entre las simulaciones ideales y las con ruido de características mas realistas, incluyendo los ejemplos con datos reales. Para reglas de Metropolis-Hasting el porcentaje de aceptación óptimo para que el proceso sea eficiente debe ser de alrededor de 23 % (Bédard, 2008), lo que se cumple en gran medida para las tres primeras simulaciones, de 1, 2 y 5 particiones respectivamente, cuyos porcentajes de aceptación fueron de aproximadamente 27 %, 22 % y 22 %. Sin embargo, dicho porcentaje alcanza valores muy bajos para el resto de los casos, teniendo en su mayoría valores cercanos al 2 %, siendo un 5 % el mayor porcentaje de aceptación para casos mas realistas.

Una posible razón de esto radica principalmente en la formulación de nuestro

problema, para nuestro *likelihood* se escogió una PDF normal multi variada, esto nos dice indirectamente que el ruido de nuestros datos se debe netamente a procesos aleatorios e independientes, es decir, no existen correlaciones en el ruido de nuestros datos (ver ecuación 3.3.4). Sin embargo, si hacemos un análisis de los residuales de nuestras predicciones, sin correlaciones, con los modelos seleccionados para las estaciones GNSS, podremos ver que efectivamente existen dichas correlaciones a posteriori. Notar que este proceso se utiliza típicamente para evaluar la consistencia del *likelihood* propuesto.

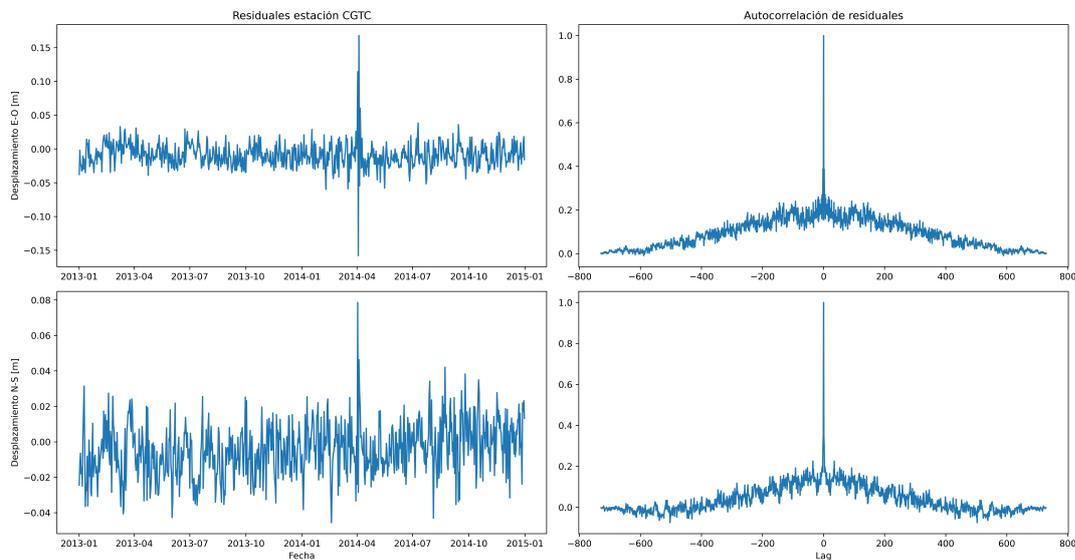


Figura 5.0.1: Residuales para componentes E-O y para componente N-S de la estación CGTC a partir de los modelo seleccionados. Las figuras de la segunda columna corresponden a una auto correlación del residual obtenido

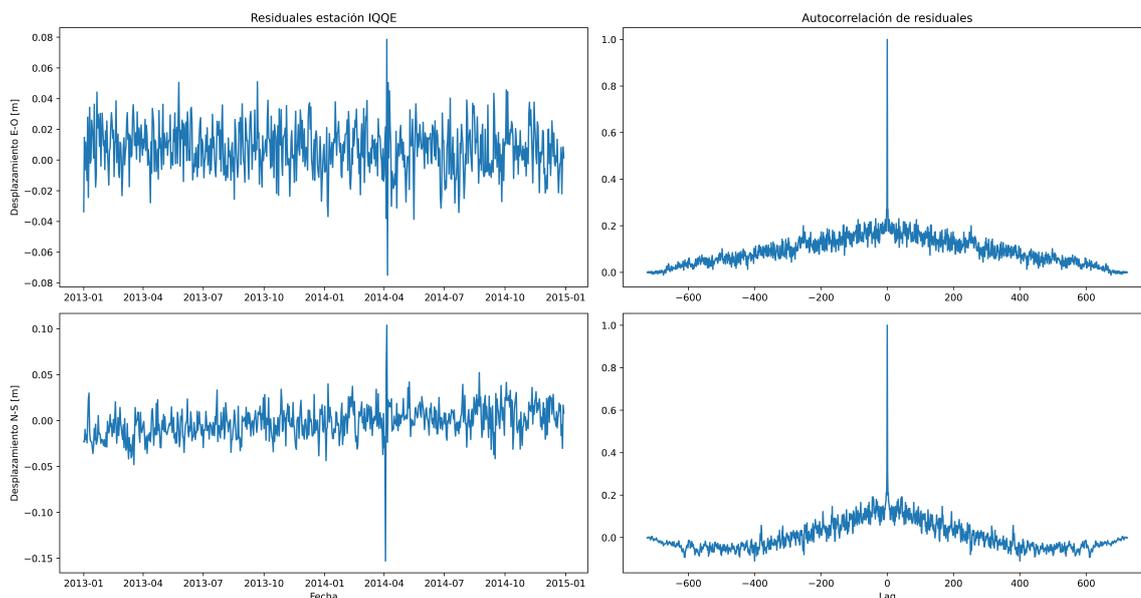


Figura 5.0.2: Residuales para componentes E-O y para componente N-S de la estación IQQE a partir de los modelo seleccionados. Las figuras de la segunda columna corresponden a una auto correlación del residual obtenido

