



Facultad de Ingeniería Económica, Estadística y Ciencias
Sociales

Escuela de Ingeniería Estadística
Universidad Nacional de Ingeniería
Trabajo de investigación para grado de Bachiller

Implementación del Modelo ETAS para la Simulación de
Réplicas en la placa de Nazca

Estudiante : Lin Chiu Chen Yang
Docente : Grados Paredes Daniel

"Año de la recuperación y consolidación de la economía peruana"
July 15, 2025

Thesis Validation Letter

Dedicatoria

A mi familia, por su apoyo incondicional y amor constante.

A mis amigos, por su compañía y motivación.

A mis docentes, por su guía y conocimientos compartidos.

Agradecimientos

Agradezco profundamente a mi familia por su apoyo incondicional a lo largo de este proyecto. A mis profesores y tutores, por su orientación y paciencia. También quiero expresar mi gratitud a mis compañeros y amigos, quienes con sus ideas y ánimos hicieron posible este trabajo. Finalmente, un reconocimiento especial a todas las personas que, de una u otra forma, contribuyeron a la realización de este informe.

Abstract

Este estudio tiene como objetivo desarrollar y aplicar el modelo de secuencia de réplicas de tipo epidémico (ETAS) para simular y analizar las secuencias de réplicas originadas entre la placa de Nazca y la Sudamericana, permitiendo así una evaluación más precisa del riesgo y una mejor planificación de la respuesta a emergencias en la región de Perú. La investigación comienza con una revisión teórica y metodológica de los modelos en sismología, centrándose en el modelo ETAS, su formulación matemática y los supuestos subyacentes. Se desarrolla una implementación computacional del modelo ETAS utilizando la técnica de máxima verosimilitud para la estimación de parámetros, la cual se valida mediante simulaciones y comparación con datos históricos de secuencias de réplicas. Luego, el modelo se aplica a un conjunto de datos reales de una región específica obtenidos del Instituto Geofísico del Perú (IGP), con un análisis de su precisión para predecir el momento y la ubicación de las réplicas, obteniendo así un mapa de calor de zonas con más riesgo sísmico. Finalmente, la capacidad predictiva del modelo sintético generado se evalúa comparando sus resultados con los datos reales, identificando las fortalezas y limitaciones del modelo ETAS en diferentes escenarios de actividad sísmica.

Keywords: etas model, point process, hawkes process, earthquake forecasting.

Table of Contents

Thesis Validation Letter	i
Dedicatoria	ii
Agradecimientos	iii
Abstract	iv
0.1 Introducción	1
0.1.1 Contexto	1
0.1.2 Problema de Investigación	4
0.1.3 Motivación	4
0.1.4 Objetivos de la Investigación	5
0.1.5 Hipótesis de Investigación	6
0.2 Marco Teórico	7
0.2.1 Antecedentes	7
0.2.2 Marco Conceptual	8
0.2.3 Base Teórica	9
0.3 Metodología	25
0.3.1 Comprensión de Datos	25
0.3.2 Preparación de Datos	26
0.3.3 Ajuste del Modelo	29
0.4 Resultados	31
0.5 Conclusión	31
References	37

Table of Figures

Figure 0-1	Observamos la distribución de la función de intensidad $\Lambda(S)$ (derecha) del proceso puntual. Gráfico generado mediante spatstat de R	15
Figure 0-2	Simulación del proceso de Hawkes con distintos valores de parámetros λ , α y β en R	15
Figure 0-3	Mapa Sísmico del Perú 1960 - 2023 [1].	27
Figure 0-4	Diagrama de dispersión de magnitudes de eventos con el tiempo.	28
Figure 0-5	Gráfico de epicentros de terremotos en un mapa base coloreado por profundidad y con tamaño proporcional a la magnitud. Los colores cálidos indican profundidades bajas y los colores fríos eventos más profundos. Se puede encontrar el código para producir este gráfico en los anexos.	29
Figure 0-6	Example figure that covers the width of the page obtained from PGFPlots [2].	30
Figure 0-7	se demuestra que se cumple la ley de G-R Gutenberg-Richter.	34
Figure 0-8	Eventos provocados en la simulación.	35
Figure 0-9	Análisis de Residuos.	36

Table of Tables

Table 0-1	Ejemplo de un catálogo de terremotos26
Table 0-2	Resutados del ajuste de parámetros31

0.1 Introducción

0.1.1 Contexto

Los terremotos representan uno de los fenómenos naturales más devastadores y difíciles de predecir en el mundo. A lo largo de la historia, estos eventos han causado pérdidas humanas y materiales significativas, generando la necesidad imperante de desarrollar métodos efectivos para su predicción y mitigación de riesgos. La predicción de terremotos es una rama de la ciencia de la sismología que estudia los terremotos, disciplina que llegó a ser científica a partir de la segunda mitad del siglo XIX. Es cuando el deseo de encontrar herramientas que permitan estudiar el fenómeno natural llevó en 1880 a John Milne, famoso ingeniero británico, a inventar el sismógrafo ([3]). Se han planteado distintos métodos de predicción desde anomalías en el campo electromagnético de la tierra o perturbaciones de la ionosfera ([4]) hasta el comportamiento animal al sentir las ondas P momentos previos al fenómeno ([5]) pero no es hasta la década de los setenta que empezaron los estudios estadísticos representándolos como procesos estocásticos asimismo esta disciplina se ha extendido no solo en la geofísica sino también a través de múltiples campos de la ciencia como la sismología fotónica donde últimas investigaciones sugieren el uso de tecnologías emergentes a base de fibra óptica para su estudio ([6]).

¿Qué es la predicción de un terremoto? Según el Diccionario Oxford, las palabras predicción y pronóstico son sinónimos, mientras que para la sismología son

términos que marcan diferencia. El pronóstico de terremotos se refiere a estimaciones a largo plazo durante un período de tiempo prolongado, que puede ir desde meses hasta décadas, en cambio, la predicción de terremotos generalmente se define como la especificación del tiempo, lugar y magnitud de un futuro terremoto dentro de límites establecidos a corto plazo pueden ocurrir en días, semanas o meses ([7]); por su naturaleza estocástica la capacidad de predecir con precisión queda corta por no decir imposible, por eso ahora en adelante la investigación se centra netamente en el pronóstico de eventos sísmicos más no predicción.

Se han planteado varios métodos de pronóstico ([8]) propone estudiar los eventos de forma epidemiológica planteando así el modelo ETAS de carácter temporal mas tarde en conjunto con ([9]) se desarrolló el modelo de tipo espacio-temporal que se detallará en las siguientes secciones del documento.

El Perú se encuentra ubicado en una región sísmicamente activa exactamente el cinturón de fuego del pacifico debido a la convergencia de dos placas tectónicas entre la Placa de Nazca y la Sudamericana. Este proceso llamado subducción, donde la Placa de Nazca se hunde debajo de la Sudamericana acumulando tensiones, es responsable de la generación de terremotos frecuentes y actividad volcánica en las costas del Pacifico incluyendo Colombia, Ecuador, Perú, Bolivia y Chile. En el artículo de ([10]) se detallan los grandes terremotos ocurridos a lo largo de la zona de subducción peruana en los segmentos centro y sur, como Lima-Callao de 1746 ($8.8M_W$) y el terremoto de Pisco de 2007 ($8.0M_W$) respectivamente.

Al estar en una área con alta frecuencia sísmica, lo que genera un peligro continuo de desastres naturales que impactan notablemente a su población, infraestructura y economía. De acuerdo con información del Instituto Nacional de Defensa Civil, más del 70% de los peruanos reside en regiones sumamente susceptibles a sismos. En estas circunstancias, la creación de herramientas científicas que permitan anticipar estos eventos se establece como una prioridad para la nación.

El pronóstico de sismos, a diferencia de la predicción determinista, se fundamenta en modelos estadísticos que analizan la probabilidad de que ocurran eventos futuros basándose en patrones históricos y en estructuras geológicas conocidas. Este enfoque capacita a los responsables de la toma de decisiones para desarrollar planes de contingencia, fortalecer infraestructuras vitales y elaborar políticas de desarrollo urbano más resistentes. La implementación del modelo ETAS en Perú brinda la oportunidad de crear mapas de riesgo que sean dinámicos y adaptables, constituyendo un progreso notable en comparación con los métodos convencionales.

Asimismo, la adopción de modelos como ETAS fomenta la integración de áreas como estadística, informática y geofísica, promoviendo un enfoque interdisciplinario en la administración del riesgo de desastres. Esto ayuda a mejorar las habilidades técnicas en entidades como el IGP, el CENEPRED y el INDECI, realzando la importancia de la ciencia en el desarrollo de políticas públicas.

Previo a la implementación del modelo ETAS, se han propuesto distintos métodos para representar la aparición de sismos y sus réplicas. Entre estos métodos se encuentran los modelos de Poisson homogéneo, que son populares por su sencillez matemática, aunque son deficientes al no poder reflejar la agrupación temporal de los eventos sísmicos. Luego se presentaron modelos de renovación y modelos de tiempo variable como el proceso de Cox, que añaden un elemento aleatorio a la tasa de ocurrencia, pero que todavía tienen limitaciones al no incluir adecuadamente la interacción entre diferentes eventos.

Más recientemente, se han analizado modelos que utilizan redes neuronales profundas, aprendizaje supervisado e enfoques híbridos con metodologías bayesianas. A pesar de que estas técnicas tienen el potencial de identificar patrones complejos, su implementación sigue siendo restringida por la falta de datos etiquetados y la dificultad de interpretar los resultados. Por otro lado, el modelo ETAS se muestra como un equilibrio entre complejidad estadística y aplicabilidad práctica, proporcionando una representación que es comprensible y ha sido validada empíricamente

respecto a la actividad sísmica.

Por lo tanto, esta investigación tiene como objetivo no solo aplicar el modelo ETAS en el ámbito nacional, sino también enriquecer el debate metodológico sobre la viabilidad de los modelos estocásticos en regiones geológicas específicas, tales como la Placa de Nazca.

0.1.2 Problema de Investigación

La comprensión de terremotos sigue siendo un desafío primario de la sismología; estas traen importantes implicaciones sociales y económicas. A pesar de décadas de investigación, aún no es posible determinar con exactitud cuándo, dónde y con qué intensidad ocurrirá un sismo. Esto se debe a la naturaleza compleja de los fenómenos sísmicos y a la interacción de múltiples factores geológicos, físicos u otros que lo preceden; lo cual se realiza el StatSei (The International Statistical Seismology conference) evento bienal que reúne a investigadores y expertos en el campo de la sismología estadística, esta investigación se apoya en varios de los artículos propuestos donde se abordará principalmente la posibilidad de su pronóstico mediante modelos estadísticos que simulan la sismicidad, acercándonos a la posición de ([11]) "¿Es posible la predicción?".

