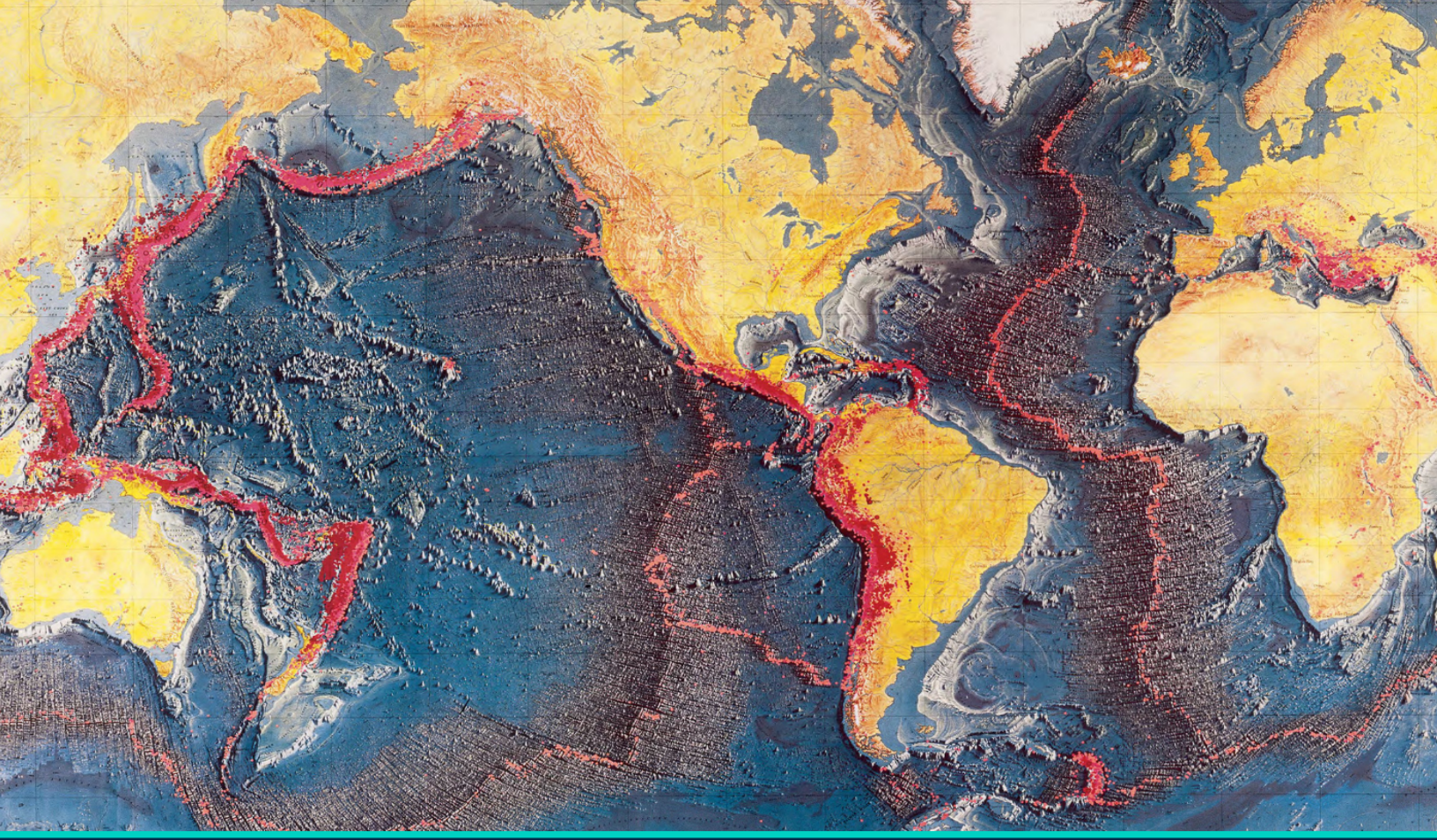

INFORME ASESORÍA II

ANÁLISIS GLOBAL DE GRANDES TERREMOTOS:

Distribución espacial y temporal de eventos extremos

Tetrametric

INSTITUCIÓN
Universidad de Santiago de Chile

FECHA
13 de julio de 2026

PRESENTADO POR
María José Valderrama Nuñez
Tamar Isai Rojas Castillo

PRESENTADO PARA
Dr. Luis Felipe Figueroa
Dr. Francisco Plaza Vega

Tabla de contenidos

1. Resumen ejecutivo

Este informe evalúa inferencialmente los principales patrones descritos en el Informe Asesoría 1 sobre 1.186 terremotos con $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025. Se confirma que la principal diferencia entre zonas corresponde a la frecuencia. El Cinturón de Fuego concentra 892 eventos y una tasa de 34,3 por año, muy por encima de las demás zonas, y mantiene la mayor concentración después de considerar la superficie. Las fuentes y los métodos de reporte no condicionan materialmente esta comparación.

También se mantiene la similitud práctica de las magnitudes, cuyas distribuciones permanecen ampliamente superpuestas. La profundidad diferencia mejor a las zonas, especialmente por el carácter superficial de la Dorsal Meso-Atlántica y el rango más amplio presente en el Cinturón de Fuego. En cambio, las fluctuaciones observadas entre períodos no constituyen evidencia de un cambio sostenido en la tasa de ocurrencia. Tampoco se identifica estacionalidad mensual ni una composición de severidad diferente entre zonas.

La magnitud y la profundidad, consideradas conjuntamente, no permiten clasificar de manera confiable la zona de un evento. El modelo asigna todos los terremotos al Cinturón de Fuego y obtiene una exactitud de 75,21 %, igual a clasificar siempre en la zona mayoritaria. En consecuencia, los patrones descriptivos de frecuencia y profundidad quedan respaldados, mientras que las diferencias temporales y de severidad requieren una interpretación más acotada. Los resultados se limitan al catálogo analizado y no permiten estimar amenaza, riesgo sísmico ni probabilidades futuras de ocurrencia.

Se adjuntan los siguientes enlaces como productos derivados del informe.

- Repositorio GitHub: MariaJoseVN/USGS-earthquake-analysis.git
- Mapa Georreferenciado Interactivo: Abrir mapa interactivo
- Dashboard Interactivo: Abrir dashboard interactivo
- Informe Asesoría 1: Abrir Informe Asesoría 1
- Guía Metodológica de Delimitación de Zonas en QGIS: Abrir guía metodológica

2. Formulación del problema y pregunta estadística

2.1. Continuidad del estudio y pregunta estadística

El Informe Asesoría 1 construyó y caracterizó el catálogo de 1.186 terremotos con magnitud $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025. En esa etapa se definieron la unidad de análisis, las zonas sísmicas, las categorías de magnitud y profundidad y el tratamiento de las variables. Su alcance descriptivo permitió identificar patrones, pero no determinar si estos superaban la variabilidad propia del catálogo, cuál era su relevancia práctica ni si la magnitud y la profundidad aportaban información conjunta para distinguir las zonas.

El presente informe mantiene esas definiciones y aborda las preguntas pendientes mediante comparaciones inferenciales, modelos de conteo y evaluación de la capacidad clasificatoria. El estudio se articula en torno a la siguiente pregunta:

> **¿Difieren formalmente las zonas sísmicas en la tasa de ocurrencia y en las distribuciones de magnitud y profundidad de los terremotos $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025, y permiten la magnitud y la profundidad, consideradas en conjunto, clasificar la zona sísmica a la que pertenece un evento?**

La primera parte compara la ocurrencia y las características de los eventos entre zonas, considerando la superficie y la estabilidad temporal del período. La segunda evalúa si la magnitud y la profundidad permiten identificar la zona de un terremoto. La comparabilidad de los procedimientos de reporte se revisa antes de interpretar estas diferencias. Las conclusiones se limitan al catálogo definido y no representan predicción, causalidad ni estimación de amenaza sísmica.

3. Antecedentes

Los antecedentes conceptuales del Informe Asesoría 1 se mantienen vigentes y no se repiten en este documento: la distribución global de la sismicidad en tres grandes cinturones, la categorización de los eventos según magnitud y la descripción del catálogo USGS se encuentran documentadas en aquel informe y en sus anexos. Esta sección presenta únicamente los antecedentes propios de la etapa inferencial.

La ocurrencia de grandes terremotos a escala global se estudia habitualmente mediante modelos de conteo y procesos puntuales. La literatura respalda el uso de catálogos globales con umbrales altos de magnitud para analizar tasas de ocurrencia *(Kagan y Jackson 2016)* y entrega la base general de los procesos puntuales aplicados a sismicidad *(Ogata 1988)*. Cuando los conteos presentan una variabilidad mayor que la esperada bajo el supuesto Poisson, el modelo binomial negativo permite incorporar esa sobredispersión, alternativa contemplada también en el instructivo de la asesoría *(Instructivo de encargo 2026)*.

El catálogo analizado está truncado en $M \geq 6{,}5$ por diseño, de modo que todas las zonas comparten el mismo piso de magnitud. Esta condición explica que las medias difieran poco como consecuencia del umbral común y orienta las comparaciones de magnitud hacia la distribución completa y sus cuantiles, en lugar del nivel medio. La completitud del catálogo y la homogeneidad del registro son condiciones reconocidas en la literatura para la validez de estas comparaciones *(Wiemer y Wyss 2000; Cabello et al. 2025)*.

Finalmente, la reformulación de la comparación entre zonas como un problema de clasificación responde a la orientación recibida durante la revisión docente del Informe Asesoría 1. En lugar de limitar el análisis a contrastes de vectores de medias, se evalúa si la magnitud y la profundidad permiten identificar la zona sísmica del evento, mediante regresión logística multinomial complementada con análisis de correspondencia para las variables nominales.

4. Objetivos

4.1. Objetivo general

Evaluar inferencialmente si los patrones de ocurrencia, magnitud, profundidad y severidad de los terremotos $M \geq 6{,}5$ del catálogo USGS 2000-2025 difieren entre zonas sísmicas, y si la magnitud y la profundidad permiten clasificar la zona de un evento, considerando la comparabilidad del registro y la robustez de las conclusiones.

4.2. Objetivos específicos

Para alcanzar el objetivo general, se plantean los siguientes objetivos específicos:

- **OE1. Verificar** la comparabilidad del registro entre zonas y períodos, considerando las fuentes, los métodos de magnitud y el ajuste de localización.
- **OE2. Comparar** la ocurrencia de terremotos entre zonas y a través del tiempo, mediante tasas anuales, densidades por superficie y evaluación de la estabilidad temporal.
- **OE3. Contrastar** las distribuciones de magnitud, profundidad y severidad entre zonas, considerando tanto la significación estadística como la relevancia práctica de las diferencias.
- **OE4. Evaluar** si la magnitud y la profundidad, consideradas conjuntamente, permiten clasificar la zona sísmica de un evento.
- **OE5. Examinar** la robustez de las conclusiones ante cambios en el umbral de magnitud y en la definición de las zonas de comparación.

5. Alcance del informe

El alcance de este informe comprende la comparación inferencial de la ocurrencia, la magnitud, la profundidad y la severidad entre zonas, junto con la evaluación de la estabilidad temporal y de la capacidad de clasificación. Las conclusiones se aplican únicamente a los eventos, variables y delimitaciones evaluados mediante los modelos utilizados.

La heterogeneidad se aborda de manera diferenciada. Las zonas se comparan mediante tasas anuales y densidades por superficie; las magnitudes comparten un umbral mínimo y se analizan mediante sus distribuciones completas; y las fuentes y métodos de reporte se revisan antes de interpretar las diferencias espaciales y temporales. Resto del mundo se considera una categoría residual y heterogénea, no una unidad tectónica equivalente a los tres cinturones definidos.

La superficie permite estandarizar la frecuencia por extensión, pero no representa exposición humana, vulnerabilidad ni riesgo sísmico. Los cambios en el umbral de magnitud y la exclusión de Resto del mundo se utilizan como escenarios de sensibilidad para examinar cuánto dependen las conclusiones de estas decisiones. Los resultados no se extienden a relaciones o escenarios que no hayan sido evaluados mediante un modelo o análisis específico.

Quedan fuera del alcance la predicción de terremotos, la alerta temprana, la estimación de amenaza, la evaluación de daños y la modelación física del proceso tectónico. Los modelos de clasificación examinan si la magnitud y la profundidad permiten distinguir las zonas dentro de los eventos observados, pero no constituyen herramientas para identificar la ubicación de futuros terremotos.

5.1. Declaración y lineamientos para el uso responsable de inteligencia artificial

El equipo utilizó herramientas de inteligencia artificial generativa como apoyo en búsquedas metodológicas, revisión documental y aspectos editoriales. Estas consultas surgieron a partir de diagnósticos realizados por los integrantes y se orientaron a localizar alternativas en la documentación de R y en la literatura estadística. Las funciones, paquetes y fuentes sugeridas se revisaron antes de incorporarse al análisis. La selección metodológica, la ejecución de los scripts, la interpretación y las decisiones finales permanecieron bajo responsabilidad del equipo.

El procedimiento de búsqueda, verificación bibliográfica y uso complementario de NotebookLM, Zotero y QGIS se documenta en el anexo correspondiente. (*véase anexo de uso de inteligencia artificial*).

6. Metodología

La metodología se organiza en seis análisis encadenados. El primero revisa la comparabilidad del registro, porque si los procedimientos de reporte difirieran entre zonas, las comparaciones posteriores mezclarían diferencias sísmicas con diferencias de medición. Establecida esa base, tres análisis abordan la primera componente de la pregunta estadística desde enfoques complementarios. La frecuencia compara cuántos eventos registra cada zona, la comparación de magnitud y profundidad examina cómo son esos eventos, y la estructura temporal evalúa si la tasa se mantuvo estable durante el período, condición que respalda la agregación de años empleada por las anteriores. El quinto análisis cruza ambas miradas al examinar la composición de severidad dentro de cada zona, y el sexto invierte la pregunta para responder la segunda componente, es decir, si la magnitud y la profundidad, en conjunto, identifican la zona de un evento.

6.1. Comparabilidad de los procedimientos de reporte

Esta sección evalúa si los procedimientos de reporte varían entre zonas o períodos y, por tanto, podrían condicionar las comparaciones posteriores. Para ello, examinamos la relación de la zona y el período con el tipo de magnitud, la fuente de magnitud y el ajuste residual de localización. Una asociación fuerte con la zona obligaría a incorporar controles en los modelos posteriores, mientras que una asociación con el período restringiría las comparaciones temporales de magnitud. Las categorías de baja frecuencia se agruparon solo cuando no alteraban la pregunta sustantiva; el detalle de esas decisiones se documenta en el anexo metodológico. (*véase anexo de especificación de dócimas y modelos*).

En esta sección, las dócimas se formularon como contrastes de independencia para las variables categóricas de reporte y como contrastes de igualdad distribucional para el ajuste residual de localización. Por ejemplo, para el tipo de magnitud entre zonas:

$$H_0 : F_{\text{tipo, zona}_1} = F_{\text{tipo, zona}_2} = \cdots = F_{\text{tipo, zona}_k} \quad \text{v/s} \quad H_1 : F_{\text{tipo, zona}_h} \neq F_{\text{tipo, zona}_{h'}} \mid h \neq h'$$

El mismo criterio se aplicó a las asociaciones entre zona y fuente de magnitud, y entre período y tipo de magnitud. Para las comparaciones del ajuste residual de localización, la hipótesis se expresa en términos de igualdad de distribuciones; por ejemplo, entre zonas:

$$H_0 : F_{\mathrm{RMS},\,\mathrm{zona}_1} = F_{\mathrm{RMS},\,\mathrm{zona}_2} = \cdots = F_{\mathrm{RMS},\,\mathrm{zona}_k} \quad \text{v/s} \quad H_1 : F_{\mathrm{RMS},\,\mathrm{zona}_h} \neq F_{\mathrm{RMS},\,\mathrm{zona}_{h'}} \mid h \neq h'$$

La comparación del RMS entre períodos se formuló de forma análoga.

Las asociaciones categóricas se evaluaron mediante chi-cuadrado de independencia (*véase anexo del contraste chi-cuadrado y frecuencias esperadas*). Como algunas tablas tenían frecuencias esperadas bajas, especialmente en zonas con pocos eventos, descartamos fusionar categorías relevantes y usamos valores p obtenidos con 10.000 simulaciones de Monte Carlo (*Hope 1968*; *Patefield 1981*). Esta decisión permite mantener la estructura original de reporte y evita depender de una aproximación chi-cuadrado poco estable en celdas escasas. La precisión de la simulación se presenta fuera del cuerpo principal. (*véase anexo de precisión de la simulación Monte Carlo*).

Para medir la intensidad de esas asociaciones se utilizó la V de Cramér corregida por sesgo. La corrección es relevante porque las tablas tienen tamaños de grupo desiguales y categorías poco frecuentes, condiciones que pueden inflar la asociación aparente aun cuando la relación real sea débil. Por ello, la versión corregida permite distinguir una asociación sustantiva de una discrepancia producida por la estructura muestral (*Bergsma 2013*). (*véase anexo de corrección por sesgo de la V de Cramér*).

El ajuste residual de localización (`rms_imp`) se comparó por zona y período mediante Kruskal-Wallis, manteniendo el criterio no paramétrico usado para variables continuas con valores atípicos y tamaños grupales desiguales. Cada contraste se acompañó con epsilon cuadrado (ε^2), entendido como un análogo del R^2 aplicado a rangos, para separar significación estadística de importancia práctica (*Mangiafico 2026*). (*véase anexo de cálculo de epsilon cuadrado*).

Los cinco contrastes de comparabilidad se trataron como una familia exploratoria. Para no aumentar artificialmente el riesgo de falsos positivos, ajustamos sus valores p mediante Benjamini-Hochberg, que controla la proporción esperada de falsos descubrimientos y conserva más capacidad de detección que una corrección orientada a evitar cualquier error tipo I (*Benjamini y Hochberg 1995*). (*véase anexo de ajuste de valores p por multiplicidad*).

6.2. Comparación de la frecuencia de eventos entre zonas

Esta sección evalúa si la frecuencia de eventos difiere entre zonas, tanto por año como por unidad de superficie. A diferencia del análisis temporal, aquí se comparan zonas resumiendo el período completo; la estabilidad de esa agregación se evalúa posteriormente.

La frecuencia de terremotos es un conteo de eventos por zona y año, por lo que la comparación se realizó con un modelo lineal generalizado de conteos y no con una comparación de medias. La unidad de análisis fue una combinación de zona y año, con 104 observaciones correspondientes a 26 años y cuatro zonas. Las combinaciones sin eventos se conservaron con conteo cero, porque representan años observados sin terremotos del catálogo y no datos faltantes. Se estimaron dos medidas complementarias: la tasa anual, comparable con el Informe Asesoría 1, y la densidad por superficie. Las zonas tienen tamaños muy distintos, de modo que comparar solo los conteos favorecería a las más extensas; la densidad corrige esto al expresar los eventos por unidad de superficie. Para ello, la superficie de cada zona se incorpora al modelo como un offset: un término conocido que se suma con un coeficiente fijo, no estimado, y que traslada la comparación de conteos absolutos a conteos por kilómetro cuadrado. De este modo, la diferencia descriptiva del Informe Asesoría 1 se convierte en una comparación inferencial que cuantifica cuánto difieren las zonas y con qué precisión. (*véase anexo de áreas de zonas y definición del offset*).

El punto de partida fue un modelo Poisson con enlace logarítmico, apropiado para conteos, pero que supone que la varianza condicional iguala a la media. Los datos mostraron sobredispersión, condición que subestima los errores estándar e infla la significación si no se corrige. Se evaluó mediante la dispersión de Pearson, la prueba unilateral de Cameron-Trivedi y una comparación de verosimilitudes entre Poisson y binomial negativa (*Cameron y Trivedi 1990*). Las tres vías evalúan la misma condición, la igualdad entre la varianza y la media frente a una varianza mayor:

$$H_0 : \ \mathrm{Var}(Y) = \mathbb{E}(Y) \quad \text{v/s} \quad H_1 : \ \mathrm{Var}(Y) > \mathbb{E}(Y)$$

Se consideró también la especificación quasi-Poisson, que mantiene la media del modelo Poisson y corrige los errores estándar multiplicándolos por un factor de escala (*McCullagh y Nelder 1989*). Esta especificación, sin embargo, no define una distribución de probabilidad completa, sino solo la media y la varianza, por lo que carece de una verosimilitud que permita comparar modelos de manera formal. Se escogió la binomial negativa porque sí es una distribución completa, con verosimilitud explícita, necesaria tanto para el contraste global de zona mediante razón de verosimilitudes como para obtener las razones de tasa con sus intervalos de confianza. El ajuste se realizó con `glm.nb` de `MASS`, siguiendo su documentación oficial (*Venables y Ripley 2002*). La justificación del modelo se presenta en el anexo correspondiente y el detalle de los tres diagnósticos, con sus definiciones, valores y la corrección por frontera del contraste, se presenta

fuera del cuerpo principal. (*véanse anexo del modelo binomial negativo para conteos con sobredispersión y anexo de diagnósticos de los modelos*).

El efecto global de zona se evaluó mediante una razón de verosimilitudes con tres grados de libertad, que compara el modelo con zonas frente a un modelo sin zonas (*véase anexo de verosimilitud y estimación de los modelos*). Su hipótesis contrasta la igualdad de las cuatro tasas frente a la existencia de al menos una diferencia:

$$H_0 : \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 \quad \text{v/s} \quad H_1 : \lambda_h \neq \lambda_{h'} \text{ para algún } h \neq h'$$

Una vez establecido el efecto global, se examinaron las seis comparaciones posibles entre zonas. Sus valores p se ajustaron mediante el procedimiento de Holm, que controla en 5 % la probabilidad de obtener al menos un falso positivo dentro de cada familia de seis contrastes (*Holm 1979*). El tamaño del efecto (cuán grande es la diferencia entre zonas, no solo si existe) se expresó como razón de tasa y de densidad. Un valor de 1 representa igualdad, un valor mayor que 1 indica mayor frecuencia en la primera zona y un valor menor que 1 indica menor frecuencia (*véase anexo de medidas de tamaño de efecto*). Cada razón se acompañó de un intervalo de confianza del 95 %, obtenido por verosimilitud perfilada en las comparaciones con la zona de referencia, que no supone simetría y es preferible en zonas con pocos eventos, y por el método de Wald en las comparaciones por pares (*véanse el anexo de intervalos de confianza y el anexo del contraste de Wald*). La decisión combinó el valor p ajustado, la magnitud de la razón y la amplitud del intervalo, con cautela adicional para las zonas de menor conteo. Los desarrollos formales de estas comparaciones se presentan fuera del cuerpo principal. (*véase anexo de diagnósticos de los modelos*).