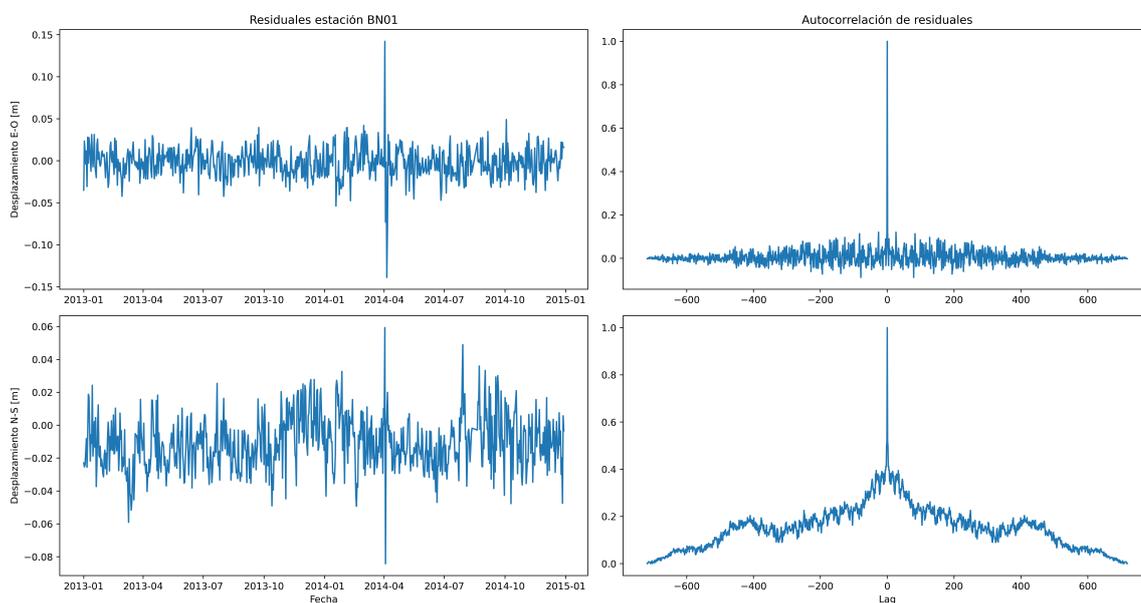


Figura 5.0.3: Residuales para componentes E-O y para componente N-S de la estación BN01 a partir de los modelo seleccionados. Las figuras de la segunda columna corresponden a una auto correlación del residual obtenido

Podemos ver de las figuras (5.0.1, 5.0.2, 5.0.3) que al hacer una auto correlación de los residuales, quedan en evidencia correlaciones temporales muy grandes en el ruido asociado a los datos para todas las estaciones. Lo anterior, puede provocar

que al momento de aplicar a dichas series de tiempo, el rjMCMC con un *likelihood* gaussiano, se encuentren particiones donde no debería, tal como se mostró con el ejemplo sintético de un terremoto lento. Otro punto importante a considerar es el alto número del factor C_σ obtenido en los casos realistas, lo que indicaría una sobre estimación del nivel de ruido al momento de hacer la inversión y que sumado a todo lo anterior, nos indicaría alguna inconsistencia en como abordamos este tema.

Pese a lo anterior, si analizamos en detalle nuestras predicciones podemos encontrar ciertas similitudes entre ellas, principalmente en el análisis en conjunto de las componentes horizontales donde se pueden ver eventos de similares características en intervalos de tiempo relativamente similares. Se muestra de forma notoria que en todas las series de tiempo se encuentra con bastante precisión un multi transiente para la fecha 01/04/2014 (± 1 día) que correspondería al terremoto de Iquique previamente mencionado. También cabe mencionar que en todas las series de tiempo se identifica un segundo multi transiente pocos días después, el cual podría estar asociado a la replica de magnitud Mw 7.6 ocurrida dos días después del evento principal.

Ahora bien, si analizamos el periodo previo a este terremoto nos encontramos con múltiples multi transientes en distintos intervalos de tiempo, que podrían (o no) estar asociados a fenómenos como terremotos lentos. Siguiendo esta idea, se tiene que numerosas investigaciones se encargaron de estudiar ciertas señales previas al terremoto de Iquique, principalmente sísmicas (e.g [Brodsky and Lay \(2014\)](#), [Kato et al. \(2016\)](#), [León-Ríos et al. \(2016\)](#)), por lo que resulta interesante comparar esos resultados con lo estudiado en este trabajo. En [Ruiz et al. \(2014\)](#), identifican un *cluster* de sismos de baja magnitud en su mayoría asísmicos, al sur de la ruptura principal del terremoto de Iquique, durante un periodo de tiempo entre el 4 y 24 de enero, además encuentran otro de características similares durante febrero. [Schurr et al. \(2014\)](#) y [Kato and Nakagawa \(2014\)](#) también analizan este grupo de sismos, y en particular, [Kato and Nakagawa \(2014\)](#) menciona que estos repetitivos sismos ubicados en el contacto de la placa de Nazca con la Sudamericana, indicarían la ocurrencia de un terremoto lento durante este periodo de tiempo. Si analizamos nuestras series de tiempo en conjunto (ver fig. 5.0.4), tendremos que para las 3 estaciones se encuentran multi transientes durante esas fechas, en particular en la estación CGTC se detectan múltiples particiones durante el mes de enero.

Notar que dicho comportamiento también que se puede ver gráficamente como un desplazamiento multi-transiente leve a partir de finales de enero en la estación BN01.