0.1.3 Motivación

Posterior al evento sísmico surgen las réplicas, también conocidas como aftershocks, ocurren después de un terremoto principal y pueden tener una magnitud menor, pero aún representan un riesgo significativo para la población y las infraestructuras afectadas ([12]) las cuales pueden permanecer en pie frente al terremoto principal pero la llegada de oleadas de réplicas consecutivas sobrepasan el límite de estrés de la infraestructura haciendo que colapse sin previo aviso es por eso que el estudio de los modelos predictivos es indispensable para la sociedad, y su precisión conlleva al dilema de ¿alarmar? ¿O no alarmar? Advertir de un terremoto

que no ocurre también conlleva un costo: no solo el costo civil y económico. Entidades públicas como el IGP realizan investigaciones científicas con finalidad de la prevención, reducción de riesgos de desastres orientando al Estado promover políticas para proteger la vida de la población. Por otro lado en ([13]) y ([14]) ambos autores proponen la importancia de la completitud y consistencia de los datos para el análisis de sismicidad pero factores no controlables como la calibración instrumental, el ruido externo disminuyen la calidad de estos por eso la realización de simulaciones pueden ayudar a mejorarlos; también, en la revisión de literatura no se encontró la implementación del modelo en la región de la placa de Nazca por consiguiente esta investigación intenta replicar un modelo descriptivo propuesto por ([15]) para comparar la sismicidad de la región y los estudios anteriores.

0.1.4 Objetivos de la Investigación

Objetivos Generales

Desarrollar y replicar el modelo ETAS (Epidemic-Type Aftershock Sequence) para simular réplicas sísmicas seleccionando como región la costa del Pacífico.

Objetivos Específicos

- **Revisión Teórica y Metodológica:** Realizar una revisión exhaustiva de la literatura sobre modelos temporales y de procesos espacio-puntuales aplicados a sismología, con énfasis en el modelo ETAS. Describir los fundamentos teóricos del modelo ETAS y sus variantes, incluyendo la formulación matemática y los supuestos subyacentes.
- **Implementación del Modelo ETAS:** Desarrollar una implementación computacional del modelo ETAS utilizando técnicas paramétricas para la estimación de parámetros. Validar la implementación mediante simulaciones y

comparación con datos históricos de secuencias de réplicas.

- **Aplicación a Datos Reales:** Aplicar el modelo ETAS a un conjunto de datos reales de secuencias de réplicas sísmicas de una región específica. Analizar la precisión del modelo en la predicción de la ocurrencia de réplicas en términos de tiempo y ubicación.
- **Evaluación y Comparación:** Evaluar la capacidad predictiva del modelo ETAS comparando sus resultados con otros modelos existentes. Identificar las fortalezas y limitaciones del modelo ETAS en diferentes escenarios de actividad sísmica.

0.1.5 Hipótesis de Investigación

Se analizará si el modelo ETAS proporcionará simulaciones de réplicas sísmicas coherentes con los datos históricos en la región de la costa del Pacífico, permitiendo una mejor identificación de focos de riesgo. También si la inclusión de datos de sismicidad simulada mejorará la calidad y la completitud de los datos utilizados en el análisis de sismicidad, superando las limitaciones impuestas por factores no controlables.

0.2 Marco Teórico

0.2.1 Antecedentes

La pregunta es como se podrá desarrollar un modelo capaz de simular fenómenos tan caóticos como los sismos, gracias a su naturaleza estocástica se puede empezar planteando desde un modelo simple como poisson o de renovación pasando a modelos agrupados , a base de la fórmula empírica de Omori-Utsu nos conduce al modelo ETAS temporal propuesto por ([8]) sin considerar el espacio todos estos son detallados por ([16]). A base de este se plantean variantes incluyendo el espacio transformándolos en modelos espacio-temporales todos detallados en ([9]) se detallará la forma de ellas en la siguiente sección.

Este modelo ha sido ampliamente aplicado en varios estudios como: análisis de sismicidad sur de California donde ([17]) usa el modelo para desagrupar réplicas; ([18]) plantea la estimación de parámetros incluso con la incompletitud de datos; ([19]) describe la sismicidad en la región de Italia; ([12]) propone la estimación de parámetros de forma bayesiana analizando la sismicidad de Olfus Islandia ; ([20]) aplica en la región de Kermanshah Irán; ([21]) incluyo al modelo el efecto de eventos de deslizamiento de la placa SSE; ([22]) propone método para estimar parámetros frente a la incompletitud de datos.

0.2.2 Marco Conceptual

Glosario de Términos

- **Aftershock o Réplica** es un terremoto más pequeño que sigue a un terremoto más grande.
- **Catálogo sísmico** Es una base de datos que contiene todos los parámetros que caracterizan a un sismo como el momento, la ubicación y la magnitud.
- **ETAS Modelo de secuencia de réplicas de tipo epidémico** : Es un modelo probabilístico utilizado para describir la ocurrencia de terremotos y réplicas (aftershocks). Este modelo se basa en la idea de que cada terremoto puede generar sus propias réplicas, y estas réplicas pueden, a su vez, generar más réplicas, creando una cadena de eventos similar a una epidemia.
- **Función de intensidad condicional** Concepto fundamental en la teoría de procesos puntuales es la probabilidad de que ocurra un evento bajo el historial del proceso hasta ese momento.
- **Hawkes process** Un tipo de proceso puntual autoexcitado, utilizado para modelar eventos que tienen la propiedad de incrementar la probabilidad de futuros eventos durante algún período de tiempo después de su ocurrencia.
- **MLE Estimación de máxima verosimilitud** Es un método estadístico utilizado para estimar los parámetros de un modelo probabilístico. Este enfoque se basa en maximizar la función de verosimilitud, que mide la probabilidad de observar los datos dados los parámetros del modelo.
- **Point process** Es un modelo matemático utilizado para describir y analizar la ocurrencia de eventos que se producen de forma aleatoria en el tiempo o en el espacio. Estos eventos se representan como puntos en una línea de tiempo, en un plano, o en un espacio de mayor dimensión.

0.2.3 Base Teórica

Procesos Puntuales

Los procesos puntuales constituyen modelos matemáticos esenciales en la probabilidad y la estadística espacial. Se utilizan para representar la aparición aleatoria de eventos en el tiempo, en el espacio, o en ambas dimensiones simultáneamente. Estos eventos son vistos como “puntos” que se producen de manera aleatoria dentro de un área específica, que puede ser una línea temporal, un área geográfica, o incluso un marco de espacio-tiempo. La relevancia de estos procesos radica en su habilidad para representar fenómenos en campos tan variados como la física, la biología, la ingeniería, la economía y, de manera destacada, la estadística aplicada a datos tanto espaciales como temporales.

Para iniciar nuestro entendimiento de esta teoría, examinamos el caso más sencillo: el proceso puntual en una dimensión temporal. Este es conocido como proceso puntual temporal, y uno de los modelos más investigados y aplicados es el de Poisson. Este modelo, también considerado como proceso de conteo, se define por el conteo de eventos que suceden hasta un momento específico t , que se representa como $(N(t))$. El proceso de Poisson tiene tres propiedades clave.

- Incrementos independientes: las cantidades de eventos que suceden en períodos de tiempo no conectados son autónomas unas de otras.
- Incrementos estacionarios: la posibilidad de que un determinado número de eventos suceda en un período depende únicamente de la duración de ese período, sin importar el instante específico en que sucede.
- Distribución exponencial entre eventos: los intervalos de tiempo entre llegadas sucesivas cumplen con una distribución exponencial con parámetro λ , lo que sugiere que los eventos ocurren de manera aleatoria en el tiempo, manteniendo una tasa de aparición constante.

Cuando la frecuencia de ocurrencia λ es uniforme a través del tiempo, se clasifica como un proceso homogéneo o estacionario. No obstante, en numerosas situaciones del mundo real, la tasa de ocurrencia puede fluctuar temporalmente, lo que resulta en un proceso de Poisson que no es homogéneo o estacionario, donde $\lambda(t)$ se convierte en una función del tiempo. Este tipo de procesos es útil para modelar situaciones en las que la probabilidad de que ocurran eventos varía con el paso del tiempo, como la llegada de clientes a un negocio en diferentes horas del día, la frecuencia de terremotos en un periodo específico o la tasa de fallos en un sistema.

El concepto de procesos puntuales no se restringe únicamente al tiempo; también se puede extender al ámbito espacial. En vez de enfocarse en un momento específico t se analiza un área del espacio $A \subset S$ donde S es el espacio total bajo análisis (por ejemplo, un plano, una superficie geográfica, o incluso un volumen tridimensional). En este contexto, $N(A)$ indica cuántos eventos se han registrado dentro del área A . Así, el proceso puntual espacial describe cómo ocurren los eventos en el espacio, como la distribución de árboles en un bosque, el surgimiento de enfermedades en una población dispersa geográficamente, o la posición de estrellas en una sección del firmamento. En el ámbito espacial, una función fundamental del proceso es la función de intensidad $\lambda(s)$ donde $s \in S$. Esta función expresa la frecuencia esperada de eventos por unidad de área (o volumen) en el punto s . Para una área determinada A , se puede determinar la media esperada de eventos en esa área de la siguiente manera:

$$\mu(A) = \int_A \lambda(s) ds$$

Este valor refleja la probabilidad de ocurrencias en A , y si la función $\lambda(s)$ permanece constante a lo largo del espacio, el fenómeno es espacialmente uniforme, similar al proceso de Poisson que es uniforme en el tiempo. Por el contrario, si $\lambda(s)$ fluctúa según la ubicación, se trata de un proceso espacial no uniforme, capaz de reflejar cambios en la densidad de ocurrencias a través del espacio.

Es importante mencionar que $\lambda(s)$ puede ser tanto determinista como estocástico. En numerosos modelos sofisticados, la intensidad se representa como variable aleatoria espacial $\Lambda(S)$, lo que origina lo que se denomina un proceso Cox o un proceso de Poisson mixto. En estos tipos de procesos, la propia intensidad enfrenta incertidumbre, permitiendo representar fenómenos más elaborados, donde no solo la localización de las ocurrencias es aleatoria, sino también la estructura subyacente que rige su aparición. Por ejemplo, en el ámbito de la epidemiología, se puede asumir que el nivel de riesgo de infección no es completamente claro, sino que se distribuye aleatoriamente en el espacio en función de elementos socioambientales.

Una característica fundamental de los procesos puntuales es su habilidad para integrar información contextual a través de covariables espaciales o temporales, lo que permite modelar la intensidad $\lambda(s)$ como una función de variables explicativas. Esta estructura es frecuentemente utilizada en modelos log-lineales como los de la forma:

$$\lambda(s) = \exp(\beta_0 + \beta_1 x_1(s) + \dots + \beta_p x_p(s))$$

donde $x_1(s), x_2(s), \dots, x_p(s)$ representan las variables observadas en el punto s , y los parámetros β_i muestran la conexión entre estas variables y la intensidad de los eventos. Estos enfoques son comúnmente utilizados en campos como geociencias, criminología y evaluación ambiental.