La validez de estas conclusiones descansa en cuatro supuestos. Primero, que los conteos de cada zona y año son independientes entre sí una vez descontado el efecto de zona. Segundo, que cada fila cubre el mismo tiempo de observación, un año, lo que permite compararlos como tasas anuales. Tercero, que la superficie de cada zona se mantiene constante durante todo el período. Cuarto, que la forma de varianza de la binomial negativa describe adecuadamente la variabilidad de los conteos. Los supuestos evaluables a partir del ajuste se revisaron mediante la dispersión residual, los residuos de Pearson, la distancia de Cook y la autocorrelación residual dentro de cada zona. Por último, la categoría "Resto del mundo" se trata como un cierre residual y heterogéneo del catálogo, no como un cinturón tectónico, por lo que sus resultados no se interpretan como los de una zona sísmica coherente.

6.3. Comparación de magnitud y profundidad entre zonas

Mientras la comparación de frecuencia examina cuántos eventos registra cada zona, esta sección evalúa si sus distribuciones de magnitud y profundidad difieren. El Informe Asesoría 1 mostró magnitudes medias casi idénticas (6,90; 6,92 y 6,91) salvo la Dorsal Meso-Atlántica (6,69), y una profundidad con composición claramente distinta por zona; aquí evaluamos si esas diferencias descriptivas son estadísticamente defendibles.

La comparación de magnitud y profundidad se realizó sobre `zona`, `mag` y `depth`, manteniendo la zonificación del Informe Asesoría 1. Antes de contrastar diferencias, evaluamos si era defendible usar una ruta paramétrica multivariante. Esta revisión era necesaria porque MANOVA supone un comportamiento conjunto aproximadamente normal y matrices de dispersión comparables; si esas condiciones no se cumplen, los valores p pueden reflejar colas, empates, valores extremos o heterogeneidad de varianzas más que diferencias reales entre zonas.

La evaluación combinó diagnóstico gráfico y pruebas formales, incluyendo QQ-plots univariados, QQ multivariante con distancias de Mahalanobis, Mardia para normalidad multivariada (*Mardia 1970*), Shapiro-Wilk para normalidad marginal (*Shapiro y Wilk 1965*) y M de Box para homogeneidad de matrices de varianza-covarianza (*Box 1949*). Bartlett y Spearman se usaron solo como diagnósticos complementarios de la relación entre magnitud y profundidad, no como criterios de normalidad (*Bartlett 1950*). Las hipótesis formales se presentan en el anexo metodológico. (*véase anexo de especificación de dócimas y modelos*).

La decisión de ruta se organizó mediante las siguientes dócimas principales. Para la normalidad multivariada del vector formado por magnitud y profundidad:

$$H_0 : \mathbf{X}_{\text{zona}_h} \sim N_2(\mu_h, \Sigma_h) \quad \text{v/s} \quad H_1 : \mathbf{X}_{\text{zona}_h} \nsim N_2(\mu_h, \Sigma_h) \mid h = 1, \ldots, k$$

Para la homogeneidad de matrices de varianza-covarianza entre zonas:

$$H_0 : \Sigma_{\text{zona}_1} = \Sigma_{\text{zona}_2} = \cdots = \Sigma_{\text{zona}_k} \quad \text{v/s} \quad H_1 : \Sigma_{\text{zona}_h} \neq \Sigma_{\text{zona}_{h'}} \mid h \neq h'$$

Como los supuestos no se sostuvieron, no se utilizó MANOVA como procedimiento principal. En su lugar, comparamos las distribuciones de magnitud y profundidad mediante Kruskal-Wallis (*Kruskal y Wallis 1952*), prueba basada en rangos que no exige normalidad y mantiene la pregunta central de comparación entre zonas. Cada contraste se acompañó con epsilon cuadrado (ε^2) para evaluar relevancia práctica (*Mangiafico 2026*).

Para esta comparación no paramétrica de magnitud y profundidad, la dócima se formuló como:

$$H_0 : F_{Y,\,\text{zona}_1} = F_{Y,\,\text{zona}_2} = \cdots = F_{Y,\,\text{zona}_k} \quad \text{v/s} \quad H_1 : F_{Y,\,\text{zona}_h} \neq F_{Y,\,\text{zona}_{h'}} \mid h \neq h'$$

donde $Y \in \{\text{magnitud}, \text{profundidad}\}$.

Se consideró PERMANOVA como alternativa no paramétrica multivariante, pero se descartó en favor de Kruskal-Wallis porque la pregunta requería interpretar magnitud y profundidad por separado. Además, PERMANOVA puede confundirse con diferencias de dispersión entre grupos, precisamente una de las heterogeneidades observadas en el catálogo.

Cuando Kruskal-Wallis indicó diferencias globales, aplicamos comparaciones post-hoc de Dunn con ajuste de Holm (*Dunn 1964*; *Holm 1979*). Esta elección cumple el mismo rol que tendrían las comparaciones múltiples después de un ANOVA, pero en una ruta no paramétrica. Permite identificar qué pares de zonas explican la diferencia global usando rangos, sin volver a supuestos de normalidad. El ajuste de Holm controla el error tipo I acumulado al realizar seis comparaciones por pares dentro de cada variable. La tabla completa se presenta en el *anexo de comparaciones Dunn-Holm*.

Para complementar esas comparaciones, se calculó delta de Cliff en pares sustantivamente relevantes (*Cliff 1993*). Esta medida permite cuantificar no solo si dos zonas difieren, sino también cuán separadas están sus distribuciones y en qué dirección. A diferencia de una comparación basada en medias, no depende de normalidad ni de desviaciones estándar, sino que estima la probabilidad relativa de que un evento de una zona presente valores mayores que un evento de la zona comparada. Por ello, es coherente con la ruta no paramétrica adoptada.

Los resultados de esta sección se retoman en la clasificación de zonas, donde magnitud y profundidad se evalúan en conjunto como predictores de la zona del evento.

6.4. Estructura temporal de la ocurrencia

Esta sección evalúa si la tasa de ocurrencia del catálogo se mantiene estable durante 2000-2025. A diferencia de la comparación de frecuencia, que contrasta zonas durante todo el período, aquí se agregan las zonas para estudiar los cambios en el tiempo. El análisis formaliza el hallazgo descriptivo del Informe Asesoría 1, según el cual la ocurrencia anual fluctúa entre años, pero no muestra una tendencia visual uniforme. Revisamos si los conteos fluctúan alrededor de un nivel estable (estacionariedad), si un período con muchos eventos tiende a ser seguido por otro igual de alto (autocorrelación), si algún tramo del período o algún mes concentra sistemáticamente más eventos (cambio de nivel y estacionalidad) y si los tiempos entre eventos consecutivos son compatibles con una tasa constante (ajuste exponencial). Para ello usamos dos series construidas a partir del mismo catálogo. La serie anual permite comparar con la descripción previa y con los períodos de análisis, mientras que la serie mensual entrega más observaciones para diagnosticar estabilidad temporal, autocorrelación y estacionalidad.

Evaluamos la estacionariedad mediante Dickey-Fuller aumentada (ADF) y KPSS, interpretadas en conjunto porque sus hipótesis nulas son opuestas (*Said y Dickey 1984*; *Kwiatkowski et al. 1992*). Aplicamos la dócima ADF con `adfTest()` del paquete `fUnitRoots`, usando la especificación sin constante ni tendencia (`type = "nc"`), ya que de esta forma se parte desde el contraste más parsimonioso y se evita imponer componentes determinísticos adicionales antes de verificar si son necesarios. Usamos un rezago para la serie anual y doce rezagos para la serie mensual, con el fin de incorporar la dependencia inmediata entre años y, en la escala mensual, cubrir un ciclo anual completo. ADF parte desde la hipótesis de raíz unitaria, por lo que rechazarla entrega evidencia a favor de estacionariedad. KPSS parte desde la hipótesis de estacionariedad en nivel, por lo que no rechazarla respalda que la serie no presenta una desviación sistemática respecto de un nivel estable. Usar ambas evita depender de una sola prueba y permite distinguir conclusiones concordantes de situaciones ambiguas, especialmente en la serie anual, donde solo hay 26 observaciones (*Banerjee et al. 1993*).

Las dócimas se expresan como:

$$H_0^{ADF} : \text{raíz unitaria} \quad \text{v/s} \quad H_1^{ADF} : \text{estacionariedad}$$

$$H_0^{KPSS} : \text{estacionariedad en nivel} \quad \text{v/s} \quad H_1^{KPSS} : \text{no estacionariedad}$$

Revisamos la independencia serial mediante Ljung-Box, porque una serie puede ser estacionaria y aun así presentar autocorrelación (*Ljung y Box 1978*). Esta revisión permite evaluar si los conteos de un período dependen sistemáticamente de los períodos anteriores; si no hay autocorrelación relevante, las fluctuaciones observadas se interpretan como variación propia del catálogo y no como rachas temporales persistentes. Esta prueba evalúa si las autocorrelaciones hasta un rezago definido son conjuntamente nulas. Como diagnóstico gráfico complementario, examinamos la ACF y la PACF de la serie mensual; priorizamos la escala mensual porque entrega más observaciones que la serie anual y permite observar rezagos específicos con mayor estabilidad.

$$H_0^{LB} : \rho_1 = \rho_2 = \cdots = \rho_m = 0 \quad \text{v/s} \quad H_1^{LB} : \rho_j \neq 0 \mid j \leq m$$

Además de estas pruebas, comparamos la tasa media anual de eventos entre tres tramos del período observado: 2000-2009, 2010-2019 y 2020-2025. Esta comparación complementa la lectura descriptiva del Informe Asesoría 1. El gráfico permite observar fluctuaciones en los conteos, mientras que el modelo evalúa si las diferencias entre períodos superan la variabilidad esperable del catálogo. Como la variable de respuesta corresponde a conteos anuales de eventos, usamos un GLM Poisson, que es el punto de partida natural para modelar números de ocurrencias en intervalos de exposición comparables. En este caso, cada observación corresponde a un año del catálogo, por lo que el modelo permite estimar y contrastar tasas medias de ocurrencia entre períodos sin tratar los conteos como si fueran variables continuas normales. Su resultado indica si algún tramo presenta una tasa anual estadísticamente distinta de los demás; si no se detectan diferencias, la lectura temporal respalda el uso de una tasa agregada para todo el período (*véase el anexo del modelo Poisson para conteos temporales*). Luego evaluamos si los conteos varían más de lo esperado bajo Poisson; cuando esto ocurre, la decisión inferencial se basa en quasi-Poisson, que mantiene la comparación de tasas pero corrige la varianza estimada. En esta parte usamos esa corrección en lugar de la binomial negativa aplicada en la comparación de frecuencia entre zonas, porque aquí el objetivo es únicamente docimar la significación de un factor temporal, para lo cual basta reescalar la varianza y aplicar un contraste F (*McCullagh y Nelder 1989*). La binomial negativa queda reservada para aquella comparación, donde se requieren razones de tasa con intervalos de confianza y comparación formal de modelos, lo que exige una verosimilitud completa. La regla es la misma en ambos casos. Todo modelo de conteos se diagnostica por sobredispersión antes de concluir, y la herramienta de corrección se elige según el objetivo del análisis.

Para la comparación entre períodos, la dócima se formula como:

$$H_0 : \lambda_{\text{2000-2009}} = \lambda_{\text{2010-2019}} = \lambda_{\text{2020-2025}} \quad \text{v/s} \quad H_1 : \lambda_a \neq \lambda_b \mid a \neq b$$

Con el mismo criterio, evaluamos la estacionalidad mensual mediante un GLM Poisson aplicado a los conteos mensuales, usando el mes calendario como factor. Este contraste revisa si algún mes concentra sistemáticamente más eventos que otros dentro del período del catálogo. La inferencia también se verifica con quasi-Poisson cuando la variabilidad observada supera la esperada bajo Poisson (*véase el anexo del modelo Poisson para conteos temporales*).

Para la estacionalidad mensual, la dócima se expresa como:

$$H_0 : \lambda_1 = \lambda_2 = \cdots = \lambda_{12} \quad \text{v/s} \quad H_1 : \lambda_m \neq \lambda_{m'} \mid m \neq m'$$

Por último, analizamos los tiempos transcurridos entre eventos consecutivos para el grupo $M \geq 7{,}0$ y para cada categoría de magnitud. Esta revisión complementa los modelos de conteo. Mientras Poisson evalúa cuántos eventos ocurren en un intervalo fijo, el ajuste exponencial evalúa cuánto tiempo pasa entre un evento y el siguiente. Si la ocurrencia fuera compatible con una tasa constante, los tiempos de espera deberían aproximarse a una distribución exponencial. Contrastamos el ajuste mediante Kolmogorov-Smirnov (*Conover 1971*; *Marsaglia et al. 2003*) con bootstrap paramétrico, porque la tasa de la distribución se estima con los mismos datos y esa estimación modifica la distribución nula del estadístico (*Durbin 1973*) (*véanse el anexo del modelo exponencial y la prueba de ajuste de los tiempos entre eventos y el anexo de bootstrap paramétrico para pruebas de ajuste*).

La dócima de ajuste se formula como:

$$H_0 : T \sim \text{Exponencial}(\lambda) \quad \text{v/s} \quad H_1 : T \nsim \text{Exponencial}(\lambda)$$

6.5. Eventos fuertes y extremos

Esta sección evalúa si la composición de severidad y la probabilidad de registrar eventos de mayor magnitud difieren entre zonas. A diferencia del análisis de frecuencia, interesa la gravedad relativa de los eventos y no su cantidad total. Para ello retomamos las categorías definidas en el Informe Asesoría 1, correspondientes a Fuerte ($6{,}5 \leq M < 7{,}0$), Mayor ($7{,}0 \leq M < 7{,}8$) y Grande o extremo ($M \geq 7{,}8$). El análisis se organizó en cuatro componentes: una tabla de contingencia entre zona y categoría de magnitud, una regresión logística para la probabilidad de que un evento alcance $M \geq 7{,}0$, una comparación entre dos representaciones de la profundidad dentro de ese modelo y una descripción de los eventos con $M \geq 7{,}8$ mediante conteos y tasas anuales.

La tabla de contingencia cruza las cuatro zonas con las tres categorías de magnitud. Su dócima evalúa si la distribución de categorías es la misma en todas las zonas:

$$H_0 : F_{\text{cat, zona}_1} = F_{\text{cat, zona}_2} = \cdots = F_{\text{cat, zona}_k} \quad \text{v/s} \quad H_1 : F_{\text{cat, zona}_h} \neq F_{\text{cat, zona}_{h'}} \mid h \neq h'$$

donde $F_{\text{cat, zona}_h}$ representa la distribución de las tres categorías de magnitud dentro de la zona h.

Como las frecuencias esperadas incluían celdas bajo 5 en las zonas de menor tamaño (*véase anexo del contraste chi-cuadrado y frecuencias esperadas*), el valor p del chi-cuadrado se obtuvo con 10.000 simulaciones de Monte Carlo, el mismo criterio definido para la comparabilidad de los procedimientos de reporte (*Hope 1968*; *Patefield 1981*). El

grado de asociación se midió con la V de Cramér corregida por sesgo, que descuenta la asociación aparente producida por la estructura muestral (*Bergsma 2013*). (*véase anexo de precisión de la simulación Monte Carlo*).

Para la regresión logística, las categorías Mayor y Grande o extremo se agruparon en un indicador binario que vale 1 cuando $M \geq 7{,}0$ y 0 en caso contrario. El modelo estima la probabilidad de superar ese umbral en función de la zona y la profundidad consideradas simultáneamente, con el Cinturón de Fuego como referencia por ser la zona de mayor tamaño muestral (*McCullagh y Nelder 1989*) (*véase anexo del modelo de regresión logística*). Incluir ambas variables en un mismo modelo pone a prueba una posibilidad que los análisis por separado no cubren: que la profundidad, con poca asociación cuando se examina sola, aporte información una vez considerada la zona. La significación de cada variable se evaluó mediante razón de verosimilitudes por bloque (*véase anexo de verosimilitud y estimación de los modelos*), porque la zona aporta tres coeficientes y debe contrastarse como una sola variable:

$$H_0 : \ \beta_{\text{zona}_2} = \beta_{\text{zona}_3} = \beta_{\text{zona}_4} = 0 \quad \text{v/s} \quad H_1 : \ \beta_{\text{zona}_h} \neq 0 \text{ para algún } h$$

donde β_{zona_h} es el coeficiente que compara la zona h con el Cinturón de Fuego.

La inferencia se organiza así en dos niveles. El contraste global decide si una variable aporta en conjunto: es la dócima anterior, con una sola respuesta por variable. Las pruebas de Wald individuales responden una pregunta más específica, cuál zona difiere de la referencia, y contrastan cada coeficiente por separado:

$$H_0 : \ \beta_j = 0 \quad \text{v/s} \quad H_1 : \ \beta_j \neq 0$$

donde β_j es, por separado, el coeficiente de cada zona frente a la referencia o el de la profundidad. Estas pruebas se reportan solo como detalle del contraste global, nunca en su reemplazo: elegir entre las tres comparaciones la más significativa, sin un veredicto global previo, aumentaría el riesgo de declarar un efecto por azar (*véase anexo del contraste de Wald*). El tamaño de efecto (cuán grande es la diferencia, no solo si existe) se expresó como razón de odds (odds ratio), donde las odds son el cociente entre la probabilidad de superar el umbral y la de no superarlo: un valor de 1 indica igualdad con la referencia y un valor menor que 1, menor probabilidad relativa (*véase anexo de la razón de odds*). Cada razón se acompañó de un intervalo de confianza del 95 % obtenido por verosimilitud perfilada, que no supone simetría y es preferible en zonas con pocos eventos (*Venables y Ripley 2002*) (*véase anexo de intervalos de confianza*). Se descartó una regresión multinomial con la categoría de magnitud como respuesta: la categoría Grande o extremo reúne pocos eventos en las zonas menores, lo que desestabiliza la estimación, y la clasificación multinomial corresponde a la sección siguiente, con la zona como respuesta.

La profundidad se evaluó en dos representaciones: continua, en kilómetros, y categórica, con las clases Superficial, Intermedio y Profundo heredadas del Informe Asesoría 1. Ambas versiones del modelo se compararon mediante la razón de verosimilitudes de cada variable y el criterio de información de Akaike (AIC), que penaliza los parámetros adicionales para equilibrar ajuste y simplicidad. Una diferencia de AIC menor que 2 se interpretó como respaldo prácticamente equivalente (*Burnham y Anderson 2004*) (*véase anexo del criterio de información de Akaike*); en ese caso se privilegió la especificación más parsimoniosa y que no pierde información por categorización.

La validez de la regresión logística descansa en tres condiciones. Primero, los eventos se tratan como observaciones independientes entre sí, tratamiento respaldado por la estructura temporal, ya que los conteos no muestran autocorrelación ni estacionalidad (*véase la sección de estructura temporal*). Segundo, el modelo asume que cada kilómetro adicional de profundidad cambia las log-odds en la misma cantidad (una relación lineal); la versión categórica, que estima un coeficiente libre por estrato, permite revisar si esa forma oculta diferencias por tramos de profundidad. Tercero, todas las zonas presentan eventos en ambos valores del indicador, condición verificada que evita problemas de separación en el umbral $M \geq 7{,}0$. Estas condiciones se complementan con un tamaño muestral holgado para las aproximaciones de muestras grandes, con 387 eventos sobre el umbral para cinco parámetros estimados.

Los eventos con $M \geq 7{,}8$ se describieron mediante conteos por zona y tasas anuales sobre los 26 años observados; a diferencia del umbral $M \geq 7{,}0$, para esta categoría no se estimó ninguna regresión. La categoría reúne 59 eventos, con 2 en el Cinturón Alpino-Himalayo y ninguno en la Dorsal Meso-Atlántica. Con esos conteos, una logística del umbral $M \geq 7{,}8$ produciría separación, porque el coeficiente de una zona sin eventos tiende a valores extremos, con errores estándar e intervalos poco fiables (*véase anexo del contraste de Wald*). La lectura de esta componente es descriptiva. Un conteo igual a cero indica ausencia en el catálogo y período analizados, no imposibilidad física. Los criterios de decisión de estas dócimas se resumen en el anexo metodológico. (*véase anexo de especificación de dócimas y modelos*).

6.6. Clasificación de zonas a partir de magnitud y profundidad

Esta sección evalúa si la magnitud y la profundidad permiten identificar la zona de un evento. A diferencia de las comparaciones anteriores, la zona deja de ser el factor de agrupación y pasa a ser la variable respuesta. El análisis tiene dos etapas: una exploración de correspondencias entre las zonas y las categorías de magnitud y profundidad, y una regresión logística multinomial confirmatoria. Una correlación parcial mide cuánto aporta una variable numérica

una vez descontada otra. Se evaluó esa vía al inicio, pero se descartó por dos razones: no admite la zona, que es categórica, como variable de interés, y el aporte de la magnitud y la profundidad por separado ya se cuantificó en la comparación entre zonas, de modo que el aporte relevante ahora es el conjunto, que el modelo multinomial estima de forma directa.

El análisis de correspondencia es una técnica descriptiva para tablas de contingencia que compara los perfiles porcentuales de filas y columnas mediante distancias chi-cuadrado y los representa en un plano de pocas dimensiones (*Greenacre 2017*) (*véase anexo de análisis de correspondencia*). La inercia total cuantifica cuánto se apartan los perfiles observados de la independencia y se reparte entre las dimensiones; las contribuciones indican qué zonas y categorías definen cada eje. Se aplicó a las tablas que cruzan las zonas con las categorías de magnitud y de profundidad; como ambas tienen cuatro filas y tres columnas, dos dimensiones representan la totalidad de su geometría. La cercanía entre puntos sugiere perfiles semejantes, pero no constituye evidencia inferencial por sí sola. Por eso la correspondencia se ubica primero, opera directamente sobre la tabla observada, sin supuestos que verificar, y se interpreta junto con las frecuencias esperadas de cada tabla. El modelo multinomial, que sí requiere condiciones de validez revisadas al final de esta sección, confirma luego de manera formal la estructura que la correspondencia solo sugiere.