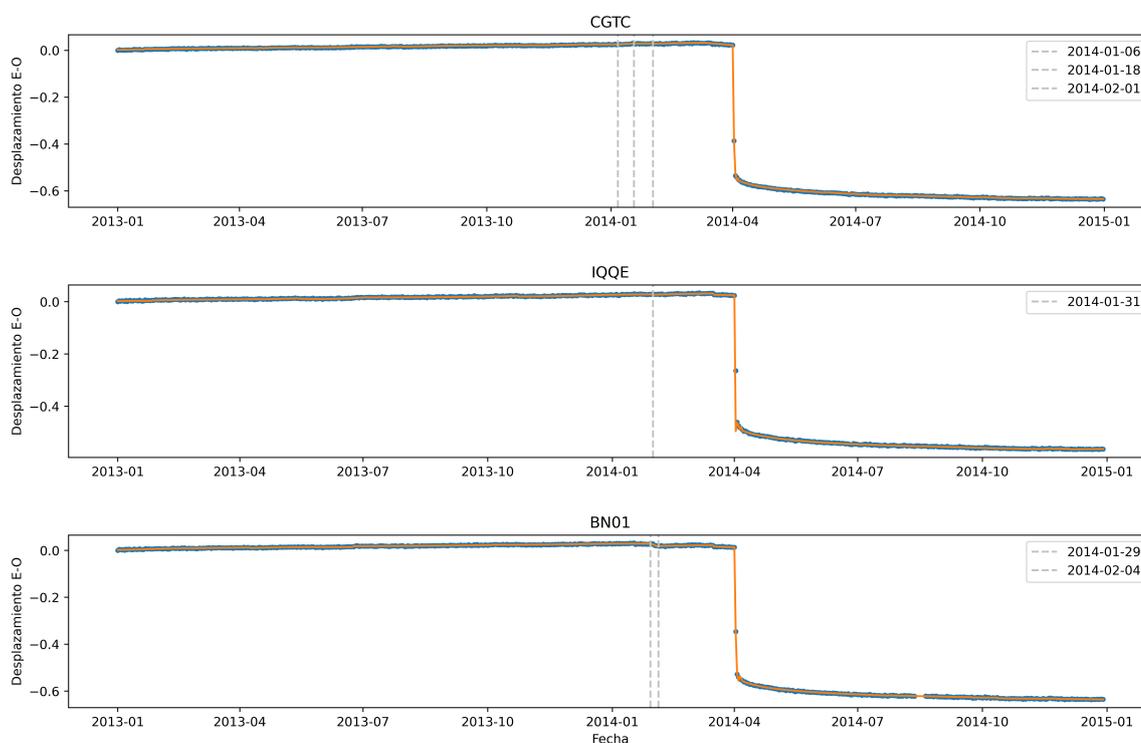


Figura 5.0.4: Multi transientes detectados ente enero y febrero del año 2014 para las estaciones CGTC, IQQE y BN01

Posterior a este *cluster* de eventos, también se identifican otra seguidilla de sismos a mediados de marzo, gatillados principalmente por un sismo Mw 6.7 ocurrido el 16 de Marzo de 2014 siendo el primer gran sismo precursor, y otro el 22 del mismo mes, de magnitud Mw 6.3. Desde la ocurrencia del primer sismo en marzo, hasta el día del terremoto del primero de abril, se tuvo que todas las estaciones continuas GNSS ubicadas en la costa entre Iquique y Pisagua comenzaron a mostrar una deformación transiente leve y a moverse lentamente hacia la fosa (Schurr et al. (2014), Ruiz et al. (2014)). Esta deformación puede ser explicada solo en un 20% por el sismo del 16 de marzo, luego el resto sería atribuido a un deslizamiento lento asísmico (i.e terremoto lento) (Ruiz et al., 2014) que bien podría venir sucediendo desde finales de enero, tal como se muestra en la serie de tiempo de la estación BN01. Si miramos nuestras series de tiempo (ver fig. 5.0.5), podemos también identificar con bastante precisión un multi transiente para fechas alrededor del 16 de Marzo en todas las estaciones analizadas.

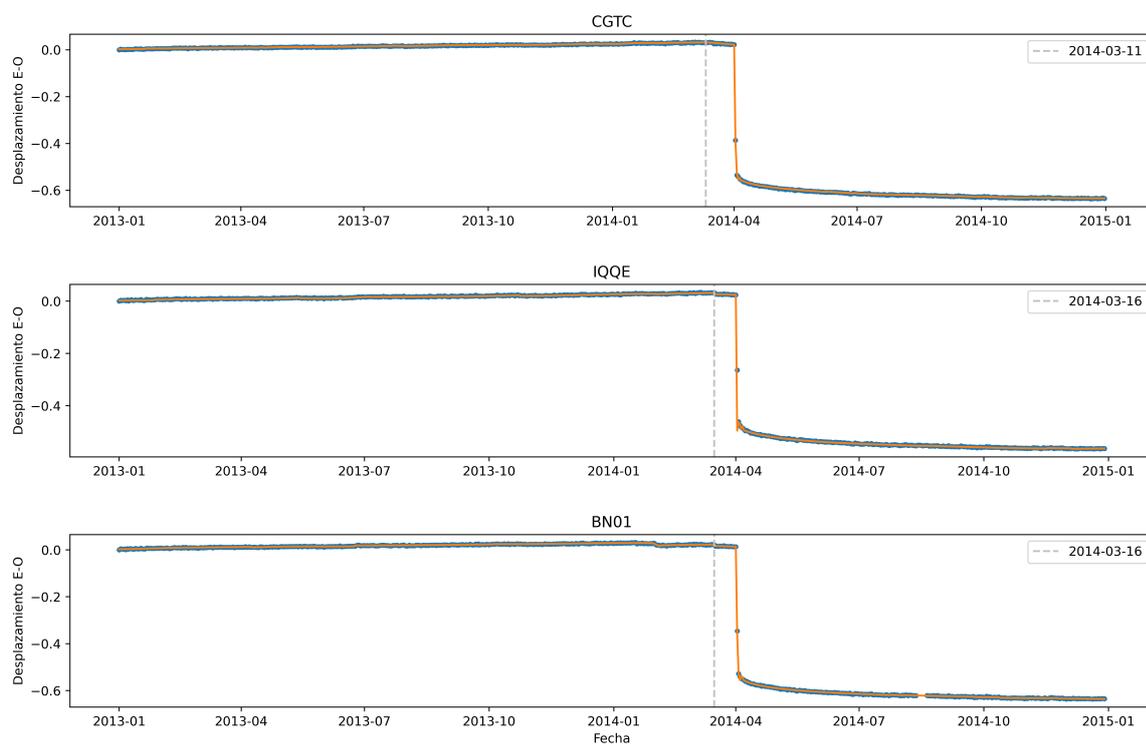


Figura 5.0.5: Multi transientes detectados en el mes de marzo/2014 para las estaciones CGTC, IQQE y BN01