En conclusión, los procesos puntuales son fundamentales para estudiar fenómenos donde los eventos suceden en ubicaciones y/o tiempos aleatorios. Desde contar simples llegadas dentro de un rango de tiempo hasta la modelación avanzada de patrones espaciales con intensidad estocástica, estos procesos facilitan el desarrollo de análisis precisos y predictivos basados en probabilidades. El proceso de Poisson, ya sea homogéneo o no homogéneo, sirve como base para muchas de estas aplicaciones, y su extensión a contextos espaciales y espacio-temporales permite una amplia gama de estudios estadísticos aplicados.

Proceso de Hawkes

Para entender el proceso de Hawkes, según lo descrito por ([23]), es importante considerar cómo en muchos fenómenos reales, la ocurrencia de un evento no es independiente de eventos pasados. En otras palabras, ciertos eventos tienen la capacidad de influir en la probabilidad de ocurrencia de eventos futuros. Esta idea da lugar a la clasificación de procesos puntuales en autoexcitados, inhibidos e independientes.

Un proceso autoexcitado es aquel en el que la ocurrencia de un evento incrementa la probabilidad de que ocurran más eventos similares en el futuro cercano. Un ejemplo común es el comportamiento del mercado financiero: una venta masiva de acciones puede desencadenar un efecto dominó, donde otros inversores, al percibir una posible crisis, también deciden vender sus activos. Este fenómeno se conoce como ventas por pánico y puede generar una cascada de eventos de venta en un corto periodo de tiempo.

En contraposición, existen los procesos inhibidos, en los cuales la ocurrencia de un evento reduce la probabilidad de eventos futuros. Un ejemplo claro se observa en fenómenos ecológicos como los incendios forestales. Después de un gran incendio, la masa combustible (vegetación) ha sido consumida, lo que disminuye significativamente la posibilidad de un nuevo incendio en la misma región durante un tiempo considerable. Así, el evento inicial tiene un efecto inhibitor sobre la ocurrencia futura.

Finalmente, tenemos los procesos independientes, donde la ocurrencia de un evento no influye en la ocurrencia de otros. Este es el caso del proceso de Poisson homogéneo, en el cual los eventos se distribuyen en el tiempo (o espacio) de forma completamente aleatoria y con una tasa constante. Estos procesos representan una base teórica importante, pero a menudo son demasiado simplistas para modelar fenómenos del mundo real donde la dependencia temporal es clave.

Para modelar estos diferentes comportamientos, se introduce el concepto de función de intensidad condicional, denotada por $\lambda^*(t)$. Esta función representa la tasa instantánea de ocurrencia de eventos en un instant t , dado todo el historial previo del proceso hasta ese momento, denotado como $H(t)$. Formalmente, se define como:

$$\lambda^*(t) = \lim_{\Delta t \rightarrow 0} \frac{E[N(t + \Delta t) - N(t)|H(t)]}{\Delta t} \quad (1)$$

esta está condicionada a la historia previa $H(t)$. El proceso de Hawkes es una variante de procesos puntuales autoexcitados planteado por Alan G. Hawkes. Utilizados para modelar eventos que están agrupados temporalmente como "terremotos" y "ventas de acciones por pánico" Definamos la función de intensidad condicional de Hawkes. Sea $0 < t_1 < t_2 < \dots < t_k$ la secuencia de tiempos de llegada del proceso puntual hasta el tiempo t , la intensidad condicional de Hawkes es:

$$\lambda^*(t) = \mu(t) + \sum_{t_i < t} \phi(t - t_i) \quad (2)$$

La constante μ se denomina intensidad de fondo y ϕ se denomina función de excitación. Si ϕ es igual a cero, este proceso puntual autoexcitado se reduce a un simple proceso de Poisson estacionario. Existen múltiples funciones de excitación según el diseño del modelo, la más común es el decaimiento exponencial.

$$\phi(t) = \alpha e^{-\beta t}$$

Los parámetros α y β son constantes. Donde α es encargado de incrementar la intensidad luego de la llegada de un evento y con el tiempo esta decae con β .

Para comprender la función de intensidad condicional, simulemos un pro-

ceso de puntos de Hawkes simple donde θ es el vector de parámetros λ , α y β . En la figura 0-4 se observan cuatro gráficos de intensidad condicional frente al tiempo. Comenzamos con $\theta = 0.1, 0.1, 0.1$ si incrementamos la intensidad λ a 0.5 en el segundo gráfico vemos que la intensidad de fondo se mantuvo por encima de 0.5. Para comprender el impacto de α , β . Aumentemos α a 0.5. En el tercer gráfico vemos que el número de eventos aumentó, porque en cada ocurrencia de un evento aumenta la intensidad de llegada del siguiente evento en α . Si aumentemos β a 0.5. En el cuarto gráfico se pueden ver decaimientos constantes.

El proceso de Hawkes ha encontrado una gran variedad de aplicaciones prácticas. En finanzas, se usa para modelar la dinámica de transacciones en los mercados, especialmente en contextos de alta frecuencia, donde las operaciones se agrupan en intervalos cortos de tiempo. También se ha empleado para estudiar eventos extremos como caídas del mercado provocadas por el contagio del miedo o la especulación.

En sismología, el modelo ha sido útil para describir réplicas de terremotos. Un sismo principal puede aumentar temporalmente la probabilidad de sismos menores en la misma región, fenómeno conocido como "aftershocks", cuya intensidad decae con el tiempo de manera similar a la función de excitación.

En el ámbito de las redes sociales, se ha utilizado para modelar la difusión de información o noticias virales, donde una publicación puede desencadenar reacciones sucesivas (comentarios, compartidos, likes), formando picos de actividad.

También en epidemiología, los modelos de Hawkes permiten estudiar la propagación de enfermedades contagiosas, donde un individuo infectado puede generar nuevos contagios cuya probabilidad disminuye conforme pasa el tiempo desde la infección inicial.

Extensiones del modelo El modelo original de Hawkes ha sido ampliado en diversas direcciones. Por ejemplo:

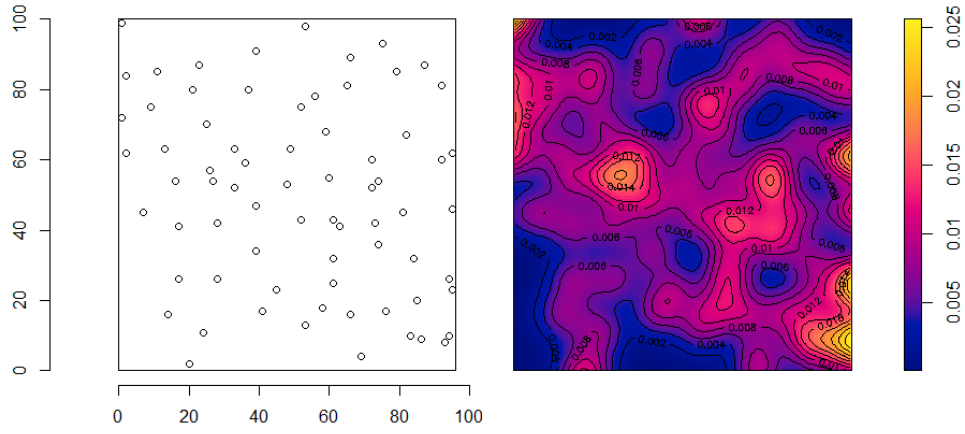


Figure 0-1: Observamos la distribución de la función de intensidad $\Lambda(S)$ (derecha) del proceso puntual. Gráfico generado mediante spatstat de R

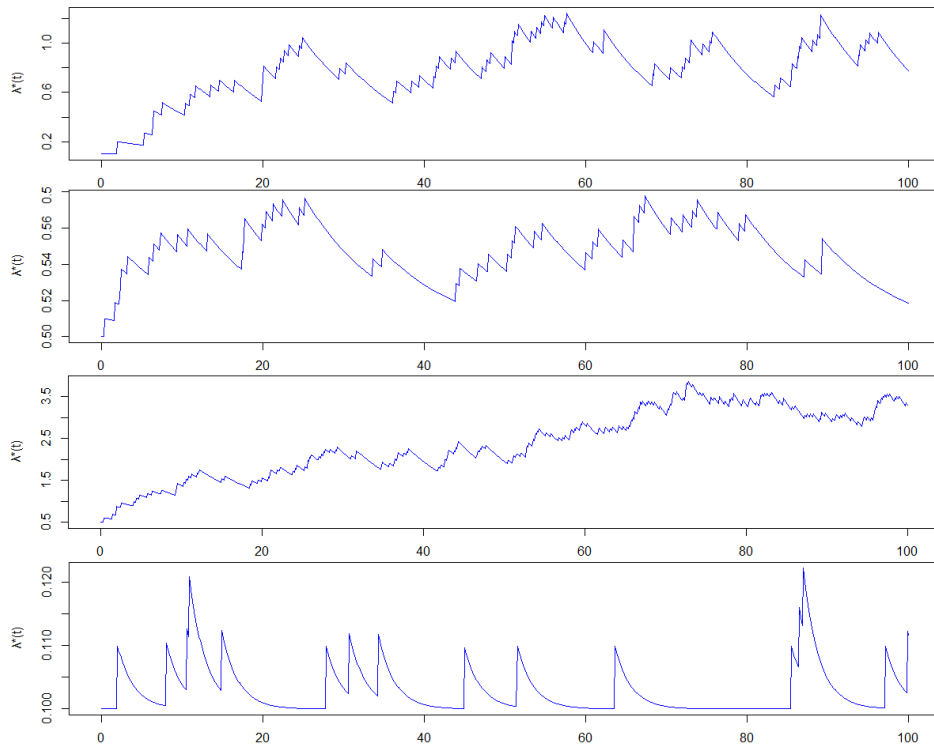


Figure 0-2: Simulación del proceso de Hawkes con distintos valores de parámetros λ , α y β en R

- En los procesos de Hawkes multivariados, se consideran múltiples tipos de eventos que pueden autoexcitarse entre sí o cruzadamente. Por ejemplo, en un mercado financiero, una venta en un activo puede aumentar la probabilidad de venta en otro.
- Los procesos espacio-temporales de Hawkes incorporan coordenadas geográficas además del tiempo, siendo aplicables en criminología, vigilancia epidemiológica y propagación de incendios.
- También existen modelos no paramétricos, donde la función de excitación se estima directamente de los datos sin asumir una forma funcional específica, lo que aumenta la flexibilidad del modelo.

