



ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

PROCESOS PUNTUALES ESPACIALES Y SUS APLICACIONES

PROCESOS PUNTUALES DE NEYMAN – SCOTT

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERA
MATEMÁTICA**

DEBBIE ELIZABETH ECHANIQUE MUÑOZ

debbie.echanique@epn.edu.ec

DIRECTORA: YANDIRA DENISSE CUVERO CALERO

yandira.curvero@epn.edu.ec

DMQ, JULIO 2024

CERTIFICACIONES

Yo, DEBBIE ELIZABETH ECHANIQUE MUÑOZ, declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

Debbie Elizabeth Echanique Muñoz

Certifico que el presente trabajo de integración curricular fue desarrollado por Debbie Elizabeth Echanique Muñoz, bajo mi supervisión.

Yandira Denisse Cuvero Calero
DIRECTORA

DECLARACIÓN DE AUTORÍA

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como los productos resultantes del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

Debbie Elizabeth Echanique Muñoz

Yandira Denisse Cuvero Calero

RESUMEN

Ecuador es uno de los países con alto riesgo sísmico, ya que es parte del Cinturón de Fuego del Pacífico y se encuentra sobre varios sistemas de fallas geológicas, las cuales al chocar, generan movimientos telúricos, que dependiendo de su profundidad e intensidad, pueden generar graves daños. En el presente trabajo se estudian las características espaciales de los sismos desde el punto de vista de procesos puntuales, para lo cual, se consideran pruebas de hipótesis, K función y el conteo de cuadrantes. Luego se ajusta un modelo espacio temporal, ETAS, el cual nos permite identificar las ubicaciones con más actividad sísmica en la región.

Palabras clave: Sismos, procesos puntuales, ETAS.

ABSTRACT

Ecuador is one of the countries with high seismic risk, since it is part of the Ring of Fire and its located under several geological fault systems, which when colliding produce telluric movements, that depending on their depth and intensity can cause serious damage. In this work the spatial characteristics of earthquakes are studied from a spatial point process point of view, for which we consider hypothesis tests, the K-function and the quadrant test. Then a spatial-temporal model, ETAS, is fitted which allow us to identify the locations with most seismic activity in the region.

Keywords: earthquakes, spatial point process, ETAS.

AGRADECIMIENTOS

A mi familia; a mi madre Elizabeth y mi hermano Martín, por su apoyo incondicional y por enseñarme a seguir adelante a pesar de todo.

A Yandira Cuvero por toda la ayuda y comprensión.

A mis amigos, principalmente a Eli y Gabriel, por escucharme y apoyarme en todo momento.

A Loyde por su guía, palabras de aliento y por compartir su tiempo conmigo para escucharme.

Dedicatoria

*A mi papi y abuelitas
por todas sus enseñanzas
y recuerdos que nunca olvidaré*

Índice general

1. Descripción del componente desarrollado	1
1.1. Objetivo general	1
1.2. Objetivos específicos	1
1.3. Alcance	1
1.4. Marco teórico	2
1.4.1. Proceso puntual	2
1.4.2. Procesos Clusters	4
1.4.3. Proceso puntual espacio-temporal	5
1.4.4. Sismos	5
2. Metodología	7
2.1. Proceso Neyman-Scott	7
2.2. Catálogo de sismos	10
2.2.1. Base de datos	10
2.2.2. Datos Faltantes	15
2.3. Análisis Temporal	15
2.3.1. Variable Mw	15
2.3.2. Variable Profundidad	18
2.4. Análisis Espacial	20
2.4.1. Proceso puntual	20

2.4.2. Intensidad	23
2.4.3. La K-función	25
2.4.4. Conjuntos Clusters	26
2.4.5. Modelos Espaciales	27
2.4.6. Modelo ETAS	29
2.4.7. ETASFLP	35
3. Resultados, conclusiones y recomendaciones	38
3.1. Resultados	38
3.1.1. Catálogo	38
3.1.2. Modelos Temporales	39
3.1.3. Modelos Espaciales	41
3.1.4. Modelo ETAS	42
3.1.5. Modelo ETAS-FLP	44
3.2. Conclusiones	46
3.3. Recomendaciones	47
Bibliografía	48
A. Resultados modelos ETAS	50
A.0.1. Cuando $m_0 = 4.5$	50
A.0.2. Cuando $m_0 = 5$	53

Índice de figuras

2.1. Simulación de un proceso de Puntual de Neyman-Scott . . .	8
2.2. K-función Simulación	8
2.3. Conteo de Cuadrantes Simulación	9
2.4. sigma 0.57	10
2.5. sigma 0.375	10
2.6. sigma 0.15	10
2.7. sigma 0.01	10
2.8. Número de sismos 1963-2009	11
2.9. Número de sismos en un solo Día	12
2.10.Magnitud	13
2.11.Clasificación sismos según su Profundidad	14
2.12.Descomposición Serie de Tiempo Máxima Mw Mensual	16
2.13.EACF Serie Máximo Mw	16
2.14.Residuos	17
2.15.ACF serie de Tiempo Mínima Profundidad Mensual	18
2.16.Descomposición de la serie de Tiempo	18
2.17.ACF residuos ETS(A,N,N)	19
2.18.Ventana Espacial	20
2.19.Número de sismos en las provincias de Ecuador	21

2.20	Sismos en ubicaciones repetidas	21
2.21	Proceso Puntual	22
2.22	Conteo de sismos por Cuadrante	22
2.23	Default	24
2.24	Diggle	24
2.25	Verosimilitud	24
2.26	Scott	24
2.27	Gráfico de la K-función	25
2.28	Gráfico de la función G	25
2.29	Domain	26
2.30	Clasificación	26
2.31	Residuos suavizados modelo 2	28
3.1.	Gráfico catálogo $m_0=3$	38
3.2.	Predicción a 10 meses Máximo Mw	39
3.3.	Predicción vs Valores reales	40
3.4.	Predicción a 10 meses	40
3.5.	Predicción vs Valores reales ETS(A,N,N)	41
3.6.	Residuos suavizados modelo 2	42
3.7.	Resultados modelo ETAS $m_0 = 4$	42
3.8.	Residuos ETAS $m_0=4$	43
3.9.	Clasificación de sismos $m_0 = 4$ con probabilidad mayor a 0.95	44
3.10	Intensidad réplicas	45
3.11	Intensidad Total	45
3.12	Fondo	45
3.13	Todo el modelo	45
3.14	Intensidad total con los puntos observados	46
A.1.	Catálogo $m_0 = 4,5$	50

A.2. Estimación parámetros	51
A.3. Estimaciones ETAS 4.5	51
A.4. ETAS 4.5	52
A.5. Clasificación de sismos $m_0=4.5$	52
A.6. Catálogo $m_0 = 5$	53
A.7. Estimación parámetros ETAS 5	53
A.8. Estimaciones ETAS 5	54
A.9. Residuos modelo ETAS 5	54
A.10 Clasificación de sismos $m_0=4.5$	55

Capítulo 1

Descripción del componente desarrollado

1.1. Objetivo general

Estudiar las propiedades y características de un Proceso Puntual de Poisson inhomogéneo.

1.2. Objetivos específicos

1. Simular un proceso puntual de Neyman-Scott.
2. Ajustar un modelo utilizando un Proceso Puntual no Homogéneo a datos reales y validarlo.
3. Estudiar el tratamiento de valores perdidos en una serie temporal.
4. Ajustar un modelo espacio-temporal a un catálogo sísmico.

1.3. Alcance

El presente trabajo se enfoca en analizar los sismos que se encuentran registrados desde 1901 hasta 2009 en el Catálogo Homogenizado del Instituto Geofísico de la Escuela Politécnica Nacional del Ecuador, el cual puede ser descargado en <https://www.igepon.edu.ec/descarga-de-datos>.

1.4. Marco teórico

1.4.1. Proceso puntual

Un proceso puntual espacial X es un subconjunto finito numerable y aleatorio de puntos que se encuentran dentro de una ventana acotada W subconjunto de un espacio S , donde $S \subseteq \mathbb{R}^d$ [12].

Intensidad La intensidad de un proceso puntual se define como el número esperado de puntos por unidad de área.

Si $S \subset \mathbb{R}^2$, la intensidad es igual a :

$$\lambda(x) = \lim_{|d_x| \rightarrow 0} \frac{E(N(d_x))}{|d_x|}$$

Donde d_x corresponde al valor del área infinitesimal en donde se puede encontrar al punto x [11].

Proceso de Poisson Un proceso se dice que es un proceso de Poisson si cumple con las siguientes propiedades:

- El número de eventos que se encuentran dentro de cualquier región A , lo cual se denota por $N(A)$, siguen una distribución de Poisson, con media igual a:

$$\mu(A) = \int_A \lambda(x) d_x$$

de donde se tiene que:

$$P(N(A) = n) = \exp(-\mu(A)) \times \frac{\mu(A)^n}{n!}$$

- Dado que $N(A)=n$, la ubicación de los n eventos en la región de A forman una muestra aleatoria independiente en A , tal que la función de densidad es proporcional a la intensidad $\lambda(\cdot)$ [11].

Proceso espacial completamente aleatorio Un proceso espacial completamente aleatorio o también conocido como proceso puntual homogé-

neo de Poisson (CSR) con función de intensidad constante $\lambda > 0$ es un proceso puntual localmente finito que cumple con las siguientes propiedades:

1. El número de puntos $N(X \cap B)$ sigue una distribución de Poisson.
2. El número esperado de puntos que se encuentran en B es igual a :

$$E(N(X \cap B)) = \lambda * \text{área}(B) \quad \forall B \subset D$$

3. Si $B_i \cap B_j = \emptyset \quad \forall i \neq j$ entonces $N(X \cap B_i)$ y $N(X \cap B_j)$ son variables aleatorias independientes.
4. Dado que $N(X \cap B) = n$, se tiene que los n puntos son independientes y se encuentran uniformemente distribuidos en B [11].

K-función La K-función es una técnica que fue propuesta por Brian Ripley en el año 1981 para analizar la correlación espacial en patrones puntuales, donde un patrón puntual es una realización del proceso puntual, y se la define como el número promedio acumulado de puntos que se encuentran a una distancia $r \in \mathbb{R}^+$ de un punto típico, es decir de un punto que pertenece al proceso puntual [4].

De donde se sigue que:

$$K(s) = \frac{E \left[\text{número de r-vecinos de u} \mid X \text{ tiene un punto en u} \right]}{\lambda}$$

Por otra parte, si tenemos un proceso CSR la K-función va a ser igual a

$$K(s) = \pi \times s^2$$

mientras que si $K(s) > \pi \times s^2$ esto nos indica que existe agrupamiento (clusters) y si es menor tenemos un proceso de inhibición [11].

Función del vecino más cercano La función G se define como la función de distribución acumulativa de las distancias de los vecinos más cercanos, tal que si $x_i \in x$, donde x es un patrón puntual, se tiene que:

$$d_i = d(x_i, x|x_i)$$

representa la distancia más pequeña desde x_i al patrón $x|x_i$ el cual consiste de todos los puntos menos x_i [4].

Para un proceso de Poisson tenemos que la función G es igual a

$$G_{\text{pois}}(r) = 1 - \exp(-\lambda\pi r^2) \quad \text{con } r \in \mathbb{R}^+$$

1.4.2. Procesos Clusters

Un proceso cluster es una modificación de un proceso de Poisson el cual incorpora aleatoriedad al solo poder observar la descendencia de los puntos padre, en donde la ubicación de cada uno de ellos depende de la localización del punto de origen formando de esta manera un patrón cluster [4].

Proceso cluster de Poisson En un proceso cluster de Poisson, se tienen un conjunto de puntos padre Y provenientes de un proceso de Poisson, en donde cada punto $y_i \in Y$ genera un conjunto finito Z_i de hijos, de tal forma que el conjunto de todos los descendientes constituyen un proceso puntual X [4].

Neyman-Scott Los procesos de Neyman-Scott es un caso particular de los procesos cluster de Poisson, en donde se asume que los clusters resultantes (hijas) son de Poisson. En este caso se tiene que el proceso puntual final esta conformado solo por los descendientes y por lo tanto el conjunto de puntos padre no son observables los cuales cumplen con las siguientes propiedades:

- Los puntos padre constituyen un proceso de Poisson.
- Los Clusters son independientes.
- Clusters diferentes que son desplazados a la misma ubicación del punto padre tienen la misma distribución.
- Los puntos hijos dentro de un mismo cluster son independientes [4].

1.4.3. Proceso puntual espacio-temporal

Un proceso espacio temporal puntual se define como un subconjunto X de $\mathbb{R}^{d-1} \times \mathbb{R}$ de donde se tiene que un punto $(s, t) \in X$ corresponde a un evento en $s \in \mathbb{R}^{d-1}$ el cual ocurre al instante $t \in \mathbb{R}$ [2].

1.4.4. Sismos

Un sismo se define como un movimiento o vibración del suelo producido por la liberación súbita de energía ocasionada por un desplazamiento de las masas rocosas en una falla tectónica. Para medir el tamaño y la fuerza del mismo se utiliza la magnitud y la intensidad. Se considera un terremoto si este ha provocado daños severos en las edificaciones o si ha dejado víctimas [8]. Se conoce como enjambre sísmico a la ocurrencia de varios sismos de magnitudes aproximadamente similares en un tiempo de lapso corto en localizaciones cercanas, los cuales pueden durar días o semanas [6].

Catálogo Sísmico Un catálogo sísmico es una lista ordenada cronológicamente de registros de datos históricos de sismos dentro de una región S durante un periodo de tiempo. Usualmente contiene las variables: tiempo, magnitud y profundidad focal de todos aquellos sismos iguales o mayores a un umbral, próximos a S . Desde un punto de vista estadístico, se define como una base de datos dentro de una ventana espacio temporal [1].

Por otra parte, existen diferentes tipos de magnitudes. Por ejemplo: magnitud de duración (Md), magnitud de ondas internas (m_b), magnitud de momento (M_w), entre otras, por lo que para la creación del catálogo se suele definir una sola escala, lo cual se conoce como Homogenizar.

Ley de Gutenberg-Richter La ley de Gutenberg-Richter establece que a mayor magnitud menos sismos deben de haber, por lo cual, se fija una proporción constante de sismos tanto para magnitudes superiores e inferiores en una determinada zona geográfica, de donde se asume que el número acumulado de sismos conforme aumenta la magnitud varía

linealmente , lo cual se expresa mediante la siguiente fórmula:

$$\log_{10}(N_m) = a - bm$$

con a y b constantes. Donde N_m representa el número de eventos del catálogo cuya magnitud es mayor a m. Esta expresión es equivalente a decir que m sigue una distribución exponencial [3]. El tener datos que no se ajustan bien, implicaría que el catálogo es incompleto.

Por otra parte, la existencia de estacionariedad en el catálogo se tiene si se cumple que:

$$N_t = \lambda_0 \times t$$

donde $\lambda_0 > 0$ y N_t es el número de sismos que ocurrieron hasta el instante t.

Capítulo 2

Metodología

2.1. Proceso Neyman-Scott

Para realizar la simulación de Neyman-Scott utilizaremos la función `rNeymanScott` de la librería `spatstat` en R, la cual tiene los siguientes parámetros:

- `kappa`: El cual corresponde al valor de la intensidad del proceso cluster de Poisson padre.
- `expand`: Corresponde al valor de la expansión de la ventana simulada para generar los puntos padre, el cual debe ser positivo.
- `rcluster`: Hace referencia a una función que genere aleatoriamente clusters que reciba como parámetros la ubicación del punto padre.
- `win`: Es la ventana en donde se va a realizar la simulación.
- `n`: Número de simulaciones a realizar.