La mayoría de las investigaciones mencionadas anteriormente abarcan el periodo de tiempo comprendido desde enero de 2014 hasta el 1 de abril del mismo año, mostrando evidencia de la ocurrencia de terremotos lentos previos al terremoto principal, sin embargo, si nos remontamos al año 2013 también podemos encontrar evidencia de dichos fenómenos. [Schurr et al. \(2014\)](#) y [Socquet et al. \(2017\)](#) encuentran, en julio de 2013, un *cluster* de sismos de características similares a los de enero y marzo de 2014 de aproximadamente 3 semanas de duración. En particular [Socquet et al. \(2017\)](#) compara estas señales sísmicas, con datos geodéticos y concluye que aproximadamente 8 meses antes del terremoto ocurrió una aceleración en estaciones GNSS ubicadas en la costa norte de Chile, en un área de alrededor de 100km al sur de la ruptura principal del terremoto de Iquique. Dicha aceleración vendría a ser provocada por un deslizamiento lento en la interfaz de las placas, nuevamente, si observamos nuestras series de tiempo (ver fig. 5.0.6), también podemos identificar de manera bastante clara en todas las estaciones, algunos multi-transientes al rededor del mes de julio.

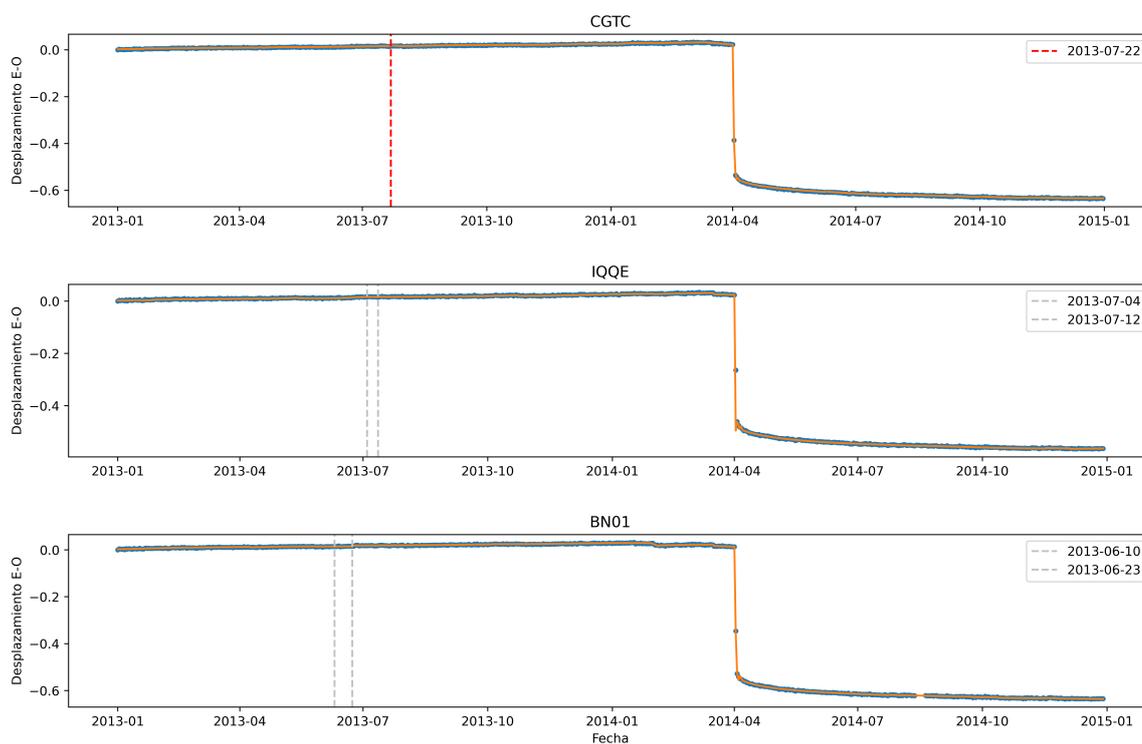


Figura 5.0.6: Multi transientes detectados entre junio y finales de junio de 2013 para las estaciones CGTC, IQQE y BN01

Además, llama la atención que en las estaciones BN01 e IQQE se encuentran también, multi transientes en fechas al rededor del mes de octubre (ver fig. 5.0.7), aproximadamente 6 meses antes del terremoto de Iquique. Estos eventos son interesantes puesto que es posible ver de manera gráfica en la componente N-S que desde esa fecha comienza un incremento notorio en la velocidad del desplazamiento. Esto puede estar asociado al terremoto lento anteriormente mencionado, sin embargo, no se encontró información que respaldara dicha hipótesis. De igual manera, se pueden encontrar multi transientes al rededor del mes de abril de 2013 (ver fig. 5.0.8) en las 3 estaciones, pero nuevamente no se encontró información con la cual contrastar este resultado.



Figura 5.0.7: Multi transientes detectados al rededor del mes de octubre de 2013 para las componentes N-S de las estaciones IQQE y BN01