Distribución de Réplicas

En todos los catálogos de hipocentros sísmicos incluyen la cinco parámetros principales las cuales son tiempo de ocurrencia, longitud y latitud del epicentro, profundidad focal y magnitud; en nuestro estudio se analiza la actividad sísmica en relación de cuatro dimensiones (t_n, x_n, y_n, M_n) , $n = 1, 2, \dots, N$, ignorando la profundidad debido a la baja precisión en comparación con las otras coordenadas.

La probabilidad de que ocurra un terremoto en un tiempo t , una ubicación (x, y) de magnitud M , condicionado a los eventos ocurridos en el pasado $H_t = \{(t_i, x_i, y_i, M_i); t_i < t\}$ se representa en la siguiente ecuación:

$$\frac{P(\text{evento en } [t, t + \Delta t] \times [x, x + \Delta x] \times [y, y + \Delta y] \times [M, M + \Delta M] | H_t)}{\Delta t \Delta x \Delta y \Delta M} \approx \lambda(t, x, y, M | H_t) \quad (3)$$

donde λ es función de intensidad condicional del proceso puntual que intentamos

modelar que representa la probabilidad de que ocurra eventos de magnitudes ΔM en el intervalo de tiempo t y una región definida por Δxy . Las tasas y magnitudes de las réplicas siguen varias leyes empíricas bien establecidas; como la de Gutenberg-Richter donde $\lambda(M)$ es la intensidad de terremotos de magnitud menor o igual a M en una región en un período de tiempo con constantes a y $b = \beta$ que varían según la región de estudio.

$$\lambda(M) = 10^{a-bM} = Ae^{-\beta M} \quad (4)$$

También la ley de [24] y (Utsu, 1961) proponen el decaimiento de las réplicas según el tiempo después del terremoto principal expresado como:

$$n(t) = \frac{k}{(c+t)^p} \quad (5)$$

teniendo como constantes k , c y p .

Modelo Espacio Temporal ETAS

El modelo ETAS es un proceso de puntos de ramificación (también conocidos como Hawkes o procesos de puntos autoexcitantes) la ocurrencia de un terremoto excita la actividad sísmica en curso, aumentando la probabilidad de ocurrencia de nuevos terremotos en el futuro cercano donde la sismicidad es el resultado de la suma de dos componentes: la sismicidad de fondo (eventos independientes) y la sismicidad desencadenada (eventos agrupados). Matemáticamente, esto conduce a la siguiente expresión para la intensidad sísmica condicional:

$$\lambda(t, x, y, M | H_t) = S(M) \left[\mu(x, y) + \sum_{i: t_i < t} k(M_i) g(t - t_i) f(x - x_i, y - y_i) \right] \quad (6)$$

donde

- H_t es la historia sísmica hasta el tiempo t ;

- μ representa la tasa de sismicidad de fondo
- $S(M)$ función de magnitud G-R representado por la ley exponencial de Gutenberg-Richter
- $k(M) = Ae^{\alpha m}$ productividad directa de Utsu
- $g(t, M)$ se define como la función de excitación en el contexto de los procesos de Hawkes

Estimación de Parámetros

- **Estimación mediante MLE**

La estimación de parámetros mediante el método de máxima verosimilitud (MLE) es una técnica estadística muy común que se usa para deducir los parámetros de un modelo probabilístico basándose en datos observados. Este método se centra en elegir los valores que optimizan la función de verosimilitud, es decir, aquellos que hacen que la observación de los datos sea lo más probable posible según el modelo considerado. La función de verosimilitud se formula como la probabilidad conjunta de observar los datos dados ciertos parámetros, o, en el caso de variables en continuo, como la densidad de probabilidad conjunta. En términos matemáticos, si tenemos una colección de datos x_1, x_2, \dots, x_n que se supone derivan de una distribución con densidad $f(x|\theta)$, donde θ representa los parámetros del modelo, la función de verosimilitud se define como $L(\theta) = \prod_{i=1}^n f(x_i|\theta)$. En la práctica, debido a la dificultad numérica de trabajar con productos, comúnmente se maximiza el logaritmo de la verosimilitud, denominado log-verosimilitud, expresado como $l(\theta) = \sum_{i=1}^n \ln f(x_i|\theta)$, dado que el logaritmo es una función que siempre aumenta, y maximizar $l(\theta)$ es lo mismo que maximizar $L(\theta)$.

El proceso de estimación utilizando MLE incluye resolver un problema de optimización con el objetivo de encontrar θ que maximice $l(\theta)$. Esto puede

llevarse a cabo de forma analítica derivando la log-verosimilitud respecto a los parámetros, estableciendo que las derivadas sean cero y solucionando el sistema de ecuaciones que se obtiene, conocidas como ecuaciones de verosimilitud. No obstante, en muchas ocasiones, las soluciones no son factibles analíticamente, por lo que se utilizan métodos numéricos como el algoritmo de Newton-Raphson, descenso por gradiente o técnicas de optimización más complejas. Una de las principales ventajas del MLE es su consistencia: bajo ciertas condiciones de regularidad, el estimador de máxima verosimilitud se aproxima al valor real del parámetro a medida que se incrementa el tamaño de la muestra. Además, el MLE es asintóticamente eficiente, lo que implica que alcanza el límite inferior de Cramér-Rao, ofreciendo estimadores con la varianza más baja posible en muestras extensas. También posee la propiedad de invarianza, lo que significa que si $\hat{\theta}$ es el estimador de máxima verosimilitud de θ , entonces $g(\hat{\theta})$ será el estimador de máxima verosimilitud de $g(\theta)$. Para cualquier función (g). No obstante, el MLE presenta ciertas limitaciones. Por ejemplo, requiere que el modelo esté bien definido; si la distribución asumida no se ajusta correctamente a los datos, los estimadores podrían resultar sesgados o ineficaces. Adicionalmente, en situaciones de muestras pequeñas, los estimadores de MLE pueden carecer de robustez, y ocasionalmente, la función de verosimilitud puede presentar múltiples máximos o puede no ser diferenciable, complicando su optimización. También es relevante señalar que el MLE puede ser costoso computacionalmente en modelos complejos, como aquellos que contienen varios parámetros o datos de alta dimensionalidad. En esos escenarios, se pueden explorar otros métodos como la estimación bayesiana o los métodos de momentos. A pesar de estas limitaciones, el MLE sigue siendo una herramienta crucial en la estadística, usada en una amplia gama de disciplinas, desde la econometría hasta el aprendizaje automático, donde se aplica para ajustar modelos como regresiones logísticas, redes neuronales y modelos de mezcla gaussiana. Su sólido fundamento teórico y su flexibilidad lo establecen

como un componente vital para la inferencia estadística contemporánea.

Los parámetros a estimar son μ, K, α, c, p donde K, α son constantes de la ley de productividad de las réplicas de Utsu y c y p son constantes de la ley de Omori-Utsu modificada. Los cinco parámetros se pueden estimar mediante el método de estimación de máxima verosimilitud (MLE) la probabilidad logarítmica se puede expresar de la siguiente manera:

$$\log L(\theta) = \sum_{i=1}^N \log \lambda(t_i, \theta) - \int_0^T \lambda(t, \theta) dt \quad (7)$$

donde θ es el conjunto de parámetros a estimar y $[0, T]$ es el intervalo de tiempo para el cual se deben estimar los parámetros. Encontrar el máximo de una función es un problema de optimización típico, que en la mayoría de los casos debe resolverse numéricamente por métodos Quasi-Newton.

- **Estimación por inferencia Bayesiana**

La estimación por inferencia bayesiana es un enfoque estadístico que permite estimar los parámetros de un modelo probabilístico incorporando tanto la información proporcionada por los datos observados como el conocimiento previo sobre los parámetros, utilizando el teorema de Bayes como fundamento. Este método contrasta con enfoques frecuentistas, como la estimación de máxima verosimilitud, al tratar los parámetros como variables aleatorias con distribuciones de probabilidad, en lugar de valores fijos desconocidos. En la inferencia bayesiana, el proceso comienza con la especificación de una distribución previa, $p(\theta)$, que refleja las creencias o conocimientos iniciales sobre los parámetros θ antes de observar los datos. Cuando se obtienen los datos x_1, x_2, \dots, x_n , la verosimilitud, $p(x|\theta)$, describe la probabilidad de los datos dado un conjunto de parámetros. El teorema de Bayes combina estas dos componentes para obtener la distribución posterior, $p(\theta|x)$, que se calcula como $p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}$, donde $p(x) = \int p(x|\theta)p(\theta)d\theta$ actúa como una constante de normalización conocida como evidencia o probabilidad marginal. La distribu-

ción posterior resume toda la información actualizada sobre los parámetros después de observar los datos y es el núcleo de la inferencia bayesiana, ya que permite realizar estimaciones puntuales (como la media, mediana o moda de la posterior) o intervalos de credibilidad, que son análogos a los intervalos de confianza frecuentistas pero con una interpretación probabilística directa. Una de las principales fortalezas de la inferencia bayesiana es su capacidad para incorporar conocimiento previo de manera sistemática, lo que es particularmente útil en contextos donde los datos son limitados o ruidosos. Por ejemplo, en un modelo de regresión, una distribución previa puede imponer restricciones sobre los coeficientes para evitar el sobreajuste. Además, la inferencia bayesiana es flexible y puede aplicarse a modelos complejos, como modelos jerárquicos o redes bayesianas, donde los parámetros tienen dependencias estructuradas. Sin embargo, esta flexibilidad tiene un costo computacional. Calcular la distribución posterior de manera analítica es factible solo en casos simples, como cuando la previa y la verosimilitud son conjugadas (por ejemplo, una previa normal con una verosimilitud normal para la media de una distribución). En modelos más complejos, la integral en el denominador del teorema de Bayes suele ser intratable, lo que requiere el uso de métodos computacionales como el muestreo de Monte Carlo, las cadenas de Markov (MCMC) o la inferencia variacional. Estos métodos han hecho posible la aplicación de la inferencia bayesiana en áreas como el aprendizaje automático, la bioestadística y la econometría, pero pueden ser intensivos en términos de recursos computacionales y requieren experiencia para implementarse correctamente. A pesar de sus ventajas, la inferencia bayesiana también enfrenta desafíos. La elección de la distribución previa puede ser subjetiva y, si no se selecciona cuidadosamente, puede influir significativamente en los resultados, especialmente en muestras pequeñas. Para mitigar esto, se pueden usar distribuciones previas no informativas o débilmente informativas, aunque estas también tienen limitaciones. Además, la interpretación de los resultados bayesianos, como los intervalos de

credibilidad, puede ser menos intuitiva para quienes están acostumbrados a los métodos frecuentistas. Otro aspecto a considerar es la sensibilidad de los resultados a la especificación del modelo; un modelo mal especificado puede llevar a inferencias erróneas, al igual que en los métodos frecuentistas. No obstante, la inferencia bayesiana ofrece una framework coherente para la incertidumbre, permite actualizar las estimaciones a medida que se obtienen nuevos datos y es particularmente poderosa en problemas de decisión secuencial o en contextos donde la incertidumbre paramétrica es crítica. En la práctica, la inferencia bayesiana se ha convertido en una herramienta indispensable en campos como la inteligencia artificial, donde se usa en modelos probabilísticos como los modelos gráficos y los procesos gaussianos, y en ciencias sociales, donde los modelos jerárquicos bayesianos son comunes para analizar datos multinivel. Su capacidad para integrar datos y conocimiento previo en un marco probabilístico unificado la posiciona como un enfoque fundamental en la estadística moderna.