Para simular un proceso de Neyman-Scott primero se genera un proceso de Poisson de puntos padre con intensidad κ en una ventana expandida. Luego se procede a reemplazar a cada uno de los puntos padre por un agrupación de puntos (clusters) de forma aleatoria, los cuales

son creados a partir de la función `rcluster`. Después se procede a combinar los puntos generados en el paso anterior formando un solo patrón puntual restringido a la ventana (`win`).

Se realizó una simulación, donde se tomó como ventana a un rectángulo de dimensión $[0, 3] \times [1, 5]$, intensidad igual a 10, `expand` igual a 0.15 y se retorna un patrón puntual que se puede ver en la Figura 2.1 que contiene 5 puntos aleatorios distribuidos uniformemente en un círculo de radio 0.10, para lo cual se utilizó la función `runifdisc`.

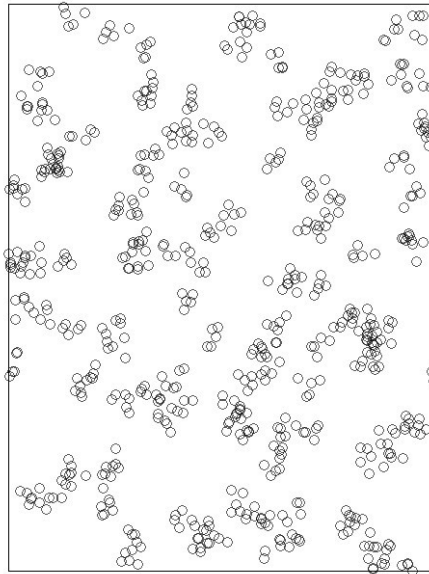


Figura 2.1: Simulación de un proceso de Puntual de Neyman-Scott

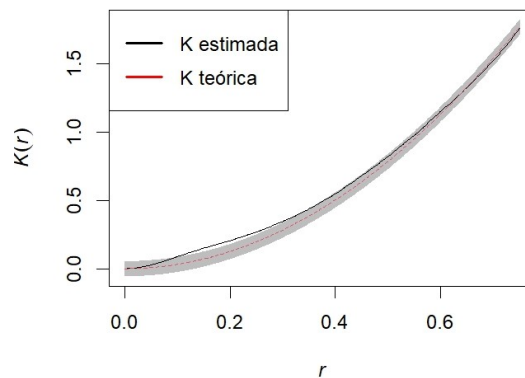


Figura 2.2: K-función Simulación

En la Figura 2.2 se presenta una comparación de la K teórica de un proceso de Poisson y la K estimada de la simulación. Al ser esta última

mayor a la K teórica concluimos que es un proceso de Cluster, lo cual es consecuente con las propiedades del proceso de Neyman-Scott.

Para analizar la estructura espacial de los datos realizaremos primero un conteo de cuadrantes, lo cual consiste en dividir la región de estudio (Figura 2.1) en B_1, B_2, \dots, B_m subregiones, que en nuestro caso, son rectángulos de igual área, los cuales no se intersecan entre sí, para luego proceder a contar el número de puntos que se encuentran en cada uno de ellos: $n_j = n(X \cap B_j), \forall j = 1, \dots, m$. De donde tenemos que n_j , corresponden a las estimaciones del valor esperado del número de puntos por subregión: $E(n(X \cap B_j))$, los cuales serían iguales en promedio si la intensidad es homogénea, caso contrario se sospecharía que es inhomogénea [11].

Para lo cual se utiliza la función `quadrantcount` de la librería `spatstat`, la cual construye un cuadrícula de dimensión $n_x \times n_y$, cuyos cuadrantes son de igual tamaño, de donde se tomo a n_x y n_y igual a 15.

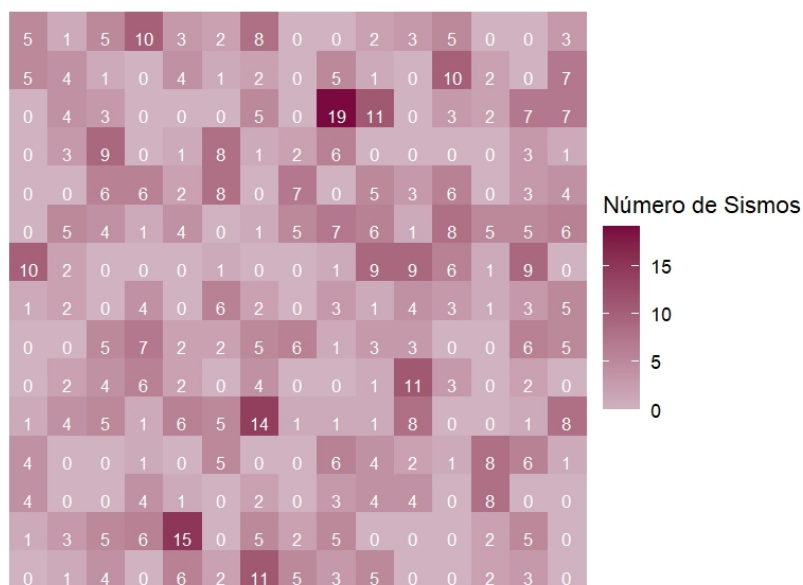


Figura 2.3: Conteo de Cuadrantes Simulación

En la Figura 2.3 se muestra la prueba de cuadrantes, en donde notamos que el número de puntos no es el mismo en cada uno ellos, lo que nos indica que es un Proceso de Poisson inhomogéneo. Por otra parte, probamos diferentes valores de sigma para la intensidad, este valor con-

trola la suavidad del ancho de banda del Kernel estimado en la función de intensidad. Como podemos observar en la Figura 2.4 a 2.7 entre menor es el valor de sigma la función de intensidad es más rugosa.

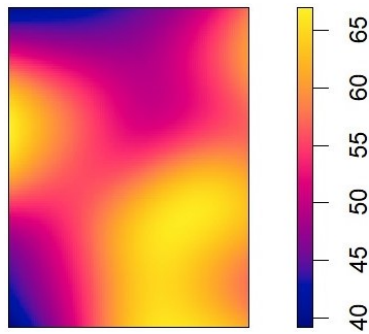


Figura 2.4: sigma 0.57

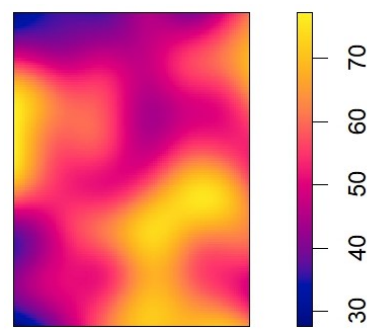


Figura 2.5: sigma 0.375

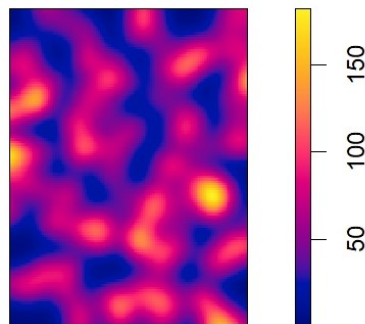


Figura 2.6: sigma 0.15

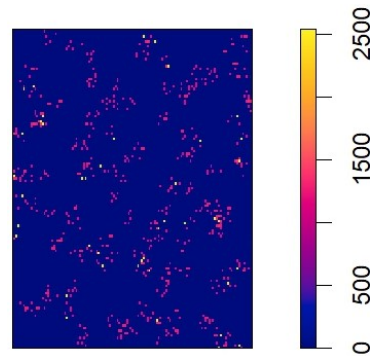


Figura 2.7: sigma 0.01

2.2. Catálogo de sismos

2.2.1. Base de datos

Ecuador al ser parte del Cinturón del Fuego del Pacífico y encontrarse entre dos placas tectónicas opuestas (la placa de Nazca y la placa Sudamericana) sufre de una alta actividad sísmica. Por ejemplo, el primer sismo del cual se tiene conocimiento ocurrió en el año 1541, en la provincia del Napo con una magnitud de 8 en la escala de Richter, mientras que el sismo del año 1906, cuya magnitud fue de 8.35 [M_w] con epicentro

en el Pacífico, frente a las costas de Ecuador y Colombia, se encuentra registrado como uno de los más fuertes en la historia a nivel mundial, el cual ocasionó hasta un tsunami, siendo el cantón de Río Verde ubicado en la provincia de Esmeraldas el más afectado.

La placa tectónica oceánica Nazca se mueve alrededor de 9 cm al año por debajo de la placa terrestre Sudamericana, a este proceso se lo conoce como subducción, las cuales al chocar liberan la energía acumulada ocasionando fuertes sismos. Por otra parte, Ecuador está sentado sobre tres sistemas de fallas geológicas, la principal es la generada por la subducción de las placas antes mencionadas, la cual va desde Puná en Guayas hasta Chingual. Estas fallas geológicas generan fracturas ocasionando movimientos telúricos, que dependiendo de su profundidad pueden ser como no percibidos por la población, por lo que el peligro sísmico en Ecuador es emitente, por ello en 1983, se funda el Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN), el cual es el encargado oficial (desde el 2003) de monitorear, determinar e informar sobre el riesgo sísmico y volcánico del Ecuador.

Para el presente estudio, se cuenta con el Catálogo Homogenizado del año 2009, el cual contiene el registro de 10823 sismos de magnitud mayores o igual a tres [M_w] desde 1901 hasta el 2009. De donde tenemos que no existe el registro de ningún sismo en un total de 28 años entre el periodo comprendido de 1901 a 1962 y además tenemos que en los primeros 62 años el número total de sismos registrados es de 68 por lo que estos años no se incluirán en el análisis. Por lo tanto, desde 1963 hasta el 2009 se tienen 10755 datos, siendo el año 2005 el año con mayor registros (1078) y el año 1976 el menor de ellos con 33 sismos, como podemos observar en la siguiente figura:

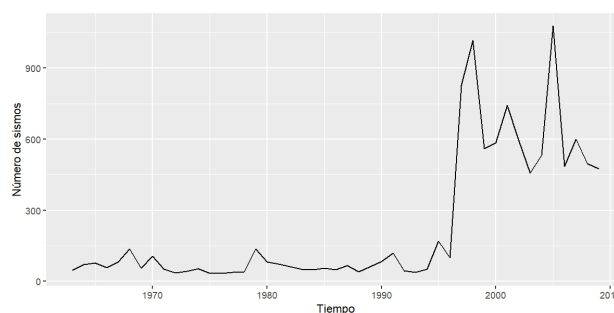


Figura 2.8: Número de sismos 1963-2009

El incremento del número de sismos en los años 1996, 1998, 2002 y 2005, se debe a que en la zona como la Isla de la Plata se registró la presencia de enjambres sísmicos, ocasionando que se libere una cantidad significativa de energía; en particular, en el año 2005 la presencia de estos enjambres fueron registrados en los meses de enero y mayo los cuáles estuvieron acompañados por sismos lentos ocasionando 4 eventos de magnitud 6 [M_w] [8]. Es así que tenemos que el mayor número de sismos que se tienen registrados en un solo día es de 61, el cual corresponde precisamente al 25 de enero del 2005, los cuales ocurrieron en el oceano Pacífico, cerca de Manabí y el segundo mayor valor es de 43 (como podemos observar en la Figura 2.9) el cual ocurrió el 30 del mismo mes.

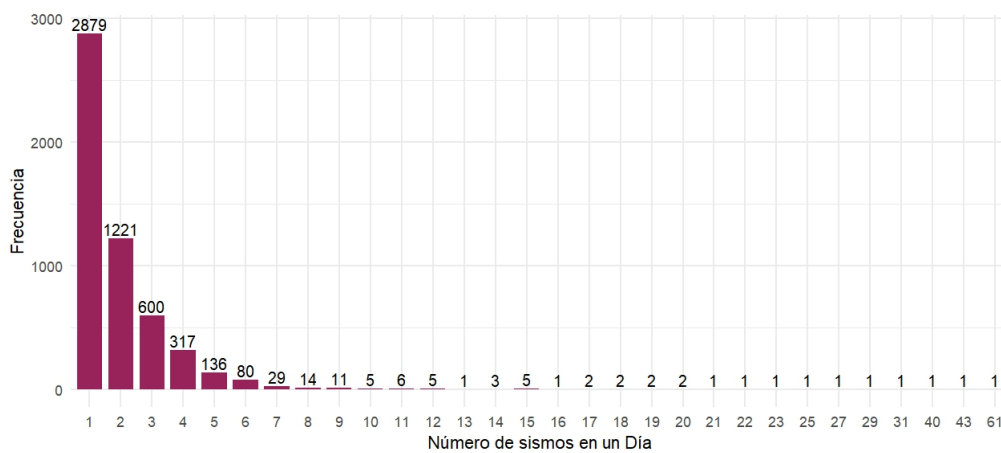


Figura 2.9: Número de sismos en un solo Día

Por otra parte, tenemos 4 fechas en las cuales 2 sismos fueron registrado en el mismo día y al mismo tiempo (hora, minuto,segundo), en las siguientes fechas: 19 de mayo 1968, 29 de diciembre 1970, 8 de diciembre de 1990 y 2 de marzo del 2001.

Cómo fuente del catálogo se tiene a : IGEPN, CENTENNIAL, International Seismological Center (ISC) y United States Geological Survey-National Earthquake Information Center (USGS-NEIC) y se cuenta con las siguientes variables: Fecha , la cual tiene el día, mes, año, hora, minuto y segundo, Latitud y Longitud del epicentro, Profundidad, Magnitud (M_w) y Fuente de cada sismo. Donde tenemos que M_w corresponde a la escala logarítmica sismológica de magnitud de momento que va de 0 a 10, la cual cuantifica la fuerza del sismo mediante la medición de la energía libera-

da, proporcionándonos el tamaño de la falla geológica que se rompió tras el sismo [9], fue introducida por Thomas C. Hanks y Hiroo Kanamori en el año 1979 y es la escala más confiable para comparar el tamaño entre sismos.

De esta forma tenemos que la magnitud más alta entre los sismos de 1963 al 2009 fue de 8.1 [M_w], la cual ocurrió el 12 de diciembre de 1979. Por otra parte, tenemos un total de 63 sismos cuya magnitud va desde 6 [M_w] a 8 [M_w] y 1360 sismos con Magnitud igual a 3 [M_w], los cuales como podemos observar en la Figura 2.10, han ido en aumento en los últimos años, mientras que en los primeros años podemos observar que la mayoría de los sismos tuvieron magnitud desde 4 [M_w] a 5 [M_w], los cuales conforman el 0.21 % de los datos.