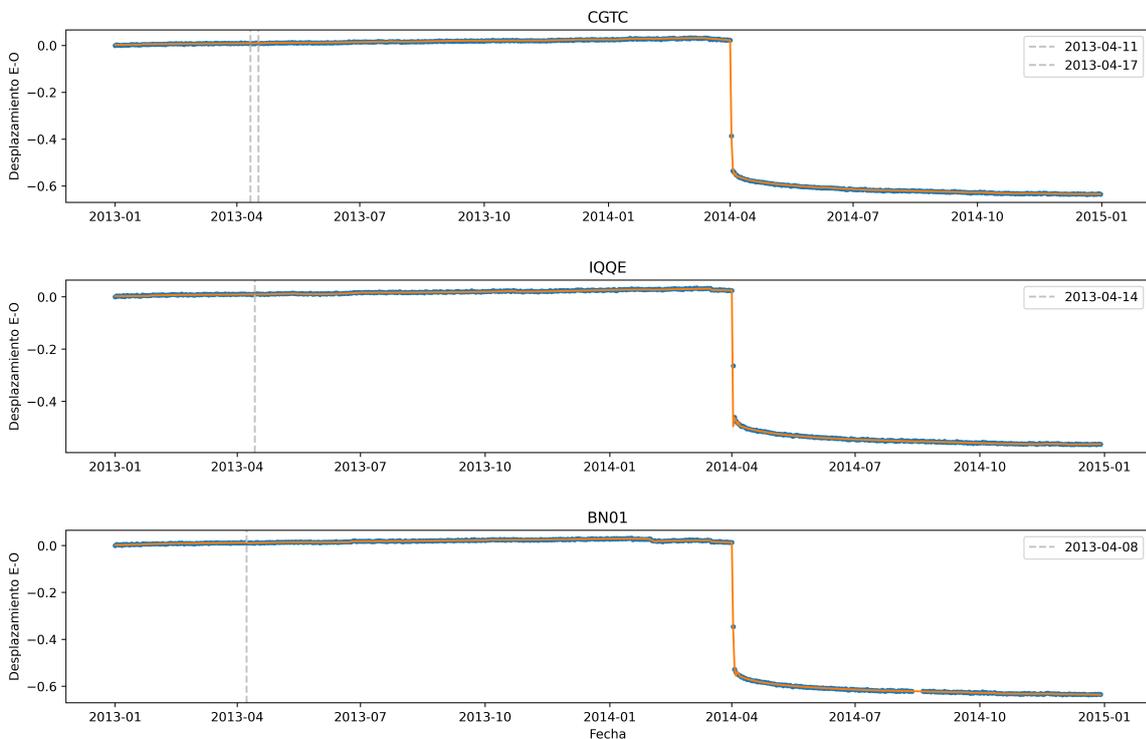


Figura 5.0.8: Multi transientes detectados al rededor del mes de abril de 2013 para las estaciones CGTC, IQQE y BN01

El hecho de que se haya tenido resultados, que de alguna forma se puedan contrastar con investigaciones previas, nos muestra que el método, pese a la simplificaciones consideradas en su formulación (e.g matriz de covarianza diagonal, ruido gaussiano) podría llegar a ser mucho mas efectivo y preciso al incorporar mayor información sobre el ruido. Esto nos permitiría mejorar considerablemente el muestreo de nuestro modelo y por consecuencia mejorar también las predicciones obtenidas. Naturalmente una forma de corregir esto, es incorporando correlaciones a nuestra matriz de covarianza, sin embargo añadir dicha información no es del todo trivial.

En primera instancia, incorporar correlaciones a nuestra matriz de covarianza podría traernos problemas al momento de cambiar de dimensiones durante el muestreo del rjMCMC. La razón de esto se debe a que en nuestra formulación, para el calculo del *proposal* se exige que la matriz de covarianza sea compatible con la descomposición de Cholesky (ver sección 3.3.3), es decir, debe ser definida positiva, lo que no necesariamente puede llegar a cumplirse, por lo que debe ser escogida cuidadosamente. Otro punto que podría verse perjudicado por la adición de correlaciones, es la elección de nuestro *likelihood*, debido a que no se estaría cumpliendo una de las hipótesis del teorema del límite central, que indica que nuestro ruido sería originado solo por causas aleatorias e independientes.

Capítulo 6

Conclusión

Con los resultados obtenidos para series de tiempo reales, tenemos que fue posible recuperar un número determinado de multi transientes de distinta complejidad a partir de observaciones GNSS, tal cómo se planteó en la hipótesis. Sin embargo, es complicado establecer que los tiempos encontrados (i.e particiones encontradas) correspondan efectivamente a una deformación multi transiente real o solo a un ajuste arbitrario producto de una inconsistencia en la definición del ruido. La razón de tener esta consideración recae principalmente en el ejemplo sintético para un terremoto lento, en donde si bien se encontró de manera relativamente precisa la ubicación de ambos terremotos modelados, también se encontraron otros desplazamientos multi-transientes de características similares en la serie, que en un principio no estaban ahí. Sin embargo, al hacer un análisis en conjunto de las componentes horizontales para cada estación, se lograron encontrar similitudes entre los multi transientes encontrados, permitiendo contrastar dichos resultados con investigaciones previas. Lo anterior tiene un peso importante, puesto que nos indicaría que de alguna forma el método está entregando resultados que pueden llegar a ser verídicos y que por medio de una mejor caracterización del ruido se podrían encontrar resultados mas consistentes.

Para reforzar esta última idea se mostró que para las simulaciones de casos ideales (i.e ruido gaussiano), se pudo recuperar con gran precisión todas las características y parámetros del modelo inicial utilizado para generar los datos sintéticos, como por ejemplo la cantidad y la complejidad de multi transientes presentes en la serie de tiempo. Cabe mencionar que en estos casos, se contaba con una correcta

caracterización del ruido, puesto que al momento de generar los datos sintéticos solo se agregó ruido blanco, lo que se reflejó en los resultados obtenidos, mostrando que en estas condiciones el método funcionaría de manera muy efectiva. Lo anterior refuerza la idea de que con una correcta caracterización del problema, el enfoque bayesiano transdimensional fuese una opción viable para resolver el problema inverso asociado a series de tiempo GNSS.

Poder utilizar esta metodología en este tipo de problemas resulta bastante interesante, puesto que la inferencia bayesiana transdimensional presenta una serie de propiedades que podrían simplificar y mejorar la caracterización de las series de tiempo de GNSS. Entre estas propiedades tenemos que se sigue naturalmente el principio de parsimonia, es decir, entre dos soluciones con igual nivel de ajuste, siempre se favorecerá la solución mas simple. Esto permitió modificar el modelo de trayectoria propuesto por [Bedford and Bevis \(2018\)](#), de modo que fuera posible ajustar multi transientes de distinta dimensión y no necesariamente una combinación fija de tres exponenciales como se plantea en su trabajo. Lo anterior tiene una implicancia muy importante puesto que la dimensión de los multi-transientes encontrados nos indicaría de alguna manera su escala temporal, asociada a su constante de decaimiento. Esta selección de modelo natural es contrastable con expresiones analíticas tales como el calculo de la evidencia, en donde ambas metodologías entregaban resultados muy similares.