Segun ([25]) el enfoque bayesiano de la estimación de parámetros ETAS el teorema de Bayes dice, la distribución de probabilidad posterior $p(\theta|X)$ del conjunto de parámetros θ dado los datos X se puede expresar utilizando la probabilidad $L(\theta|X)$ y la distribución de probabilidad previa $p(\theta)$ del conjunto de parámetros $\theta|X$ de la siguiente manera:

$$p(\theta|X) \propto L(\theta|X)p(\theta) \tag{8}$$

Selección del Modelo

Criterios de selección:

- **Criterio de Información Akaike (AIC):**

El Criterio de Información Bayesiano (BIC, por sus siglas en inglés, Bayesian

Information Criterion) es una herramienta estadística ampliamente utilizada para la selección de modelos, que busca equilibrar la bondad de ajuste de un modelo con su complejidad, penalizando aquellos modelos que incorporan un número excesivo de parámetros. Introducido por Gideon Schwarz en 1978, el BIC se deriva de un enfoque bayesiano y está basado en una aproximación a la probabilidad marginal de los datos, también conocida como evidencia, que se calcula integrando la función de verosimilitud sobre la distribución previa de los parámetros. Matemáticamente, el BIC para un modelo (M) con datos x_1, x_2, \dots, x_n se define como $BIC = -2 \ln(\hat{L}) + k \ln(n)$ es la máxima verosimilitud del modelo (es decir, el valor de la función de verosimilitud evaluada en los estimadores de máxima verosimilitud de los parámetros), (k) es el número de parámetros del modelo, y (n) es el tamaño de la muestra. El término $-2 \ln(\hat{L})$ refleja la bondad de ajuste del modelo, mientras que el término $k \ln(n)$ actúa como una penalización por la complejidad del modelo, que aumenta con el número de parámetros y el tamaño de la muestra. En la práctica, el BIC se utiliza para comparar modelos, donde el modelo con el valor más bajo de BIC se considera preferible, ya que representa un compromiso óptimo entre ajuste y simplicidad. El fundamento teórico del BIC radica en su conexión con la inferencia bayesiana, específicamente en la aproximación de Laplace a la probabilidad marginal ($p(x|M)$), que integra la verosimilitud sobre la distribución previa de los parámetros. Esta aproximación es válida bajo ciertas condiciones de regularidad y cuando el tamaño de la muestra es grande, lo que permite simplificar la integral compleja en una expresión que depende del máximo de la verosimilitud y una penalización por la dimensionalidad del modelo. A diferencia de otros criterios, como el Criterio de Información de Akaike (AIC), el BIC impone una penalización más fuerte por la complejidad del modelo debido al término $\ln(n)$, lo que lo hace más conservador en la selección de modelos con muchos parámetros, especialmente en muestras grandes. Esta característica hace que el BIC tienda a favorecer modelos más

simples que el AIC, lo que puede ser ventajoso en contextos donde se desea evitar el sobreajuste, pero también puede llevar a seleccionar modelos demasiado simplificados si la penalización es excesiva. Además, el BIC tiene una interpretación bayesiana directa: bajo ciertas condiciones, la diferencia en los valores de BIC entre dos modelos se relaciona con el logaritmo del factor de Bayes, una medida de la evidencia relativa a favor de un modelo sobre otro. A pesar de sus ventajas, el BIC tiene limitaciones importantes. Su cálculo asume que la verosimilitud está bien especificada y que los datos provienen de una distribución que pertenece a la familia del modelo considerado, lo que no siempre es cierto en aplicaciones prácticas. Si el modelo está mal especificado, el BIC puede llevar a selecciones subóptimas. Además, el BIC depende de la aproximación asintótica, lo que significa que su desempeño puede ser menos confiable en muestras pequeñas, donde la penalización $k \ln(n)$ puede no reflejar adecuadamente la incertidumbre en los parámetros. Otro aspecto crítico es que el BIC no incorpora explícitamente la incertidumbre de la distribución previa, lo que puede ser una desventaja en contextos bayesianos puros donde las prioras tienen un impacto significativo. En la práctica, el BIC se utiliza ampliamente en disciplinas como la econometría, la bioestadística y el aprendizaje automático para comparar modelos lineales, modelos de regresión logística, modelos de series temporales o modelos de mezcla, entre otros. Por ejemplo, en el análisis de series temporales, el BIC puede ayudar a seleccionar el orden de un modelo ARIMA, mientras que en aprendizaje automático, se usa para elegir entre diferentes configuraciones de modelos probabilísticos. Aunque el BIC no es la única herramienta para la selección de modelos, su simplicidad y su fundamento teórico bayesiano lo convierten en un criterio robusto y popular, especialmente en contextos donde se busca un equilibrio entre interpretabilidad y precisión predictiva.

Evalúa la calidad del ajuste del modelo penalizando por el número de parámetros

ros. Se calcula como:

$$AIC = 2k2\log(L) \quad (9)$$

donde k es el número de parámetros y L es la verosimilitud máxima.

- **Criterio de Información Bayesiano (BIC):** Similar al AIC, pero penaliza más fuertemente modelos con más parámetros. Se calcula como:

$$BIC = k\log(n)2\log(L) \quad (10)$$

donde n es el número de datos.

0.3 Metodología

En el presente proyecto de investigación esta orientado al desarrollo de modelos estadísticos lo que convalida el uso de la metodología metodología CRISP-ML(Q) ([26]) un proceso estructurado que guía el desarrollo de modelos de aprendizaje automático de manera eficiente y confiable basado en el modelo CRISP-DM (Cross-Industry Standard Process for Data Mining), pero extendido para incluir tareas y metodologías de aseguramiento de la calidad específicas. Esta se divide en 6 fases:

0.3.1 Comprensión de Datos

Un catálogo sísmico es un registro de eventos sísmicos dentro de una región específica o a nivel mundial. Funciona como una base de datos histórica que recopila información sobre terremotos. Esta información normalmente incluye: fecha y hora del evento, ubicación en coordenadas geográficas (latitud y longitud), magnitud en la escala de Richter y profundidad; algunos catálogos también pueden incluir detalles

adicionales; se puede consultar en [27].Podemos un ejemplar en la Tabla 0-1

Los datos sismicos ([28]) se adquirieron del Centro Sismologico Nacional (CENSIS) servicio desarrollado por el Instituto Geofisico del Perú (IGP) institución oficial del Estado que se encarga de monitorear la ocurrencia de sismos en el territorio nacional para ello, se dispone de datos provenientes de la Red Sísmica Nacional (RSN) cuyos sensores se encuentran distribuidos en todo el territorio nacional. La base de datos sísmicos para el Perú ha sido dividida en Sismos Históricos para los eventos sísmicos ocurridos entre los años 1471 y 1959; y Sismos Instrumentales para los ocurridos después del año 1960, La elección del ultimo se basa en las siguientes razones: la precisión de los instrumentos de medicion, completitud y consistencia, el incremento notable de estaciones sismicas a partir del año 2000 nos dan la disponibilidad de datos mas robustos. El catálogo nos provee datos de los sismos ocurridos en una region rectangular que abarca el territorio y el mar peruano, con latitud sur desde 1.396° hasta 25.701° y longitud oeste desde 87.382° hasta 65.624° .

0.3.2 Preparación de Datos

Procedemos a realizar un análisis exploratorio de los datos contenidos en el catálogo tomando como referencia a ([29]), lo primero que se analizará es la magnitud. Si bien la magnitud más conocida es la escala Richter esta tambien tiene varios tipos, no confundir con la escala de intensidad representado por la escala de Mercalli, como: Richter Magnitude Scale (M_L); Body Wave Magnitude Scale (m_b); Duration Magnitude Scale (M_d); Moment Magnitude (M_W); [30] Describe

Table 0-1: Ejemplo de un catálogo de terremotos

Event	Date	Time	Longitude	Latitude	Depth	Magnitude
1	1998-01-06	15:39:31.00	-75.82	-14.15	49	4.20
2	1998-01-06	20:01:50.90	-70.34	-18.10	188	4.80
3	1998-01-10	17:02:56.00	-80.28	-4.19	36	5.10
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
100	1998-07-25	08:57:26.90	-69.14	-16.83	67	4.90
101	1998-07-29	13:34:14.00	-74.65	-14.98	96	4.70

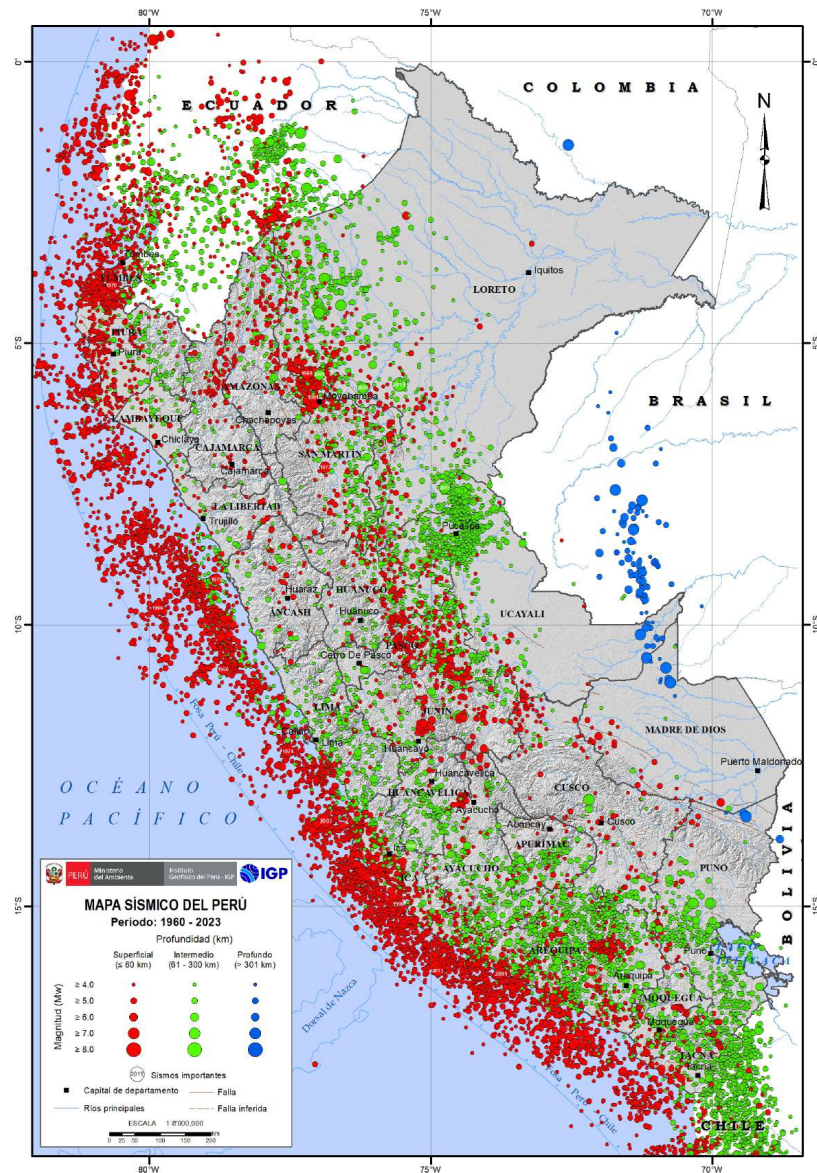


Figure 0-3: Mapa Sísmico del Perú 1960 - 2023 [1].