La regresión logística multinomial modela la probabilidad de pertenecer a cada zona en función de la magnitud y la profundidad, con el Cinturón de Fuego como referencia por ser la zona de mayor tamaño muestral. El modelo estima tres ecuaciones, una por cada zona alternativa frente a la referencia (*Venables y Ripley 2002*) (*véase anexo del modelo de regresión logística multinomial*). La significación de cada variable se evaluó mediante razón de verosimilitudes por bloque, con tres grados de libertad por variable, porque cada una aporta un coeficiente en cada ecuación:

$$H_0 : \ \beta_{Y,\text{zona}_2} = \beta_{Y,\text{zona}_3} = \beta_{Y,\text{zona}_4} = 0 \quad \text{v/s} \quad H_1 : \ \beta_{Y,\text{zona}_h} \neq 0 \text{ para algún } h$$

donde β_{Y,zona_h} es el coeficiente de la variable $Y \in \{\text{magnitud}, \text{profundidad}\}$ en la ecuación que compara la zona h con el Cinturón de Fuego.

Los predictores se estandarizaron a puntajes Z, es decir, se expresaron en desviaciones estándar en lugar de sus unidades originales (kilómetros de profundidad y unidades de magnitud) (*véase anexo de estandarización de los predictores*). Es un cambio de escala, no de modelo; el ajuste se mantiene idéntico (misma devianza y AIC), pero los coeficientes de magnitud y profundidad quedan en una escala común comparable y, con ello, un coeficiente extremo que en kilómetros pasa inadvertido se vuelve evidente. El diagnóstico posterior reveló separación cuasi perfecta en la Dorsal Meso-Atlántica, sus 28 eventos son superficiales, de modo que la profundidad la distingue de manera casi determinista; en esa condición los estimadores de máxima verosimilitud divergen hacia valores extremos y la prueba de Wald sobre ese coeficiente deja de ser fiable (*Albert y Anderson 1984*; *Hauck y Donner 1977*) (*véase anexo del contraste de Wald*). Por ello, los coeficientes y las razones de odds finales se obtuvieron con un ajuste de reducción de sesgo, que produce estimaciones finitas ante separación (*Kosmidis y Firth 2011*) (*véase anexo de separación y reducción de sesgo*). Los intervalos de confianza del 95 % de las razones de odds se calcularon por el método de Wald sobre los coeficientes corregidos, adecuado para las estimaciones finitas; la profundidad de la Dorsal Meso-Atlántica, que permanece en el borde por la separación, se reporta sin un intervalo utilizable. Las pruebas globales de razón de verosimilitudes se mantienen sobre el modelo de máxima verosimilitud, porque comparan modelos anidados bajo el mismo criterio de ajuste.

La capacidad clasificatoria se evaluó primero con la matriz de confusión bajo la regla de máxima probabilidad, que asigna cada evento a la zona con mayor probabilidad ajustada. A partir de ella se calcularon tres métricas (*véase anexo de métricas de clasificación*). La exactitud global se comparó con la base trivial de clasificar siempre la zona mayoritaria (75,2 % de los eventos); la exactitud balanceada se obtuvo como promedio de las sensibilidades por zona, decisiva ante el desbalance porque pondera igual a las cuatro zonas; y la sensibilidad de cada zona se definió como la proporción de eventos de esa zona correctamente identificados.

Como complemento, se evaluó la discriminación probabilística mediante curvas ROC uno contra el resto (*Robin et al. 2011*) (*véase anexo de curvas ROC y área bajo la curva*). Aunque la respuesta del modelo es multinomial, cada zona puede examinarse por separado como una comparación binaria: pertenecer a esa zona frente a pertenecer a cualquiera de las demás. En esa lectura, la sensibilidad corresponde a la proporción de eventos de la zona evaluada que el modelo logra reconocer como tales cuando se fija un umbral de probabilidad; por ejemplo, para la Dorsal Meso-Atlántica indica qué proporción de sus eventos queda capturada al exigir cierto nivel de probabilidad asignada a Dorsal. La especificidad corresponde a la proporción de eventos que no pertenecen a esa zona y que el modelo deja correctamente fuera. Por eso el eje horizontal se presenta como 1 − especificidad: representa los falsos positivos, es decir, eventos de otras zonas que serían incluidos erróneamente como si fueran de la zona evaluada.

El área bajo la curva ROC (AUC) resume el equilibrio entre sensibilidad y falsos positivos a través de todos los umbrales posibles. Un AUC cercano a 0,5 indica que las probabilidades del modelo ordenan los eventos de forma similar al azar; valores más altos indican que los eventos de la zona evaluada tienden a recibir probabilidades mayores que los eventos del resto. Esta métrica no reemplaza la matriz de confusión, porque no entrega una clase final; permite

detectar señal discriminante en las probabilidades aun cuando la regla de máxima probabilidad no reasigne eventos a esa zona. Dado el fuerte desbalance entre zonas, especialmente en la Dorsal Meso-Atlántica, los AUC se interpretaron como evidencia complementaria y con cautela cuando la clase positiva tuvo pocos eventos. Además, el aporte de cada variable se contrastó con modelos anidados (solo magnitud, solo profundidad y conjunto), comparados por razón de verosimilitudes y por AIC con el criterio ya definido de diferencia menor que 2 (*Burnham y Anderson 2004*) (*véase anexo del criterio de información de Akaike*).

La validez del modelo descansa en tres condiciones. Los eventos se tratan como observaciones independientes, las zonas son categorías excluyentes y los predictores continuos entran con relación lineal en la escala de log-odds. La condición adicional de no separación se revisó de manera explícita y no se cumple para la Dorsal; ese incumplimiento se declara y se aborda con la corrección descrita. Los criterios de decisión de estas dócimas se resumen en el anexo metodológico. (*véase anexo de especificación de dócimas y modelos*).

7. Resultados y discusión

7.1. Comparabilidad del registro

Esta subsección evalúa si las comparaciones entre zonas y períodos pueden estar condicionadas por diferencias en las fuentes y procedimientos de reporte del catálogo. Esta verificación es necesaria porque la magnitud informada depende del método y de la institución que la estima, mientras que el RMS resume el ajuste residual de la localización. Si estas condiciones variaran de forma importante entre zonas, las diferencias observadas podrían atribuirse parcialmente al proceso de medición y no al comportamiento sísmico; asimismo, sus cambios temporales podrían generar tendencias aparentes. Para examinar este riesgo, se analiza la distribución del tipo y la fuente de magnitud, junto con `rms_imp`. Esta versión completada incorpora nueve valores imputados y permite conservar los 1.186 eventos; el Informe Asesoría 1 mostró que la imputación tuvo un efecto descriptivo prácticamente nulo. Los resultados permiten decidir si las comparaciones posteriores pueden interpretarse directamente o requieren ajustes, restricciones o advertencias adicionales.

7.1.1. Comparabilidad entre zonas

Para evaluar si las comparaciones espaciales podían estar condicionadas por los procedimientos de reporte, la asociación de la zona con el tipo y la fuente de magnitud se examinó mediante pruebas chi-cuadrado de Pearson con valores p obtenidos a partir de 10.000 simulaciones de Monte Carlo. Estas pruebas comparan las frecuencias observadas en cada combinación de categorías con las que se esperarían si las variables no estuvieran asociadas (*véase anexo del contraste chi-cuadrado y frecuencias esperadas*). El diagnóstico previo identificó frecuencias esperadas inferiores a cinco en 3 de las 16 celdas de la tabla zona por tipo de magnitud y en 4 de las 16 celdas de la tabla zona por fuente, lo que fundamentó el uso de la simulación. (*véase tabla de diagnóstico de frecuencias esperadas*). Adicionalmente, el RMS de localización se comparó entre las cuatro zonas mediante la prueba no paramétrica de Kruskal-Wallis, que determina si sus distribuciones presentan diferencias globales.

En la tabla siguiente, χ^2 representa el estadístico chi-cuadrado. Cuanto mayor es la discrepancia entre las frecuencias observadas y esperadas, mayor es su valor. El símbolo $H(3)$ corresponde al estadístico de Kruskal-Wallis y el número entre paréntesis indica sus tres grados de libertad, obtenidos al comparar cuatro zonas. Los valores p presentados incorporan el ajuste conjunto de los cinco contrastes de comparabilidad. Finalmente, la V de Cramér corregida por sesgo cuantifica la magnitud de la asociación entre variables categóricas, mientras que ε^2 aproxima la proporción de variabilidad de los rangos del RMS atribuible a la zona. Ambos tamaños de efecto se interpretan directamente a partir de su valor numérico, sin asignarles categorías mediante puntos de corte rígidos. (*véase anexo de cálculo de epsilon cuadrado*).

Contraste	Estadístico	Valor p	Tamaño de efecto	Lectura
Zona por tipo de magnitud	$\chi^2 = 6{,}736$	0,671	$V_c = 0{,}000$	Asociación nula.
Zona por fuente de magnitud	$\chi^2 = 10{,}761$	0,351	$V_c = 0{,}022$	Asociación mínima.
RMS por zona	$H(3) = 10{,}348$	0,026	$\varepsilon^2 = 0{,}009$	Efecto menor al 1 %.

Tabla 1. Comparabilidad de los procedimientos de reporte entre zonas sísmicas.

El tipo y la fuente de magnitud no presentaron asociaciones relevantes con la zona sísmica. Sus valores de V de Cramér corregida fueron 0,000 y 0,022, respectivamente. Aunque el RMS mostró una diferencia estadísticamente detectable, $\varepsilon^2 = 0{,}009$ indica que la zona se asocia con menos del 1 % de la variabilidad de sus rangos. En conjunto, no se identificó evidencia de que las comparaciones entre zonas estén materialmente condicionadas por diferencias en los procedimientos de reporte o en el ajuste residual de la localización.

7.1.2. Cambios temporales en los procedimientos de reporte

Los dos contrastes temporales completan la familia de cinco pruebas ajustadas conjuntamente. El RMS se comparó entre los períodos 2000-2009, 2010-2019 y 2020-2025 mediante Kruskal-Wallis. Por su parte, la relación entre el período y el tipo de magnitud se evaluó mediante chi-cuadrado con simulación Monte Carlo. Esta tabla no presentó frecuencias esperadas inferiores a cinco; sin embargo, se mantuvo la simulación para aplicar el mismo procedimiento a los tres contrastes categóricos. (*véase tabla de diagnóstico de frecuencias esperadas*). En ambos casos, los tamaños de efecto permiten determinar si las diferencias detectadas tienen importancia práctica.

Contraste	Estadístico	Valor p	Tamaño de efecto	Resultado
RMS por período	$H(2) = 261{,}930$	$< 0{,}001$	$\varepsilon^2 = 0{,}221$	22,1 % de variabilidad de los rangos.
Período por tipo de magnitud	$\chi^2 = 882{,}090$	$< 0{,}001$	$V_c = 0{,}608$	Asociación marcadamente alejada de cero.

Tabla 2. Cambios temporales en los procedimientos de reporte del catálogo.

La dirección de esta asociación se observa en la participación de `mww`, que aumentó desde 1,3 % en 2000-2009 hasta 83,0 % en 2010-2019 y 97,1 % en 2020-2025. En sentido contrario, `mwc`, que representaba 72,1 % de los registros durante el primer período, disminuyó hasta 0,4 % en el último. Estos cambios respaldan la transición temporal hacia `mww` descrita en el Informe Asesoría 1.

Los procedimientos de reporte no permanecen constantes durante todo el período. En el RMS, $\varepsilon^2 = 0{,}221$ indica que el período se asocia con aproximadamente el 22,1 % de la variabilidad de sus rangos. (*véase anexo de cálculo de epsilon cuadrado*). Para el tipo de magnitud, $V_c = 0{,}608$ se encuentra marcadamente alejado de cero y es coherente con la transición hacia `mww`. Por ello, las comparaciones espaciales pueden interpretarse sin una restricción importante asociada con la fuente o el método, mientras que las comparaciones temporales de magnitud requieren cautela, porque parte de su variación puede reflejar cambios en el procedimiento de estimación y no en la sismicidad.

7.2. Frecuencia de eventos entre zonas

El Informe Asesoría 1 mostró que el Cinturón de Fuego concentra la mayor cantidad de eventos, con tasas descriptivas de 34,3, 7,8, 2,4 y 1,1 eventos por año. Esta subsección evalúa si esa diferencia es estadísticamente defendible y no un rasgo del conteo bruto. Para separar la cantidad de eventos de la extensión de cada zona, se comparan dos medidas: la tasa anual, expresada como eventos por año, y la densidad, expresada como eventos por millón de kilómetros cuadrados y año.

El análisis se realizó con un modelo de conteos. El modelo Poisson inicial presentó sobredispersión, es decir, los conteos variaron más de lo que ese modelo admite. Por ello, las estimaciones se obtuvieron mediante una distribución binomial negativa. En el modelo final, la dispersión residual fue próxima a la esperada y no se detectó dependencia temporal significativa. Aunque cuatro observaciones zona-año superaron el umbral exploratorio de influencia, sus distancias de Cook fueron reducidas y ninguna dominó individualmente el ajuste. El detalle del diagnóstico se presenta fuera del cuerpo. (*véase anexo de diagnósticos de los modelos*).

Una vez comprobada la adecuación del modelo, la primera pregunta fue si la zona modifica la frecuencia. Se respondió con el estadístico de razón de verosimilitudes (LR), que compara el ajuste de un modelo con zonas frente a uno sin zonas: un valor alto, poco probable si las zonas compartieran la misma frecuencia, se traduce en un valor p pequeño. Sus grados de libertad, en la columna `gl`, son tres, uno por cada zona comparada con la referencia.

Modelo	LR	gl	Valor p	Lectura
Tasa anual	220,37	3	$< 0{,}001$	La zona modifica la tasa de ocurrencia.
Densidad por superficie	225,85	3	$< 0{,}001$	La zona modifica la densidad por km^2.

Tabla 3. Contraste global del efecto de zona sobre la frecuencia.

Ambos contrastes resultaron significativos ($p < 0{,}001$), de modo que las cuatro zonas no comparten una misma tasa de ocurrencia ni una misma densidad por superficie. Establecida esa diferencia, la *Tabla 4* cuantifica su magnitud. La razón de tasa mide cuánto difiere cada zona del Cinturón de Fuego, tomado como referencia: un valor de 1 indica igual frecuencia y un valor menor que 1, una frecuencia menor. La razón de densidad cumple la misma función para la densidad por superficie. El intervalo de confianza del 95 % indica con qué precisión se conoce cada razón, es decir, el rango de valores compatibles con los datos.

Zona	Eventos	Tasa anual	Razón de tasa (IC 95 %)	Densidad	Razón de densidad (IC 95 %)
Cinturón de Fuego	892	34,31	1,000 (referencia)	0,720	1,000 (referencia)
Resto del mundo	203	7,81	0,228 [0,191; 0,270]	0,018	0,025 [0,021; 0,030]
Cinturón Alpino-Himalayo	63	2,42	0,071 [0,053; 0,092]	0,167	0,232 [0,176; 0,302]
Dorsal Meso-Atlántica	28	1,08	0,031 [0,021; 0,045]	0,057	0,079 [0,053; 0,114]

Tabla 4. Frecuencia anual bruta y densidad de eventos por zona.

El Cinturón de Fuego presentó la mayor frecuencia en ambas medidas. Frente a él, las razones de tasa de las demás zonas son 0,228 (Resto del mundo), 0,071 (Cinturón Alpino-Himalayo) y 0,031 (Dorsal Meso-Atlántica), todas menores

que 1: el Cinturón de Fuego las supera 4,39, 14,16 y 31,86 veces. Por densidad, el orden se altera: Resto del mundo cae del segundo al último lugar, con razón 0,025, y el Cinturón Alpino-Himalayo pasa al segundo, con 0,232, porque abarca casi toda la superficie considerada. Los seis intervalos de confianza excluyen el 1, de modo que ninguna zona iguala al Cinturón de Fuego. El catálogo es desigual: 892 eventos en el Cinturón de Fuego y 28 en la Dorsal Meso-Atlántica.

La *Tabla 5* detalla las seis comparaciones por pares en ambas métricas. La razón compara la primera zona del par con la segunda: un valor mayor que 1 indica que la primera presenta mayor frecuencia. Como se compararon seis pares a la vez, cada valor p se ajustó mediante el procedimiento de Holm, que mantiene en 5 % la probabilidad de declarar alguna diferencia falsa dentro de la familia. Un valor p ajustado bajo indica que la diferencia entre dos zonas difícilmente se debe al azar, pero no dice cuán grande es: eso lo indican la razón y su intervalo de confianza.

Comparación (zona 1 / zona 2)	Razón de tasa (IC 95 %)	Razón de densidad (IC 95 %)	Lectura
C. Fuego / C. Alp.-Him.	14,16 [10,82; 18,53]	4,30 [3,29; 5,63]	Fuego mayor en ambas.
C. Fuego / Dorsal M.-Atl.	31,86 [21,67; 46,83]	12,64 [8,60; 18,58]	Fuego mayor en ambas.
C. Fuego / Resto	4,39 [3,69; 5,23]	39,57 [33,26; 47,09]	Fuego mayor en ambas.
C. Alp.-Him. / Dorsal M.-Atl.	2,25 [1,43; 3,54]	2,94 [1,87; 4,62]	Alpino mayor en ambas.
C. Alp.-Him. / Resto	0,31 [0,23; 0,42]	9,20 [6,85; 12,35]	Resto mayor en tasa, menor en densidad.
Dorsal M.-Atl. / Resto	0,14 [0,09; 0,21]	3,13 [2,09; 4,69]	Resto mayor en tasa, menor en densidad.

Tabla 5. Comparaciones por pares entre zonas en tasa anual y densidad.

El hallazgo principal es que el orden de las zonas depende de la medida utilizada, la tasa anual o la densidad por superficie, y el cambio se concentra en Resto del mundo: presenta más eventos por año que el Cinturón Alpino-Himalayo y la Dorsal Meso-Atlántica, pero una densidad menor que ambos. Las demás comparaciones son consistentes en ambas medidas: el Cinturón de Fuego supera a las tres zonas y el Cinturón Alpino-Himalayo supera a la Dorsal.

Las seis comparaciones se mantuvieron significativas tras el ajuste de Holm en ambas métricas. Todos los valores p ajustados quedaron por debajo de 0,001, con un máximo de 0,00045. Los valores exactos, junto con el estadístico de cada contraste, se presentan fuera del cuerpo. (*véase anexo de comparaciones por pares de frecuencia*).

La *Figura 1* muestra ambas medidas lado a lado y hace visible el reordenamiento de Resto del mundo.

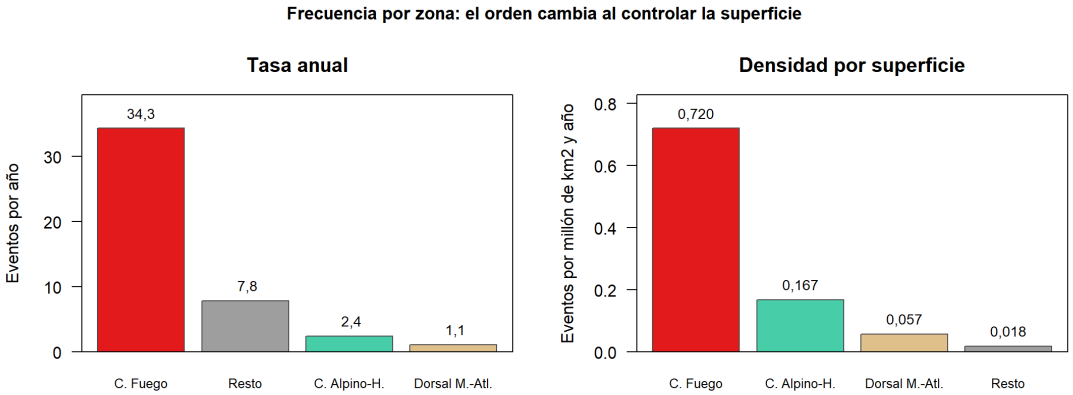

Figura 1. Frecuencia de eventos por zona: tasa anual y densidad por superficie.

Para los terremotos de magnitud $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025, existe evidencia de que la frecuencia de ocurrencia difiere entre las zonas adoptadas. El Cinturón de Fuego presenta la mayor frecuencia tanto por año como por unidad de superficie. La posición de Resto del mundo depende de la medida utilizada: es segundo por conteo bruto, pero último por densidad. Estas razones resumen cada zona con una tasa única para 2000-2025; la subsección de estructura temporal de la ocurrencia evalúa si esa agregación es defendible, es decir, si la tasa se mantuvo estable entre años, tramos del período y meses. Esta conclusión se limita al catálogo, el período, el umbral de magnitud y la delimitación espacial analizados; no constituye una comparación de amenaza sísmica ni una explicación causal.

7.3. Magnitud y profundidad entre zonas

El Informe Asesoría 1 mostró que las zonas presentan magnitudes similares, mientras que la profundidad exhibe diferencias más claras en su distribución. Esta subsección formaliza ambos hallazgos y evalúa si la magnitud y la profundidad permiten distinguir estadísticamente a las zonas sísmicas. Antes de realizar las comparaciones se examinan la normalidad conjunta de ambas variables y la homogeneidad de sus matrices de covarianza. Esta evaluación funciona como una compuerta metodológica: si los supuestos se sostienen, sería posible utilizar una comparación multivariante paramétrica; si se incumplen, corresponde adoptar una ruta no paramétrica basada en rangos. De esta forma, el procedimiento inferencial se selecciona a partir de las características observadas en los datos y no únicamente por preferencia metodológica.

7.3.1. Diagnóstico gráfico de la magnitud

El primer diagnóstico revisó si la magnitud de cada zona se aproxima a una distribución normal. En un QQ-plot, una normalidad razonable se observa cuando los puntos siguen la línea punteada sin curvaturas sistemáticas.

Los QQ-plots de magnitud muestran desviaciones sistemáticas en las cuatro zonas. En C. Fuego y Resto se observa acumulación en $M = 6{,}5$ y curvatura ascendente en la cola superior, patrón coherente con el truncamiento del catálogo, los valores repetidos y la presencia de eventos de mayor magnitud.
En C. Alp.-Him. y Dorsal M.-Atl. la desviación es menos extrema, pero persisten empates y separaciones en los extremos.