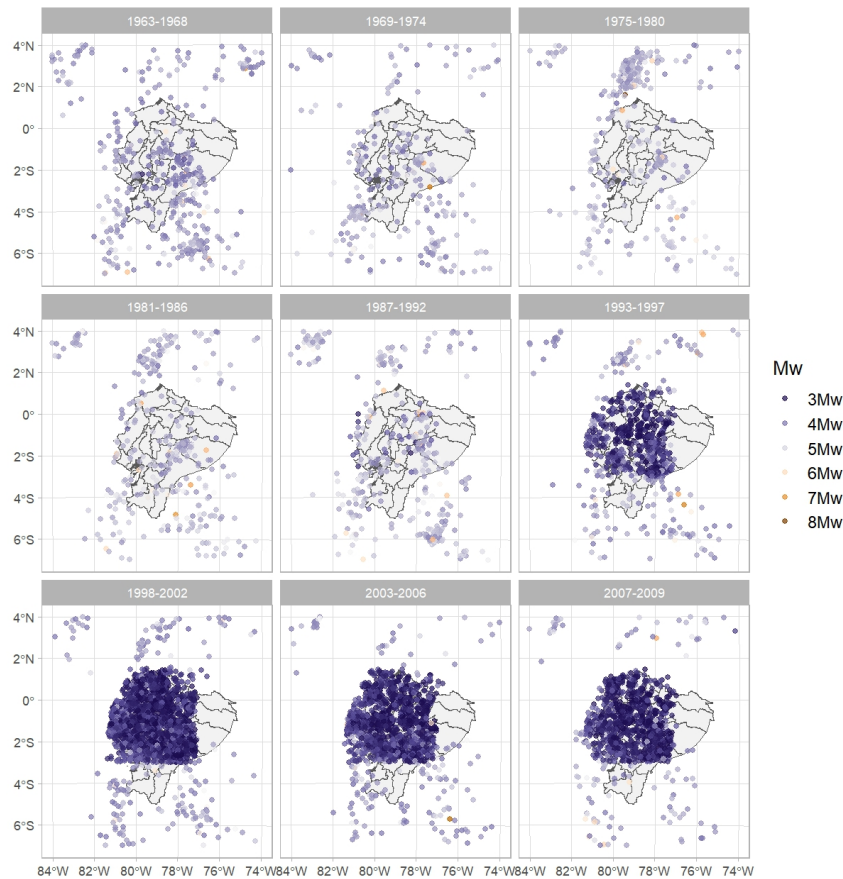


Figura 2.10: Magnitud

Por otra parte, la variable profundidad hace referencia a la distancia entre el hipocentro (lugar en donde se origina el sismo) y el epicentro (sitio en donde se manifiesta el sismo), la cual se mide en kilómetros y a partir de esto se pueden clasificar los sismos en tres tipos: Superficial, si la profundidad es menor a 70 km, Intermedia, entre 70 y 300 km y Profunda, mayor a 300 km.

La profundidad del hipocentro afecta la intensidad del sismo, al generarse un movimiento telúrico, la energía liberada al interior de la tierra se transmite a la superficie a través de las ondas sísmicas, las cuales son mayores cerca del hipocentro y disminuyen conforme aumenta la distancia hasta el epicentro, de donde a partir de este punto se mueven a una velocidad de entre 2 a 3 km/s las cuales pueden ocasionar una deformación en la superficie además de daños en las diferentes estructuras. De esta forma se tiene que a menor profundidad mayores daños pueden ocasionar.

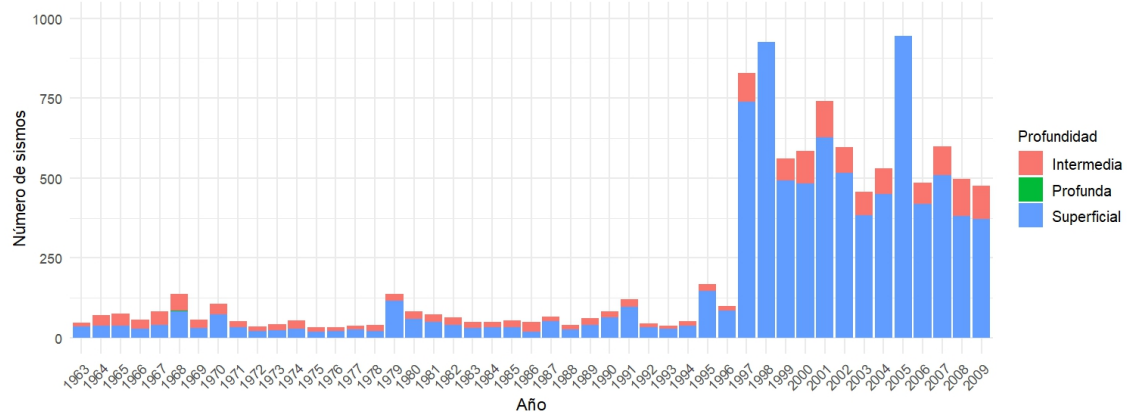


Figura 2.11: Clasificación sismos según su Profundidad

De acuerdo a la Figura 2.11 tenemos que la mayoría de los sismos se encuentran entre Superficial e Intermedia, siendo los Superficiales los más frecuentes.

2.2.2. Datos Faltantes

Desde 1963 hasta el 2009 no tenemos el registro de un sismo en nueve meses, indicados en la siguiente tabla :

Mes	Año
Septiembre	1973
Enero	1975
Junio	1976
Diciembre	1977
Febrero	1978
Octubre	1978
Julio	1979
Noviembre	1987
Mayo	1992

Para el análisis temporal, se les imputará a las variables Profundidad con valor de cero y para Magnitud con la media en dichos períodos.

2.3. Análisis Temporal

2.3.1. Variable Mw

Para analizar esta variable se realizó una serie de tiempo mensual tomando la mayor magnitud $[M_w]$ de cada uno de los meses, para lo cual fue necesario primero completar los meses faltantes utilizando el comando `fill_gaps` y reemplazando los valores NA por la media de los mayores valores de donde podemos observar al realizar la descomposición de la misma, en la Figura 2.12, que entre los años 1963 y 1964 se tiene una tendencia creciente y luego decreciente hasta el año 1967, mientras que desde el año 1970 al 1980 se mantiene constante, luego vuelve a repetirse este patrón hasta el año 1990 de donde se puede observar una mayor disminución hasta el año 2000 a partir del cual se puede ver una tendencia creciente. Por otra parte, tenemos que existe estacionalidad y que los residuos no siguen el comportamiento de un ruido blanco, por lo que se

puede ajustar un modelo para esta serie.

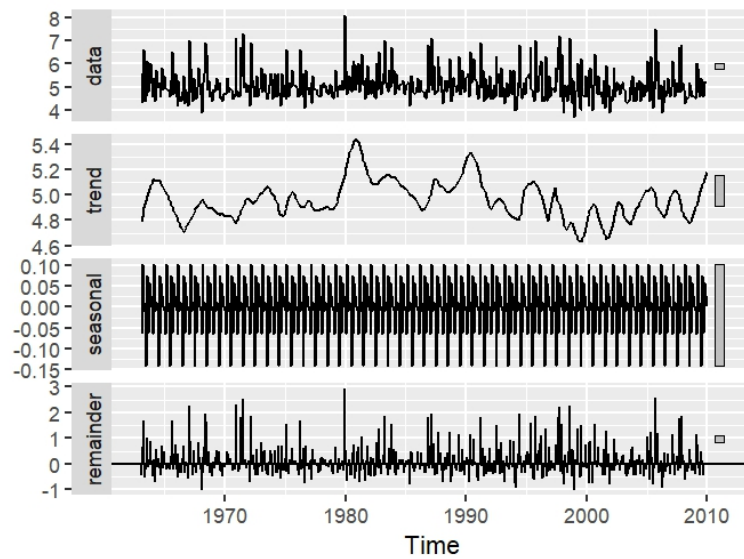


Figura 2.12: Descomposición Serie de Tiempo Máxima Mw Mensual

Para determinar que modelo sería adecuado, se realizó el gráfico del EACF, el cual se lo puede observar a continuación, por lo que se propone utilizar el modelo ARIMA(2,0,2).

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	o	o	o	o	o	o	x	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	x	o	o	o	o
3	x	x	x	o	o	o	o	o	o	x	o	o	o	o
4	o	x	x	x	o	o	o	o	o	o	o	o	o	o
5	x	o	x	x	x	o	o	o	o	o	o	o	o	o
6	x	x	x	o	o	o	o	o	o	o	o	o	o	o
7	x	x	x	o	o	o	x	o	o	x	o	o	o	o

Figura 2.13: EACF Serie Máximo Mw

De donde se procede a verificar si los residuos de dicho modelo son independientes y estacionarios, es decir si siguen el comportamiento de un ruido blanco:

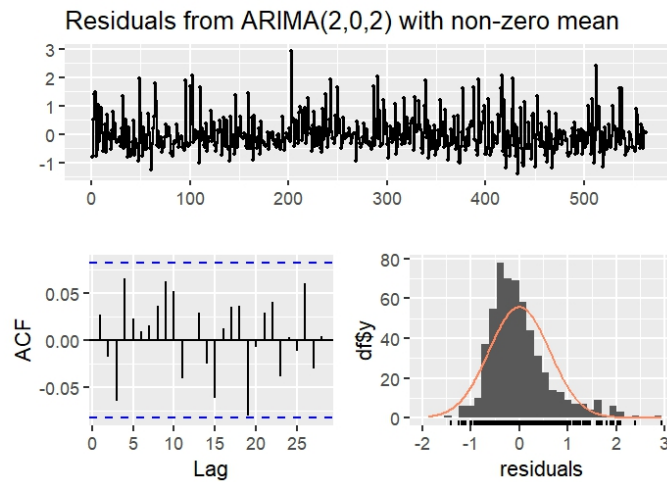


Figura 2.14: Residuos

- Independencia:

Como podemos observar en la Figura 2.14, en el gráfico del ACF, tenemos que todos los valores de las autocorrelaciones se encuentran dentro de las barras de confianza, por lo que se tiene que los residuos son independientes.

- Estacionariedad:

Como podemos observar en la gráfica, la mayoría de los puntos se encuentran alrededor de 0, por lo que se tendría que los residuos son estacionarios, lo cual podemos verificar por medio del test de Kwiatkowski-Phillips-Schmidt-Shin (kpss), el cual tiene como hipótesis:

H_0 : La serie es estacionaria

H_1 : La serie no es estacionaria

De donde se obtuvo un p-valor de 0.1, por lo que se acepta la hipótesis nula, es decir que se tiene que los residuos son estacionarios y al ser también independientes, siguen el comportamiento de un ruido blanco.

2.3.2. Variable Profundidad

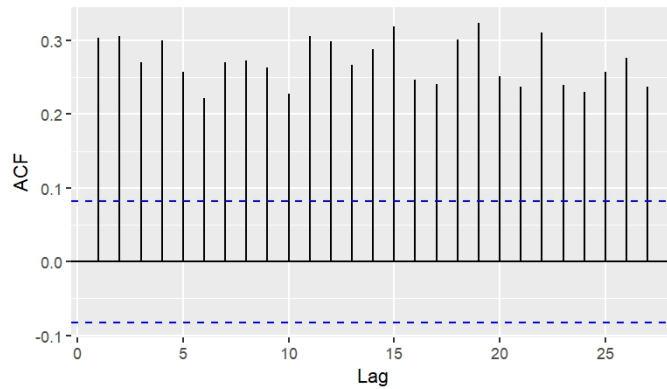


Figura 2.15: ACF serie de Tiempo Mínima Profundidad Mensual

Al igual que para la variable de Magnitud de Momento se realizó una serie de tiempo mensual pero se tomó esta vez la mínima profundidad cuyos valores NA fueron reemplazados por cero, de donde podemos ver en la Figura 2.15 que existe una autocorrelación significativa entre los datos. Por otra parte, al realizar la descomposición de la serie, podemos observar en la Figura 2.16, que a partir del año 1997 la profundidad de los sismos ha ido disminuyendo y que sigue una tendencia decreciente, por otra parte tenemos que existe estacionalidad y que los residuos no se asemejan a un ruido blanco por lo que se propone utilizar el modelo de ETS(A,N,N) de donde se procede a verificar si los residuos cumplen con las condiciones de un ruido blanco.

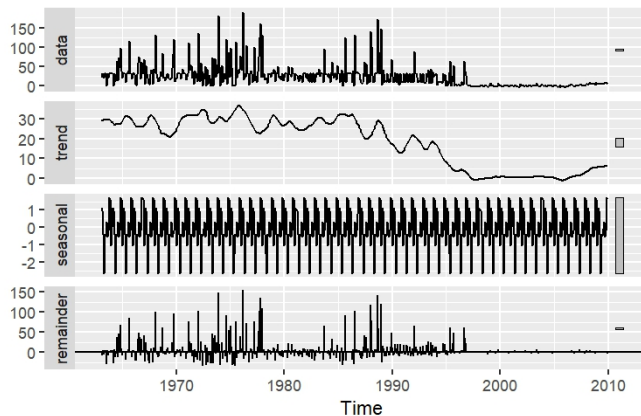


Figura 2.16: Descomposición de la serie de Tiempo

- Independencia:

Para verificar esta característica realizamos el gráfico del ACF.

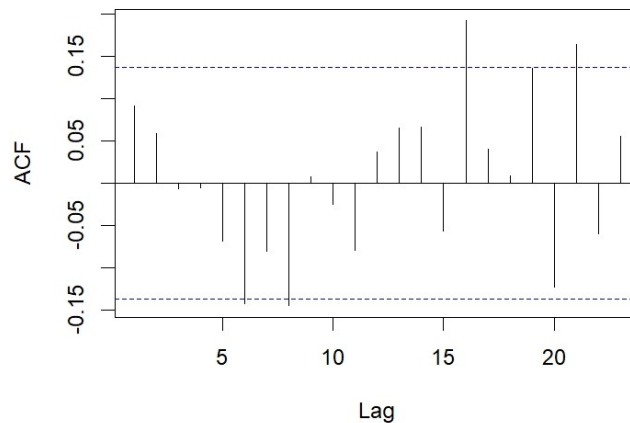


Figura 2.17: ACF residuos ETS(A,N,N)

como podemos observar en la gráfica no todos los valores de las autocorrelaciones se encuentran dentro de las líneas de confianza por lo que se procede a verificar si son o no independientes por medio de la prueba de hipótesis de Ljung Box:

H_0 = Los datos son independientes entre sí

H_1 = Los datos son dependientes

A partir del cual se obtuvo un p-valor de 0.3861, lo que quiere decir que no hay autocorrelación significativa y por lo tanto se puede considerar que los residuos del modelo son independientes.

- Estacionariedad

Para verificar si los residuos son estacionarios realizamos la prueba de Dickey Fuller , de donde se obtuvo un p-valor menor a 0.01, por lo que se tiene que la serie es estacionaria.

2.4. Análisis Espacial

2.4.1. Proceso puntual

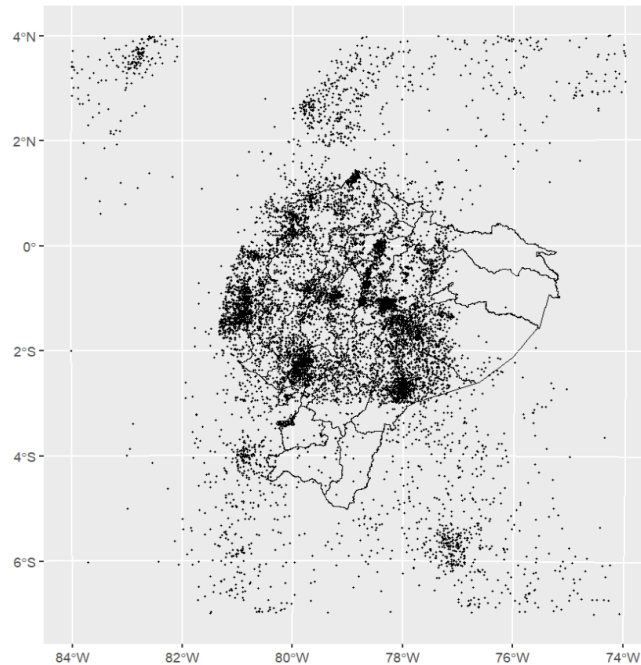


Figura 2.18: Ventana Espacial

Para la creación del proceso puntual, tenemos que la ventana espacial que se considera es un rectángulo que va desde -84.25° a -73.75° de longitud y de -7.25° a $+4.25^{\circ}$ de latitud, la cual cubre el área de aquellos eventos sísmicos que podrían ocasionar un significativo movimiento telúrico en Ecuador [5] los cuales se pueden observar en la Figura 2.18, donde tenemos que la mayoría de sismos ocurren en la zona central del país y en la zona de subducción, mientras que en la Figura 2.19, podemos ver que las provincias que han tenido más sismos de magnitud iguales o mayores a 3 [M_w] corresponden a Napo y Morona Santiago, de donde se tiene que los sismos cuyos epicentros se encuentran en dichas provincias suelen sentirse en las provincias de la Costa, ya que las ondas sísmicas se propagan con facilidad por la interfaz formada entre las placas de Nazca y Sudamericana [7].