En general, la mayoría de los problemas en geofísica son mal condicionados, es decir, las mediciones no tienen información suficiente para determinar un modelo de manera precisa, por lo que es común utilizar métodos de regularización para obtener soluciones del problema inverso. Esta práctica en algunos casos puede presentar ciertas complicaciones, debido a que dicha regularización podría introducir algún tipo de sesgo en la solución final. Esto es otro de los motivos por el cual se planteo un enfoque bayesiano para resolver el problema inverso asociado al modelo de trayectoria, de modo que los mismos datos sean los que nos entreguen toda la información sobre el modelo. Ahora bien, como el problema se trataba de uno donde la dimensión no es conocida (no conocemos la cantidad de multi transientes) se recurrieron a técnicas bayesianas transdimensionales, las cuales presentan una serie de ventajas, entre ellas, el hecho de que no sea necesario tener buena información a priori ([Bretthorst, 1988](#)) como también el hecho de que se seleccionarán naturalmente los modelos mas simples ya que

técnicas transdimensionales siguen el principio de parsimonia (Sambridge et al., 2006).

Para poder abordar el problema desde una perspectiva bayesiana transdimensional se utilizó una técnica llamada *reversible jump Markov Chain Monte Carlo*, la cual consiste en una técnica de muestreo que a diferencia de la *Markov Chain Monte Carlo* común, permite caracterizar el *posterior PDF* por medio del muestreo en espacios de distinta dimensión, luego la dimensión del espacio mas visitado corresponderá a la dimensión del modelo seleccionado. Como este método esta basado en una cadena de Markov, los modelos propuestos serán generados a partir de uno inmediatamente anterior siendo comparados mediante una regla de aceptación de tipo *Metropolis-Hasting* (Sambridge et al., 2006) con ciertas modificaciones asociadas al cambio de dimensión entre los espacios visitados.

Si verificamos los objetivos planteados al inicio de este trabajo, tendremos que en su mayoría se cumplieron, siendo el principal, lograr aplicar inferencia bayesiana transdimensional a series de tiempo tanto sintéticas como reales. Y por consecuencia encontrar soluciones al modelo de trayectoria mediante rjMCMC, como también el uso de aplicaciones del método tales como el análisis de mas de una componente. Ahora bien, como ha sido la tónica principal en la sección de discusiones, hay que tener una especial atención en la caracterización del ruido para problemas mas realistas de modo de obtener buenos resultados.

Ahora bien, la metodología propuesta en este trabajo aún se encuentra en desarrollo y en fases experimentales, por lo que los resultados obtenidos sirven como una línea de base para futuras aplicaciones usando inferencia bayesiana transdimensional, tanto para series de tiempo como para el estudio de otros fenómenos de interés (como por ejemplo estimación de *b-value*). Esto también nos indica las directrices a seguir para lograr una mejor implementación del método, enfocadas principalmente a una mejor caracterización del ruido. Incorporando esto a nuestra formulación, se debería mejorar considerablemente el ajuste a los datos, permitiendo un mayor entendimiento y mejor caracterización de los fenómenos que pudieran estar afectando nuestras series de tiempo. Esto con un enfoque principal en el periodo intersísmico, de manera que se pueda lograr un mayor entendimiento de las señales precursoras transientes que puedan llegar a tener lugar en series de tiempo de GNSS, previo a un gran terremoto.

Bibliografía

- Angermann, D., Klotz, J., and Reigber, C. (1999). Space-geodetic estimation of the nazca-south america euler vector. *Earth and Planetary Science Letters*, 171(3):329 – 334.
- Bedford, J. and Bevis, M. (2018). Greedy automatic signal decomposition and its application to daily gps time series. *Journal of Geophysical Research: Solid Earth*, 123(8):6992–7003.
- Bedford, J., Moreno, M., Deng, Z., Oncken, O., Schurr, B., John, T., Baez, J., and Bevis, M. (2020). Months-long thousand-kilometre-scale wobbling before great subduction earthquakes. *Nature*, 580:628–635.
- Bilek, S. L. and Lay, T. (2018). Subduction zone megathrust earthquakes. *Geosphere*, 14(4):1468–1500.
- Bretthorst, G. L. (1988). *The likelihood function (Chapter 2)*, pages 14 – 18. Springer-Verlag New York, United Kingdom.
- Brodsky, E. E. and Lay, T. (2014). Recognizing foreshocks from the 1 april 2014 chile earthquake. *Science*, 344(6185):700–702.
- Bédard, M. (2008). Optimal acceptance rates for metropolis algorithms: Moving beyond 0.234. *Stochastic Processes and their Applications*, 118(12):2198–2222.
- Carr Agnew, D. (2013). Realistic Simulations of Geodetic Network Data: The Fakenet Package. *Seismological Research Letters*, 84(3):426–432.
- Dong, D., Fang, P., Bock, Y., Cheng, M. K., and Miyazaki, S. (2002). Anatomy of apparent seasonal variations from gps-derived site position time series. *Journal of Geophysical Research: Solid Earth*, 107(B4):ETG 9–1–ETG 9–16.
- Ferreira, V., Montecino C., H. D., Ndehedehe, C., Rio, R., Cuevas, A., and De Freitas, S. (2019a). Determining seasonal displacements of earth’s crust in south america using observations from space-borne geodetic sensors and surface-loading models. *Earth Planets and Space*, 71.
- Ferreira, V. G., Montecino, H. D., Ndehedehe, C. E., del Rio, R. A., and Cuevas, Aharon, d. F. S. R. C. (2019b). Determining seasonal displacements of earth’s crust in south america using observations from space-borne geodetic sensors and surface-loading models. *Earth, Planets and Space*, 71(84).