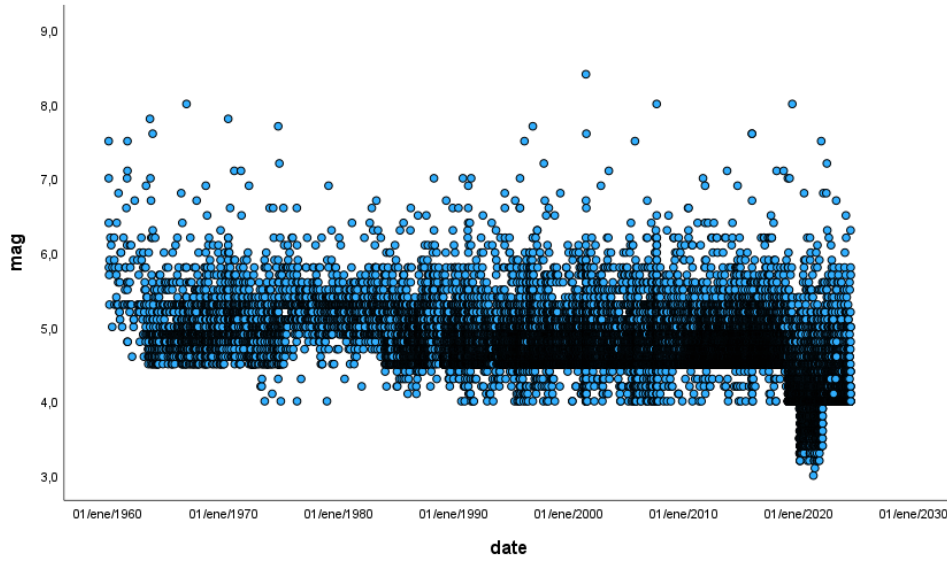


Figure 0-4: Diagrama de dispersión de magnitudes de eventos con el tiempo.

la metodología desarrollada en los EE.UU. por el U.S. Geological Survey (USGS) para la elaboración de catálogos sísmicos; con este mismo criterio el IGP opto por homogenizar los valores a magnitud momento (M_w) de forma propuesta por ([31])

El catálogo consta de 23 944 eventos una magnitud media de 4.71. En la figura 0-4 podemos ver que casi no hay eventos de magnitud menor a 4.0 esto se debe a la sencibilidad de los instrumentos o la depuración humana y hay pocos eventos de gran magnitud, la región intermedia es la más representativa donde se visualiza mejor en una escala log-lineal como en la figura 0-5 Esto muestra claramente una relación lineal, entre las magnitudes 4 y 7, cumpliendose la ley de Gutenberg-Richter.

Se empleará la libreria *ETAS* de ([32]) en *R* puede adaptar el modelo para un catálogo de terremotos utilizando el enfoque de desagrupación estocástica propuesto por Zhuang y Ogata. Definimos nuestra ventana de estudio tomando una porción de nuestro catálogo; esta será nuestro evento principal, independiente o de fondo; los datos fuera de esta ventana o externo al poligono de estudio, que se explicará más adelante, son eventos secundarios o de ruido, `time.begin` como el comienzo del lapso de tiempo del catálogo, `time.start` es el inicio del período de estudio, `time.end` es el final del período de estudio, tener en cuenta el formato de

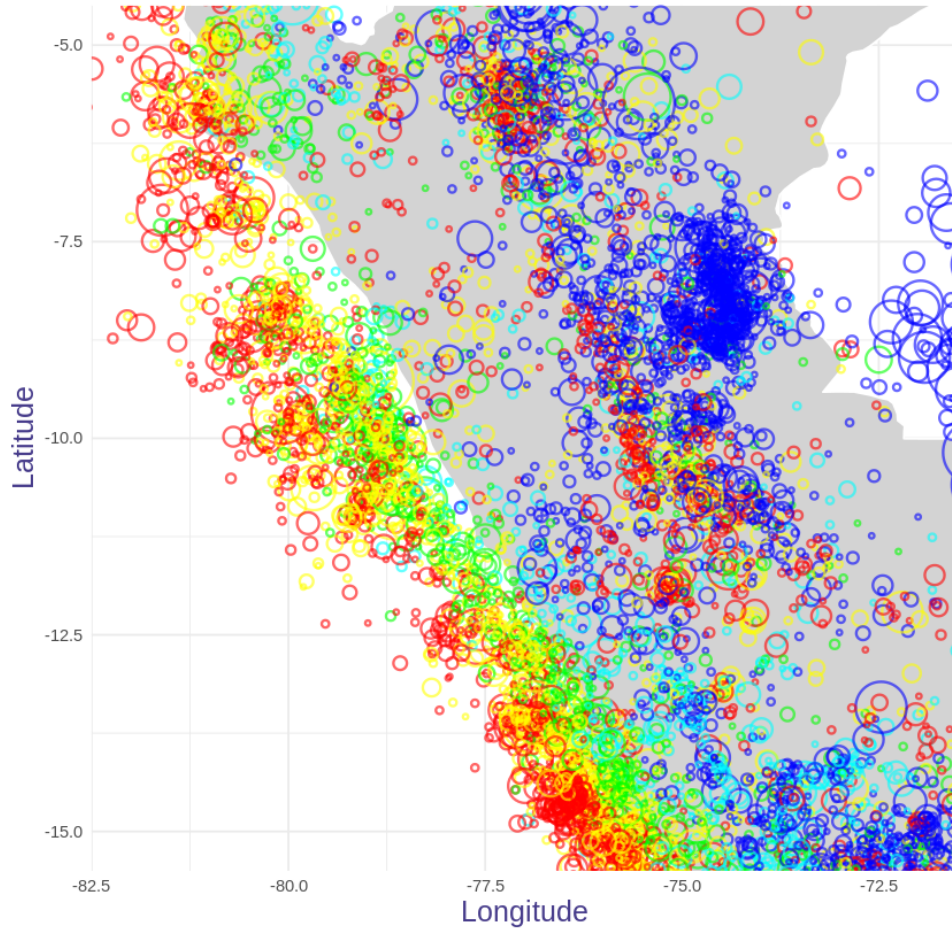


Figure 0-5: Gráfico de epicentros de terremotos en un mapa base coloreado por profundidad y con tamaño proporcional a la magnitud. Los colores cálidos indican profundidades bajas y los colores fríos eventos más profundos. Se puede encontrar el código para producir este gráfico en los anexos.

las fechas como yyyy-mm-dd.

Lo siguiente es definir el región de estudio puede ser rectangular o poligonal en nuestro caso se opto por el ultimo, se puede especificar mediante el argumento `region.poly` que contiene coordenadas de los vértices del polígono, ver en la Tabla?? deben enumerarse en orden antihorario

0.3.3 Ajuste del Modelo

Sea θ el conjunto de parámetros a estimar en el intervalo de tiempo $[0, T]$. En la libreria de ([32]) se utiliza el método Davidon-Fletcher-Powell DFS el algoritmo genera iterativamente una aproximación local de la función de log-verosimilitud por

una función cuadrática hasta que converge a su máximo.

Algoritmo 1 Davidon-Fletcher-Powell algorithm.

A1.entradas

$\widehat{\theta}_0$ parámetros iniciales del modelo

$\widehat{\mu}_0(x_j, y_j)$ $j = 1, 2, \dots, N$ tasa de fondo inicial

H_0 la inversa de la matriz de Hessiana

A2. $k \leftarrow 0$

A3.repetir

$d_k \leftarrow -H_k \times \nabla \xi(\widehat{\theta}_k)$ dirección ascendente

$s_k \leftarrow \text{argmin}_s \xi(\widehat{\theta}_k + s d_k)$

$\widehat{\theta}_{k+1} \leftarrow \widehat{\theta}_k + s_k d_k$ actualizar la estimación

$\eta_k \leftarrow \nabla \xi(\widehat{\theta}_{k+1}) - \nabla \xi(\widehat{\theta}_k)$

$c1 \leftarrow s_k / (d_k^\top \times \eta_k)$

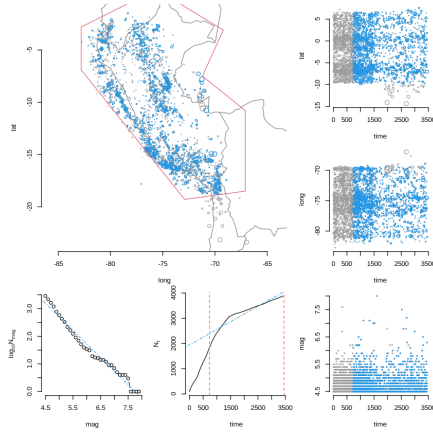
$c2 \leftarrow 1 / (\eta_k^\top \times H_k \times \eta_k)$

$H_{k+1} \leftarrow H_k + c1 d_k \times d_k^\top - c2 H_k \times \eta_k \times \eta_k^\top \times H_k$ actualizar el hessiano inverso

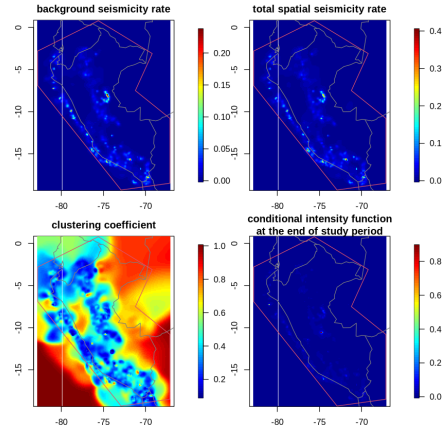
$k \leftarrow k + 1$

A4.hasta $\|\widehat{\theta}_k - \widehat{\theta}_{k-1}\| < \epsilon$ o $\|\nabla \xi(\widehat{\theta}_k)\| < \epsilon$

A5.retornar $\widehat{\theta}_k$ y H_k



(a) Example left figure.