Por otra parte, tenemos 1915 datos cuya ubicación se repite al menos 1 vez, los cuales se pueden apreciar en la Figura 2.20, donde además ve-

mos que estos se encuentran agrupados en las zonas de Ambato, Manta, Portoviejo y Morona Santiago.

Para poder crear nuestro proceso puntual vamos a utilizar el comando `ppp` el cual recibe como parámetros a las coordenadas correspondientes a longitud y latitud y la ventana espacial, que debe de estar en formato `owin` y dado que existen puntos que se repiten vamos agregar ruido a las coordenadas por medio del comando `jitter`, donde la Figura 2.21 corresponde a nuestro proceso puntual.

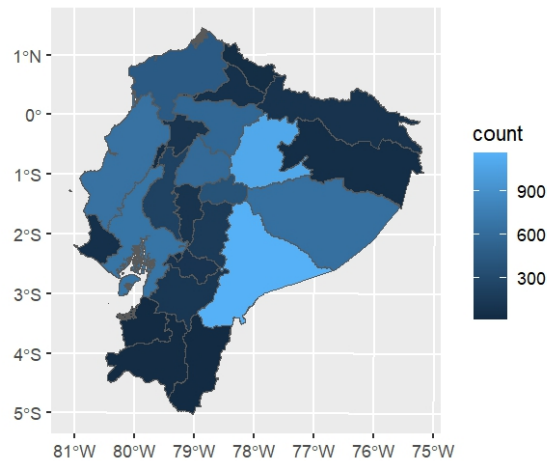


Figura 2.19: Número de sismos en las provincias de Ecuador

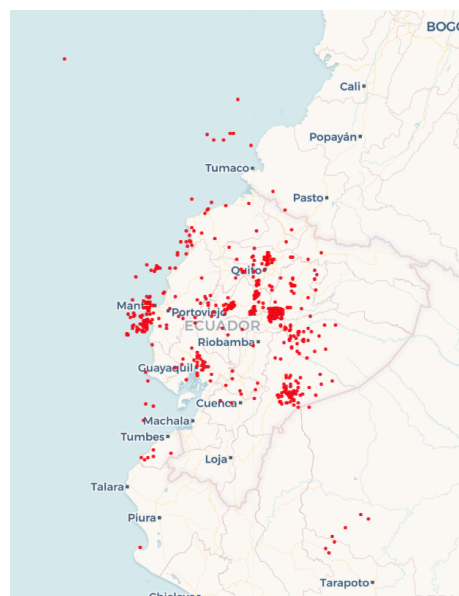


Figura 2.20: Sismos en ubicaciones repetidas

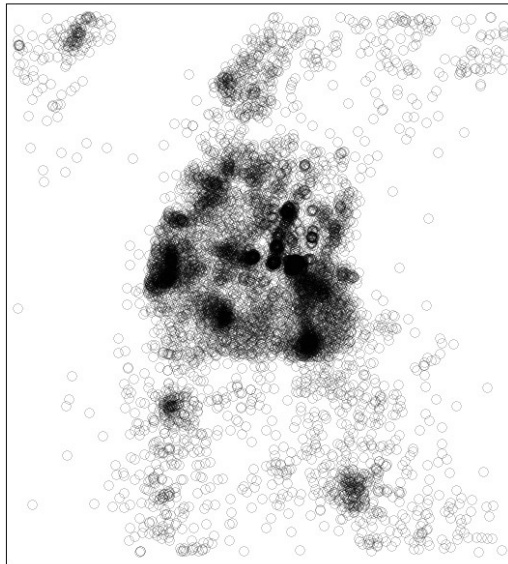


Figura 2.21: Proceso Puntual

Para analizar la estructura espacial primero realizaremos un conteo de cuadrantes con n_x y n_y igual a 10, el cual se puede observar en la Figura 2.22, en donde tenemos que el número de sismos no es el mismo en cada uno de los cuadrantes, por lo tanto se puede suponer que la intensidad es inhomogénea.

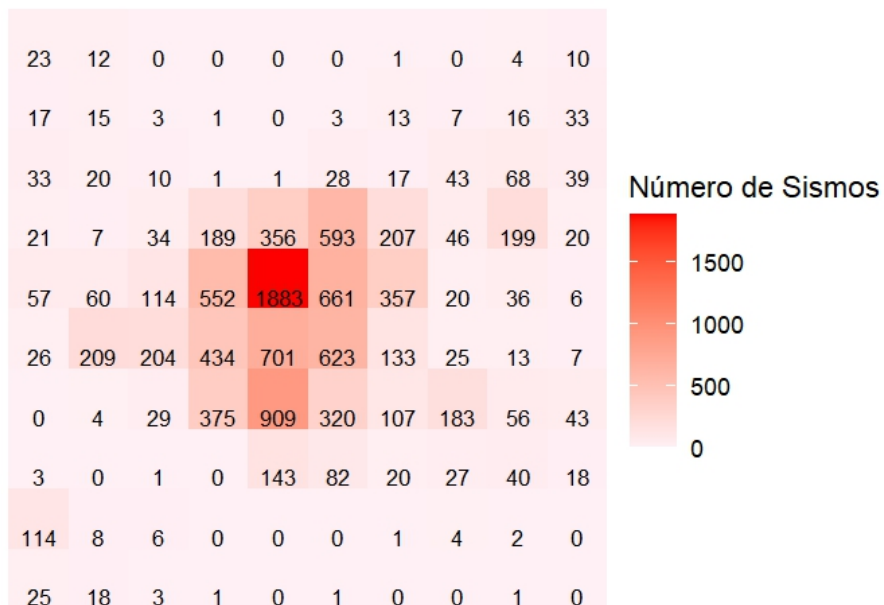


Figura 2.22: Conteo de sismos por Cuadrante

Además tenemos que los cuadrantes con altos valores se encuentran agrupados en la zona central de la ventana, donde el mayor de ellos, el de color rojo, corresponde a la región sierra- oriente donde se encuentra Tungurahua, Napo, Morona Santiago y Pastaza. Por otra parte, se tiene diecisiete cuadrantes en los que no se tiene ningún sismo.

Ahora se procede a verificar si el proceso puntual es o no completamente aleatorio (CSR), para lo cual, primero realizaremos el test de cuadrantes mediante el comando `quadrant.test`, el cual consiste en la siguiente prueba de hipótesis:

H_0 = El proceso puntual es CSR

H_1 = El proceso puntual no es CSR (regular o cluster)

en donde se compara el estadístico observado con la distribución Chi-cuadrado (χ^2), el cual es igual a:

$$X^2 = \sum_{i=1}^m \frac{n_i - n^*}{n^*}$$

en donde $n^* = \frac{n}{m}$ y n corresponde al número total de eventos observados [11]. De esta forma se tiene que el número esperado de sismos en cada cuadrante (n^*) es igual a 107.55 y el valor del estadístico es igual a 57933 el cual sigue una distribución Chi-cuadrado con 99 grados de libertad, de donde se obtiene un p-valor < 2.2e-16, por lo que se rechaza la hipótesis nula y por lo tanto se tiene que el proceso espacial no es completamente aleatorio.

2.4.2. Intensidad

Para estimar la intensidad de nuestro proceso puntual de forma no paramétrica vamos a utilizar el comando `density()`, el cual utiliza por defecto el Kernel Gaussiano y tiene como parámetro al ancho de banda (sigma), el cual controla el grado de suavidad del Kernel y nos retorna un objeto de tipo `im` .

Por defecto el ancho de banda se calcula mediante el tamaño de la ventana espacial, pero existen otros métodos que nos permiten identificar

un adecuado valor del mismo, como son el método de validación cruzada del error al cuadrado de Diggle y Bermans, el cual se lo puede utilizar con el comando `bw.diggle()` y asume que se tiene un proceso de Cox; otro método es el de verosimilitud `bw.ppl()`, el cual asume que el proceso es Poisson inhomogéneo, de donde obtenemos las siguientes Figuras

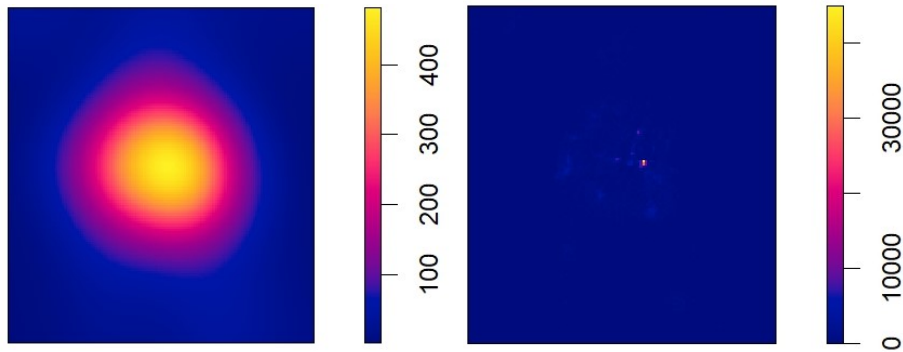


Figura 2.23: Default

Figura 2.24: Diggle

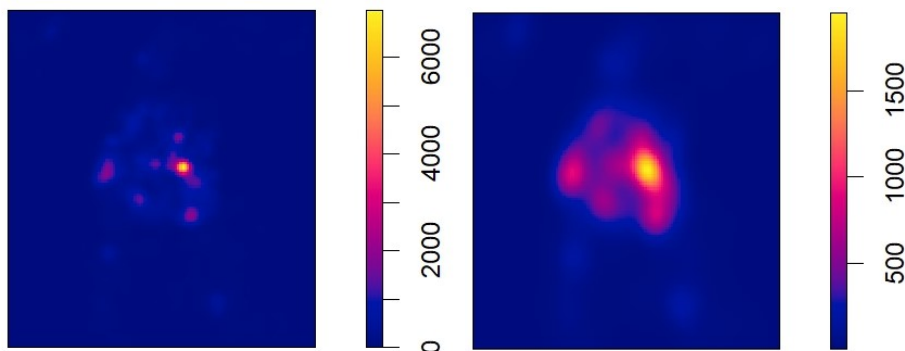


Figura 2.25: Verosimilitud

Figura 2.26: Scott

Vemos en la Figura 2.23 que el valor más alto de la intensidad se encuentra localizada en la mitad de la ventana y conforme nos vamos alejando esta va disminuyendo. Mientras que al utilizar los otros métodos podemos observar que logran identificar de mejor manera las regiones en las que se encuentran un mayor número de sismos, siendo el método de Diggle el más específico.

2.4.3. La K-función

Para estimar la K-función de nuestro proceso puntual utilizamos el comando Kest de donde obtenemos el siguiente gráfico:

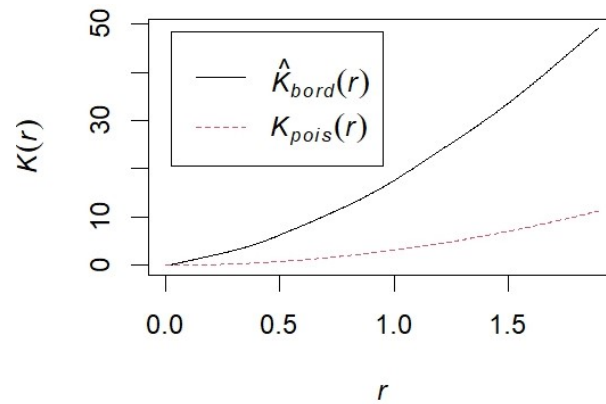


Figura 2.27: Gráfico de la K-función

Como podemos observar la estimación de la K-función, que es la función que se encuentra de color negro, está por encima de la K-función que se obtiene de un proceso de Poisson homogéneo, lo cual nos indica que existe agrupamiento (clusters). Esto también se puede verificar al realizar el gráfico de la función del vecino más cercano, de donde podemos observar que la G estimada se encuentra por encima de la G de un proceso de Poisson homogéneo.

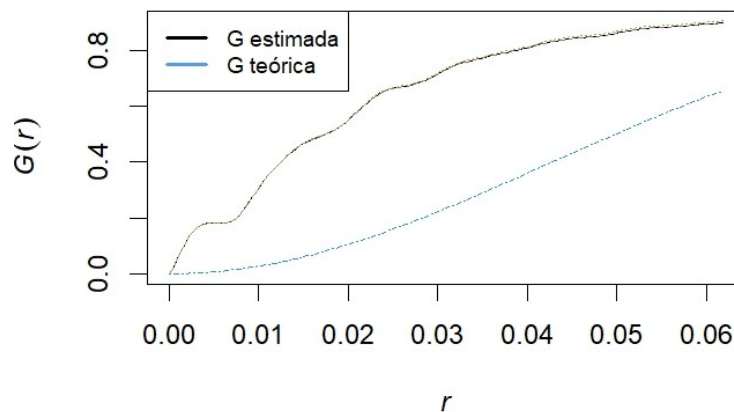


Figura 2.28: Gráfico de la función G

2.4.4. Conjuntos Clusters

Para poder identificar las características de las agrupaciones de un patrón puntual, podemos utilizar la función `clusterset`, la cual mediante el estimador de Allard-Fraley identifica aquellas regiones con una alta densidad de puntos. Luego utiliza la teselación de Dirichlet para dividir el espacio proporcionado en celdas, las cuales contienen aquellos puntos más cercanos a su punto central y se selecciona aquellas celdas de Dirichlet más pequeñas. El número total de celdas a seleccionar, m , se determina utilizando el criterio de máxima verosimilitud.



Figura 2.29: Domain

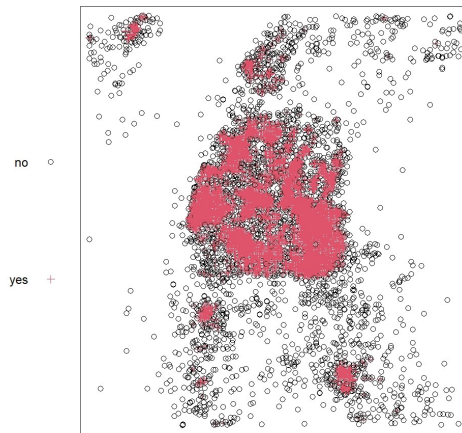


Figura 2.30: Clasificación

En la Figura 2.29 se presenta la unión de las celdas de Dirichlet, donde se tiene los subconjuntos clusters. Mientras que en la Figura 2.30, los puntos en rosado son aquellos sismos que son parte de un cluster y los círculos representan a los que no forman parte de uno. Además tenemos que las agrupaciones más grandes se encuentran al sur-oriente del país y en el océano Pacífico cerca de la costa ecuatoriana. Por otra parte, vemos que la mayoría de puntos que forman parte de un cluster se encuentran en la zona continental, mientras que la mayoría de puntos que no forman parte, se encuentran dispersos en la región de Perú y Colombia.