- Freed, A. M., Bürgmann, R., and Herring, T. (2007). Far-reaching transient motions after mojave earthquakes require broad mantle flow beneath a strong crust. *Geophysical Research Letters*, 34(19).
- Fukuda, J. and Johnson, K. M. (2008). A Fully Bayesian Inversion for Spatial Distribution of Fault Slip with Objective Smoothing. *Bulletin of the Seismological Society of America*, 98(3):1128–1146.
- Fukuda, J. and Johnson, K. M. (2010). Mixed linear—non-linear inversion of crustal deformation data: Bayesian inference of model, weighting and regularization parameters. *Geophysical Journal International*, 181(3):1441–1458.
- Govers, R., Furlong, K. P., van de Wiel, L., Herman, M. W., and Broerse, T. (2018). The geodetic signature of the earthquake cycle at subduction zones: Model constraints on the deep processes. *Reviews of Geophysics*, 56(1):6–49.
- Green, P. (2003). *Trans-dimensional Markov chain Monte Carlo (Chapter 6)*, pages 179 – 198. Oxford University Press, United Kingdom.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.
- Heki, K. (2001). Seasonal modulation of interseismic strain buildup in northeastern japan driven by snow loads. *Science (New York, N.Y.)*, 293((5527)):89–92.
- Jiang, Y., Dixon, T., and Wdowinski, S. (2010). Accelerating uplift in the north atlantic region as an indicator of ice loss. *Nat. Geosci.*, 3.
- Kato, A., Fukuda, J., Kumazawa, T., and Nakagawa, S. (2016). Accelerated nucleation of the 2014 iquique, chile mw 8.2 earthquake. *Scientific Reports*, 6:24792.
- Kato, A. and Nakagawa, S. (2014). Multiple slow-slip events during a foreshock sequence of the 2014 iquique, chile mw 8.1 earthquake. *Geophysical Research Letters*, 41.
- Khan, S. A., Wahr, J., Bevis, M., Velicogna, I., and Kendrick, E. (2010). Spread of ice mass loss into northwest greenland observed by grace and gps. *Geophysical Research Letters*, 37(6).
- Klein, E., Bock, Y., Xu, X., Sandwell, D. T., Golriz, D., Fang, P., and Su, L. (2019a). Transient deformation in california from two decades of gps displacements: Implications for a three-dimensional kinematic reference frame. *Journal of Geophysical Research: Solid Earth*, 124(11):12189–12223.
- Klein, E., Bock, Y., Xu, X., Sandwell, D. T., Golriz, D., Fang, P., and Su, L. (2019b). Transient deformation in california from two decades of gps displacements: Implications for a three-dimensional kinematic reference frame. *Journal of Geophysical Research: Solid Earth*, 124(11):12189–12223.

- Knowles, L. A., Bennett, R. A., and Harig, C. (2020). Vertical displacements of the amazon basin from grace and gps. *Journal of Geophysical Research: Solid Earth*, 125(2):e2019JB018105. e2019JB018105 10.1029/2019JB018105.
- Kumazawa, T., Ogata, Y., and Toda, S. (2010). Precursory seismic anomalies and transient crustal deformation prior to the 2008 mw = 6.9 iwate-miyagi nairiku, japan, earthquake. *Journal of Geophysical Research: Solid Earth*, 115(B10).
- León-Ríos, S., Ruiz, S., Maksymowicz, A., Leyton, F., Fuenzalida, A., and Madariaga, R. (2016). Diversity of the 2014 iquique’s foreshocks and aftershocks: clues about the complex rupture process of a mw 8.1 earthquake. *Journal of Seismology*, 20.
- Lin, Y.-n. N., Sladen, A., Ortega-Culaciati, F., Simons, M., Avouac, J.-P., Fielding, E. J., Brooks, B. A., Bevis, M., Genrich, J., Rietbrock, A., Vigny, C., Smalley, R., and Socquet, A. (2013). Coseismic and postseismic slip associated with the 2010 maule earthquake, chile: Characterizing the arauco peninsula barrier effect. *Journal of Geophysical Research: Solid Earth*, 118(6):3142–3159.
- Melbourne, T. I. and Webb, F. H. (2002). Precursory transient slip during the 2001 mw = 8.4 peru earthquake sequence from continuous gps. *Geophysical Research Letters*, 29(21):28–1–28–4.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.
- Michael Bevis, A. B. (2014). Trajectory models and reference frames for crustal motion geodesy. *Journal of Geodesy*, 88:283–311.
- Mitsuhata, Y., Uchida, T., Murakami, Y., and Amano, H. (2001). The Fourier transform of controlled-source time-domain electromagnetic data by smooth spectrum inversion. *Geophysical Journal International*, 144(1):123–135.
- Pollitz, F. F., Peltzer, G., and Bürgmann, R. (2000). Mobility of continental mantle: Evidence from postseismic geodetic observations following the 1992 landers earthquake. *Journal of Geophysical Research: Solid Earth*, 105(B4):8035–8054.
- Reddy, C., Arora, S., Sunil, P. S., and Prajapati, S. (2011). Earthquake related deformation cycle: Perspectives from 2004 sumatra and 2010 chile mega-earthquakes. *Disaster Advances*, 4:13–21.
- Rikitake, T. (1976). Recurrence of great earthquakes at subduction zones. *Tectonophysics*, 35(4):335–362.
- Ruff, L. J. (1996). *Large Earthquakes in Subduction Zones: Segment Interaction and Recurrence Times*, pages 91–104. American Geophysical Union (AGU).
- Ruiz, S., Metois, M., Fuenzalida, A., Ruiz, J., Leyton, F., Grandin, R., Vigny, C., Madariaga, R., and Campos, J. (2014). Intense foreshocks and a slow slip event preceded the 2014 iquique mw 8.1 earthquake. *Science*, 345(6201):1165–1169.