(b) Example right figure.

Figure 0-6: Example figure that covers the width of the page obtained from PGFPlots [2].

0.4 Resultados

Luego de ajustar nuestros parámetros el modelo ETAS: ajustado mediante el método iterativo de desagrupación estocástica convergió después de 5 iteraciones tiempo de ejecución transcurrido 18,4 minutos procesado paralelamente en un procesador octacore Intel Core i5 decima generación:

Table 0-2: Resultados del ajuste de parámetros

ML estimates of model parameters:									
	β	μ	A	c	α	p	D	q	γ
$\widehat{\theta}$	3.3243	0.5497	1.8560	0.0453	0.8895	1.0132	0.0023	1.2261	0.0117
SE	0.0053	0.0084	0.0276	0.0662	0.0213	0.0003	0.0732	0.0092	1.3761
Declustering Probabilities									
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.			
	0.0000	0.5188	0.7382	0.6490	0.8512	1.0000			
log-likelihood:			-10361.76		AIC:	20739.53			

0.5 Conclusión

De los puntos planteados anteriormente se pudo desarrollar con éxito el modelo ETAS pero según los resultados la prueba de bondad de ajuste de Kolmogorov-Smirnov indica que U_j no sigue una distribución $U(0, 1)$ con un p-value de $8.438e-12$ el modelo ajustado aun no es capaz de describir con detalle la compleja agrupación espacial de los terremotos podría existir anomalía en la sismicidad de la región se sospecha que son originados por los datos depurados. En el punto de vista de la literatura la revisión teórica y metodologías empleados por otros investigadores se aconseja modificar método de estimación de parámetros como la bayesiana o encontrar otra variante de función de activación. Por último la sismicidad de cada región es distinta por lo cual existe múltiples métodos para modelarla.

Anexos

Codigos

Se ejecutó las siguientes funciones en lenguaje R versión 4.4.1 con un procesador octacore Intel i5 decima generación

Los codigos empleados para el desarrollo del modelo.

Listing 1: Codigos empleados.

```
# importando datos

dataf <- read.csv("IGP_peru.csv", sep = ";")
head(dataf)

# adecuando nuestros datos

df <- dataf[, c("date", "time", "long", "lat", "mag")]
df$long <- as.numeric(gsub(",", ".", df$long))
df$lat <- as.numeric(gsub(",", ".", df$lat))
df$date <- as.Date(df$date, format = "%d/%m/%Y")
df$date <- as.factor(df$date)

head(df)

# instalar los paquetes necesarios

install.packages('ETAS')
library(ETAS)

# specifying a polygonal geographical region

ppoly <- list(
  long=c(-82.8, -82.8, -72.9, -67.1, -67.1, -71.2, -69.2, -75.6),
```

```

lat=c(-2.8, -6.9, -19.3, -18.5, -10.8, -7.5, -3.1, 0.8)
)

# creating a catalog with polygonal study region
peru.cat <- catalog(df, time.begin="2015-01-01",
study.start="2017-01-01", study.end="2024-06-16",
region.poly=ppoly, mag.threshold=4.5)
plot(peru.cat)

param01 <- c(0.5, 0.2, 0.05, 2.7, 1.2, 0.02, 2.3, 0.03)

# ajustando el modelo
peru.fit <- etas(peru.cat, param0 = param01)

```

los graficos generados son:

Agradecimientos

Al docente del curso por impartir conocimientos necesarios para desarrollar el presente trabajo. A CORSSA: the Community Online Resource for Statistical Seismicity Analysis por proporcionar recursos valiosos en el análisis estadístico de la sismicidad. A Kowshik chilamkurthy por proporcionar una vista más simple de los procesos Hawkes. Finalmente al IGP por proporcionar datos para el presente estudio.

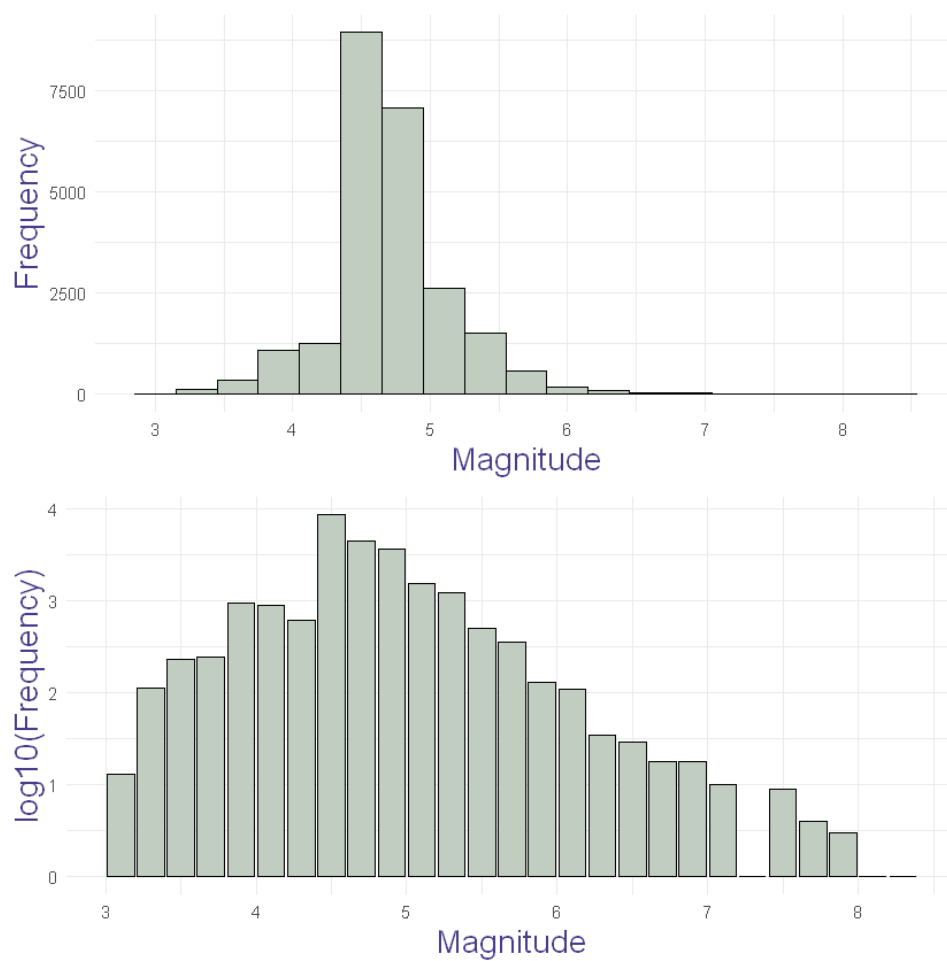


Figure 0-7: se demuestra que se cumple la ley de G-R Gutenberg-Richter.

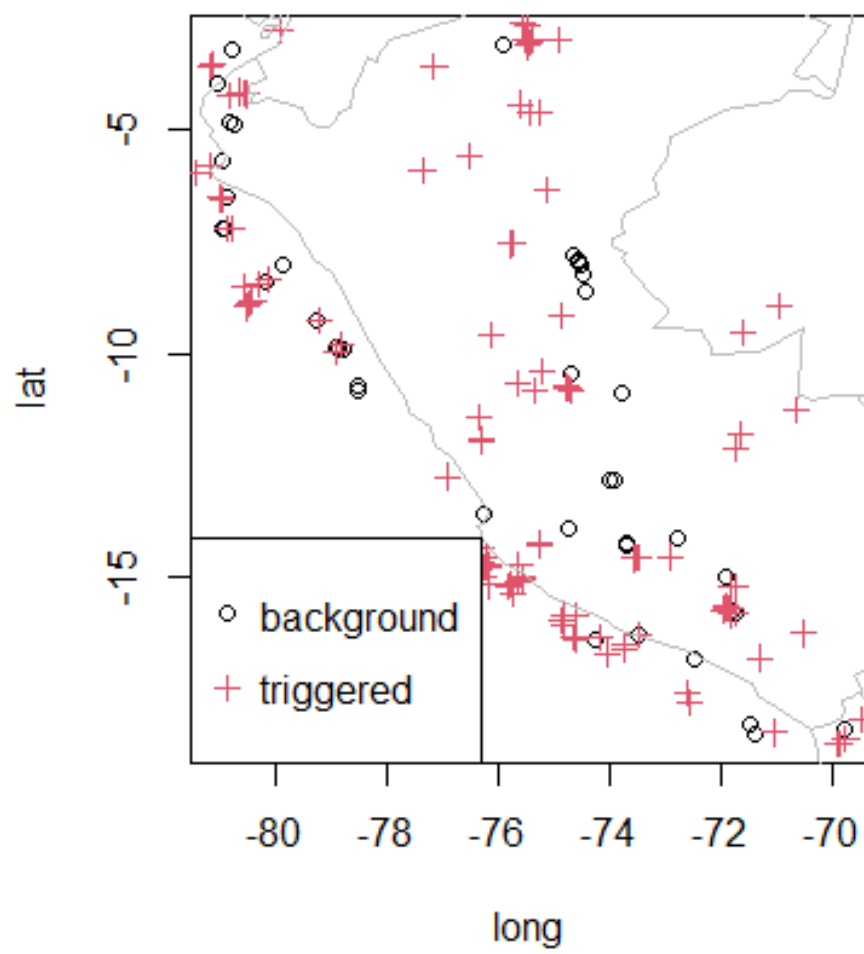


Figure 0-8: Eventos provocados en la simulación.