2.4.5. Modelos Espaciales

Como nuestro proceso puntual no es CSR, vamos a ajustar un Proceso de Poisson Inhomogéneo, con intensidad igual a [4]:

$$\lambda_{\theta}((x, y)) = \exp(B(u) + \theta^T Z(u))$$

Donde $Z(u)$ corresponde al vector de covariables, que en nuestro caso es 0 y $B(u)$ es una función que depende de las coordenadas u .

Para ello vamos a utilizar el comando `ppm()`, el cual tiene la siguiente sintáxis:

$$\text{ppm}(X \sim \text{trend}, \dots)$$

Donde X corresponde al proceso puntual y `trend` es el logaritmo de la función de intensidad.

Se ajustaron los siguientes modelos:

1. $B(u) = \theta_0 + \theta_1 y$

Tenemos que:

$$\lambda_{\theta}(x, y) = \exp(4,5171 + 0,0199y)$$

donde

	Estimación	S.E	CI95.lo	CI95.hi	Zval
Intercepto	4.5171	0.01034	4.4968	4.53735	436.98178
y	0.01992	0.00291	0.01422	0.02562	6.84810

y un valor de AIC igual a $-75100,03$.

2. $B(u) = \theta_0 + \theta_1 x + \theta_2 y + \theta_3 x^2 + \theta_4 y^2 + \theta_5 xy$

De donde se obtuvo:

$$\lambda(u) = \exp(-1668,326 - 42,3x - 5,614y - 0,2672x^2 - 0,0667xy - 0,13623y^2)$$

con

	Estimación	S.E	CI95.lo	CI95.hi	Zval
Intercepto	-1668.326	23.0371	-1713.48	-1623.18	-72.42
x	-42.30	0.582	-43.445	-41.163	-72.6899
y	-5.614	0.3012	-6.2045	-5.024	-18.641
I(x²)	-0.2672	0.00368	-0.2744	-0.25998	-72.6953
I(x*y)	-0.0667	0.0038	-0.0741	-0.05922	-17.5621
I(y²)	-0.13623	0.0020	-0.1402	-0.1323	-67.8756

y el valor de AIC es igual a: -96496,78

Por lo tanto, entre estos dos modelos, el mejor es el segundo de ellos al tener un menor AIC, por lo que se procede a validarlo.

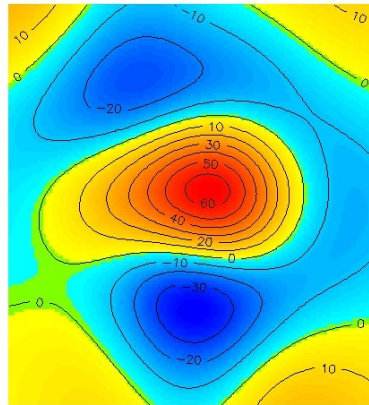


Figura 2.31: Residuos suavizados modelo 2

En la Figura 2.31 se muestra los residuos suavizados del modelo. Vemos que no existe un patrón marcado de ellos, por lo que se tendría que el modelo es adecuado. Por otra parte, al realizar el *quadrant.test*, con la siguiente prueba de hipótesis:

H_0 = El patrón de puntos sigue una distribución de Poisson inhomogénea, con función de intensidad $B(u)$.

H_1 = El patrón de puntos no sigue una distribución de Poisson inhomogénea, con función de intensidad igual a $B(u)$.

se obtuvo un p-valor menor a $2.2e-16$, lo que nos indica que el modelo no logra capturar la estructura espacial de nuestros datos. Por lo que,

se propone ajustar el modelo Cluster de Neyman-Scott, por medio del comando `kppm()`, el cual recibe como parámetro el proceso puntual, `trend`, el cual por defecto toma el valor de 1, lo que indicaría que el proceso es homogéneo y `cluster()`, en donde se debe especificar el proceso que se desea implementar, los cuales pueden ser: Thomas, MatClust, Cauchy, VarGamma y LGCP. En donde tenemos que el proceso de Cauchy ajusta un modelo de Neyman- Scott utilizando el kernel de Cauchy, cuya función de densidad de probabilidad es igual a :

$$f(r) = \frac{1}{\pi\gamma \left[1 + \left(\frac{r}{\gamma} \right)^2 \right]}$$

donde r representa la distancia desde el centro del cluster y γ es un parámetro de escala.

El algoritmo primero ajusta un proceso de Poisson con el logaritmo de la intensidad especificada en `trend` para luego obtener los parámetros del modelo por medio del método de mínimo contraste.

2.4.6. Modelo ETAS

El modelo de secuencia de sismos de tipo epidémico o ETAS por sus siglas en inglés, es un modelo estadístico estándar de sismos, el cual nos permite describir los catálogos e identificar las características de los clusters sísmicos. Usualmente es utilizado como punto de partida para la predicción de sismos a corto plazo, el cual fue introducido por Ogatha en 1988 y se basa en la idea de que los sismos tienden a agruparse tanto en tiempo como en espacio.

Por otra parte, tenemos que el modelo ETAS es un proceso puntual de auto-excitación, lo que quiere decir que la ocurrencia de un evento aumenta la probabilidad de ocurrencia de otros eventos, el cual puede ser interpretado como un Proceso Cluster de Poisson generalizado en donde se tiene que a cada uno de los puntos padre, con intensidad λ , se le asocia un proceso de ramificación de descendientes, cuya función de intensidad condicional es igual a:

$$\lambda_{\theta}(t, x, y | H_t) = \tilde{u}(x, y) + \sum_{i:t_i < t} k_{A,\alpha}(m_i) g_{c,p}(t - t_i) f_{D,\gamma,q}(x - x_i, y - y_i; m_i)$$

donde se tiene que:

- H_t es la historia previa del proceso.
- $\tilde{u}(x, y)$ representa la tasa de sismicidad de fondo, es decir la tasa de ocurrencia de sismos que se consideran normales en una región en particular, la cual es independiente del tiempo e igual a:

$$\tilde{u}(x, y) = \mu u(x, y)$$

Donde $\mu > 0$ es desconocido y $u(x, y)$ es la función de intensidad espacial.

- $k_{A,\alpha}$ corresponde al número esperado de eventos desencadenados o réplicas generadas por un evento de magnitud m_i , la cual es igual a:

$$k_{A,\alpha} = A \exp\{\alpha(m - m_0)\}$$

Donde $A > 0$ y $\alpha > 0$, los cuales son parámetros desconocidos.

- $g_{c,p}(t - t_i)$ representa la función de densidad de probabilidad de ocurrencia en el tiempo de que un evento sea desencadenado por un evento de magnitud m_i en el instante t_i , donde se tiene que:

$$g_{c,p}(t - t_i) = \begin{cases} \frac{p-1}{c} \left(1 + \frac{t-t_i}{c}\right) & \text{si } t - t_i > 0 \\ 0 & \text{si } t - t_i \leq 0 \end{cases}$$

Donde $c > 0$ y $p > 1$ son desconocidos, y p es el exponente de la ley de Omori para la tasa de decrecimiento de réplicas.

- $f_{D,\gamma,q}(x - x_i, y - y_i; m_i)$ es la función de densidad de probabilidad de la ocurrencia de un evento producto de un sismo de magnitud m_i en la posición (x_i, y_i) , la cual depende de la magnitud y ubicación del principal sismo y es igual a:

$$f_{D,\gamma,q}(x - x_i, y - y_i; m_i) = \frac{q - 1}{\pi D \exp[\gamma(m_i - m_0)]} \left(1 + \frac{(x - x_i)^2 + (y - y_i)^2}{D \exp[\gamma(m_i - m_0)]} \right)^{-q}$$

donde $D > 0$, $\gamma > 0$ y $q > 1$ son desconocidos.

Por otra parte, tenemos que $\tilde{u}(x, y)$ se estima por medio de:

$$\hat{u}(x, y) = \frac{1}{T} \sum_{j=1}^N (1 - \hat{p}_j) \phi(x - x_j, y - y_j; h_j)$$

donde:

$$\phi(x, y; h) = \frac{1}{2\pi h^2} \exp\left(-\frac{x^2 + y^2}{2h^2}\right)$$

y h_j corresponde al valor de ancho de banda.

Además tenemos que la función total espacial de intensidad se la define como:

$$\Lambda(x, y) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \lambda_\theta(t, x, y | H_t) dt$$

Mientras que el coeficiente de agrupamiento (desencadenamiento) es igual a:

$$w(x, y) = 1 - \frac{u(x, y)}{\Lambda(x, y)}$$

Para estimar los parámetros: $\tilde{\theta} = (\mu, k_0, c, p, d, q, \alpha)$ el modelo utiliza el método de máxima verosimilitud y el método de Davidon-Fletcher-Powell, el cual es parte de los métodos Quasi-Newton, mientras que para realizar el ajuste del modelo emplea un método estocástico de desagrupamiento, propuesto por Zhuang, el cual plantea un enfoque probabilístico para desagrupar un catálogo sísmico, en donde se tiene que la probabilidad de que el evento j sea provocado por el evento i , se define como:

$$p_{ij} = \begin{cases} \frac{k_{A,\alpha}(m_i) g_{c,p}(t_j - t_i) f_{D,\gamma,q}(x_j - x_i, y_j - y_i; m_i)}{\lambda_\theta(t_j, x_j, y_j | H_{t_j})} & , t_j > t_i \\ 0 & , t_j \leq t_i \end{cases}$$

De esta forma se tiene que:

$$p_j = \sum_{i:t_i < t_j} p_{ij}, j = 1, \dots, N$$

Mientras que:

$$1 - p_j = \frac{\mu u(x_j, y_j)}{\lambda(t_j, x_j, y_j | H_{t_j})}$$

representa la probabilidad de que el evento j sea un evento de fondo (padre) [1].

Para poder aplicar el modelo ETAS, primero debemos cambiar nuestra base de datos (data.frame) a tipo catalog, para ello debemos asegurarnos que nuestra base de datos contenga solo 5 columnas, con las siguientes variables: Fecha, hora, latitud, longitud y magnitud. Los cuales deben ser nombradas de la siguiente forma respectivamente: date, time, lat, long y mag. Luego se debe definir el polígono que encierra el área geográfica que deseamos estudiar, para ello existen 2 opciones, la primera si se trata de un cuadrado debemos especificar el rango que toma la variable longitud y latitud respectivamente con los comandos long.range y lat.range; por otra parte, si no se trata de un rectángulo, debemos utilizar el comando region.poly, el cual puede ser una lista o un data.frame que contenga las coordenadas de los vértices del polígono ordenadas de forma antihoraria donde ningún vértice debe de repetirse. Por otra parte, tenemos que definir la fecha de inicio y final de nuestro estudio y el valor del umbral de la magnitud. Estos datos son ingresados a la función catalog que pertenece a la librería ETAS.

Al graficar el catálogo se obtienen seis gráficas las cuales corresponden a:

- El proceso puntual cuya ventana se muestra en color rojo.
- La variación de la latitud, longitud y magnitud en función del tiempo.
- La relación que existe entre $\log_{10}(N_m)$ y la magnitud.
- N_t vs tiempo.

Para estimar el modelo ETAS, utilizamos el comando `etas` el cual recibe los siguientes parámetros:

- Catálogo.
- El parámetro inicial $\theta = (\mu, k_0, c, p, d, q, \alpha)$ el cual si no se ingresan valores se toma por defecto a: $\mu = N/4T|S|$, $A = 0,01$, $c = 0,01$, $\alpha = 1$, $p = 1,3$, $D = 0,01$, $q = 2$ y $\gamma = 1$, los cuales fueron propuestos por Ogatha(1988).
- `no.iter`: Número de iteraciones a realizar , el cual por defecto es igual a 11.
- `bwd` y `nnp`: donde `bwd` corresponde al valor del ancho de banda y `nnp` el número de vecinos cercanos, por defecto se toma `nnp=5` y `bwd=0.05`

Una vez ingresados los datos, el modelo internamente utiliza las funciones `decluster` y `etasfit`, para estimar la tasa sísmica de fondo $u(x, y)$ y θ , donde se tiene que el modelo es muy sensible a los parámetros iniciales, ya que dependiendo de ellos puede como no converger el algoritmo. Una vez que haya finalizado, se puede obtener un resumen del modelo ajustado por medio de `plot` y `print`. Por otra parte, la función `rates` nos permite obtener las estimaciones de la intensidad total espacial, la tasa de sismicidad de fondo y el coeficiente de agrupamiento (clustering) del modelo ajustado. Mientras que la función `probs` y `resid.etas`, nos dan las estimaciones de las probabilidades p_j de desagrupamiento y la segunda función nos da los residuos del modelo incluidas sus gráficas.

La función `etas` asume que el catálogo está completo y que es estacionario en el tiempo.

Se tomo como parámetros iniciales para $m_0 = 5$ los cuales se presentan en el artículo 'ETAS: An R Package for Fitting the Space-Time ETAS Model to Earthquake Data ' , para el catálogo sísmico de Italia:

mu	A	c	alpha	p	D	q	gamma
1	3.031117	0.005	1.05	1.01	1.1	1.52	0.6

Una vez obtenidas las estimaciones del modelo, los parámetros se utilizarán para estimar el modelo con $m_0 = 4,5$. Este proceso se repite para obtener las estimaciones correspondientes a $m_0 = 4$.

El modelo ETAS fue aplicado a los catálogos con m_0 igual a 4, 4.5 y 5, en donde se obtuvieron las siguientes estimaciones para θ por el método MLE ;

	mu	A	c	alpha	p	D	q	gamma
4	0.95934	0.04434	0.69141	0.68854	1.32485	0.013587	1.27265	1.5e-17
4.5	0.92251	0.68554	0.85012	0.45995	1.39713	0.00734	1.15304	2.7e-19
5	0.83791	0.52704	1.50640	0.42166	1.39193	0.00536	1.1479	8.66e-22

de donde se tiene que:

- μ : Representa la intensidad general de fondo.
- c : La actividad sísmica de la región.
- p : Tasa de decaimiento de las réplicas.
- A : Representa la reproductividad de un evento.
- α y γ : Representa la eficacia con la que un sismo genera réplicas.
- d y q : representan la influencia espacial del sismo principal.

De donde tenemos que la intensidad general de fondo es mayor cuando $m_0 = 4$, de donde además se tiene un valor pequeño de tasa de productividad sísmica (0.04), mientras se tiene un valor alto en c , lo que nos indica que probablemente se generen una gran cantidad de réplicas inmediatamente tras el sismo original con una eficacia igual a 0.6885. Por otra parte, tenemos que p es igual a 1.32485, lo que quiere decir que el número de réplicas decrece lentamente conforme transcurre el tiempo. Mientras que para $m_0 = 4,5$ tenemos que existe una alta probabilidad de que se generen réplicas al tener un valor de A de 0.68554 con eficacia igual a 0.45995, las cuales disminuyen lentamente en una tasa de 1.39713. Por otra parte, cuando $m_0 = 5$ tenemos que la tasa de generación de clusters es menor en comparación con los otros dos umbrales mientras que el valor de c es mayor.

Además tenemos que el algoritmo convergió en 7 iteraciones en cada uno de los valores correspondientes a m_0 .