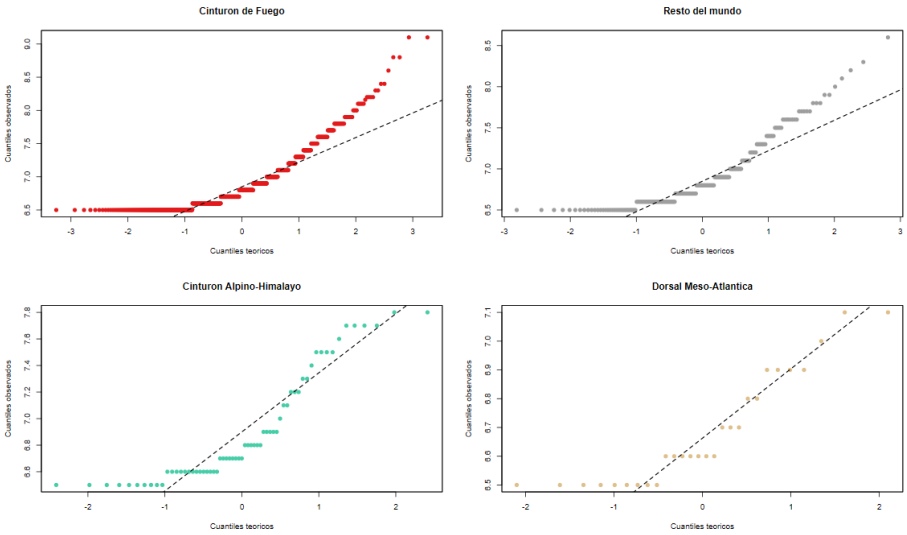

Figura 2. QQ-plots de magnitud por zona sísmica.

Por tanto, la magnitud no muestra un comportamiento compatible con normalidad dentro de las zonas.

7.3.2. Diagnóstico gráfico de la profundidad

El segundo diagnóstico evaluó la profundidad. Esta variable presenta mayor rango y heterogeneidad entre zonas, por lo que el QQ-plot permite revisar si su distribución mantiene una forma aproximadamente normal.

Los QQ-plots de profundidad presentan desviaciones más marcadas que los de magnitud. En C. Fuego y Resto aparece una fuerte curvatura ascendente, asociada con eventos intermedios y profundos que generan colas derechas extensas.
En C. Alp.-Him. también se observa cola derecha, aunque menos extrema. En Dorsal M.-Atl. predomina la concentración en torno a 10 km, lo que produce una línea de referencia casi horizontal y revela escasa variabilidad, empates y algunos valores superiores aislados.

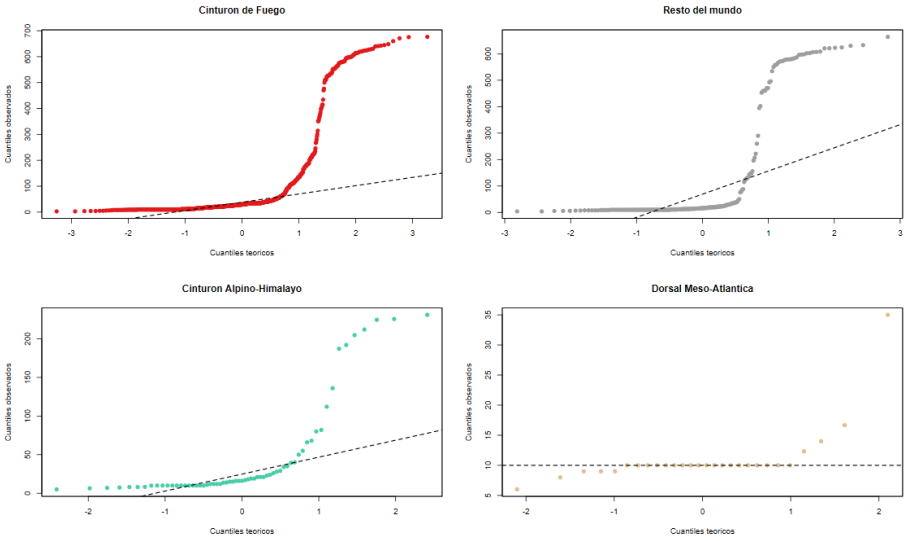

Figura 3. QQ-plots de profundidad por zona sísmica.

En conjunto, la profundidad no presenta normalidad gráfica en ninguna zona. Este resultado refuerza la necesidad de evitar una comparación paramétrica multivariada.

7.3.3. Diagnóstico gráfico multivariante

El diagnóstico multivariante se realizó con distancias de Mahalanobis. Esta distancia mide qué tan alejado está cada evento del centro conjunto de magnitud y profundidad, considerando la dispersión y covariación de ambas variables. Bajo normalidad multivariada, las distancias ordenadas deberían aproximarse a los cuantiles teóricos de una distribución chi-cuadrado con dos grados de libertad.

Los QQ-plots multivariantes también se separan de la línea de referencia. En C. Fuego la desviación es marcada en los cuantiles superiores; Resto se ajusta parcialmente en la zona central, pero se separa en la cola; C. Alp.-Him. muestra desviaciones crecientes hacia valores altos; y Dorsal M.-Atl. queda dominada por baja variabilidad y pocos eventos alejados.

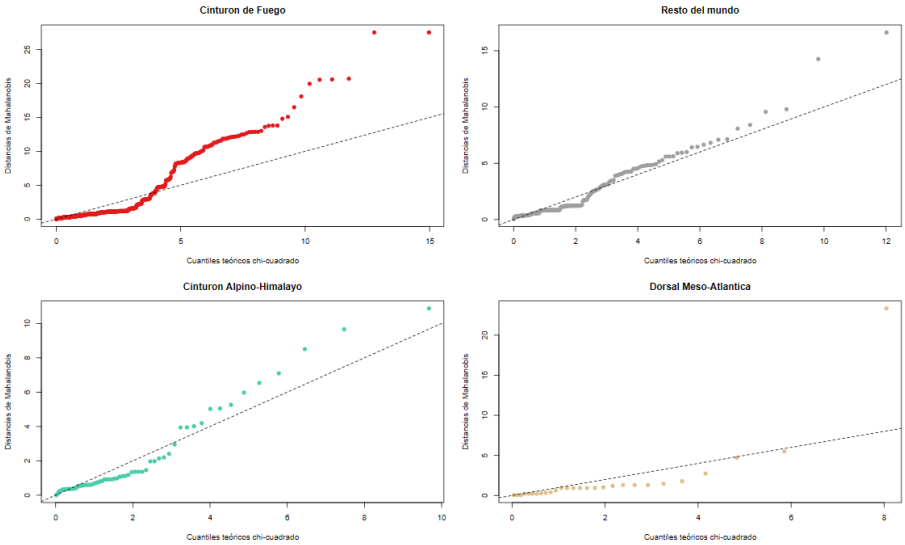

Figura 4. QQ-plots multivariantes de Mahalanobis por zona sísmica.

Por tanto, el diagnóstico multivariante tampoco respalda la normalidad conjunta de magnitud y profundidad, resultado coherente con los QQ-plots univariados.

7.3.4. Pruebas formales de normalidad

Los diagnósticos gráficos se complementaron con pruebas formales de normalidad. Mardia evalúa la normalidad conjunta de magnitud y profundidad; Shapiro-Wilk evalúa la normalidad de cada variable por separado. La *tabla de normalidad* resume ambos resultados por zona.

Zona	Mardia asim.	p	Mardia curt.	p	SW mag.	p	SW prof.	p
Cinturón de Fuego	1.356,483	$< 0{,}001$	30,724	$< 0{,}001$	0,835	$< 0{,}001$	0,537	$< 0{,}001$
Resto del mundo	130,684	$< 0{,}001$	3,435	$< 0{,}001$	0,854	$< 0{,}001$	0,602	$< 0{,}001$
Cinturón Alpino-Himalayo	49,331	$< 0{,}001$	2,059	0,039	0,840	$< 0{,}001$	0,617	$< 0{,}001$
Dorsal Meso-Atlántica	86,180	$< 0{,}001$	10,549	$< 0{,}001$	0,849	$< 0{,}001$	0,431	$< 0{,}001$

Tabla 6. Diagnóstico de normalidad univariada y multivariada de magnitud y profundidad por zona.

Mardia rechaza la normalidad multivariada en las cuatro zonas. La única excepción parcial es la curtosis de C. Alp.-Him. ($p = 0{,}039$), que igualmente rechaza normalidad con $\alpha = 0{,}05$. Shapiro-Wilk también rechaza la normalidad de magnitud y profundidad en todas las zonas ($p < 0{,}001$). Por tanto, la evidencia formal confirma lo observado en los QQ-plots.

7.3.5. Homogeneidad de covarianzas y relación entre variables

Después de evaluar normalidad, se revisó la homogeneidad de matrices de varianza-covarianza mediante M de Box. Bartlett y Spearman se usaron solo como diagnósticos complementarios de la relación entre magnitud y profundidad.

Diagnóstico	Estadístico	Valor p
M de Box	$\chi^2(9) = 286{,}503$	$< 0{,}001$
Bartlett	$\chi^2(1) = 0{,}265$	0,607
Spearman	$\rho = 0{,}130$	$< 0{,}001$

Tabla 7. Diagnósticos complementarios para magnitud y profundidad.

La M de Box rechazó la igualdad de matrices de varianza-covarianza entre zonas ($p < 0{,}001$). Aunque esta prueba es sensible al incumplimiento de normalidad y al desbalance muestral, el resultado es coherente con la heterogeneidad descriptiva observada en profundidad y dispersión. Por tanto, MANOVA no se justifica como procedimiento principal.

Bartlett no detectó asociación lineal relevante entre magnitud y profundidad ($p = 0{,}607$). Spearman sí detectó una asociación monotónica positiva, pero débil ($\rho = 0{,}130$). Así, las variables no parecen redundantes y pueden aportar información distinta en análisis posteriores.

Con base en estos resultados, la comparación formal entre zonas continúa mediante procedimientos no paramétricos basados en rangos.

7.3.6. Contrastes de distribución

Dado que los supuestos de normalidad multivariada y homogeneidad de covarianzas no se sostienen, la comparación global entre zonas se realizó mediante Kruskal-Wallis. Esta prueba compara los rangos de la variable entre grupos, no sus medias.

Variable	H	gl	Valor p	ε^2	Lectura del efecto
Magnitud	8,324	3	0,0398	0,007	Prácticamente nulo
Profundidad	69,498	3	$< 0{,}001$	0,059	Pequeño

Tabla 8. Comparación global de magnitud y profundidad entre zonas mediante Kruskal-Wallis y tamaño de efecto.

Por ello, es adecuada para distribuciones asimétricas, con empates y tamaños muestrales desiguales. Se aplicó por separado a magnitud y profundidad. Kruskal-Wallis detectó diferencias globales entre zonas tanto en magnitud como en profundidad. Sin embargo, los tamaños de efecto muestran que estas diferencias no tienen la misma relevancia práctica. En magnitud, la evidencia es estadísticamente detectable ($p = 0{,}0398$), pero el efecto global es prácticamente nulo ($\varepsilon^2 = 0{,}007$). En profundidad, la evidencia es más clara ($p < 0{,}001$) y el efecto global es mayor, aunque todavía pequeño ($\varepsilon^2 = 0{,}059$). Este patrón coincide con el Informe Asesoría 1: las magnitudes eran similares entre zonas, mientras que la profundidad mostraba diferencias más marcadas.

Como la prueba global no identifica qué pares de zonas difieren, el análisis continúa con comparaciones post-hoc de Dunn-Holm.

Variable	Comparación relevante	Z	p_{aj}	Lectura
Magnitud	C. Fuego - Dorsal M.-Atl.	2,713	0,033	Dorsal con rangos menores
Magnitud	Resto - Dorsal M.-Atl.	2,877	0,024	Dorsal con rangos menores
Profundidad	C. Fuego - Resto	3,965	$< 0{,}001$	C. Fuego con rangos mayores
Profundidad	C. Fuego - C. Alp.-Him.	3,552	$< 0{,}001$	C. Fuego con rangos mayores
Profundidad	C. Fuego - Dorsal M.-Atl.	7,020	$< 0{,}001$	C. Fuego con rangos mayores
Profundidad	Resto - Dorsal M.-Atl.	5,154	$< 0{,}001$	Dorsal con rangos menores
Profundidad	C. Alp.-Him. - Dorsal M.-Atl.	3,893	$< 0{,}001$	Dorsal con rangos menores

Tabla 9. Comparaciones post-hoc relevantes mediante Dunn-Holm.

En magnitud, las únicas diferencias significativas involucran a Dorsal M.-Atl. frente a C. Fuego y Resto. Las comparaciones entre C. Fuego, Resto y C. Alp.-Him. no presentan diferencias, y C. Alp.-Him. frente a Dorsal M.-Atl. queda bajo el umbral convencional ($p_{aj} = 0{,}070$). En profundidad, Dorsal M.-Atl. difiere de todas las demás zonas, y C. Fuego también difiere de Resto y C. Alp.-Him. La única comparación no significativa es Resto frente a C. Alp.-Him. ($p_{aj} = 0{,}283$). La tabla completa se presenta en el *anexo de comparaciones Dunn-Holm*.

Para complementar las comparaciones post-hoc, se calculó delta de Cliff en los pares principales definidos para contrastar C. Fuego con zonas de referencia. Este tamaño de efecto no paramétrico mide la dirección y magnitud práctica de la separación entre dos distribuciones.

Variable	Comparación	δ	IC 95 %	Lectura
Magnitud	C. Fuego - Dorsal M.-Atl.	0,298	[0,114; 0,463]	Pequeña
Magnitud	C. Fuego - C. Alp.-Him.	-0,011	[-0,159; 0,137]	Despreciable
Profundidad	C. Fuego - Dorsal M.-Atl.	0,783	[0,679; 0,856]	Grande
Profundidad	C. Fuego - C. Alp.-Him.	0,269	[0,109; 0,416]	Pequeña

Tabla 10. Tamaños de efecto por pares mediante delta de Cliff.

Valores positivos indican que C. Fuego tiende a presentar valores mayores que la zona comparada; valores cercanos a cero indican escasa separación. En magnitud, C. Fuego presenta valores algo mayores que Dorsal M.-Atl., pero el efecto es pequeño ($\delta = 0{,}298$). Frente a C. Alp.-Him., la diferencia es despreciable ($\delta = -0{,}011$), coherente con la ausencia de diferencia post-hoc. En profundidad, la separación entre C. Fuego y Dorsal M.-Atl. es grande ($\delta = 0{,}783$), mientras que frente a C. Alp.-Him. es pequeña ($\delta = 0{,}269$).

En conjunto, la magnitud separa poco a las zonas, mientras que la profundidad muestra una capacidad de diferenciación más clara, especialmente entre C. Fuego y Dorsal M.-Atl. Esta conclusión es consistente con el análisis descriptivo del Informe Asesoría 1 y con los contrastes globales de esta fase. Ninguna de las dos variables discrimina por sí sola a las cuatro zonas; queda abierta la posibilidad de que ganen capacidad al actuar en conjunto, hipótesis que se retoma en la clasificación de zonas.

7.4. Estructura temporal de la ocurrencia

El Informe Asesoría 1 mostró que, para el catálogo USGS 2000-2025 de eventos $M \geq 6{,}5$, la serie presenta oscilaciones alrededor de la media, sin una tendencia sostenida claramente identificable a partir del gráfico. La subsección de frecuencia respondió dónde se concentra la ocurrencia: las tasas difieren entre zonas. Esta subsección responde la pregunta restante: ¿la tasa del catálogo completo se mantuvo estable durante 2000-2025? Para ello los conteos se agregan sobre las cuatro zonas, porque la pregunta ya no distingue territorios sino momentos, y se examina en cuatro pasos: estabilidad del nivel, dependencia entre períodos consecutivos, cambios entre tramos y meses, y ritmo de los tiempos entre eventos. La *Figura 5* retoma la lectura del Informe Asesoría 1 en escala mensual.

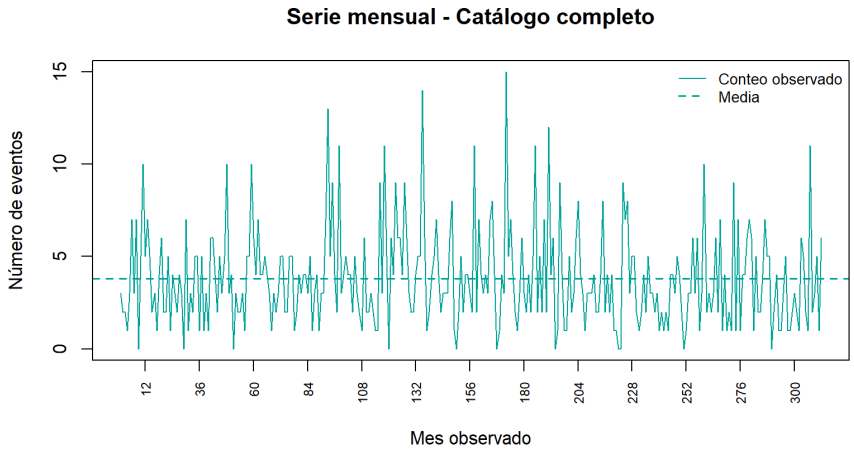

Figura 5. Serie mensual de eventos del catálogo completo.

Para complementar la lectura visual, se aplicaron pruebas de raíz unitaria, estacionariedad e independencia serial. En este contexto, evaluar estacionariedad significa revisar si los conteos fluctúan alrededor de un nivel relativamente estable durante el período del catálogo, o si existen señales de un cambio persistente en la ocurrencia. La *Tabla 11* resume los resultados.

Escala	Prueba	Rezagos	Estadístico	Valor p
Anual	ADF	1	-0,466	0,460
Anual	KPSS	2	0,250	$> 0{,}10$
Anual	Ljung-Box	5	3,815	0,576
Mensual	ADF	12	-0,969	0,308
Mensual	KPSS	5	0,270	$> 0{,}10$
Mensual	Ljung-Box	12	16,883	0,154

Tabla 11. Resultados de estacionariedad e independencia serial para las series anual y mensual de conteos.

En ambas escalas, KPSS no entrega evidencia de que los conteos se alejen de un nivel estable. ADF, en cambio, no permite descartar formalmente un comportamiento no estacionario al 5 %. En conjunto, los resultados no muestran una ruptura clara de estabilidad temporal en el período 2000-2025, pero la conclusión debe leerse con cautela porque las dos pruebas no entregan una señal completamente concordante.

La independencia serial se revisó de forma conjunta mediante Ljung-Box y el gráfico ACF/PACF mensual. Esto permite contrastar formalmente la presencia de autocorrelación y, al mismo tiempo, observar si existen rezagos específicos con señales relevantes.

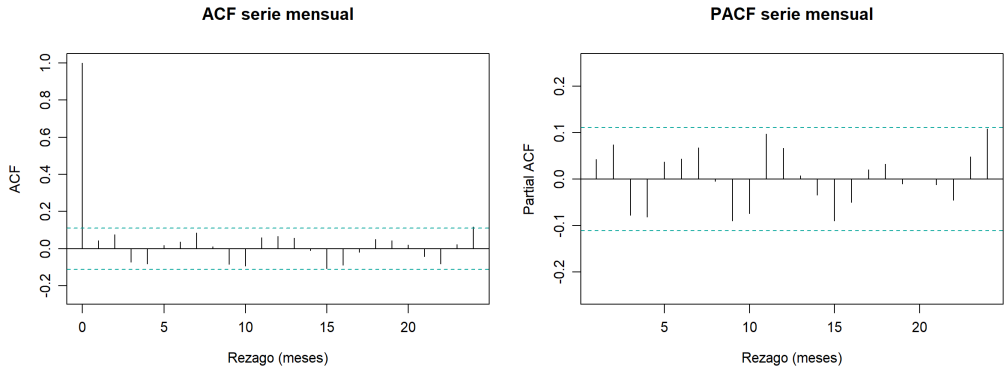

Figura 6. ACF y PACF de la serie mensual de conteos del catálogo.

En conjunto, Ljung-Box no detectó autocorrelación relevante en la serie anual ($p = 0{,}58$) ni en la mensual ($p = 0{,}15$ a rezago 12). La ACF y la PACF mensual refuerzan esta lectura, ya que no muestran un patrón persistente entre rezagos. En términos del catálogo, esto sugiere que un año o mes con más eventos no queda seguido sistemáticamente por otro período también alto, ni uno bajo por otro igualmente bajo.

Una vez revisada la estabilidad general de la serie, se evaluó si el nivel de ocurrencia anual cambió entre tramos del catálogo. Esta pregunta no se resuelve solo con una comparación descriptiva, porque dos períodos pueden tener promedios distintos y aun así diferir únicamente por fluctuaciones propias del registro. Se utilizó un modelo Poisson porque la variable analizada corresponde a un conteo de eventos por año; este tipo de modelo se emplea habitualmente para contrastar tasas de ocurrencia en intervalos definidos de tiempo. Así, permite evaluar si las diferencias descriptivas entre períodos reflejan un cambio formal en la tasa anual de eventos o solo variaciones propias del catálogo. Luego se revisó si la variabilidad de los datos era compatible con ese modelo o si requería la corrección quasi-Poisson.

Período	Media anual	Mín.	Máx.
2000-2009	45,5	36	58
2010-2019	48,9	33	61
2020-2025	40,3	28	53

Tabla 12. Conteos anuales por período del catálogo.

Evaluación	Estadístico	Valor	Valor p
Poisson inicial	LRT	6,13	0,047
Dispersión	ϕ	1,57	-
Quasi-Poisson	F	1,91	0,170

Tabla 13. Contraste del efecto período en los conteos anuales.

Las medias muestran un aumento leve entre 2000-2009 y 2010-2019, seguido de una disminución en 2020-2025. El modelo Poisson inicial sugería diferencias entre períodos, pero el diagnóstico de dispersión indicó que los conteos anuales variaban más de lo esperado bajo ese modelo. Por ello, la decisión se basó en quasi-Poisson. Al considerar esta mayor variabilidad entre años, la diferencia entre períodos ya no resulta suficientemente clara como para afirmar que la tasa anual de eventos haya cambiado. En otras palabras, el catálogo muestra años con más y menos eventos, pero no entrega evidencia suficiente para concluir que un período tenga una tasa anual de ocurrencia distinta a los demás.