- Sambridge, M., Bodin, T., Gallagher, K., and Tkalčić, H. (2013). Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1984):20110547.
- Sambridge, M., Gallagher, K., Jackson, A., and Rickwood, P. (2006). Transdimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, 167(2):528–542.
- Schatte, P. (1988). On strong versions of the central limit theorem. *Mathematische Nachrichten*, 137(1):249–256.
- Schurr, B., Asch, G., Hainzl, S., Bedford, J., Hoechner, A., Palo, M., Wang, R., Moreno, M., Bartsch, M., Zhang, Y., Oncken, O., Tilmann, F., Dahm, T., Victor, P., Barrientos, S., and Vilotte, J.-P. (2014). Gradual unlocking of plate boundary controlled initiation of the 2014 iquique earthquake. *Nature*, 512.
- Sdrolias, M. and Müller, R. D. (2006). Controls on back-arc basin formation. *Geochemistry, Geophysics, Geosystems*, 7(4).
- Sobrero, F. S., Bevis, M., Gómez, D. D., and Wang, F. (2020). Logarithmic and exponential transients in gnss trajectory models as indicators of dominant processes in postseismic deformation. *Journal of Geodesy*, 94(9).
- Socquet, A., Valdes, J. P., Jara, J., Cotton, F., Walpersdorf, A., Cotte, N., Specht, S., Ortega-Culaciati, F., Carrizo, D., and Norabuena, E. (2017). An 8 month slow slip event triggers progressive nucleation of the 2014 chile megathrust. *Geophysical Research Letters*, 44(9):4046–4053.
- Yabuki, T. and Matsu'ura, M. (1992). Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. *Geophysical Journal International*, 109(2):363–375.
- Ye, N., Roosta-Khorasani, F., and Cui, T. (2019). *Optimization Methods for Inverse Problems*, chapter 2. Springer, Cham.

Apéndice A

Test

A1. Notación

Durante todo el documento se utilizará la siguiente notación:

- Para denotar vectores se ocupará una letra minúscula en negrita, luego:

$$\mathbf{m} = \begin{bmatrix} m_1 \\ \cdot \\ \cdot \\ m_n \end{bmatrix}_{n \times 1} \quad (\text{A1.1})$$

- En el caso de matrices, se denotarán por una letra mayúscula también en negrita, por ejemplo:

$$\mathbf{G} = \begin{bmatrix} G_{1,1} & \cdot & \cdot & G_{1,k} \\ \cdot & \cdot & & \\ \cdot & & \cdot & \\ G_{n,1} & & & G_{n,k} \end{bmatrix}_{n \times k} \quad (\text{A1.2})$$

- El producto vectorial entre matrices y/o vectores se omitirá, por lo que

escribir dos vectores o matrices juntos se interpretara como dicho producto.

$$\begin{bmatrix} G_{1,1} & \cdot & \cdot & G_{1,k} \\ \cdot & \cdot & & \\ \cdot & & \cdot & \\ G_{n,1} & & & G_{n,k} \end{bmatrix}_{n \times k} \times \begin{bmatrix} m_1 \\ \cdot \\ \cdot \\ m_n \end{bmatrix}_{n \times 1} = \mathbf{Gm} \quad (\text{A1.3})$$

- En gran parte del documento se utilizarán probabilidades condicionales para expresar funciones de distribución de probabilidad (PDF), cuya notación sigue la forma:

$$P(\mathbf{A}|\mathbf{B}) \quad (\text{A1.4})$$

Tenemos que el termino $P(\mathbf{A}|\mathbf{B})$ se puede interpretar como la probabilidad de suceda A, sabiendo que también sucede B, es decir "la probabilidad de A dado B". Análogamente, el término $P(\mathbf{B}|\mathbf{A})$ se lee "la probabilidad de B dado A". Por otro lado, el termino $P(\mathbf{A})$ se lee como "la probabilidad de A", lo mismo para $P(\mathbf{B})$.

- El término $P(\mathbf{A}, \mathbf{B})$ corresponderá a una probabilidad conjunta y se lee como "la probabilidad de que ocurra A y B"
- Ahora bien, también se pueden mezclar probabilidades condicionales con probabilidades conjuntas, luego:

$$P(\mathbf{A}, \mathbf{I}|\mathbf{B}), \quad (\text{A1.5})$$

Lo que se puede interpretar como "la probabilidad de que ocurra A e I, dado B". La expresión [A1.5](#) puede escribirse también como una multiplicación de una probabilidad condicional con una probabilidad en conjunto, luego:

$$P(\mathbf{A}, \mathbf{I}|\mathbf{B}) = P(\mathbf{A}|\mathbf{B}, \mathbf{I}) \times P(\mathbf{B}|\mathbf{I}) \quad (\text{A1.6})$$

A2. Greedy automatic signal decomposition (GrAtSiD) method

A grandes rasgos, el método Greedy Automatic Signal Decomposition (GrAtSiD) consiste en un método optimal utilizado por [Bedford and Bevis \(2018\)](#) para la resolución del problema inverso para el modelo de trayectoria descrito en 2.1.4. Este método busca recuperar un número mínimo de funciones multi transientes o heaviside que puedan ser ajustadas como una combinación lineal con los términos globales del modelo, de modo que se minimicen los residuales.

Estas funciones se llamarán *transient onset* (TO), luego en cada iteración del algoritmo ajustará el residual con diferentes combinaciones de dos TO para encontrar uno nuevo TO como candidato. Cada TO corresponderá entonces a una matriz de tres exponenciales (multi transiente) o una función Heaviside.

Un algoritmo “greedy” consiste en buscar iterativamente una función entre un diccionario de funciones (en este caso TO) que mejor pueda ajustar los residuales en la iteración actual, luego cada iteración del algoritmo consistirá en multiples inversiones, que tienen la forma:

$$\mathbf{m} = (\mathbf{G}^T \mathbf{G} + I \cdot \epsilon^2)^{-1} \mathbf{G}^T \mathbf{d} \quad (\text{A2.1})$$

donde \mathbf{m} es el vector de los parámetros del modelo, \mathbf{G} representa las funciones de base, \mathbf{d} corresponderá a la señal que está siendo ajustada, \mathbf{I} corresponde a una matriz identidad y ϵ^2 un parámetro de regularización. Cabe mencionar que como el método consiste en uno de optimización, si se tiene información previa de la ubicación de algún TO, es conveniente considerar dicha información, de modo que la cantidad de incógnitas a encontrar sea mínima.

Como nuestro modelo de trayectoria está basado en el modelo utilizado por [Bedford and Bevis \(2018\)](#) y que fue abordado usando este método optimal, es que buscaremos comparar las predicciones realizadas por el modelo obtenido tanto por inferencia bayesiana transdimensional como por éste método.