Por otra parte, analizando las probabilidades de desagrupamiento, las cuales nos ayudan a clasificar los sismos en sismos de fondo o réplicas, se obtuvo los siguientes cuantiles:

m_0	Mín	1 Quantil	Mediana	Media	3 Quantil	Máy
4	0	0.2223	0.5905	0.5690	0.9418	1
4.5	0	0.1798	0.6295	0.5650	0.9525	0.9999
5	0	0.06333	0.26347	0.41294	0.82162	0.99990

2.4.7. ETASFLP

A diferencia del modelo ETAS el modelo ETASFLP utiliza el método FLP (Forward likelihood predictive) para estimar el ancho de banda del kernel de la tasa de sismicidad de fondo y además la función condicional de intensidad es igual a [1]:

$$\lambda_{\theta}(t, x, y | H_t) = \mu u(x, y) + \sum_{i:t_i < t} \frac{k_0 \exp((\alpha - \gamma)(m_i - m_0))}{(t - t_i + c)^p} \left\{ D + \frac{(x - x_i)^2 + (y - y_i)^2}{\exp[\gamma(m_i - m_0)]} \right\}^{-q}$$

donde :

$$k_0 = \frac{1}{\pi} A(p - 1) c^{p-1} (q - 1) D^{q-1}$$

Para poder aplicar este modelo, primero debemos asegurarnos que la base de datos contenga cinco variables: tiempo, latitud, longitud, profundidad y magnitud. Las cuales deben ser nombradas respectivamente: time, lat, long, z y magn1. Luego, se utiliza la función etasclass(), la cual recibe los siguientes parámetros :

- Un data.frame con los datos antes señalados.
- magn.threshold, el cual corresponde al valor del umbral de la magnitud.

- `magn.threshold.back`, corresponde al valor del umbral de la magnitud de los eventos a utilizar para la estimación del kernel de la tasa de sismicidad de fondo.
- Datos iniciales para los parámetros: μ , k_0 , c , p , a , γ , q y d .

Este modelo se encuentra implementado en el paquete `etasFLP`. Para determinar que parámetros iniciales tomar para μ , k_0 , c , p , a , γ , q y d se puede utilizar la función `etas.starting()`, la cual recibe como parámetro el `data.frame` con las variables y los nombres antes descritos. Sino se incluyen estos parámetros iniciales, la función `etas.class()`, llama a la función `etas.starting()`, para estimar los valores de c y d a partir de la distribución empírica de las de las diferencias en el tiempo y de las distancias en el espacio, se estiman los valores de μ y k_0 por medio del método de máxima verosimilitud. Se suele tomar $p=1$ y $q=2$.

La función `etas.class`, luego de obtener los parámetros iniciales, realiza dos etapas, la primera se denomina ML, en donde se calcula el vector de parámetros θ maximizando la función de log-verosimilitud la cual es igual a :

$$\log L(\theta; H_{t_n}) = \sum_{i=1}^n \log \lambda(x_i, y_i, t_i; \theta) - \int_{T_0}^{T_{max}} \int \int_{\Gamma(x,y)} \lambda(x, y, t; \theta) dx dy dt$$

A través de la función `nlm`, el cual utiliza un algoritmo de tipo Newton además de utilizar derivadas numéricas para aproximar el gradiente y la matriz Hessiana . Luego de tener las estimaciones, realiza la segunda etapa en la cual se estima la intensidad de fondo al maximizar FLP, la cual es igual a:

$$FLP_{k_1, k_2}(\hat{\phi}) = \sum_{k=k_1}^{n-1} \delta_{k, k+1}(\hat{\phi}(H_{t_k}); H_{t_{k+1}})$$

donde:

- $k_1 = \frac{n}{2}$
- $k_2 = n - 1$
- $\delta_{k, k+1}(\hat{\phi}(H_{t_k}); H_{t_{k+1}})$ es la información predictiva de las primeras k observaciones dadas en la $k+1$ -ésima observación, de donde se tiene

que:

$$\delta_{k,k+1}(\hat{\phi}(H_{t_k}); H_{t_{k+1}}) = \log L(\hat{\phi}(H_{t_k}; H_{t_{k+1}})) - \log L(\hat{\phi}(H_{t_k}; H_{t_k}))$$

Donde H_k es la historia del proceso hasta el instante t_k y $\hat{\phi}(H_{t_k})$ es la estimación obtenida hasta la k -ésima observación [10].

Capítulo 3

Resultados, conclusiones y recomendaciones

3.1. Resultados

3.1.1. Catálogo

Una vez realizado el catálogo se obtienen las siguientes gráficas:

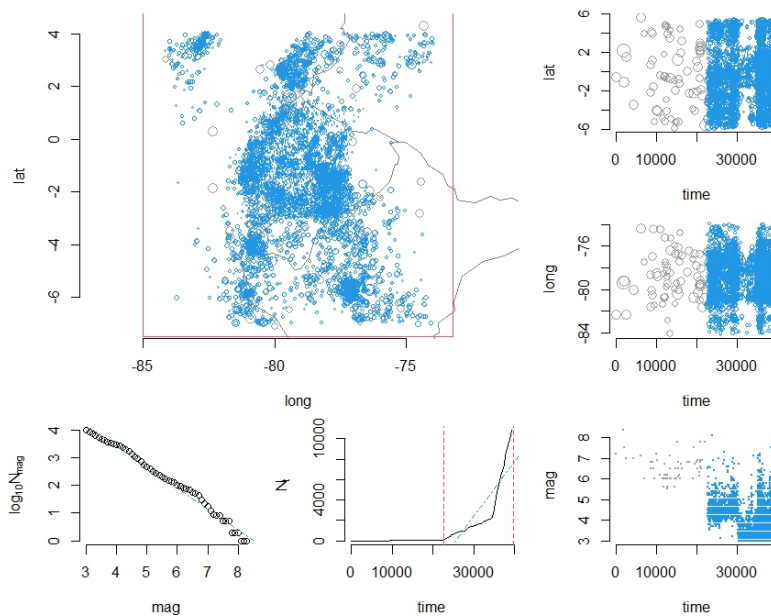


Figura 3.1: Gráfico catálogo $m_0=3$

Podemos observar en el gráfico de magnitud vs $\log_{10}N_m$, ubicado en la

esquina inferior izquierda de la Figura 3.1 que nuestros datos se ajustan bien a la ley de Gutenberg-Richter, por lo tanto tenemos que el catálogo es completo. Por otra parte, en el gráfico N_t vs tiempo podemos observar que no se tiene una tendencia lineal, ya que existe un aumento significativo de sismos en los últimos años, sin embargo, vemos que al aumentar el umbral de la magnitud se aproxima a dicha tendencia .

Por otra parte, tenemos que los círculos sin color corresponden a aquellos eventos complementarios, que en nuestro caso corresponden a los 68 sismos que ocurrieron desde el 1 de julio de 1901 hasta el 20 de noviembre de 1960, mientras que los círculos azules son los puntos dentro del tiempo de estudio.

3.1.2. Modelos Temporales

Para la serie de tiempo del máximo valor mensual M_w , se ajustó el modelo ARIMA(2,0,2) y se realizó la predicción a 10 meses, correspondientes al año 2010. En la Figura 3.2 ,se muestran los valores predichos en color azul junto con sus intervalos de confianza.

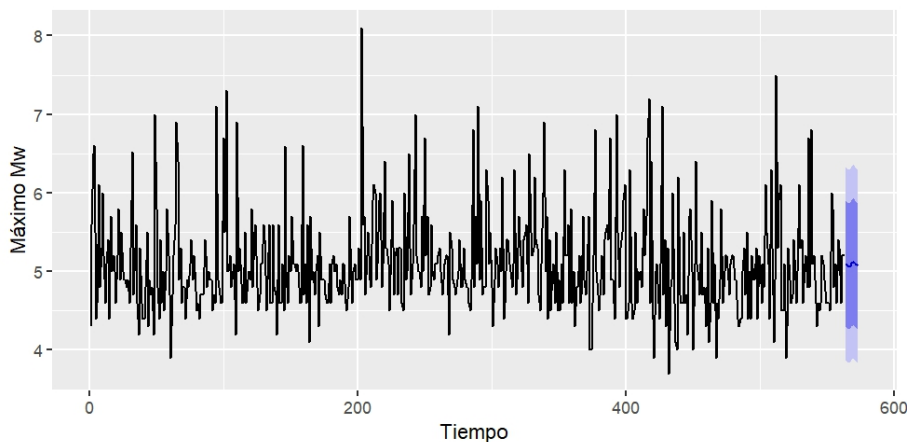


Figura 3.2: Predicción a 10 meses Máximo Mw

De donde vemos que se espera que el sismo de mayor magnitud en los siguientes 10 meses sea en junio con magnitud 5.122 [M_w] y el menor de ellos sea en marzo con 5.065 [M_w].

Por otra parte, en la Figura 3.3, se comparan los valores obtenidos con

los valores reales, donde los valores reales se encuentran de color negro y los obtenidos de color azul.

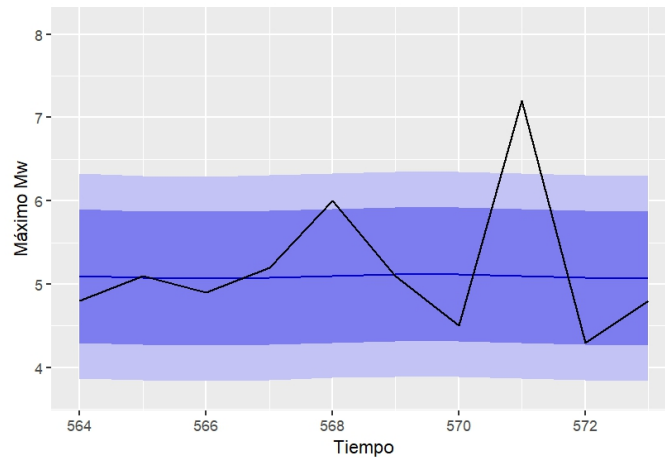


Figura 3.3: Predicción vs Valores reales

Vemos que los valores coinciden en los meses de febrero y junio. Mientras que en mayo y agosto se tienen una gran diferencia, la cual se debe a que en el año 2010 hubo la presencia de enjambres sísmicos frente a la costa de Manabí, en Jama, además de que el 12 de agosto hubo un sismo de magnitud 7.2 el cual generó 7 réplicas con magnitudes entre 3.7 y 4.3 a los 5 días posteriores al evento principal [8].

Por otra parte, para la serie de mínima profundidad mensual se ajustó el modelo ETS(A,N,N) y se realizó de igual forma la predicción a 10 meses, la cual se puede observar en la Figura 3.4.

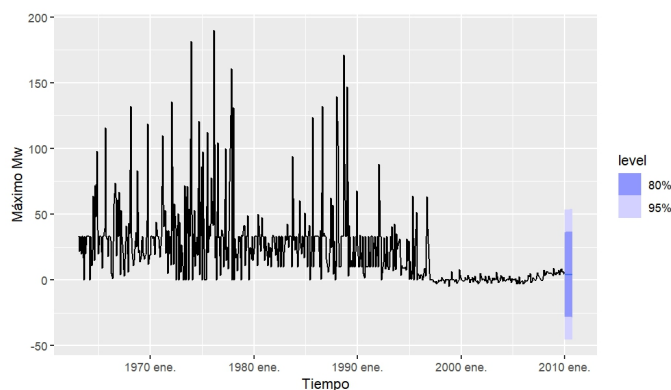


Figura 3.4: Predicción a 10 meses

De donde tenemos que se espera que la profundidad mínima mensual de los sismos sea de 4.12 y en la Figura 3.5, vemos que estos valores se aproximan bien a los valores reales hasta el mes de julio, mientras que en agosto tenemos que los sismos registrados tuvieron mayor profundidad.

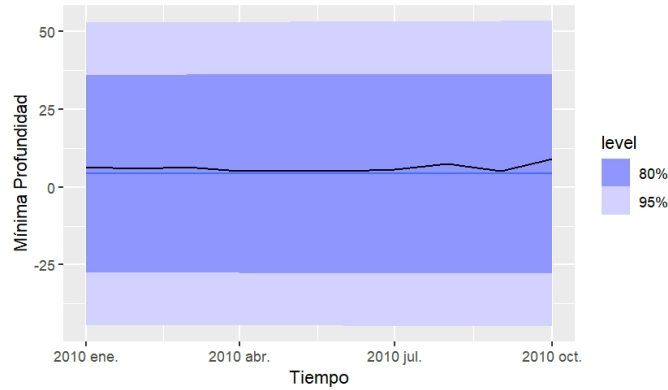


Figura 3.5: Predicción vs Valores reales ETS(A,N,N)

3.1.3. Modelos Espaciales

Modelo Cluster

Se ajustó el modelo de Neyman-Scott con distribución de Cauchy dentro de los cluster, con trend igual a: $x+y$, de donde se obtuvieron los siguientes valores de los parámetros:

θ_0	θ_1	θ_2
4.44246	-0.00094	0.01992

y

kappa	scale	phi	psib
0.02009	0.45187	9.312	0.903

Donde kappa y scale corresponden a los parámetros de agrupamiento, mientras que psib es la probabilidad de hijos y phi es la fuerza de agrupamiento. De esta forma tenemos que en promedio hay 0.02009 sismos por unidad de área, los cuales se encuentran dispersos en una escala

de 0.45187 dentro de los clusters. Por otra parte, en la Figura 3.6 tenemos que la función K estimada es mayor que la K de un Proceso de Poisson lo que nos indica que en un radio de 0.45187 unidades los datos se encuentran agrupados.

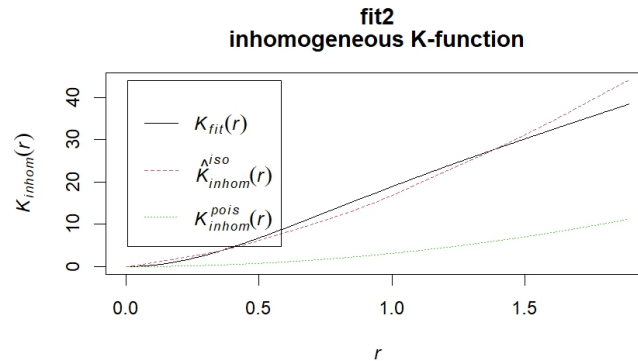


Figura 3.6: Residuos suavizados modelo 2

3.1.4. Modelo ETAS

Se considera como modelo base a estudiar a $m_0 = 4$ debido al costo computacional del modelo y con este umbral se tiene un mayor número de datos que con $m_0 = 4,5$ y $m_0 = 5$.

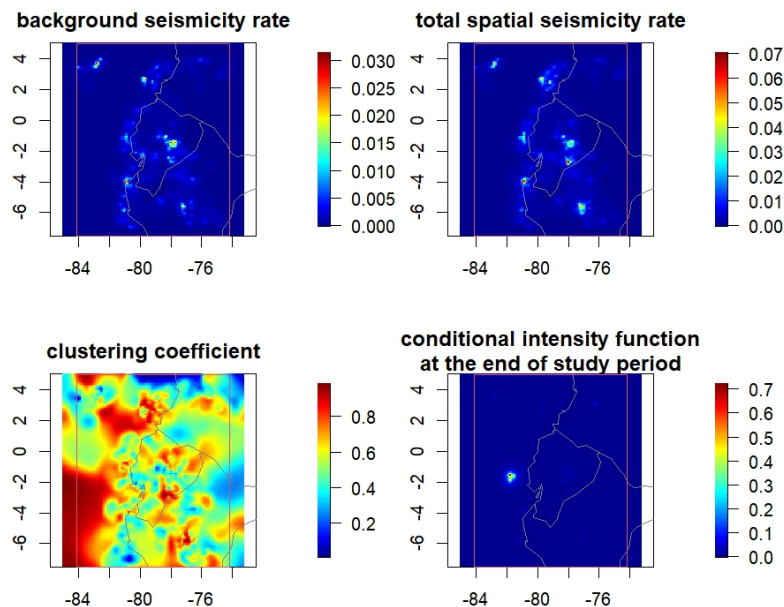


Figura 3.7: Resultados modelo ETAS $m_0 = 4$

En la Figura 3.7 podemos ver que la tasa de sismicidad de fondo es mayor en Napo y en Tumbes, Perú. Por otra parte, existe un mayor coeficiente de agrupamiento de sismos en Morona Santiago, Sucumbios, Orellana, Imbabura y Pichincha. La función de intensidad condicional indica que toma mayor valor en el Pacífico frente a Manabí.