De forma complementaria, se evaluó si la ocurrencia mensual presentaba una señal de estacionalidad, es decir, si ciertos meses del calendario concentraban sistemáticamente más eventos que otros.

Evaluación	Estadístico	Valor	Valor p
Poisson inicial	LRT	19,65	0,050
Dispersión	ϕ	1,79	-
Quasi-Poisson	F	1,03	0,416

Tabla 14. Contraste de estacionalidad mensual en los conteos del catálogo.

El contraste Poisson inicial fue limítrofe, pero los conteos mensuales variaron más de lo esperado bajo ese modelo. Al corregir esta variabilidad mediante quasi-Poisson, no se observa evidencia suficiente para afirmar que la tasa de ocurrencia difiera entre meses calendario. En términos del catálogo, los eventos no muestran un patrón mensual claro; las diferencias entre meses pueden explicarse por fluctuaciones propias del registro observado.

Como cierre de la revisión temporal, se retomaron los tiempos transcurridos entre eventos consecutivos descritos en el Informe Asesoría 1.
Esta mirada complementa el análisis de conteos: mientras los modelos Poisson evalúan cuántos eventos ocurren en un intervalo fijo, el ajuste exponencial evalúa cuánto tiempo pasa entre un evento y el siguiente. Ambas miradas están conectadas: si la ocurrencia temporal fuera compatible con una tasa constante, entonces los conteos por intervalo se describen mediante un modelo Poisson y los tiempos de espera entre eventos mediante una distribución exponencial. Por ello, se contrastó si los intervalos observados eran compatibles con una distribución exponencial usando Kolmogorov-Smirnov con bootstrap paramétrico.

Grupo	n	Tasa diaria	Mediana días	D	Valor p
$M \geq 7{,}0$	386	0,0408	15,6	0,0778	0,001
Fuerte	798	0,0841	7,3	0,0636	$< 0{,}001$
Mayor	327	0,0345	19,6	0,0543	0,102
Grande o extremo	58	0,0063	122,8	0,0884	0,531

Tabla 15. Ajuste exponencial de los tiempos entre eventos consecutivos.

Los resultados muestran que el grupo combinado $M \geq 7{,}0$ no es compatible con un único ajuste exponencial. Sin embargo, al separar por categoría de magnitud, los eventos "Mayor" y "Grande o extremo" sí resultan compatibles con este comportamiento, mientras que los eventos "Fuerte" lo rechazan. Esto sugiere que la mezcla de eventos con ritmos de ocurrencia distintos puede ocultar patrones más específicos. Los eventos de mayor magnitud son menos frecuentes y presentan intervalos más largos, por lo que no deberían interpretarse junto con eventos más habituales como si compartieran una misma tasa temporal.

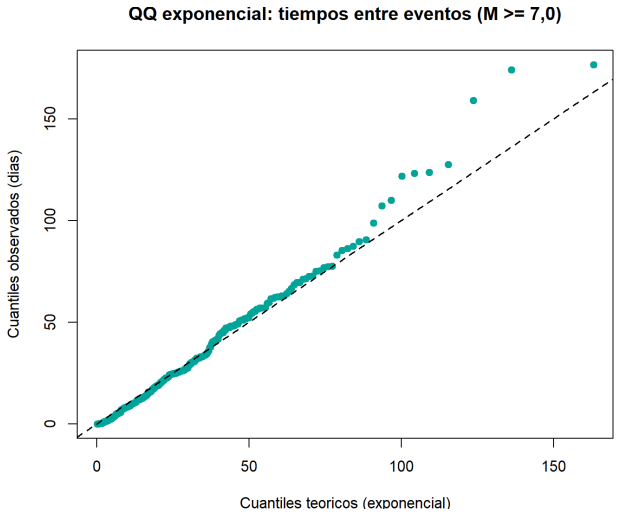

Figura 7. QQ exponencial de los tiempos entre eventos para $M \geq 7{,}0$.

Con este resultado quedan cubiertas las dos caras de la ocurrencia del catálogo, dónde y cuándo. La pregunta espacial tuvo respuesta afirmativa, ya que las tasas difieren entre zonas. La pregunta temporal resultó compatible con la estabilidad, sin evidencia de cambios de nivel, autocorrelación ni estacionalidad durante 2000-2025. La lectura temporal no es una pregunta adicional del informe, sino el respaldo de la primera componente de la pregunta estadística. En concreto, confirma que una tasa anual única por zona es un resumen adecuado del período.

7.5. Eventos fuertes y extremos

El Informe Asesoría 1 definió las categorías Fuerte, Mayor y Grande o extremo y describió su composición por zona con alcance exploratorio. Esta subsección ejecuta la transición inferencial propuesta allí. Primero contrasta si la composición de categorías depende de la zona, luego evalúa mediante regresión logística si la zona y la profundidad explican la probabilidad de que un evento alcance $M \geq 7{,}0$ y cierra con la descripción de los eventos con $M \geq 7{,}8$.

La *Tabla 16* presenta los conteos observados y la composición porcentual dentro de cada zona. Las tres primeras zonas presentan composiciones semejantes. Sus porcentajes por categoría no difieren en más de 2,3 puntos entre sí (Fuerte entre 66,0 % y 68,3 %, Mayor entre 27,8 % y 29,1 %, Grande o extremo entre 3,2 % y 5,3 %).

Zona	Fuerte	Mayor	Grande o extremo	Total
C. de Fuego	597 (66,9 %)	248 (27,8 %)	47 (5,3 %)	892
R. del mundo	134 (66,0 %)	59 (29,1 %)	10 (4,9 %)	203
C. Alpino-Himalayo	43 (68,3 %)	18 (28,6 %)	2 (3,2 %)	63
D. Meso-Atlántica	25 (89,3 %)	3 (10,7 %)	0 (0,0 %)	28
Total	799	328	59	1.186

Tabla 16. Composición de categorías de magnitud por zona: conteos y porcentajes por fila.

La Dorsal Meso-Atlántica se aparta de ese patrón, ya que concentra 89,3 % de sus eventos en la categoría Fuerte y no registra ninguno Grande o extremo, aunque reúne solo 28 eventos. El contraste chi-cuadrado produjo $X^2 = 7{,}122$ con valor p simulado de 0,307 (10.000 réplicas), por lo que no se rechaza la independencia entre zona y categoría de magnitud. Dos de las doce frecuencias esperadas quedaron bajo 5 (3,13 y 1,39, ambas en la columna Grande o extremo), lo que confirma la pertinencia de la simulación; la matriz completa se detalla en el anexo (*véase la Tabla 30*). La V de Cramér corregida fue 0,022, un grado de asociación prácticamente nulo (*véase anexo de corrección por sesgo de la V de Cramér*). La lectura conjunta es consistente. El catálogo no entrega evidencia de que la composición de severidad difiera entre zonas, y la diferencia descriptiva de la Dorsal no basta para establecer una asociación, en parte por su tamaño muestral reducido. Este resultado no contradice la diferencia global detectada en la magnitud continua (*véase la Tabla 8*). La categorización agrupa valores y pierde detalle, y ambos análisis coinciden en que la relevancia práctica de la magnitud es reducida.

La regresión logística evaluó si la zona y la profundidad, consideradas simultáneamente, explican la probabilidad de alcanzar $M \geq 7{,}0$; es decir, si la probabilidad de que un evento sea de los más fuertes se relaciona con la zona y la profundidad en que ocurre. Para ello, las tres categorías de severidad se redujeron a una distinción de sí o no: alcanzar al menos magnitud 7,0 (categorías Mayor y Grande o extremo) o no alcanzarla (categoría Fuerte). Bajo ese criterio, 387 de los 1.186 eventos superan el umbral: 295 en el Cinturón de Fuego, 69 en el Resto del mundo, 20 en el Cinturón Alpino-Himalayo y 3 en la Dorsal Meso-Atlántica.

Los resultados se leen en dos niveles: primero el contraste global de cada variable, que decide si la zona o la profundidad aportan en conjunto, y luego los contrastes individuales, que detallan qué zona se aparta de la referencia. La *Tabla 17* presenta el primer nivel: el contraste de razón de verosimilitudes por variable.

Variable	LR	gl	Valor p	Lectura
Zona	7,358	3	0,0613	No significativa; queda apenas sobre el umbral
Profundidad	0,312	1	0,5766	No significativa; sin aporte tras controlar por zona

Tabla 17. Contraste de razón de verosimilitudes por variable en la regresión logística.

Ambos valores p superan el criterio de decisión del 5 %, por lo que ninguna de las dos variables muestra un efecto estadísticamente detectable. Con estos datos no se puede afirmar que la probabilidad de alcanzar $M \geq 7{,}0$ cambie según la zona ni según la profundidad. La diferencia entre ambas es de grado. La zona quedó apenas sobre el umbral ($p = 0{,}0613$), con la señal concentrada en la Dorsal Meso-Atlántica, como se verá en el segundo nivel; el criterio se fijó antes de observar los resultados y no se modifica después, de modo que esa cercanía se describe pero no cambia el veredicto. La profundidad quedó lejos del umbral ($p = 0{,}5766$). Tras conocer la zona, la profundidad de un evento no añade información sobre si alcanza $M \geq 7{,}0$.

La *Tabla 18* presenta el segundo nivel, con las razones de odds, sus intervalos y los contrastes específicos de Wald (*véase anexo del contraste de Wald*). En estas razones, un valor de 1 indica igual probabilidad relativa que el Cinturón de Fuego, un valor menor que 1 una probabilidad menor de alcanzar $M \geq 7{,}0$ y uno mayor que 1, una probabilidad mayor (*véase anexo de la razón de odds*).

Término	OR	IC 95 %	p de Wald
Resto del mundo	1,032	[0,744; 1,423]	0,848
Cinturón Alpino-Himalayo	0,950	[0,538; 1,625]	0,855
Dorsal Meso-Atlántica	0,247	[0,058; 0,712]	0,023*
Profundidad (por km)	1,0002	[0,9995; 1,0009]	0,575

* Significativa al 5 %.

Tabla 18. Razones de odds de alcanzar $M \geq 7{,}0$ respecto del Cinturón de Fuego.

La única señal específica corresponde a la Dorsal, la fila marcada con asterisco. Su razón de odds es 0,247: sus odds de alcanzar $M \geq 7{,}0$ equivalen al 24,7 % de las del Cinturón de Fuego, una reducción de 75,3 %. El intervalo [0,058; 0,712] respalda esa diferencia, porque no incluye el 1, el valor que indicaría odds iguales. Pero el intervalo es ancho: la reducción real podría ir desde 28,8 % hasta 94,2 % ($1 - 0{,}712$ y $1 - 0{,}058$, sus dos extremos). Esa imprecisión se debe a que la Dorsal registra solo 3 eventos sobre el umbral. Este contraste individual no sustituye la evaluación global de

zona, que no fue significativa. Resto del mundo y el Cinturón Alpino-Himalayo no se distinguen de la referencia. El efecto de la profundidad es un OR de 1,0002 por kilómetro, con IC 95 % [0,9995; 1,0009]: el valor es prácticamente 1 y el intervalo lo contiene, compatible con ausencia de efecto. La versión con profundidad categórica no modificó estas conclusiones: su bloque tampoco fue significativo ($LR = 2{,}063$; $p = 0{,}357$) y ambos AIC resultaron prácticamente idénticos (1500,01 continua frente a 1500,26 categórica, diferencia de 0,249, menor que 2), por lo que se conservó la especificación continua (*véase anexo del criterio de información de Akaike*). La *Tabla 19* resume los eventos Grandes o extremos.

Entre 2000 y 2025 se registraron 59 eventos con $M \geq 7{,}8$. El Cinturón de Fuego reunió 47, equivalentes a 79,7 % de los extremos, con una tasa de 1,81 por año. Esa concentración es menos llamativa de lo que parece, porque el catálogo está desbalanceado: el Cinturón de Fuego reúne el 75,2 % de todos los eventos analizados, extremos o no.

Zona	Eventos	Porcentaje	Tasa anual
Cinturón de Fuego	47	79,7 %	1,81
Resto del mundo	10	16,9 %	0,385
Cinturón Alpino-Himalayo	2	3,4 %	0,0769
Dorsal Meso-Atlántica	0	0 %	0

Tabla 19. Eventos de magnitud $M \geq 7{,}8$ por zona: conteos, porcentajes y tasas anuales.

Si la severidad no dependiera de la zona, le corresponderían de todos modos cerca de 44,4 de los 59 extremos ($59 \times 0{,}752$); los 47 observados quedan apenas 4,5 puntos porcentuales por encima de esa base (79,7 % frente a 75,2 %). La concentración refleja, sobre todo, que en esa zona ocurre la mayoría de los eventos, en concordancia con la ausencia de asociación del chi-cuadrado aplicado a la *Tabla 16* ($p = 0{,}307$). Las tasas son conteos divididos por los 26 años observados, sin normalizar por superficie. Para este umbral no se estimó ninguna regresión: con una zona sin eventos, el modelo produciría separación (*véase anexo del contraste de Wald*). La ausencia de eventos extremos en la Dorsal se limita al catálogo y período analizados.

Para los terremotos de magnitud $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025, la composición de severidad no difiere de manera estadísticamente detectable entre las zonas adoptadas y la profundidad no explica que un evento supere el umbral $M \geq 7{,}0$. La única señal específica, la menor probabilidad de eventos mayores en la Dorsal, se estima con alta imprecisión y no basta para declarar un efecto global de zona. Estos resultados delimitan la pregunta de la sección siguiente: la profundidad no indica cuán grande es un evento, pero podría ayudar a identificar en qué zona ocurre.

7.6. Clasificación de zonas a partir de magnitud y profundidad

Las subsecciones anteriores establecieron que la magnitud apenas distingue a las zonas y que la profundidad presenta diferencias más claras. Esta subsección evalúa la segunda componente de la pregunta estadística: si ambas variables, consideradas conjuntamente, permiten identificar la zona de un evento. Primero se explora la estructura de las tablas categóricas mediante análisis de correspondencia y luego se ajusta el modelo multinomial. En términos aplicados, el análisis de correspondencia funciona como un mapa de asociación: ubica cerca a las zonas y categorías con perfiles parecidos, y separa aquellas cuya composición es distinta (*véase anexo de análisis de correspondencia*). Los porcentajes que acompañan a cada dimensión en los mapas no corresponden al porcentaje de eventos, sino a la proporción de inercia explicada por cada eje; es decir, indican cuánta parte de la asociación entre zonas y categorías queda resumida horizontal y verticalmente.

El análisis se calcula sobre la composición de categorías dentro de cada zona. La de magnitud se presentó en la *Tabla 16*; la *Tabla 20* presenta la de profundidad.
Las casillas en cero (el Cinturón Alpino-Himalayo y la Dorsal sin eventos profundos, y la Dorsal tampoco con intermedios) anticipan la separación que más adelante obliga a corregir el modelo multinomial.

Zona	Superficial	Intermedio	Profundo	Total
Cinturón de Fuego	688 (77,1 %)	121 (13,6 %)	83 (9,3 %)	892
Resto del mundo	145 (71,4 %)	18 (8,9 %)	40 (19,7 %)	203
Cinturón Alpino-Himalayo	52 (82,5 %)	11 (17,5 %)	0 (0,0 %)	63
Dorsal Meso-Atlántica	28 (100,0 %)	0 (0,0 %)	0 (0,0 %)	28
Total	913	150	123	1.186

Tabla 20. Composición de categorías de profundidad por zona: conteos y porcentajes por fila.

La *Tabla 21* resume la inercia de ambas tablas. La tabla de profundidad presenta una inercia total 5,3 veces mayor que la de magnitud (0,0321 frente a 0,0060), por lo que entrega una estructura más útil para distinguir zonas. En magnitud, aunque la primera dimensión concentra 92,9 % de la inercia, esta proviene de una asociación total muy baja; por tanto, no indica una separación fuerte entre zonas. Su señal se concentra casi únicamente en la Dorsal Meso-Atlántica, donde 89,3 % de los eventos son Fuertes. En profundidad, la primera dimensión concentra 81,9 % de la asociación y separa principalmente la categoría Profundo de las zonas que no registran esos eventos. El Resto del mundo aparece relativamente asociado a

Tabla	Inercia total	Dim. 1	Dim. 2	Punto dominante
Zona x categoría magnitud	0,0060	92,9 %	7,1 %	Dorsal Meso-Atlántica (95,1 % de la dim. 1)
Zona x categoría profundidad	0,0321	81,9 %	18,1 %	Categoría Profundo (87,8 % de la dim. 1)

Tabla 21. Inercia y contribuciones dominantes del análisis de correspondencia.

eventos profundos, mientras que el Cinturón Alpino-Himalayo y la Dorsal no presentan ninguno; la Dorsal, además, es exclusivamente superficial. El Cinturón de Fuego queda cerca del origen, con un perfil próximo al promedio del catálogo. Esta lectura requiere cautela: en ambas tablas, las frecuencias esperadas bajo independencia caen bajo cinco en cuatro casillas de las zonas menores (*véase anexo del contraste chi-cuadrado y frecuencias esperadas*). En profundidad corresponden a las casillas Intermedio y Profundo de la Dorsal Meso-Atlántica, con esperadas de 3,54 y 2,90; en magnitud, a las dos de la columna Grande o extremo (3,13 y 1,39, detalladas en la *Tabla 30*). Por eso, la posición extrema de la Dorsal en los mapas refleja en parte su escaso número de eventos y no una separación robusta por sí sola.

Dicho de forma directa, al mirar la magnitud las zonas se parecen bastante entre sí, salvo la particularidad de la Dorsal; al mirar la profundidad, aparecen perfiles más distinguibles. Esa diferencia se explica principalmente por la presencia o ausencia de eventos profundos y por el carácter exclusivamente superficial de la Dorsal Meso-Atlántica. Por eso el análisis de correspondencia funciona aquí como una primera lectura visual. No concluye por sí mismo, pero muestra que la profundidad contiene más información para separar zonas que la magnitud. La *Figura 8* muestra ambos planos.

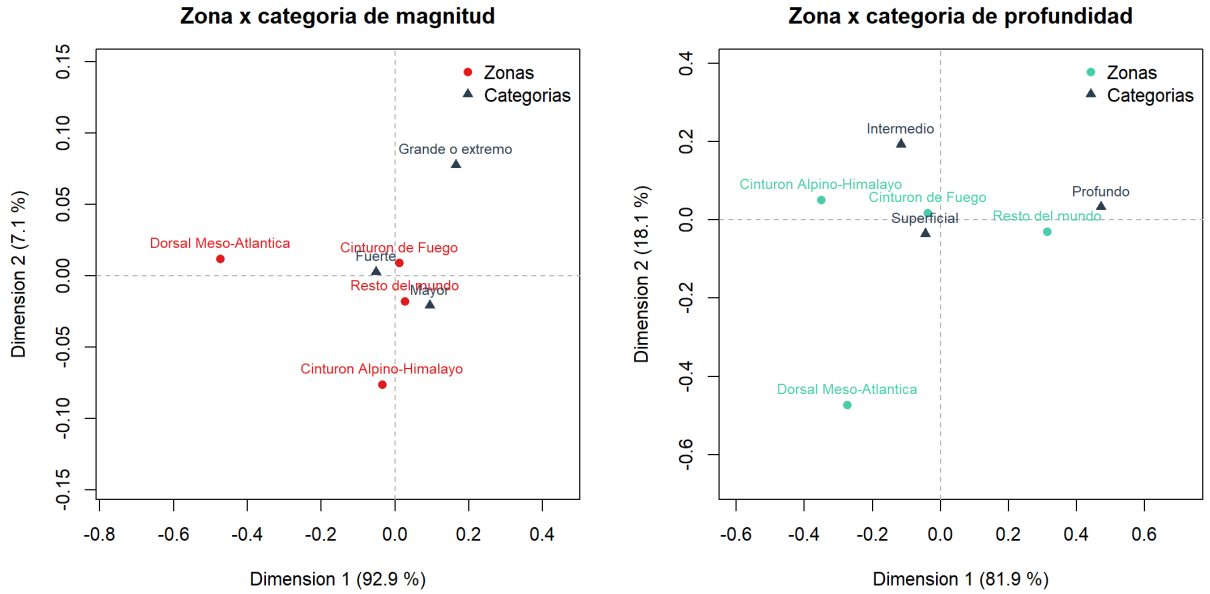

Figura 8. Mapas de correspondencia entre zonas y categorías de magnitud y profundidad.

El modelo multinomial cuantifica ese contraste de manera formal (*véase anexo del modelo de regresión logística multinomial*). La *Tabla 22* presenta la razón de verosimilitudes por variable, calculada sobre el ajuste de máxima verosimilitud; la corrección de sesgo se aplica solo a los coeficientes reportados más adelante.

Variable	LR	gl	Valor p	Lectura
Magnitud	5,102	3	0,1645	No mejora la distinción entre zonas
Profundidad	64,906	3	$< 0{,}001$	Aporta aun controlando por magnitud

Tabla 22. Contraste de razón de verosimilitudes por variable en el modelo multinomial.

La profundidad aporta información global para distinguir las zonas incluso después de controlar por magnitud; la magnitud no. Este resultado ilustra cómo una misma variable puede cambiar de relevancia según la pregunta del modelo: en la subsección anterior la profundidad no añadió información sobre si un evento alcanzaba $M \geq 7{,}0$; aquí sí distingue la zona del evento. No hay contradicción: son variables respuesta y preguntas distintas.

Establecido qué variables distinguen las zonas, el paso siguiente es cuantificar el tamaño de su efecto. Con ese fin, al estandarizar los predictores, la ecuación de la Dorsal produjo un coeficiente de profundidad de $-27{,}66$ con error estándar 7,21, un patrón de separación cuasi perfecta: como sus 28 eventos son superficiales, la profundidad la distingue de manera casi determinista y la estimación de máxima verosimilitud diverge. Los coeficientes finales se obtuvieron con la corrección de reducción de sesgo declarada en la metodología (*véase anexo de separación y reducción de sesgo*).

La *Tabla 23* presenta las razones de odds del modelo corregido; cada una corresponde a un aumento de una desviación estándar del predictor (0,417 unidades de magnitud o 164,1 km de profundidad).