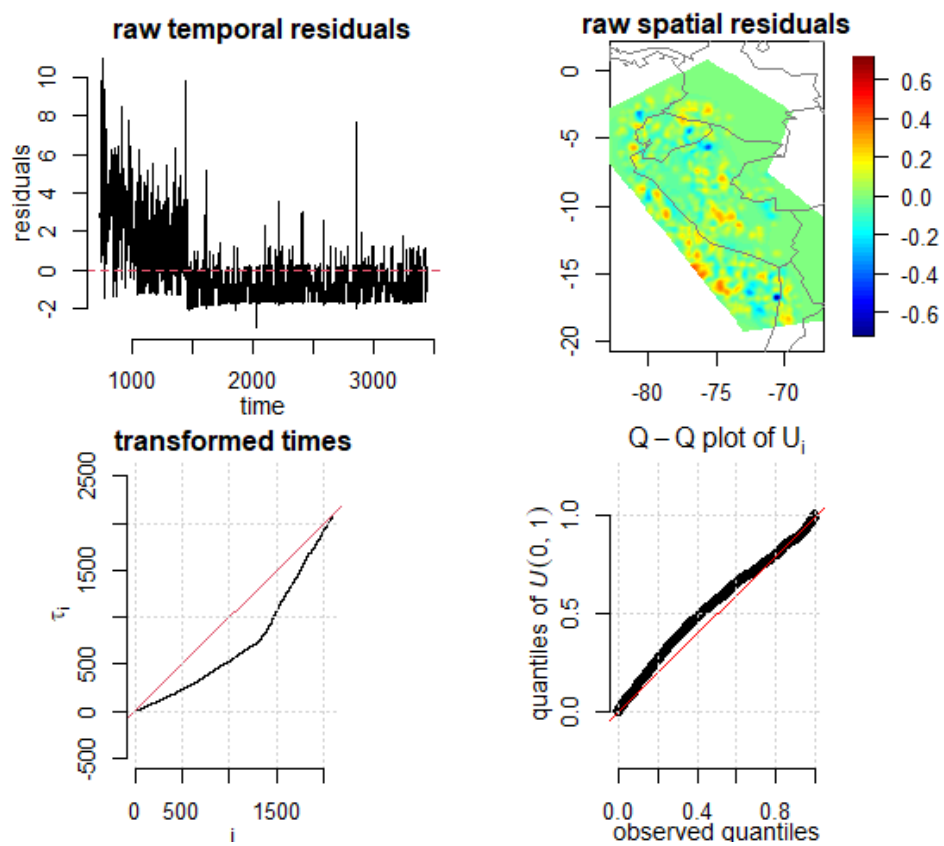


Figure 0-9: Análisis de Residuos.

References

- [1] CENSIS, “Mapasismo2023,” 2023. [Online]. Available: https://ultimosismo.igp.gob.pe/img/mapas_sismicos/mp_2023.jpg?v=3221881
- [2] “PGFPlots - A LaTeX package to create plots.” [Online]. Available: <https://pgfplots.sourceforge.net/>
- [3] V. Kossobokov and P. Shebalin, *Earthquake Prediction*. Springer Berlin Heidelberg, 2003, p. 141—207. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-05298-3_4
- [4] F. T. Freund, I. G. Kulahci, G. Cyr, J. Ling, M. Winnick, J. Tregloan-Reed, and M. M. Freund, “Air ionization at rock surfaces and pre-earthquake signals,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 71, no. 17—18, p. 1824—1834, Dec. 2009. [Online]. Available: <http://dx.doi.org/10.1016/J.JASTP.2009.07.013>
- [5] H. Woith, G. M. Petersen, S. Hainzl, and T. Dahm, “Review: Can animals predict earthquakes?” *Bulletin of the Seismological Society of America*, vol. 108, no. 3A, p. 1031—1045, Apr. 2018. [Online]. Available: <http://dx.doi.org/10.1785/0120170313>
- [6] N. J. Lindsey and E. R. Martin, “Fiber-optic seismology,” *Annual Review of Earth and Planetary Sciences*, vol. 49, no. Volume 49, 2021, pp. 309—336, 2021. [Online]. Available: <https://www.annualreviews.org/content/journals/10.1146/annurev-earth-072420-065213>
- [7] A. T. Ismail-Zadeh, *Earthquake Prediction and Forecasting*. Springer Netherlands, 2013, p. 225—231. [Online]. Available: http://dx.doi.org/10.1007/978-1-4020-4399-4_106
- [8] Y. Ogata, “Statistical models for earthquake occurrences and residual analysis for point processes,” *Journal of the American Statistical Association*,

- vol. 83, no. 401, p. 9—27, Mar. 1988. [Online]. Available: <http://dx.doi.org/10.1080/01621459.1988.10478560>
- [9] J. Zhuang, M. J. Werner, S. Zhou, D. Harte, and S. Hainzl, “Basic models of seismicity: spatiotemporal models,” 2011. [Online]. Available: <http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/Zhuang-et-al-2011-CORSSA-Spatiotemporal-models.pdf>
- [10] J. C. Villegas-Lanza, M. Chlieh, O. Cavalié, H. Tavera, P. Baby, J. Chire-Chira, and J. Nocquet, “Active tectonics of peru: Heterogeneous interseismic coupling along the nazca megathrust, rigid motion of the peruvian sliver, and subandean shortening accommodation,” *Journal of Geophysical Research: Solid Earth*, vol. 121, no. 10, p. 7371—7394, Oct. 2016. [Online]. Available: <http://dx.doi.org/10.1002/2016JB013080>
- [11] R. J. Geller, “Earthquake prediction: a critical review,” *Geophysical Journal International*, vol. 131, no. 3, p. 425—450, Dec. 1997. [Online]. Available: <http://dx.doi.org/10.1111/j.1365-246X.1997.tb06588.x>
- [12] A. Darzi, B. Halldorsson, B. Hrafnkelsson, and K. S. Vogfjörð, “Short-term bayesian etas spatiotemporal forecasting of the Ölfus 2008 earthquake sequence in iceland,” *Tectonophysics*, vol. 839, p. 229522, Sep. 2022. [Online]. Available: <http://dx.doi.org/10.1016/j.tecto.2022.229522>
- [13] A. Mignan and J. Woessner, “Estimating the magnitude of completeness for earthquake catalogs,” 2012. [Online]. Available: <http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/Mignan-Woessner-2012-CORSSA-Magnitude-of-completeness.pdf>
- [14] L. Gulia, S. Wiemer, and M. Wyss, “Catalog artifacts and quality control,” 2012. [Online]. Available: http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/gulia_et_al.pdf

- [15] Y. Ogata, “Statistics of earthquake activity: Models and methods for earthquake predictability studies,” *Annual Review of Earth and Planetary Sciences*, vol. 45, no. 1, p. 497—527, Aug. 2017. [Online]. Available: <http://dx.doi.org/10.1146/annurev-earth-063016-015918>
- [16] J. Zhuang, M. J. Werner, S. Zhou, S. Hainzl, and D. Harte, “Basic models of seismicity: temporal models,” 2012. [Online]. Available: <http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/Zhuang-et-al-2012-CORSSA-Temporal-models.pdf>
- [17] R. Console, D. D. Jackson, and Y. Y. Kagan, “Using the etas model for catalog declustering and seismic background assessment,” *Pure and Applied Geophysics*, vol. 167, no. 6—7, p. 819—830, Mar. 2010. [Online]. Available: <http://dx.doi.org/10.1007/s00024-010-0065-5>
- [18] T. Omi, Y. Ogata, Y. Hirata, and K. Aihara, “Estimating the etas model from an early aftershock sequence,” *Geophysical Research Letters*, vol. 41, no. 3, p. 850—857, Feb. 2014. [Online]. Available: <http://dx.doi.org/10.1002/2013GL058958>
- [19] A. M. Lombardi and W. Marzocchi, “The etas model for daily forecasting of italian seismicity in the csep experiment,” *Annals of Geophysics*, vol. 53, no. 3, Nov. 2010. [Online]. Available: <http://dx.doi.org/10.4401/ag-4848>
- [20] d. k. lafmejani and E. Shabani, “Modeling earthquake data using etas model to forecast aftershock subsequences applying different parameterizations in kermanshah region, iran,” Sep. 2022. [Online]. Available: <http://dx.doi.org/10.21203/rs.3.rs-1996634/v1>
- [21] T. Nishikawa and T. Nishimura, “Development of an epidemic-type aftershock-sequence model explicitly incorporating the seismicity-triggering effects of slow slip events,” *Journal of Geophysical Research: Solid Earth*,

- vol. 128, no. 5, May 2023. [Online]. Available: <http://dx.doi.org/10.1029/2023JB026457>
- [22] L. Mizrahi, S. Nandan, and S. Wiemer, “Embracing data incompleteness for better earthquake forecasting,” *Journal of Geophysical Research: Solid Earth*, vol. 126, no. 12, Dec. 2021. [Online]. Available: <http://dx.doi.org/10.1029/2021JB022379>
- [23] Kowshik.C, “Understanding point processes,” 2020. [Online]. Available: <https://kowshikchilamkurthy.medium.com/understanding-point-processes-6e3d2f6c5480>
- [24] F. Omori, “On the after-shocks of earthquakes,” *The journal of the College of Science, Imperial University, Japan = 帝國大學紀要. 理科*, vol. 7, pp. 111–200, 1894, application/pdf. [Online]. Available: <https://doi.org/10.15083/00037562>
- [25] C. Molkenhuth, C. Donner, S. Reich, G. Zöller, S. Hainzl, M. Holschneider, and M. Opper, “Gp-etas: semiparametric bayesian inference for the spatio-temporal epidemic type aftershock sequence model,” *Statistics and Computing*, vol. 32, no. 2, Mar. 2022. [Online]. Available: <http://dx.doi.org/10.1007/s11222-022-10085-3>
- [26] S. Studer, T. B. Bui, C. Drescher, A. Hanuschkin, L. Winkler, S. Peters, and K.-R. Müller, “Towards crisp-ml(q): A machine learning process model with quality assurance methodology,” *Machine Learning and Knowledge Extraction*, vol. 3, no. 2, p. 392–413, Apr. 2021. [Online]. Available: <http://dx.doi.org/10.3390/make3020020>
- [27] E. Hauksson, J. L. Hardebeck, and J. Woessner, “What is an instrumental seismicity catalog?” 2010. [Online]. Available: http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/woessner_et_al.pdf
- [28] CENSIS, “Igp_datos_sismicos,” 2024. [Online]. Available: <https://ultimosismo.igp.gob.pe/descargar-datos-sismicos>

- [29] D. H. Mark Naylor, Katerina Orfanogiannaki, “Exploratory data analysis: magnitude, space and time,” 2010. [Online]. Available: <http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/Naylor-et-al-2010-CORSSA-Exploratory-data-analysis.pdf>
- [30] C. S. Mueller, “Earthquake catalogs for the usgs national seismic hazard maps,” *Seismological Research Letters*, vol. 90, no. 1, p. 251—261, Oct. 2018. [Online]. Available: <http://dx.doi.org/10.1785/0220170108>
- [31] E. M. Scordilis, “Empirical global relations converting m_s and m_b to moment magnitude,” *Journal of Seismology*, vol. 10, no. 2, p. 225—236, Apr. 2006. [Online]. Available: <http://dx.doi.org/10.1007/s10950-006-9012-4>
- [32] A. Jalilian, “Etas: An r package for fitting the space-time etas model to earthquake data,” *Journal of Statistical Software*, vol. 88, no. Code Snippet 1, 2019. [Online]. Available: <http://dx.doi.org/10.18637/jss.v088.c01>