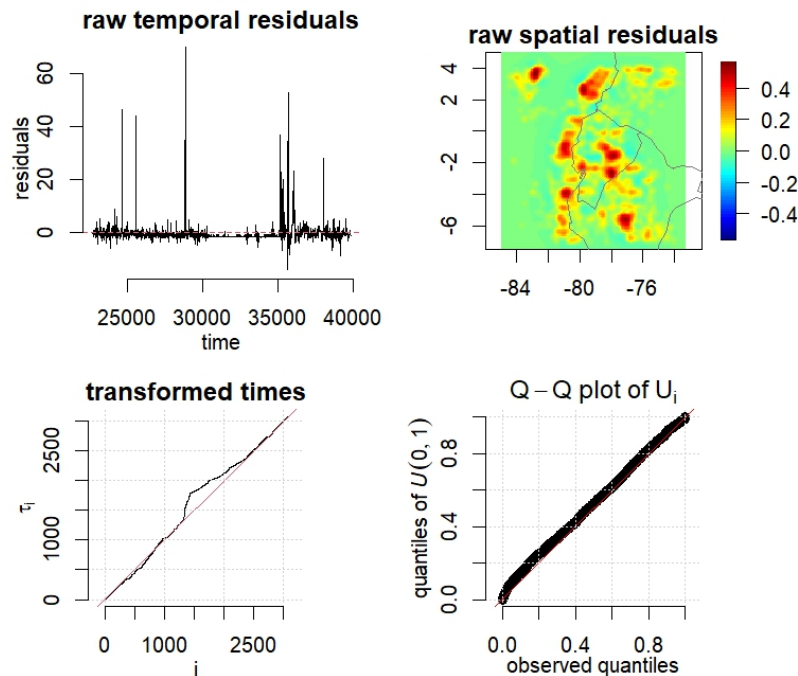


Figura 3.8: Residuos ETAS $m_0=4$

En la Figura 3.8 se muestra un análisis de los residuos. Vemos en la imagen superior izquierda que los residuos se encuentran alrededor de 0 en general. Al comparar los residuos con la distribución uniforme $U(0,1)$, en una gráfica de QQ-plot, vemos que los datos se ajustan bien. En el gráfico inferior izquierdo se muestra la relación entre τ_i y i donde $i \in J = \{i : \delta_i = 1\}$ y τ_i es el valor de la integral de la función de intensidad condicional. Podemos ver que los datos se ajustan bien a la recta.

Cabe indicar que los resultados se mantienen para los otros modelos.

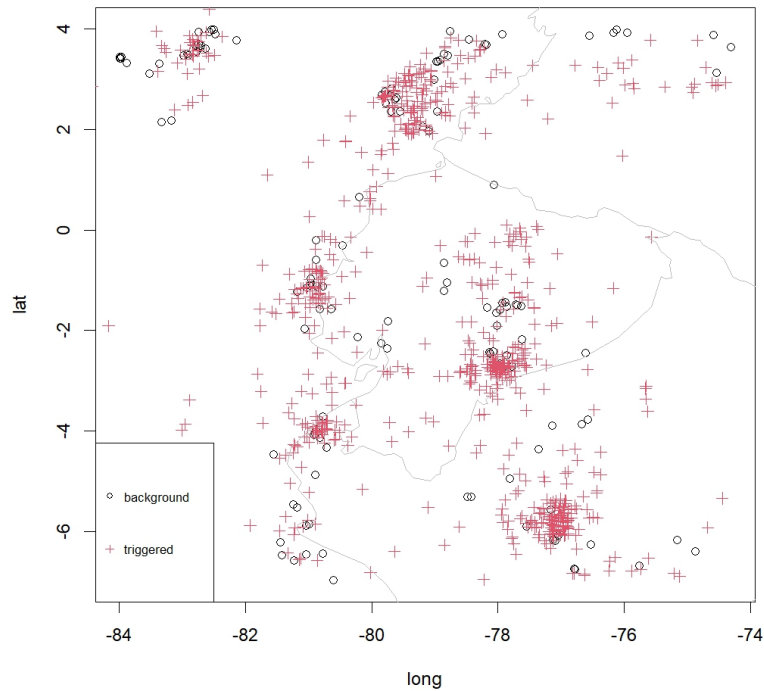


Figura 3.9: Clasificación de sismos $m_0 = 4$ con probabilidad mayor a 0.95

En la Figura 3.9, se muestra la clasificación de sismos en sismos de fondo (background) y réplicas (triggering). Podemos ver que los sismos de fondo son menos que las réplicas y que estos se encuentran principalmente en la zona costera y en la zona central del país. Mientras que se tienen sismos de fondo que se encuentran en Perú y en Colombia.

3.1.5. Modelo ETAS-FLP

El modelo ETAS fue aplicado tomando $m_0 = 4,5$, en donde se tomó como valores iniciales a los obtenidos en el modelo ETAS cuando $m_0 = 5$, ya que con estos se logró que el algoritmo converga y $magn.threshold.back = 5$. Se obtuvieron las siguientes estimaciones:

mu	k0	c	p	gamma	D	q	magnitud
0.025252	4002.020796	0.175974	1.230380	0.00000	727.653530	2.803686	0.493426

Como $k_0 = 4002,02$, representa la reproductividad de un evento, tenemos que existe una alta probabilidad de que los terremotos mayores o

iguales a 4.5 generen un número significativo de réplicas. Por otra parte, tenemos que estas disminuyen en una tasa de 1,23, mientras que tenemos que la actividad sísmica de la región es de 0,18 con intensidad general igual a 0,025.

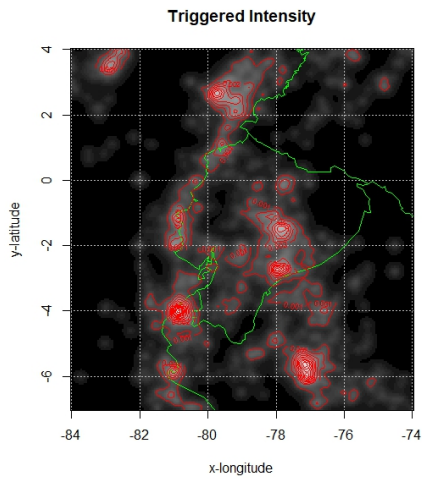


Figura 3.10: Intensidad réplicas

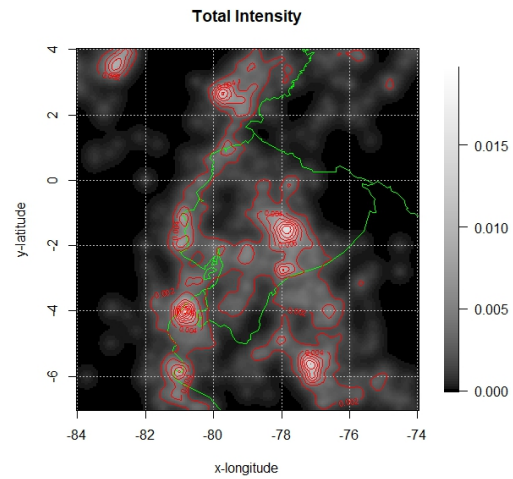


Figura 3.11: Intensidad Total

En la Figura 3.10 se muestra la intensidad desencadenada de las réplicas. Vemos que es mayor cerca de Morona Santiago y en el Pacífico cerca de la costa Ecuatoriana. En la Figura 3.11, se puede observar que la intensidad total a diferencia de la intensidad de las réplicas, es mayor en el área sur del país.

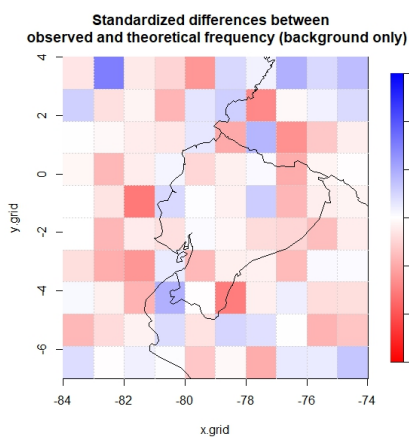


Figura 3.12: Fondo

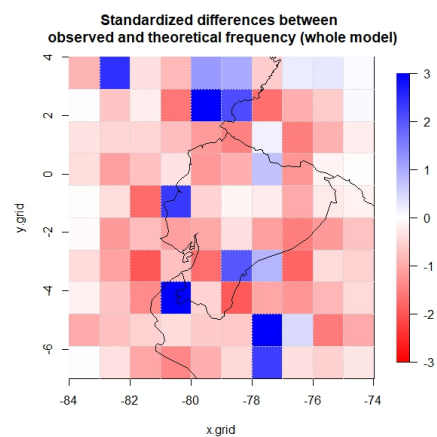


Figura 3.13: Todo el modelo

En las Figuras 3.12 y 3.13 se muestran las diferencias entre las frecuencias observadas y teóricas, si solo se consideran los sismos de fondo estas son cercanas a 0 en el área continental del territorio ecuatoriano, mientras que, al considerar todo el modelo vemos que existe una mayor diferencia en Manabí, Morona Santiago y la frontera con Perú.

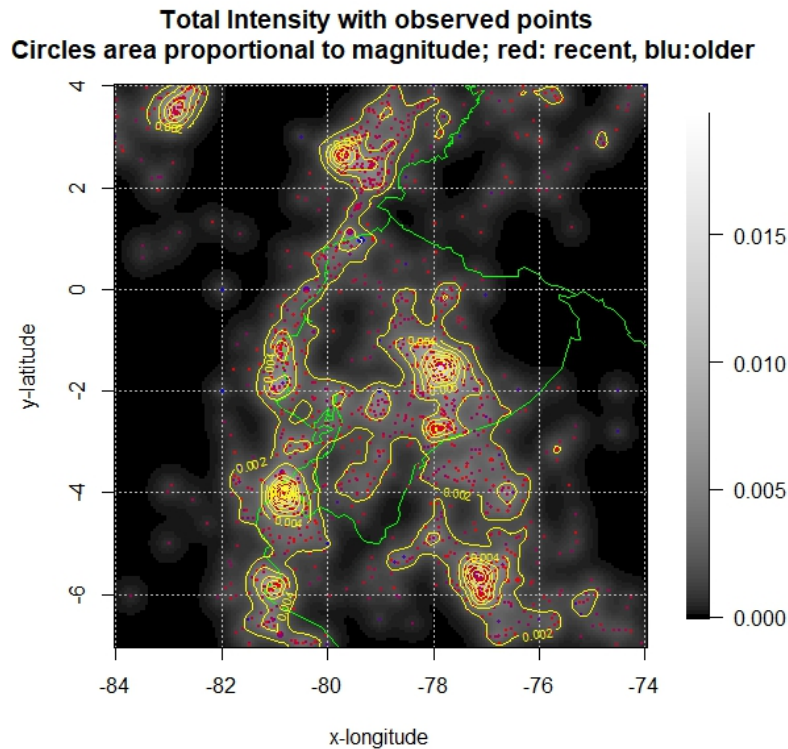


Figura 3.14: Intensidad total con los puntos observados

En la Figura 3.14 se muestra la intensidad total de los puntos observados. Vemos que la mayoría de ellos corresponden a los más recientes del catálogo.

3.2. Conclusiones

1. El test de CSR junto con las gráficas de la función K y G mostraron que el proceso puntual conformado por los sismos de 1963 hasta el 2009 siguen un Proceso Cluster Inhomogéneo.
2. El Catálogo Homogenizado del 2009 es completo y no estacionario, lo que nos indica que el número de sismos va disminuyendo confor-

me aumenta la magnitud pero estos no siguen una tendencia lineal conforme aumentan los años.

3. Los modelos temporales nos permiten identificar las tendencias de los sismos a través de los años; sin embargo, Ecuador al ser un país sísmico, eventos de mayor magnitud a la esperada o la generación de enjambres sísmicos pueden ocurrir en cualquier momento cambiando de esta forma el comportamiento de la serie, obteniendo pronósticos mayores o menores a los reales.
4. Los sismos se agrupan tanto en tiempo y en espacio, por lo que para poder modelarlos es necesario utilizar modelos en los que se incluyan estas dos variables.
5. El modelo ETAS-FLP es computacionalmente más rápido que el modelo ETAS.
6. El mayor número de sismos se encuentran en la zona de subducción de las placas Nazca y Sudamericana y en las provincias: Morona Santiago y Napo.

3.3. Recomendaciones

1. Se debe tomar en cuenta que el modelo ETAS es sensible a los valores iniciales, por lo que dependiendo de ellos el algoritmo puede o no converger.
2. Para futuros trabajos se recomienda incluir covariables que permitan explicar la posición geográfica de Ecuador y aplicar los modelos ETAS y ETAS-FLP.

Referencias bibliográficas

- [1] Jalilian Abdollah. Etas: An r package for fitting the space-time etas model to earthquake dat. *Journal of Statistical Software*, 2019.
- [2] Giada Adelfio and Marcello Chiodi. Including covariates in a space-time point process with application to seismicity. *Statistical Methods & Applications*, 2020.
- [3] Yasmin Azucena and Portillo Chang. Ley gutenbergritcher:relación magnitud-frecuencia de ocurrencia, aplicación de sismos en guatemala. Master's thesis, Universidad del Valle de Guatemala, 2011.
- [4] Adrian Baddeley, Ege Rubak, and Rolf Turner. *Spatial Point Patterns Methodology and Applications with R*. Taylor & Francis Group,LLC, 2016.
- [5] Céline Beauval, Hugo Yepes, Pablo Palacios, Monica Segovia, Yvonne Font Alexandra Alvarado, Jorge Aguilar, Liliana Troncoso, and Sandro Vaca. An earthquake catalog for seismic hazard assessment in ecuador. Technical report, IGEPN, 2013.
- [6] IGEPN. Informe sísmico especial no. 007 - 2020. Technical report, IGEPN, 2020.
- [7] IGEPN. Informe sísmico especial no. 2023-009. Technical report, IGEPN, 2023.
- [8] IGEPN. Igepn. <https://www.igepn.edu.ec/>, 2024. Accessed: 2024-07-31.

- [9] Universidad Nacional Autónoma de México Instituto Geofísica. Servicio sísmológico nacional. *Instituto Geofísica*.
- [10] Chiodi M and Adelfio G. Non-parametric and parametric estimation techniques in r package etasflp for earthquakes' description. *Journal of Statistical Software*, vol. 76.
- [11] Paula. Moraga. *Spatial Statistics for Data Science: Theory and Practice with R*. Chapman & Hall/CRC Data Science Series., 2023.
- [12] Jesper Møller and Rasmus Plenge Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, 2004.