Zona frente a C. de Fuego	Magnitud: OR (IC 95 %)	Profundidad: OR (IC 95 %)
Resto del mundo	1,05 [0,91; 1,22]	1,25 [1,10; 1,43]
Cinturón Alpino-Himalayo	1,04 [0,81; 1,33]	0,65 [0,41; 1,01]
Dorsal Meso-Atlántica	0,58 [0,31; 1,07]	≈ 0 (separación)

Tabla 23. Razones de odds del modelo corregido por un aumento de una desviación estándar.

El único intervalo que excluye el 1 es el de la profundidad para el Resto del mundo. Una desviación estándar adicional, unos 164 km, aumenta 25,2 % las odds de pertenecer a esa zona frente al Cinturón de Fuego, coherente con su mayor

proporción de eventos profundos. El intervalo del Cinturón Alpino-Himalayo queda en el límite, compatible con menores odds a mayor profundidad. La razón de la Dorsal en profundidad se muestra como aproximadamente cero, pero no se interpreta como estimación precisa, ya que refleja la separación producida por su perfil exclusivamente superficial. La magnitud no presenta intervalos que excluyan el 1 en ninguna comparación.

La *Tabla 24* resume la capacidad clasificatoria del modelo corregido bajo la regla de máxima probabilidad (*véase anexo de métricas de clasificación*). El modelo asignó los 1.186 eventos al Cinturón de Fuego. Reconoció sus 892 eventos y ningún evento de las tres zonas restantes.

Métrica	Valor	Lectura
Exactitud global	0,7521	Igual a la base trivial
Base trivial	0,7521	Clasificar siempre Cinturón de Fuego
Exactitud balanceada	0,25	Reconoce solo una de las cuatro zonas
Sensibilidad por zona	1 / 0 / 0 / 0	Fuego / Resto / Alpino / Dorsal

Tabla 24. Métricas de clasificación del modelo corregido.

El resultado no contradice la significación de la profundidad, porque la prueba global evalúa si las probabilidades relativas cambian, mientras que la clasificación exige que alguna zona supere al Cinturón de Fuego como categoría más probable. El Cinturón de Fuego parte como zona más probable porque reúne el 75,21 % del catálogo, y la magnitud y la profundidad, al distribuirse de forma similar entre zonas, no aportan evidencia suficiente para revertir esa ventaja: su probabilidad ajustada nunca baja de 0,584 y ninguna zona alternativa supera 0,353.

Para no limitar la evaluación a esa decisión final, se revisaron las curvas ROC uno contra el resto (*véase anexo de curvas ROC y área bajo la curva*). Esta herramienta transforma el problema multinomial en cuatro comparaciones binarias, cada zona frente a todas las demás. Técnicamente, el área bajo la curva ROC (AUC) mide si las probabilidades ajustadas ordenan correctamente los eventos de una zona por encima de los que no pertenecen a ella. Un valor cercano a 0,5 indica discriminación similar al azar, mientras que valores más altos indican mejor separación. En términos no técnicos, la matriz de confusión pregunta "¿qué etiqueta final asignó el modelo?", mientras que la curva ROC pregunta "¿el modelo al menos elevó la probabilidad de la zona correcta, aunque no haya sido la etiqueta ganadora?". La *Figura 9* muestra esta evaluación complementaria.

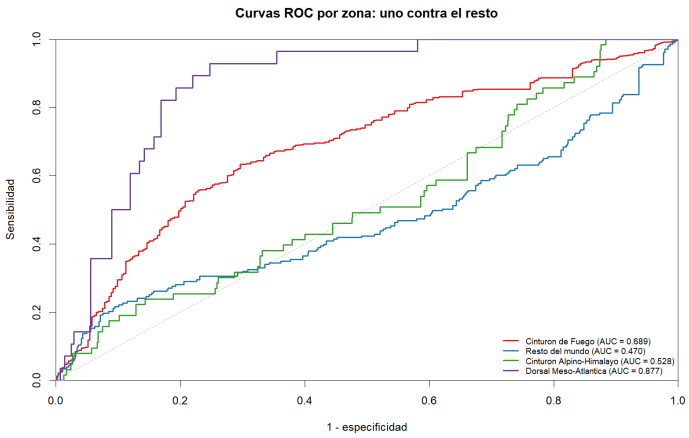

Figura 9. Curvas ROC por zona bajo comparación uno contra el resto.

La señal discriminante no es homogénea entre zonas. La Dorsal Meso-Atlántica alcanza el mayor AUC (0,877), compatible con una buena separación de sus eventos frente al resto del catálogo; esto es coherente con su perfil físico, exclusivamente superficial y de magnitud más acotada. Sin embargo, esta lectura debe tomarse con cautela, porque la clase positiva contiene solo 28 eventos. Por tanto, el AUC sugiere una señal discriminante clara asociada a su perfil superficial, pero no constituye evidencia suficiente de una clasificación estable para esa zona. El Cinturón de Fuego presenta discriminación moderada (AUC = 0,689), porque cubre buena parte del rango conjunto de magnitud y profundidad. El Cinturón Alpino-Himalayo queda cercano al azar (AUC = 0,528) y el Resto del mundo no muestra capacidad útil de discriminación (AUC = 0,470), consistente con su carácter residual y heterogéneo. En conjunto, ROC matiza, pero no revierte la matriz de confusión. El modelo contiene señal para reconocer parcialmente el perfil de la Dorsal, pero esa señal no alcanza para sustentar una clasificación operativa de las cuatro zonas bajo la regla de máxima probabilidad.

La *Figura 10* muestra el solapamiento que explica este resultado, con eventos del Cinturón de Fuego que cubren prácticamente todo el rango conjunto de magnitud y profundidad.

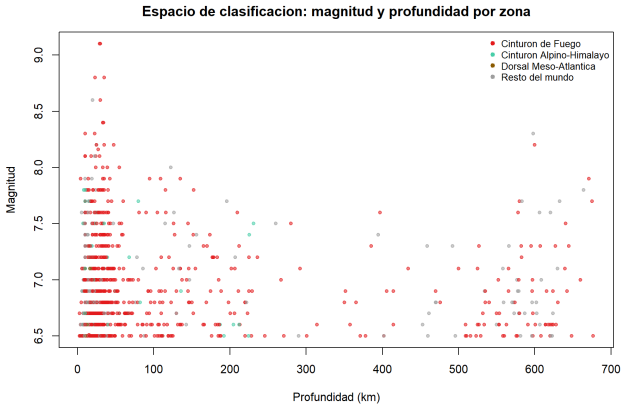

Figura 10. Espacio de clasificación: magnitud y profundidad por zona.

La *Tabla 25* compara los modelos anidados. La columna ΔAIC expresa la diferencia respecto del modelo de menor AIC, y la columna LR mide el aporte de agregar la variable faltante para llegar al modelo conjunto.

Modelo	Parámetros	AIC	ΔAIC	LR	Valor p	Lectura
Solo magnitud	6	1805,663	59,805	64,906	$< 0{,}001$	Peor ajuste
Solo profundidad	6	1745,858	0	5,102	0,1645	Mejor modelo
Magnitud y profundidad	9	1746,757	0,899	-	-	La magnitud no aporta

Tabla 25. Comparación de los modelos multinomiales anidados.

El modelo con solo profundidad obtiene el menor AIC, y agregarle la magnitud no lo mejora. La hipótesis planteada al cierre de la comparación de magnitud y profundidad entre zonas, que ambas variables ganarían relevancia al actuar conjuntamente, no queda respaldada: el aporte conjunto se reduce al aporte de la profundidad.

Para los terremotos de magnitud $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025, la magnitud y la profundidad, incluso consideradas conjuntamente, no permiten clasificar la zona sísmica de un evento con las proporciones observadas. El modelo no supera la base trivial de 75,21 %. La profundidad sí desplaza las probabilidades relativas de zona, con información distribucional real, pero no alcanza para reasignar clases. En conjunto con la sección de frecuencia, la conclusión es que el Cinturón de Fuego se distingue por cuánto y dónde ocurre su actividad, no por la magnitud o la profundidad de sus eventos individuales.

7.7. Alcance, limitaciones y sensibilidad de los resultados

Los resultados no presentan el mismo grado de estabilidad en todos los análisis. Las diferencias de frecuencia entre zonas son amplias y se mantienen al considerar tanto las tasas anuales como la superficie, por lo que constituyen el hallazgo más consistente del informe. En cambio, las estimaciones de la Dorsal Meso-Atlántica presentan mayor incertidumbre debido a su menor número de eventos. “Resto del mundo” también requiere una interpretación diferenciada, ya que corresponde a una categoría residual y no a una unidad tectónica única.

La heterogeneidad entre zonas, magnitudes y fuentes de reporte se abordó mediante procedimientos diferenciados. Las zonas se compararon utilizando tasas anuales y densidades por superficie. Para calcular estas últimas, la superficie se incorporó como exposición del modelo mediante un offset, lo que permite distinguir la frecuencia total de la concentración de eventos por unidad espacial. Este ajuste mejora la comparabilidad entre zonas de distinta extensión, pero no elimina sus diferencias tectónicas internas.

La heterogeneidad de las magnitudes se consideró mediante un umbral común de $M \geq 6{,}5$, la comparación de las distribuciones completas y la clasificación de los eventos según su severidad. La posible heterogeneidad del reporte se examinó comparando el tipo y la fuente de magnitud y el ajuste residual de localización entre zonas y períodos. No se identificaron diferencias materiales en estos procedimientos entre zonas, aunque su composición cambia a través del tiempo. Esta revisión permite reconocer posibles efectos del proceso de medición, pero no equivale a homogeneizar instrumentalmente las magnitudes del catálogo.

En la comparación de magnitud y profundidad, el incumplimiento de los supuestos multivariantes llevó a utilizar procedimientos no paramétricos. Esta decisión permite sostener las conclusiones sin imponer normalidad ni homogeneidad de las matrices de covarianza, pero restringe su interpretación a diferencias entre distribuciones. En el análisis temporal, la serie anual contiene solo 26 observaciones, por lo que las pruebas tienen una capacidad limitada para detectar cambios débiles.

La clasificación presenta la principal limitación aplicada del informe. Aunque la profundidad modifica las probabilidades relativas de zona y permite distinguir parcialmente el perfil de la Dorsal Meso-Atlántica, el fuerte desbalance del catálogo lleva al modelo a clasificar todos los eventos en el Cinturón de Fuego. Por ello, la significación de los coeficientes no se interpreta como evidencia de una capacidad clasificatoria suficiente. Las curvas ROC complementan esta evaluación, pero no modifican la conclusión general.

La sensibilidad se examina mediante dos escenarios. El primero eleva el umbral a $M \geq 7{,}0$ para determinar si las conclusiones dependen de los eventos con magnitudes entre 6,5 y 6,9. El segundo excluye “Resto del mundo” y restringe la comparación a los tres cinturones sísmicos definidos. La robustez se evalúa comparando la dirección, la magnitud y la incertidumbre de los resultados con el escenario principal, sin interpretar los escenarios como nuevas familias independientes de contrastes.

8. Conclusiones y recomendaciones

A partir de los análisis realizados, concluimos que las zonas se distinguen sobre todo por cuántos terremotos ocurren y por la profundidad a la que se originan, no por su magnitud. Esta conclusión se sustenta en 1.186 eventos con $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025. Las diferencias observadas entre zonas tampoco parecen deberse principalmente a las fuentes o a los métodos utilizados para reportar los eventos, por lo que la comparación espacial resulta adecuada para el catálogo analizado.

La diferencia más clara corresponde a la frecuencia. El Cinturón de Fuego reunió 892 eventos y registró, en promedio, 34,3 por año. En comparación, Resto del mundo presentó 7,8 eventos anuales, el Cinturón Alpino-Himalayo 2,4 y la Dorsal Meso-Atlántica 1,1. El Cinturón de Fuego también mantuvo la mayor concentración después de considerar la superficie de cada zona. Por tanto, su predominio no se explica únicamente por la extensión territorial utilizada en la delimitación.

La tasa anual y la densidad entregan lecturas complementarias. Resto del mundo ocupa el segundo lugar por cantidad de eventos, pero cae al último al expresar esa ocurrencia por unidad de superficie. Esto significa que acumula más terremotos que algunas zonas porque comprende una extensión muy amplia y heterogénea, no porque presente una mayor concentración espacial. Por esta razón, las comparaciones futuras deberían informar ambas medidas y no basarse solo en los conteos totales.

La actividad presenta años y meses con más eventos que otros, pero esas fluctuaciones no forman un aumento, una disminución ni un patrón estacional sostenido durante 2000-2025. Tampoco se observa que un período con alta ocurrencia sea seguido sistemáticamente por otro similar. En consecuencia, una tasa anual única por zona resume de manera razonable el período estudiado, aunque no debe interpretarse como una predicción de la actividad futura.

La magnitud cambia muy poco entre zonas. Aunque el análisis detectó una diferencia global, su tamaño fue tan reducido que, en términos prácticos, las magnitudes se encuentran ampliamente superpuestas. La profundidad entrega una distinción más clara. La Dorsal Meso-Atlántica se caracteriza por eventos superficiales, mientras que el Cinturón de Fuego comprende eventos superficiales, intermedios y profundos. Así, conocer la profundidad ayuda a reconocer perfiles propios de algunas zonas, pero no permite separarlas completamente.

La mayor cantidad de eventos mayores y extremos en el Cinturón de Fuego tampoco significa que sus terremotos sean proporcionalmente más severos. Esta zona concentra esos eventos principalmente porque reúne la mayor parte de todos los terremotos del catálogo. Una vez considerada esa diferencia de frecuencia, no se encontró evidencia suficiente de que la composición de severidad cambie entre zonas ni de que la profundidad determine si un evento alcanza $M \geq 7{,}0$.

La magnitud y la profundidad no bastan para identificar de manera confiable la zona donde ocurrió un terremoto. La profundidad aporta información y permite reconocer parcialmente el perfil superficial de la Dorsal Meso-Atlántica, pero existe demasiado solapamiento entre las zonas. El modelo terminó asignando todos los eventos al Cinturón de Fuego y alcanzó una exactitud de 75,21 %, el mismo resultado que se obtendría clasificando siempre en la zona mayoritaria. Por ello, este modelo no debe utilizarse como herramienta operativa de clasificación.

Los resultados de esta etapa respaldan gran parte de los patrones descritos en el Informe Asesoría 1. Se confirma que el Cinturón de Fuego concentra la mayor ocurrencia, incluso después de considerar la superficie de las zonas. También se sostiene que las magnitudes son muy similares en términos prácticos y que la profundidad entrega una diferenciación más clara, especialmente por el carácter superficial de la Dorsal Meso-Atlántica y el rango más amplio de profundidades presente en el Cinturón de Fuego. Del mismo modo, la revisión de las variables de reporte respalda que estas comparaciones espaciales no están materialmente condicionadas por la fuente o el tipo de magnitud utilizado.

Algunos hallazgos descriptivos requieren, sin embargo, una interpretación más acotada. Las diferencias observadas entre períodos en el Informe Asesoría 1 no se mantienen como evidencia de un cambio temporal después de considerar la variabilidad adicional de los conteos. Por ello, las fluctuaciones entre años se interpretan como compatibles con una tasa estable durante 2000-2025. Asimismo, la mayor cantidad de eventos mayores y extremos en el Cinturón de Fuego no implica una mayor severidad proporcional, sino que responde principalmente a que esa zona concentra la mayoría de los terremotos del catálogo. En respuesta a la pregunta estadística, las zonas sí difieren en su tasa de ocurrencia y en las distribuciones de magnitud y profundidad, pero la magnitud aporta poca separación práctica y su combinación con la profundidad no permite determinar de forma confiable la zona de un evento.

Para apoyar decisiones futuras, se recomienda informar conjuntamente la tasa anual y la densidad por superficie, interpretar Resto del mundo como una categoría residual y mantener separadas las nociones de frecuencia y severidad. Si la contraparte requiere identificar zonas o desarrollar una herramienta de clasificación, será necesario incorporar ubicación geográfica, cercanía a límites de placas, tipo de contacto tectónico u otras variables espaciales y geológicas. Los resultados de este informe permiten orientar la comparación y comunicación del catálogo, pero no estimar amenaza, riesgo sísmico ni probabilidades futuras de ocurrencia.

9. Referencias

Akaike, Hirotugu. 1974. «A New Look at the Statistical Model Identification». *IEEE Transactions on Automatic Control* 19 (6): 716-23. https://doi.org/10.1109/TAC.1974.1100705.

Albert, Adelin, y John A. Anderson. 1984. «On the Existence of Maximum Likelihood Estimates in Logistic Regression Models». *Biometrika* 71 (1): 1-10. https://doi.org/10.1093/biomet/71.1.1.

Banerjee, Anindya, Juan J. Dolado, John W. Galbraith, y David F. Hendry. 1993. *Cointegration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford University Press.

Bartlett, M. S. 1950. «Tests of Significance in Factor Analysis». *British Journal of Statistical Psychology* 3 (2): 77-85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x.

Benjamini, Yoav, y Yosef Hochberg. 1995. «Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing». *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bergsma, Wicher. 2013. «A Bias-Correction for Cramér's V and Tschuprow's T». *Journal of the Korean Statistical Society* 42 (3): 323-28. https://doi.org/10.1016/j.jkss.2012.10.002.

Berkson, Joseph. 1944. «Application of the Logistic Function to Bio-Assay». *Journal of the American Statistical Association* 39 (227): 357-65. https://doi.org/10.1080/01621459.1944.10500699.

Bland, J. Martin, y Douglas G. Altman. 2000. «Statistics Notes: The Odds Ratio». *BMJ* 320 (7247): 1468. https://doi.org/10.1136/bmj.320.7247.1468.

Box, G. E. P. 1949. «A General Distribution Theory for a Class of Likelihood Criteria». *Biometrika* 36 (3/4): 317-46. https://doi.org/10.1093/biomet/36.3-4.317.

Burnham, Kenneth P., y David R. Anderson. 2004. «Multimodel Inference: Understanding AIC and BIC in Model Selection». *Sociological Methods & Research* 33 (2): 261-304. https://doi.org/10.1177/0049124104268644.

Cabello, Catalina, Gonzalo Montalva, y Felipe Leyton. 2025. «Integrated Seismic Catalog for Chile from 1513 to 2022». *Journal of South American Earth Sciences* 164: 105639. https://doi.org/10.1016/j.jsames.2025.105639.

Cameron, A. Colin, y Pravin K. Trivedi. 1990. «Regression-Based Tests for Overdispersion in the Poisson Model». *Journal of Econometrics* 46 (3): 347-64. https://doi.org/10.1016/0304-4076(90)90014-K.

Cliff, Norman. 1993. «Dominance Statistics: Ordinal Analyses to Answer Ordinal Questions». *Psychological Bulletin* 114 (3): 494-509. https://doi.org/10.1037/0033-2909.114.3.494.

Cochran, William G. 1954. «Some Methods for Strengthening the Common χ^2 Tests». *Biometrics* 10 (4): 417-51. https://doi.org/10.2307/3001616.

Conover, William J. 1971. *Practical Nonparametric Statistics*. Wiley.

Dunn, Olive Jean. 1964. «Multiple Comparisons Using Rank Sums». *Technometrics* 6 (3): 241-52. https://doi.org/10.1080/00401706.1964.10490181.

Durbin, James. 1973. *Distribution Theory for Tests Based on the Sample Distribution Function*. Society for Industrial; Applied Mathematics. https://doi.org/10.1137/1.9781611970586.

Fisher, Ronald A. 1922. «On the Mathematical Foundations of Theoretical Statistics». *Philosophical Transactions of the Royal Society of London, Series A* 222: 309-68. https://doi.org/10.1098/rsta.1922.0009.

Greenacre, Michael. 2017. *Correspondence Analysis in Practice*. 3.ª ed. Chapman; Hall/CRC. https://doi.org/10.1201/9781315369983.

Hauck, Walter W., y Allan Donner. 1977. «Wald's Test as Applied to Hypotheses in Logit Analysis». *Journal of the American Statistical Association* 72 (360): 851-53. https://doi.org/10.1080/01621459.1977.10479969.

Holm, Sture. 1979. «A Simple Sequentially Rejective Multiple Test Procedure». *Scandinavian Journal of Statistics* 6 (2): 65-70. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/p.adjust.html.

Hope, A. C. A. 1968. «A Simplified Monte Carlo Significance Test Procedure». *Journal of the Royal Statistical Society: Series B (Methodological)* 30 (3): 582-98. https://doi.org/10.1111/j.2517-6161.1968.tb00759.x.

Instructivo de encargo. 2026. *Instructivo de encargo: Asesorías 1 y 2*. Universidad de Santiago de Chile.

Kagan, Yan Y., y David D. Jackson. 2016. «Earthquake Rate and Magnitude Distributions of Great Earthquakes for Use in Global Forecasts». *Geophysical Journal International* 206 (1): 630-43. https://doi.org/10.1093/gji/ggw161.

Kosmidis, Ioannis, y David Firth. 2011. «Multinomial Logit Bias Reduction via the Poisson Log-Linear Model». *Biometrika* 98 (3): 755-59. https://doi.org/10.1093/biomet/asr026.

Kruskal, William H., y W. Allen Wallis. 1952. «Use of Ranks in One-Criterion Variance Analysis». *Journal of the American Statistical Association* 47 (260): 583-621. https://doi.org/10.1080/01621459.1952.10483441.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, y Yongcheol Shin. 1992. «Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root». *Journal of Econometrics* 54 (1–3): 159-78. https://doi.org/10.1016/0304-4076(92)90104-Y.

Ljung, Greta M., y George E. P. Box. 1978. «On a Measure of Lack of Fit in Time Series Models». *Biometrika* 65 (2): 297-303. https://doi.org/10.1093/biomet/65.2.297.

Mangiafico, Salvatore S. 2026. *epsilonSquared: Epsilon-Squared Effect Size for the Kruskal-Wallis Test*. Comprehensive R Archive Network. https://search.r-project.org/CRAN/refmans/rcompanion/html/epsilonSquared.html.

Mardia, Kanti V. 1970. «Measures of Multivariate Skewness and Kurtosis with Applications». *Biometrika* 57 (3): 519-30. https://doi.org/10.1093/biomet/57.3.519.

Marsaglia, George, Wai Wan Tsang, y Jingbo Wang. 2003. «Evaluating Kolmogorov's Distribution». *Journal of Statistical Software* 8 (18): 1-4. https://doi.org/10.18637/jss.v008.i18.

McCullagh, Peter, y John A. Nelder. 1989. *Generalized Linear Models*. 2.ª ed. Monographs on Statistics y Applied Probability 37. Chapman; Hall/CRC. https://doi.org/10.1201/9780203753736.