Capítulo A

Resultados modelos ETAS

A.0.1. Cuando $m_0 = 4,5$

Catálogo

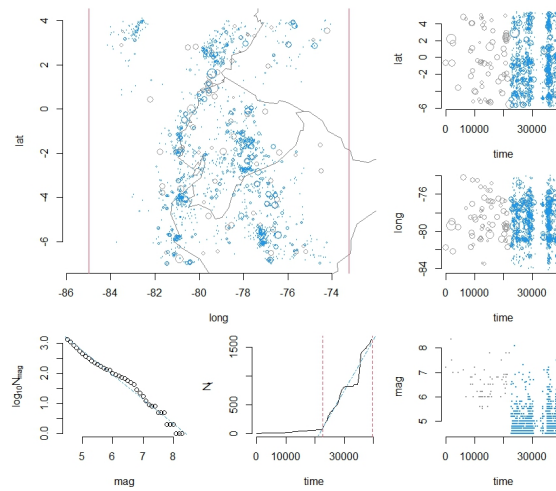


Figura A.1: Catálogo $m_0 = 4,5$

Estimaciones e iteraciones

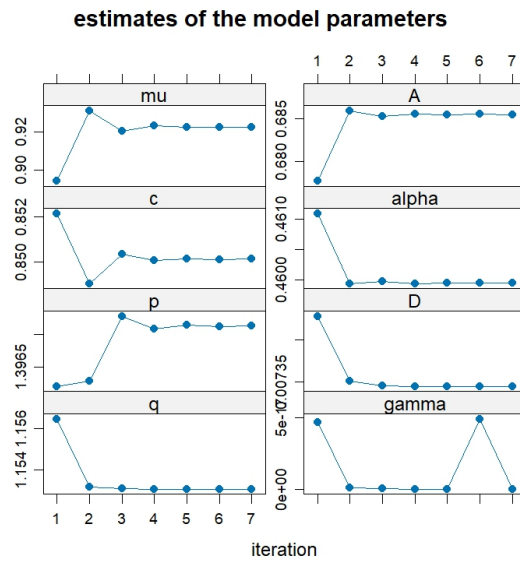


Figura A.2: Estimación parámetros

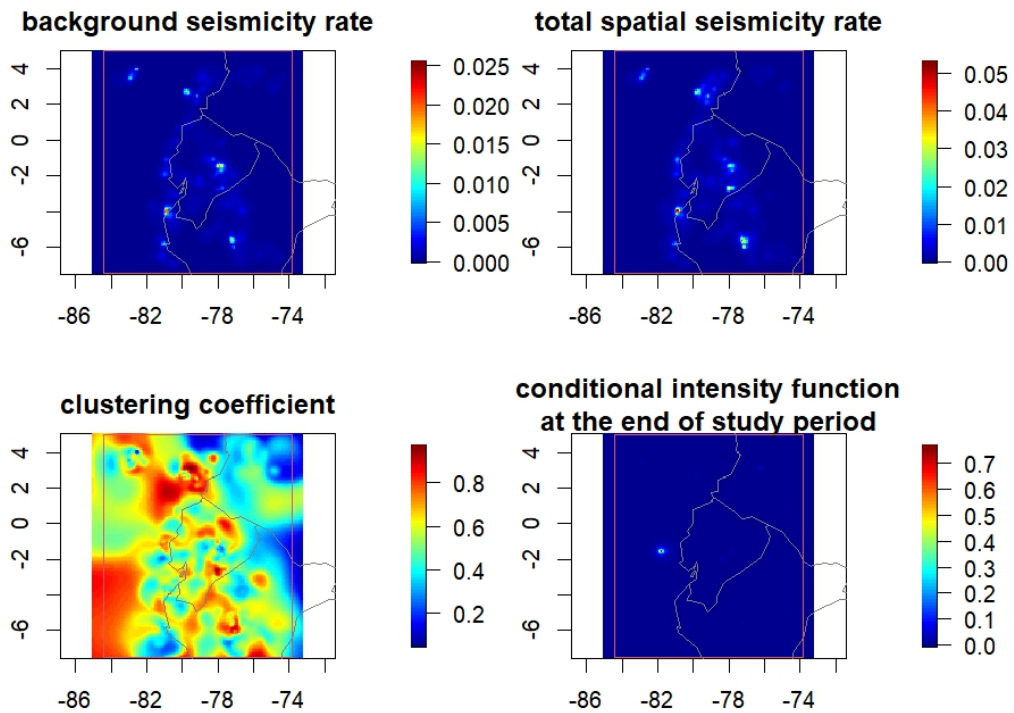


Figura A.3: Estimaciones ETAS 4.5

Residuos

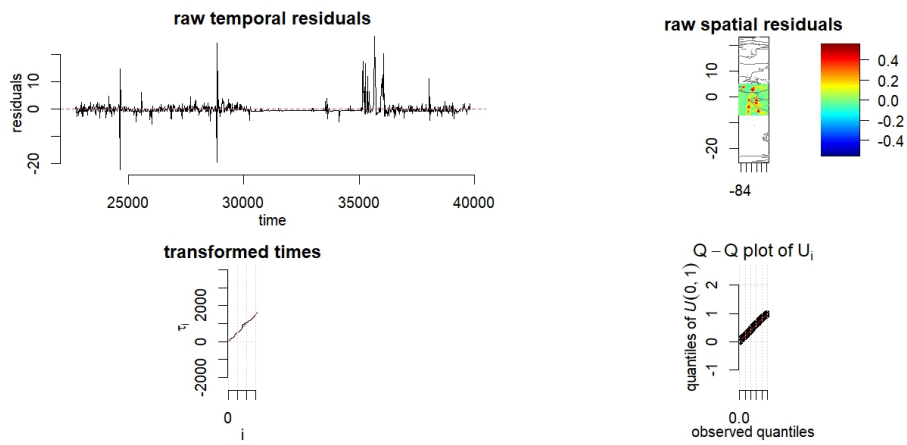


Figura A.4: ETAS 4.5

Clasificación

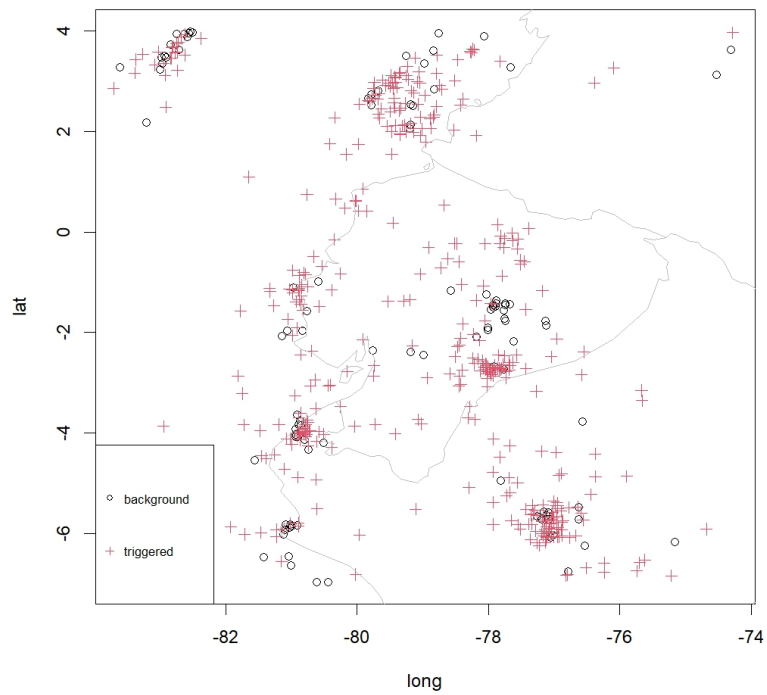


Figura A.5: Clasificación de sismos $m_0=4.5$

A.0.2. Cuando $m_0=5$

Catálogo

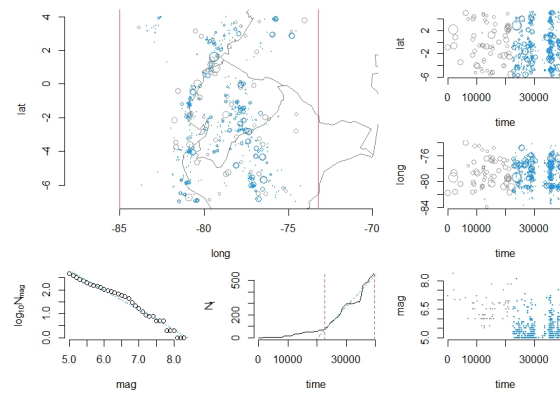


Figura A.6: Catálogo $m_0 = 5$

Estimaciones e iteraciones

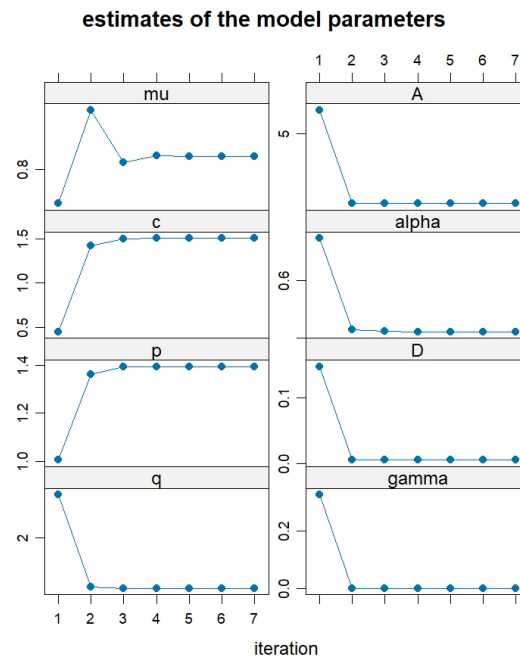


Figura A.7: Estimación parámetros ETAS 5

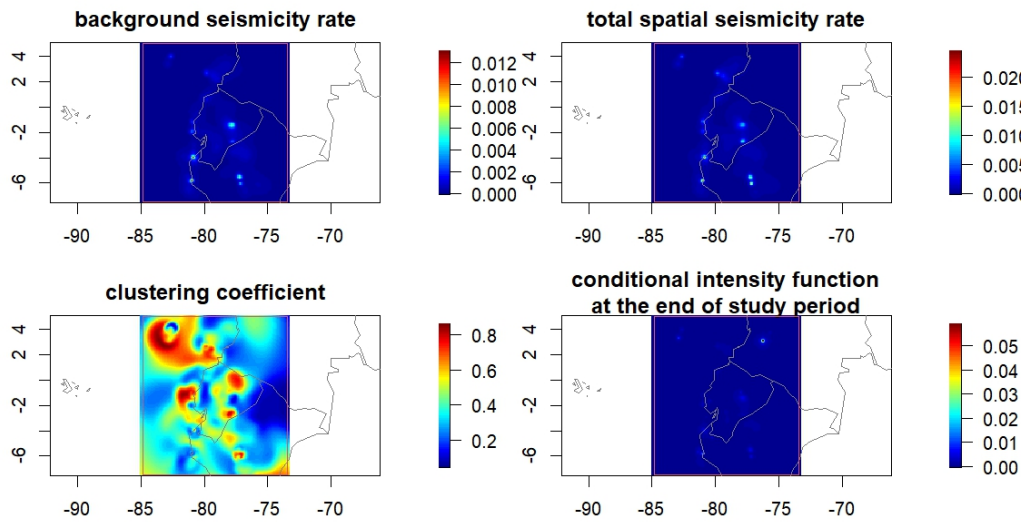


Figura A.8: Estimaciones ETAS 5

Residuos

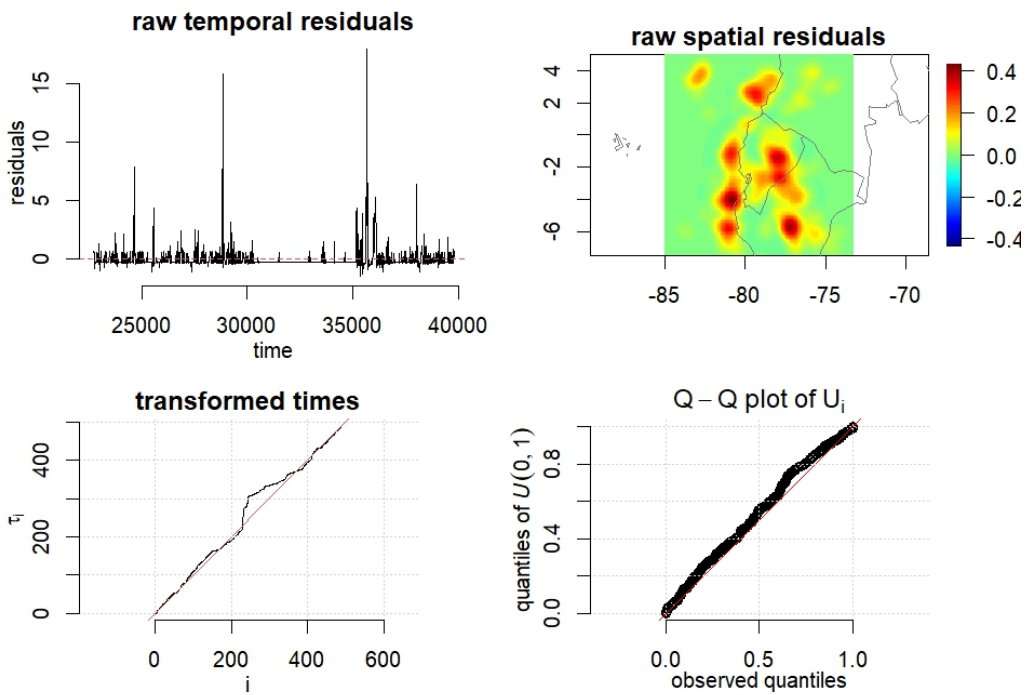


Figura A.9: Residuos modelo ETAS 5

Clasificación

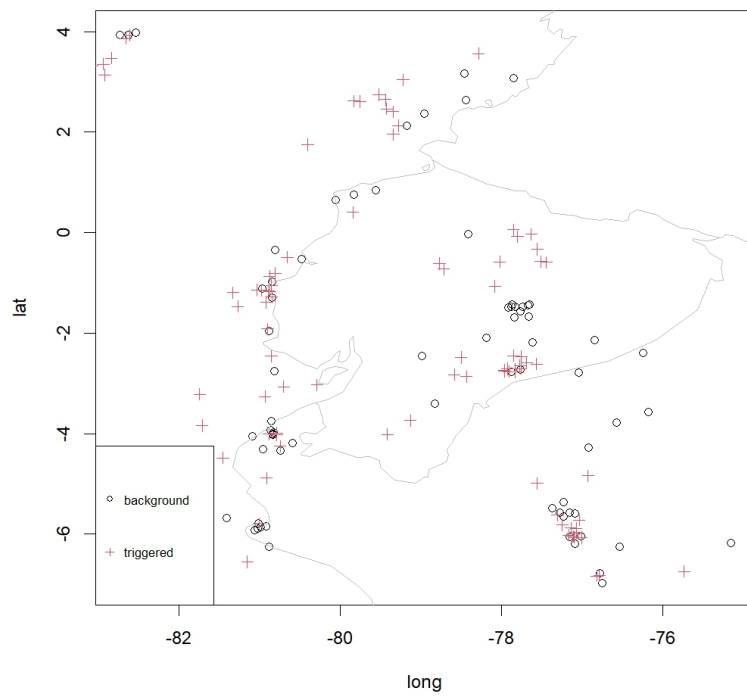


Figura A.10: Clasificación de sismos $m_0=4.5$