Neyman, Jerzy. 1937. «Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability». *Philosophical Transactions of the Royal Society of London, Series A* 236 (767): 333-80. https://doi.org/10.1098/rsta.1937.0005.

Ogata, Yosihiko. 1988. «Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes». *Journal of the American Statistical Association* 83 (401): 9-27. https://doi.org/10.1080/01621459.1988.10478560.

Patefield, W. M. 1981. «Algorithm AS 159: An Efficient Method of Generating Random R × C Tables with Given Row and Column Totals». *Applied Statistics* 30 (1): 91-97. https://doi.org/10.2307/2346669.

R Core Team. s. f.-a. *Family Objects for Models*. R Foundation for Statistical Computing. Accedido 11 de julio de 2026. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/family.html.

R Core Team. s. f.-b. *Fitting Generalized Linear Models*. R Foundation for Statistical Computing. Accedido 11 de julio de 2026. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html.

R Core Team. s. f.-c. *p.adjust: Adjust P-values for Multiple Comparisons*. R Foundation for Statistical Computing. Accedido 4 de julio de 2026. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/p.adjust.html.

R Core Team. s. f.-d. *The Poisson Distribution*. R Foundation for Statistical Computing. Accedido 11 de julio de 2026. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/Poisson.html.

Robin, Xavier, Natacha Turck, Alexandre Hainard, et al. 2011. «pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves». *BMC Bioinformatics* 12: 77. *https://doi.org/10.1186/1471-2105-12-77*.

Said, Said E., y David A. Dickey. 1984. «Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order». *Biometrika* 71 (3): 599-607. *https://doi.org/10.1093/biomet/71.3.599*.

Self, Steven G., y Kung-Yee Liang. 1987. «Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions». *Journal of the American Statistical Association* 82 (398): 605-10. *https://doi.org/10.1080/01621459.1987.10478472*.

Shapiro, S. S., y M. B. Wilk. 1965. «An Analysis of Variance Test for Normality». *Biometrika* 52 (3/4): 591-611. *https://doi.org/10.1093/biomet/52.3-4.591*.

Venables, William N., y Brian D. Ripley. 2002. *Modern Applied Statistics with S*. 4.ª ed. Springer. *https://doi.org/10.1007/978-0-387-21706-2*.

Wald, Abraham. 1943. «Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large». *Transactions of the American Mathematical Society* 54 (3): 426-82. *https://doi.org/10.1090/S0002-9947-1943-0012401-3*.

Wiemer, Stefan, y Max Wyss. 2000. «Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan». *Bulletin of the Seismological Society of America* 90 (4): 859-69. *https://doi.org/10.1785/0119990114*.

10. Anexos reproducibles

10.1. Continuidad de datos y variables

Este anexo identifica los datos y las variables heredadas del Informe Asesoría 1 y documenta las incorporaciones realizadas en esta etapa. El análisis conserva la base `sismos`, integrada por 1.186 eventos, junto con sus variables derivadas. Entre ellas se encuentra `rms_imp`, versión completada del RMS que incorpora nueve valores ausentes imputados mediante la mediana de la década correspondiente. La comparación descriptiva realizada en el informe anterior mostró que esta imputación no modificó sustancialmente la media, mediana, dispersión ni cuantiles. El RMS representa el ajuste residual entre los tiempos de llegada observados y los estimados por el modelo de localización, pero no constituye por sí solo una medida completa de precisión instrumental o espacial.

También se mantienen `magnitud_cat`, `profundidad_cat`, `zona` y `magType_grupo`. La única información nueva es la superficie de los polígonos de zona, incorporada como exposición en los modelos de conteo. La correspondencia entre variables, scripts y objetos derivados se conserva en el flujo reproducible del repositorio.

10.2. Áreas de zonas y definición del offset

Este anexo presenta las superficies de cada zona y la definición del offset logarítmico de los modelos de conteo. El área de cada zona se calculó en el software QGIS y quedó registrada en el campo `Area` de la capa `MAPA_HTML/zonas_inf2.geojson`, derivada de los tres cinturones definidos en `area_regiones.geojson` y de su complemento espacial. Las superficies se expresan en kilómetros cuadrados y, para facilitar su lectura, también en millones de kilómetros cuadrados.

Zona	Superficie (km^2)	Superficie (millones de km^2)
Cinturón de Fuego	47.639.417	47,64
Cinturón Alpino-Himalayo	14.475.681	14,48
Dorsal Meso-Atlántica	18.900.689	18,90
Resto del mundo	429.049.835	429,05
Total	510.065.622	510,07

Tabla 26. Superficie de cada zona sísmica utilizada como exposición en los modelos de conteo.

El modelo emplea un enlace logarítmico, por lo que modelar la densidad, es decir, los eventos por unidad de superficie, equivale a restar el logaritmo del área al logaritmo del conteo esperado. Ese logaritmo del área es el offset: entra con un coeficiente fijo en 1, no estimado. De esta forma, los coeficientes de zona se interpretan como razones de densidad, en eventos por unidad de superficie y año, en lugar de razones de tasa, en eventos por año. Como la superficie es constante dentro de cada zona, el modelo de tasa y el de densidad producen el mismo ajuste de las medias por zona y difieren solo en la interpretación de sus coeficientes.

Formalmente, sea Y_{jt} el número de eventos en la zona j durante el año t, $\mu_{jt} = \mathbb{E}(Y_{jt})$ su valor esperado y A_j la superficie de la zona. El modelo de densidad es:

$$\log(\mu_{jt}) = \beta_0 + \sum_{k=1}^{3} \beta_k \, z_{jk} + \log(A_j),$$

donde z_{jk} son las tres variables indicadoras de zona, con el Cinturón de Fuego como categoría de referencia, y $\log(A_j)$ es el offset, un término conocido con coeficiente fijo en 1. Al pasar ese término al lado izquierdo:

$$\log\left(\frac{\mu_{jt}}{A_j}\right) = \beta_0 + \sum_{k=1}^{3} \beta_k \, z_{jk},$$

de modo que el modelo describe directamente la densidad μ_{jt}/A_j, es decir, los eventos por unidad de superficie. El modelo de tasa anual es idéntico pero sin el término $\log(A_j)$, y sus coeficientes se leen sobre el conteo μ_{jt}.

10.3. Flujo de trabajo e implementación computacional

Este anexo documenta cómo se prepara la base, se ejecutan los análisis y se generan los resultados del Informe Asesoría 2. Su propósito es facilitar la revisión y reproducción del trabajo, sin repetir las decisiones estadísticas desarrolladas en la metodología y los anexos específicos. El código se encuentra disponible en la carpeta Scripts del repositorio GitHub.

10.3.1. Preparación y ejecución en Positron

El análisis se desarrolla en Positron mediante scripts de R organizados dentro del proyecto `USGS-earthquake-analysis`. La etapa inferencial utiliza la base `sismos` construida en el Informe Asesoría 1, que contiene las variables temporales, las categorías de magnitud y profundidad, la zona sísmica y las versiones imputadas de `nst` y `rms`.

El script `00_preparacion_base.R` permite reconstruir directamente esta base sin volver a ejecutar todos los análisis descriptivos. Para ello importa los archivos originales, crea las variables derivadas, aplica las imputaciones definidas anteriormente y asigna cada evento a una zona mediante el cruce espacial. Los módulos inferenciales pueden ejecutarse desde `Código.R` o de manera independiente, porque comprueban la existencia de la base y la preparan cuando es necesario.

10.3.2. Organización de los análisis

Cada módulo reúne la preparación específica de los datos, el ajuste de modelos, los diagnósticos y la generación de las tablas o figuras correspondientes. Los nombres de la primera columna enlazan directamente con el código utilizado.

Código	Propósito principal
`Código.R`	Ejecuta de manera secuencial la preparación de la base y los módulos del proyecto.
`00_preparacion_base.R`	Reconstruye la base analítica utilizada en esta etapa.
`06_inf_comparabilidad.R`	Revisa la comparabilidad de los procedimientos de reporte.
`07_inf_frecuencia.R`	Compara la tasa anual y la densidad de eventos entre zonas.
`08_inf_magnitud_profundidad.R`	Evalúa las diferencias de magnitud y profundidad entre zonas.
`09_inf_temporal.R`	Analiza la estabilidad, dependencia y distribución temporal de los eventos.
`10_inf_extremos.R`	Estudia los eventos fuertes y extremos.
`11_inf_clasificacion.R`	Evalúa la clasificación de las zonas mediante magnitud y profundidad.

Tabla 27. Código del flujo reproducible y propósito de cada módulo del Informe Asesoría 2.

En los procedimientos que requieren simulación se utiliza la semilla `2026`, lo que permite reproducir los valores obtenidos. La separación por módulos también facilita la revisión de cada análisis sin tener que ejecutar nuevamente el flujo completo.

10.3.3. QGIS, superficies y control de versiones

La delimitación de las zonas se conserva desde el Informe Asesoría 1. El procedimiento completo de delimitación, elaborado por el equipo, se documenta en la guía metodológica de delimitación de zonas sísmicas globales en QGIS. En esta etapa, QGIS también se utiliza para calcular la superficie de los polígonos y construir la capa `MAPA_HTML/zonas_inf2.geojson`. R importa estas superficies para expresar la frecuencia como densidad de eventos y utilizarlas como exposición en los modelos de conteo.

El proyecto se mantiene mediante Git y GitHub. Los scripts, las bases, las capas geográficas, la bibliografía y el archivo QMD permanecen en el mismo repositorio, lo que permite registrar los cambios y conservar la correspondencia entre el código ejecutado y los resultados presentados.

10.4. Especificación de dócimas y modelos

Este anexo reúne las reglas de decisión que sustentan los procedimientos descritos de forma resumida en la metodología.

Tabla 28. Especificación de los criterios de decisión estadística.

Familia de procedimientos	Rol	Criterio de decisión	Justificación
Mardia, Shapiro-Wilk, M de Box y Bartlett	Diagnóstico	$\alpha = 0{,}10$ junto con diagnóstico gráfico	Se prioriza detectar condiciones que podrían invalidar la ruta paramétrica; con una muestra grande, el resultado se interpreta junto con los gráficos
Kruskal-Wallis para `mag` y `depth`	Confirmatorio	$\alpha = 0{,}05$ por variable	Cada variable responde una pregunta sustantiva diferente; la multiplicidad posterior se controla dentro de cada familia de comparaciones por pares
Dunn, seis pares por variable	Confirmatorio	Error familiar de 0,05 con corrección de Holm	La corrección limita la probabilidad de falsos positivos producidos por las seis comparaciones simultáneas
GLM de conteos	Confirmatorio	$\alpha = 0{,}05$ e IC 95 % de las razones de tasa	La conclusión considera la magnitud y la incertidumbre de las diferencias entre tasas
Sobredispersión	Diagnóstico de especificación	$\alpha = 0{,}05$ unilateral	La alternativa relevante es que la varianza supere la prevista por el modelo Poisson
ADF (`type="nc"`) y KPSS	Diagnóstico	$\alpha = 0{,}05$ e interpretación conjunta	Sus hipótesis nulas son opuestas y la conclusión se apoya en su concordancia; ADF se aplica sin constante ni tendencia como especificación inicial, y con 26 observaciones anuales se reconoce una potencia limitada
Ljung-Box	Diagnóstico	$\alpha = 0{,}05$	Evalúa la ausencia de autocorrelación en las series anuales y mensuales y en los residuos de los modelos de conteo; su rechazo señalaría dependencia temporal no considerada
Ajuste exponencial	Diagnóstico descriptivo	Bootstrap paramétrico y gráfico QQ	El parámetro se estima con los mismos datos, por lo que no se utiliza el valor p convencional de Kolmogorov-Smirnov
Chi-cuadrado entre categorías	Confirmatorio	$\alpha = 0{,}05$; Monte Carlo cuando existen esperados inferiores a cinco	La simulación evita depender de una aproximación asintótica imprecisa
Modelos logístico y multinomial	Confirmatorio	Razón de verosimilitud por variable, $\alpha = 0{,}05$, e IC 95 %	La evaluación por bloques es más estable que las pruebas de Wald individuales ante posibles problemas de separación
Selección de modelos por AIC	Comparación	Se prefiere el menor AIC; una diferencia menor que 2 indica respaldo equivalente y se resuelve por parsimonia	Pondera ajuste y complejidad, y complementa la razón de verosimilitud, en especial cuando los modelos no son estrictamente anidados
Corrección de separación (Firth)	Reducción de sesgo	Se aplica ante separación, detectada por un coeficiente desproporcionado con error estándar inflado y probabilidades ajustadas en cero	Garantiza estimaciones finitas; los contrastes globales y el AIC se mantienen sobre el modelo sin penalizar
Comparabilidad del registro	Exploratorio	FDR de 5 % mediante Benjamini-Hochberg	El ajuste limita la proporción esperada de falsos descubrimientos dentro de los cinco contrastes

Para la comparabilidad del registro se mantuvo `magType_grupo`, definida mediante `mww`, `mwc`, `mwb` y "otros". En `magSource_grupo`, las fuentes con menos de 20 registros se reunieron en una categoría residual. Este umbral representa un promedio potencial mínimo de cinco registros por zona si la distribución fuera uniforme. Su propósito es evitar interpretaciones individuales de fuentes con respaldo insuficiente, no asegurar que todas las frecuencias esperadas superen cinco, porque los tamaños de las zonas son desiguales. Las fuentes principales se conservaron separadas para no ocultar diferencias en los procedimientos de reporte. Los períodos se definieron como 2000-2009, 2010-2019 y 2020-2025.

Para la comparación de magnitud y profundidad, las hipótesis se organizaron en dos bloques. El primer bloque define la ruta metodológica. En Shapiro-Wilk, la hipótesis nula plantea que la variable evaluada sigue una distribución normal dentro de cada zona, mientras que la alternativa indica desviación de normalidad. En Mardia, la hipótesis nula plantea normalidad multivariante del vector formado por magnitud y profundidad dentro de cada zona; la alternativa indica desviación por asimetría o curtosis multivariante. En la M de Box, la hipótesis nula plantea igualdad de matrices de varianza-covarianza entre zonas; la alternativa indica que al menos una zona tiene una matriz distinta. Estas dócimas no se usan para concluir diferencias sustantivas entre zonas, sino para decidir si la ruta paramétrica multivariante es metodológicamente defendible. Si se rechazan de forma consistente, MANOVA deja de ser el procedimiento principal porque sus valores p pueden quedar afectados por la forma de las distribuciones, la heterogeneidad de dispersión y el fuerte desbalance entre tamaños muestrales.

Bartlett y Spearman se trataron como diagnósticos complementarios de relación entre variables. En Bartlett, la hipótesis nula plantea que la matriz de correlaciones es equivalente a una matriz identidad; con dos variables, esto equivale a evaluar ausencia de correlación lineal entre magnitud y profundidad. En Spearman, la hipótesis nula plantea ausencia de asociación monotónica. Estas pruebas permiten revisar si magnitud y profundidad aportan información conjunta o si una de ellas resulta prácticamente redundante. No reemplazan las pruebas de normalidad ni justifican por sí solas una ruta paramétrica.

El segundo bloque corresponde a las comparaciones no paramétricas. En Kruskal-Wallis, la hipótesis nula plantea igualdad de distribuciones, expresada operacionalmente como igualdad de rangos entre zonas; la alternativa indica que al menos una zona presenta rangos sistemáticamente distintos. La prueba identifica una diferencia global, pero no señala qué pares de zonas explican esa diferencia. Por eso, cuando el contraste global fue significativo, se aplicó Dunn-Holm. En cada comparación de Dunn, la hipótesis nula plantea igualdad de rangos entre dos zonas; la alternativa indica diferencia entre ese par. El ajuste de Holm controla el error familiar de las seis comparaciones por variable.

Los tamaños de efecto se incluyeron para evitar conclusiones basadas solo en significación estadística. Epsilon cuadrado resume la magnitud global del efecto de zona en Kruskal-Wallis y se interpreta como un análogo no paramétrico del R^2 aplicado a rangos. Delta de Cliff compara dos zonas y se define como la diferencia entre la probabilidad de que una observación de la primera zona sea mayor que una de la segunda y la probabilidad inversa. Valores cercanos a cero indican fuerte superposición entre distribuciones; valores alejados de cero indican mayor separación práctica. Para las asociaciones categóricas, la V de Cramér corregida por sesgo cumple una función equivalente: resume el grado de asociación de una tabla de contingencia y descuenta la sobreestimación propia de tablas con categorías poco frecuentes.

10.5. Contraste chi-cuadrado de independencia y frecuencias esperadas

Este anexo explica el contraste chi-cuadrado de independencia y el papel de las frecuencias esperadas, que sustentan los contrastes categóricos del informe, la V de Cramér y la inercia del análisis de correspondencia.

La frecuencia esperada de una casilla es el conteo que tendría si las dos variables de la tabla fueran independientes. Se obtiene de los totales marginales:

$$E_{ij} = \frac{n_{i\cdot}\, n_{\cdot j}}{n},$$

donde $n_{i\cdot}$ es el total de la fila i, $n_{\cdot j}$ el total de la columna j y n el total de eventos. Es el punto de referencia de la no asociación: las desviaciones entre lo observado y lo esperado son la materia prima del contraste.

El estadístico chi-cuadrado acumula esas desviaciones, relativizadas por lo esperado en cada casilla:

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

con O_{ij} el conteo observado. Bajo independencia, el estadístico sigue aproximadamente una distribución chi-cuadrado con $(r-1)(c-1)$ grados de libertad, donde r y c son el número de filas y columnas. La aproximación es asintótica

y se vuelve imprecisa cuando alguna esperada es pequeña, porque el denominador E_{ij} amplifica la contribución de casillas con pocos eventos. La regla práctica habitual exige esperadas de al menos 5; la recomendación original es más matizada y admite un mínimo de 1 cuando son pocas las casillas bajo ese umbral (*Cochran 1954*). Cuando la condición no se cumple, el valor p asintótico se reemplaza por uno simulado (*véase anexo de precisión de la simulación Monte Carlo*).

Como ejemplo con el catálogo, en el cruce entre zona y categoría de magnitud la casilla de la Dorsal Meso-Atlántica en Grande o extremo tiene esperada

$$E = \frac{28 \times 59}{1186} = 1{,}39,$$

es decir, si la Dorsal se comportara como el catálogo completo, se esperaría poco más de un evento en esa casilla. Se observaron 0, y la contribución de la casilla al estadístico es $(0-1{,}39)^2/1{,}39 = 1{,}39$, cerca de un quinto del $\chi^2 = 7{,}1222$ total aportado por una sola casilla casi vacía. Por eso las esperadas reducidas obligan a simular el valor p y moderan la lectura de la posición extrema de la Dorsal en el análisis de correspondencia.

El mismo estadístico alimenta las dos medidas descriptivas del informe: dividido por n produce ϕ^2, la base de la V de Cramér corregida (*véase anexo de corrección por sesgo de la V de Cramér*) y de la inercia total del análisis de correspondencia (*véase anexo de análisis de correspondencia*).

10.6. Precisión de la simulación Monte Carlo

Este anexo documenta la precisión de los valores p obtenidos mediante simulación Monte Carlo y las condiciones que justifican su uso.

En cada prueba se generaron $B = 10.000$ tablas bajo la hipótesis de independencia y con los mismos totales marginales que la tabla observada (*Hope 1968*; *Patefield 1981*). Si K representa la cantidad de tablas simuladas cuyo estadístico chi-cuadrado es igual o mayor que el observado, el valor p se estima mediante:

$$\hat{p} = \frac{K + 1}{B + 1}.$$

El diagnóstico de las frecuencias esperadas fue el siguiente:

Tabla 29. Diagnóstico de frecuencias esperadas para los contrastes evaluados por simulación Monte Carlo.

Contraste	Mínima esperada	Celdas bajo 5	Total de celdas
Zona por tipo de magnitud	0,779	3	16
Zona por fuente de magnitud	0,732	4	16
Período por tipo de magnitud	6,734	0	12
Zona por categoría de magnitud	1,39	2	12

Las tres tablas que cruzan zonas justificaron el uso de simulación por sus frecuencias esperadas reducidas. En la tabla temporal no se presentó esta condición; no obstante, se mantuvo el mismo procedimiento para obtener los valores p categóricos de la comparabilidad bajo una estrategia común.

Para el contraste entre zona y categoría de magnitud de la sección de eventos fuertes y extremos, la matriz completa de frecuencias esperadas bajo independencia fue:

Tabla 30. Frecuencias esperadas bajo independencia para el cruce entre zona y categoría de magnitud.

Zona	Fuerte	Mayor	Grande o extremo
Cinturón de Fuego	600,93	246,69	44,37
Resto del mundo	136,76	56,14	10,10
Cinturón Alpino-Himalayo	42,44	17,42	3,13
Dorsal Meso-Atlántica	18,86	7,74	1,39

La corrección evita informar $p = 0$ cuando ninguna simulación produce un resultado más extremo, pues ese resultado solo indica que el evento no apareció dentro del número finito de réplicas. En consecuencia, la resolución mínima con 10.000 simulaciones es:

$$p_{\text{mín}} = \frac{1}{10.000 + 1} \approx 0{,}0001.$$

Como $\hat{p}$ es una proporción, su error estándar puede aproximarse mediante:

$$EE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{B}}.$$

Evaluando esta expresión en el nivel de significación utilizado:

$$EE(\hat{p}) = \sqrt{\frac{0{,}05(1-0{,}05)}{10.000}} \approx 0{,}0022.$$

El margen aproximado con un 95 % de confianza es:

$$1{,}96 \times 0{,}0022 \approx 0{,}0043.$$

Por tanto, si un valor p simulado se encontrara aproximadamente entre 0,046 y 0,054, se aumentaría el número de simulaciones antes de establecer una conclusión. Se fijó una semilla aleatoria para asegurar la reproducibilidad del procedimiento.

10.7. Bootstrap paramétrico para pruebas de ajuste

Este anexo explica cómo obtuvimos el valor p de la prueba de ajuste exponencial. El bootstrap paramétrico permite evaluar si la diferencia entre los tiempos observados y el modelo es mayor que la que podría aparecer por variación aleatoria. Es necesario porque la tasa de ocurrencia no se conoce previamente, sino que se calcula con los mismos tiempos que luego se comparan. En esa situación, el valor p convencional de Kolmogorov-Smirnov no representa correctamente la incertidumbre del ajuste (*Durbin 1973*).

El procedimiento reproduce muchas veces el análisis bajo un escenario en que el modelo exponencial sí fuera adecuado. En cada repetición se genera una muestra artificial con el mismo número de observaciones, se vuelve a calcular la tasa y se mide nuevamente la distancia entre los datos simulados y su distribución exponencial. Finalmente, la distancia observada se compara con las 10.000 distancias obtenidas mediante simulación.

En términos formales, si D_{obs} es la distancia observada y D_b^* la obtenida en la simulación b, el valor p bootstrap corresponde a

$$\hat{p}_{\text{boot}} = \frac{1}{B}\sum_{b=1}^{B} I(D_b^* \geq D_{\text{obs}}), \qquad B = 10.000.$$

Un valor p pequeño indica que pocas muestras generadas bajo el modelo presentan una discrepancia tan grande como la observada. Un valor mayor indica que esa diferencia puede aparecer razonablemente aun cuando el modelo exponencial sea adecuado. La semilla se fijó en 2026 para que las simulaciones puedan reproducirse.

Para el grupo $M \geq 7{,}0$, la distancia observada fue $D = 0{,}0778$ y el valor p bootstrap fue 0,001. Esto significa que solo cerca del 0,1 % de las muestras simuladas presentó una discrepancia igual o mayor. Por ello, los tiempos del grupo combinado no se describen adecuadamente mediante una única distribución exponencial. Este resultado no surge de comparar solo la media o la mediana, sino la forma completa de la distribución.

La simulación Monte Carlo y el bootstrap paramétrico utilizan repeticiones computacionales, pero responden preguntas distintas. Monte Carlo aproxima el resultado de una prueba de independencia en tablas de frecuencias. El bootstrap paramétrico reproduce el ajuste completo de una distribución y considera que su tasa fue estimada con los mismos datos. En ambos casos, las simulaciones permiten obtener una referencia más adecuada que la aproximación convencional para las condiciones específicas del análisis.

10.8. Corrección por sesgo de la V de Cramér

Este anexo explica la corrección aplicada a la V de Cramér para reducir su sobreestimación en tablas pequeñas o con categorías poco frecuentes.

La V de Cramér convencional se obtiene a partir de:

$$\phi^2 = \frac{\chi^2}{n}, \qquad V = \sqrt{\frac{\phi^2}{\min(r-1, c-1)}},$$

donde n es el número de observaciones, r y c corresponden al número de filas y columnas de la tabla y χ^2 es el estadístico del contraste de independencia (*véase anexo del contraste chi-cuadrado y frecuencias esperadas*). Para reducir su sesgo positivo se aplicó la corrección propuesta en la literatura (*Bergsma 2013*):

$$\tilde{\phi}^2 = \max\left(0,\ \phi^2 - \frac{(r-1)(c-1)}{n-1}\right),$$

$$\tilde{r} = r - \frac{(r-1)^2}{n-1}, \qquad \tilde{c} = c - \frac{(c-1)^2}{n-1},$$

$$V_{\text{corregida}} = \sqrt{\frac{\tilde{\phi}^2}{\min(\tilde{r}-1, \tilde{c}-1)}}.$$

Esta corrección descuenta la asociación positiva que puede aparecer únicamente por variabilidad muestral. Por ello, la medida corregida se utilizó para valorar la importancia práctica de las asociaciones detectadas.

A modo de ilustración, para el cruce entre zona y categoría de magnitud de la sección de eventos fuertes y extremos ($\chi^2 = 7{,}1222$, $n = 1186$, $r = 4$, $c = 3$) se obtuvo:

$$\phi^2 = \frac{7{,}1222}{1186} = 0{,}0060052, \qquad \tilde{\phi}^2 = \max\left(0,\ 0{,}0060052 - \frac{6}{1185}\right) = 0{,}0009419,$$

$$\tilde{r} = 4 - \frac{9}{1185} = 3{,}9924, \qquad \tilde{c} = 3 - \frac{4}{1185} = 2{,}9966,$$

$$V_{\text{corregida}} = \sqrt{\frac{0{,}0009419}{\min(2{,}9924;\ 1{,}9966)}} = 0{,}02172.$$

Sin la corrección, la V convencional habría sido $\sqrt{0{,}0060052/2} = 0{,}0548$, más del doble. La diferencia muestra que la mayor parte de la asociación aparente en esta tabla proviene de su estructura muestral (doce celdas, dos con frecuencias esperadas bajo 5) y no de una relación sustantiva entre zona y categoría de magnitud.

10.9. Cálculo de epsilon cuadrado

Epsilon cuadrado (ε^2) se utilizó como tamaño de efecto de Kruskal-Wallis. La medida expresa qué proporción de la variabilidad de los rangos se asocia con el factor de agrupación. Su interpretación es equivalente al R^2 de un ANOVA aplicado a los rangos, por lo que no representa varianza explicada de los valores originales del RMS (*Mangiafico 2026*).

La fórmula aplicada por el script y por la función `epsilonSquared()` de `rcompanion` es:

$$\varepsilon^2 = \frac{H}{n-1},$$

donde H es el estadístico de Kruskal-Wallis, con la corrección por empates incorporada por R, y n es el número total de observaciones. Para la comparación espacial se obtuvo:

$$\varepsilon^2_{\text{zona}} = \frac{10{,}348}{1186 - 1} = 0{,}0087 \approx 0{,}009.$$

Por tanto, la zona se asocia con menos del 1 % de la variabilidad de los rangos del RMS. Para la comparación temporal:

$$\varepsilon^2_{\text{período}} = \frac{261{,}930}{1186 - 1} = 0{,}221.$$

En este caso, el período se asocia aproximadamente con el 22,1 % de la variabilidad de los rangos. Ambos resultados se interpretaron mediante sus valores numéricos y no mediante puntos de corte rígidos.

10.10. Modelo Poisson para conteos temporales

Este anexo explica el modelo Poisson utilizado para comparar conteos entre períodos y meses. A diferencia del modelo binomial negativo, aquí el interés se concentra en los cambios temporales del catálogo. Cada observación es un año o un mes y la respuesta es el número de eventos registrados. Poisson es el punto de partida para conteos de ocurrencias en unidades de exposición comparables y resulta más adecuado que una comparación normal de medias, porque los conteos son enteros, no negativos y su varianza depende del nivel medio de ocurrencia (*McCullagh y Nelder 1989*; *Venables y Ripley 2002*).

La distribución asigna probabilidad a observar exactamente y eventos cuando la media esperada del intervalo es μ; esta es la parametrización usada por `dpois()` en R (*R Core Team, s. f.-d*):

$$P(Y = y) = \frac{e^{-\mu}\mu^y}{y!}.$$

En el GLM Poisson, la media se modela con enlace logarítmico, implementado en R como `glm(..., family = poisson(link = "log"))` (*R Core Team, s. f.-b*, *s. f.-a*):

$$\log(\mu_t) = \beta_0 + \beta_1 I_{\text{2010-2019},t} + \beta_2 I_{\text{2020-2025},t},$$

donde μ_t es el número esperado de eventos en el año t y 2000-2009 queda como período de referencia. El enlace logarítmico garantiza medias positivas, ya que $\mu_t = \exp(\eta_t)$. El contraste del factor período evalúa:

$$H_0 : \lambda_{\text{2000-2009}} = \lambda_{\text{2010-2019}} = \lambda_{\text{2020-2025}} \quad \text{v/s} \quad H_1 : \lambda_a \neq \lambda_b \mid a \neq b.$$

Si no se rechaza H_0, las diferencias entre promedios anuales son compatibles con fluctuaciones de una tasa común. Si se rechaza, al menos un tramo presenta una tasa anual distinta y no conviene resumir todo 2000-2025 con una única tasa temporal sin advertencia.

Como ejemplo con los datos del catálogo, el período 2000-2009 presenta una media de 45,5 eventos por año y en 2007 se registraron 58 eventos. Si tomamos $\mu = 45{,}5$ como tasa anual esperada para ese tramo, la probabilidad Poisson de observar exactamente 58 eventos en un año es:

$$P(Y = 58) = \frac{e^{-45{,}5}45{,}5^{58}}{58!} = 0{,}0108.$$

Así, bajo una media de 45,5 eventos anuales, observar exactamente 58 eventos tiene probabilidad cercana al 1,1 %. Este cálculo no prueba por sí solo un cambio de tasa; el contraste usa todos los años y compara los períodos completos. Su aporte es mostrar por qué revisamos sobredispersión: si varios años quedan lejos de su media esperada, la variabilidad observada puede superar la prevista por Poisson. En ese caso usamos quasi-Poisson, que conserva la comparación de tasas pero corrige errores estándar y valores p.

10.11. Modelo exponencial y prueba de ajuste de los tiempos entre eventos

Este anexo explica el modelo utilizado para evaluar los tiempos transcurridos entre terremotos consecutivos. Mientras el modelo Poisson describe el número de eventos en intervalos fijos, la distribución exponencial describe cuánto tiempo transcurre hasta el siguiente evento. Bajo un proceso de ocurrencia con tasa constante, ambas representaciones son complementarias (*Ogata 1988*).

Si T representa el tiempo de espera en días y λ la tasa diaria de ocurrencia, su función de distribución es

$$F(t) = P(T \le t) = 1 - e^{-\lambda t}, \qquad t \ge 0.$$

La media esperada es $1/\lambda$ y la mediana teórica es $\log(2)/\lambda$. Una tasa mayor implica esperas más breves; una tasa menor, intervalos más largos. En cada grupo estimamos la tasa mediante $\hat{\lambda} = 1/\overline{T}$, donde $\overline{T}$ es el promedio de los intervalos observados.

La prueba de Kolmogorov-Smirnov compara la distribución acumulada empírica de los tiempos, $\widehat{F}_n(t)$, con la distribución exponencial ajustada, $F_{\widehat{\lambda}}(t)$ (*Conover 1971*; *Marsaglia et al. 2003*). El estadístico es la mayor distancia vertical entre ambas curvas:

$$D = \sup_t \left| \widehat{F}_n(t) - F_{\widehat{\lambda}}(t) \right|.$$

Un valor de D mayor indica una discrepancia más marcada, pero su interpretación depende del tamaño muestral. Además, el valor p convencional de Kolmogorov-Smirnov supone que la distribución de referencia está completamente especificada. En este análisis λ se estima con los mismos tiempos que se contrastan, por lo que el valor p se obtuvo mediante bootstrap paramétrico (*Durbin 1973*) (*véase anexo de bootstrap paramétrico para pruebas de ajuste*).

Como ejemplo, para los 386 intervalos del grupo $M \ge 7{,}0$ estimamos una tasa de 0,0408 eventos por día. Esta tasa corresponde a una espera media aproximada de $1/0{,}0408 = 24{,}5$ días y a una mediana teórica de $\log(2)/0{,}0408 = 17{,}0$ días, cercana a la mediana observada de 15,6 días. Sin embargo, la prueba considera la distribución completa y no solo su centro. El resultado $D = 0{,}0778$ y $p_{\text{boot}} = 0{,}001$ indica que las diferencias a lo largo de la distribución son demasiado grandes para describir el grupo combinado mediante una única exponencial.

Al separar los eventos por categoría de magnitud, los grupos Mayor ($p = 0{,}102$) y Grande o extremo ($p = 0{,}531$) resultaron compatibles con el ajuste exponencial, mientras que el grupo Fuerte lo rechazó ($p < 0{,}001$). Esto no demuestra que los primeros sigan exactamente un proceso de tasa constante; indica que, con los datos disponibles, no se detectó una discrepancia suficiente para descartar ese modelo. La diferencia entre el resultado combinado y los resultados por categoría muestra que reunir eventos con ritmos distintos puede producir un ajuste inadecuado.

10.12. Modelo binomial negativo para conteos con sobredispersión

Este anexo presenta la formulación e interpretación del modelo binomial negativo utilizado para comparar la frecuencia entre zonas. La evidencia que justifica su elección se presenta en el anexo de diagnósticos (*véase anexo de diagnósticos de los modelos*). La binomial negativa se usa cuando la respuesta es un conteo, pero la variabilidad observada supera la prevista por Poisson (*McCullagh y Nelder 1989*; *Cameron y Trivedi 1990*). En esta comparación, cada observación corresponde a una combinación zona-año y la respuesta es el número de terremotos registrados. El punto de partida natural es Poisson; sin embargo, ese modelo supone $\text{Var}(Y_i) = \mu_i$, condición que no siempre se cumple en conteos sísmicos agregados.

El modelo binomial negativo mantiene la media positiva mediante enlace logarítmico, igual que un GLM de conteos, pero incorpora un parámetro de dispersión θ (*Venables y Ripley 2002*):

$$Y_i \sim \text{BN}(\mu_i, \theta), \qquad \log(\mu_i) = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi}.$$

Su diferencia central está en la varianza:

$$\text{Var}(Y_i) = \mu_i + \frac{\mu_i^2}{\theta}.$$

Cuando θ es grande, el segundo término se aproxima a cero y el modelo se parece a Poisson. Cuando θ es finito, la varianza queda por encima de la media y el modelo reconoce la sobredispersión. Esta corrección no cambia la pregunta sustantiva, pero evita que los errores estándar sean demasiado pequeños y que los intervalos de confianza queden artificialmente estrechos.

En este informe se utiliza la binomial negativa en la comparación de frecuencia entre zonas porque se requieren razones de tasa, intervalos de confianza y contraste global mediante verosimilitud. La especificación quasi-Poisson también corrige la varianza, pero no define una verosimilitud completa; por ello es suficiente para contrastes temporales simples, pero menos adecuada cuando se comparan modelos y se reportan razones de tasa.

Como ejemplo con los datos del catálogo, el Cinturón de Fuego registra 892 eventos en 26 años, equivalente a 34,3 eventos por año. El Resto del mundo registra 203 eventos, equivalente a 7,8 eventos por año. La razón descriptiva entre ambas tasas es:

$$\frac{7{,}8}{34{,}3} = 0{,}228.$$

El modelo binomial negativo formaliza esa comparación. Con el Cinturón de Fuego como referencia, una razón de tasa cercana a 0,228 indica que la tasa anual estimada del Resto del mundo equivale aproximadamente al 22,8 % de la tasa del Cinturón de Fuego. El aporte del modelo es acompañar esa razón con incertidumbre corregida por sobredispersión y permitir contrastar si las tasas difieren entre zonas.

10.13. Medidas de tamaño de efecto

Un valor p indica si una diferencia es estadísticamente detectable, pero no cuán grande es; con 1.186 eventos, diferencias irrelevantes en la práctica pueden producir valores p pequeños. Por eso cada análisis del informe acompaña su dócima con una medida de tamaño de efecto, que cuantifica la magnitud de la diferencia en una escala interpretable. La medida depende del tipo de análisis y se resume por sección:

Tabla 31. Medida de tamaño de efecto utilizada en cada sección de resultados.

Sección de resultados	Análisis	Medida de tamaño de efecto
Comparabilidad del registro	Asociaciones categóricas	V de Cramér corregida
Comparabilidad del registro	Kruskal-Wallis del RMS	Epsilon cuadrado
Frecuencia entre zonas	Modelo binomial negativo	Razones de tasa y de densidad
Magnitud y profundidad entre zonas	Kruskal-Wallis	Epsilon cuadrado
Magnitud y profundidad entre zonas	Comparaciones por pares Dunn-Holm	Delta de Cliff
Eventos fuertes y extremos	Tabla de contingencia	V de Cramér corregida
Eventos fuertes y extremos	Regresión logística	Razón de odds
Clasificación de zonas	Regresión multinomial	Razones de odds

La V de Cramér y epsilon cuadrado se mueven entre 0 (sin asociación) y 1; delta de Cliff, entre −1 y 1, con 0 indicando superposición total entre las dos distribuciones comparadas; las razones de tasa y de odds comparan cada zona con la referencia, con el 1 como valor de igualdad. La estructura temporal no requirió una medida de asociación: sus resultados se expresan directamente como tasas anuales por período. El desarrollo de cada medida se encuentra en su anexo (*véanse los anexos de corrección por sesgo de la V de Cramér, cálculo de epsilon cuadrado y razón de odds*); delta de Cliff se define en la especificación de dócimas y modelos (*véase anexo de especificación de dócimas y modelos*). Ninguna conclusión del informe se establece con el valor p aislado: la decisión combina el valor p, el tamaño de efecto y su intervalo de confianza.

10.14. Verosimilitud y estimación de los modelos

Este anexo explica cómo se estiman y comparan los modelos mediante máxima verosimilitud. Su propósito es fundamentar los coeficientes y contrastes utilizados en distintas secciones del informe.

Estimar un modelo consiste en asignar valores numéricos a sus coeficientes β a partir de los datos; esos valores cuantifican la dirección y el tamaño de cada relación, y de ellos derivan todas las cifras reportadas: razones de odds o de tasas, intervalos y contrastes. En este informe esa asignación se realizó por máxima verosimilitud (*Fisher 1922*). Para cada observación i, $P(y_i \mid \beta)$ denota la probabilidad que el modelo con coeficientes β asigna al resultado efectivamente observado y_i; en la regresión logística de eventos fuertes, y_i vale 1 si el evento alcanzó $M \geq 7{,}0$ y 0 si no. La verosimilitud del catálogo completo es:

$$L(\beta) = \prod_{i=1}^{n} P(y_i \mid \beta),$$

es decir, la probabilidad de haber observado exactamente el catálogo que se observó. La estimación escoge el vector $\hat{\beta}$ que maximiza L, y de ahí el nombre del criterio: los coeficientes bajo los cuales lo observado resulta más probable, que son los que aparecen en las tablas de resultados de cada modelo.

Existen otros criterios de estimación, como los mínimos cuadrados de la regresión lineal clásica, que escogen los coeficientes que minimizan las diferencias al cuadrado entre lo observado y lo ajustado. Se usó máxima verosimilitud porque es el criterio general para respuestas binarias, de conteo y multinomiales, donde los supuestos de los mínimos cuadrados no se sostienen (con respuesta continua y errores normales ambos criterios coinciden), y porque sus estimadores tienen propiedades conocidas en muestras grandes: errores estándar, intervalos y contrastes derivan de la misma función L, propiedades aplicables con los 1.186 eventos del catálogo.

La devianza reexpresa la verosimilitud como medida de desajuste:

$$D = -2 \log L,$$

salvo una constante que depende solo de los datos y se cancela al comparar modelos. Cumple aquí el papel que la suma de cuadrados de residuos cumple en la regresión lineal: mientras menor, mejor el ajuste. El logaritmo convierte el producto de probabilidades en una suma manejable y el factor -2 hace que las diferencias de devianza entre modelos anidados sigan una distribución chi-cuadrado, lo que habilita el contraste siguiente.

La razón de verosimilitudes (LR) contrasta una variable comparando la devianza del modelo con y sin ella:

$$LR = D_{\text{sin la variable}} - D_{\text{con la variable}},$$

que bajo la hipótesis nula sigue una distribución chi-cuadrado con tantos grados de libertad como coeficientes aporta la variable. Evaluada en la regresión logística de eventos fuertes:

$$LR_{\text{zona}} = 1497{,}37 - 1490{,}01 = 7{,}3582,$$

el valor reportado en la tabla de contrastes de esa sección. Las devianzas de los modelos involucrados fueron:

Tabla 32. Devianzas de los modelos logísticos de eventos fuertes y contrastes de razón de verosimilitudes derivados.

Modelo	Devianza	LR frente al completo (gl)	Valor p
Completo (zona + profundidad)	1490,012	-	-
Sin zona (solo profundidad)	1497,370	7,358 (3)	0,0613
Sin profundidad (solo zona)	1490,323	0,312 (1)	0,5766
Nulo (solo intercepto)	1497,995	-	-

Ante separación, la estimación de máxima verosimilitud diverge y se reemplaza por la corrección de reducción de sesgo (*Kosmidis y Firth 2011*), según se declara en la clasificación de zonas (*véase anexo de separación y reducción de sesgo*).

10.15. Modelo de regresión logística

Este anexo presenta el modelo utilizado para estimar la probabilidad de que un evento alcance $M \geq 7{,}0$.

La regresión logística modela una respuesta binaria (en este informe, si un evento alcanza $M \geq 7{,}0$) en la escala logit (*Berkson 1944*; *McCullagh y Nelder 1989*):

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k,$$

donde p es la probabilidad de la respuesta y $p/(1-p)$ son sus odds (*véase anexo de la razón de odds*). La escala logit recorre todos los números reales, de modo que cualquier combinación de coeficientes produce una probabilidad válida entre 0 y 1; una regresión lineal aplicada a la respuesta binaria, en cambio, puede producir probabilidades menores que 0 o mayores que 1 y no sostiene sus supuestos de residuos.

Al aplicar la función exponencial, cada coeficiente se convierte en una razón de odds. Esa lectura directa motivó preferir el enlace logit frente a alternativas como el probit, cuyos coeficientes carecen de una interpretación equivalente. La regresión multinomial de la clasificación de zonas es la extensión de este modelo a una respuesta con más de dos categorías (*Venables y Ripley 2002*): estima una ecuación como la anterior por cada zona alternativa frente a la referencia (*véase anexo del modelo de regresión logística multinomial*).

10.16. Razón de odds

Este anexo explica cómo se interpretan los efectos estimados por los modelos logísticos mediante razones de odds. Las odds de un suceso son el cociente entre la probabilidad de que ocurra y la de que no ocurra, y la razón de odds compara dos grupos dividiendo sus odds (*Bland y Altman 2000*):

$$\text{odds} = \frac{p}{1-p}, \qquad OR = \frac{\text{odds}_{\text{grupo}}}{\text{odds}_{\text{referencia}}}.$$

Evaluada con el catálogo para la Dorsal Meso-Atlántica frente al Cinturón de Fuego en el umbral $M \geq 7{,}0$:

$$OR = \frac{3/25}{295/597} = \frac{0{,}120}{0{,}494} = 0{,}243.$$

Esta es la razón cruda: proviene solo de los conteos de la tabla. El modelo logístico repite la misma comparación entre eventos de igual profundidad y la estima mediante la exponencial del coeficiente de la Dorsal, $e^{-1{,}3986} = 0{,}247$; esa versión ajustada es la que reporta el informe. Para las cuatro zonas:

Tabla 33. Probabilidad, odds y razones de odds de alcanzar $M \geq 7{,}0$ por zona.

Zona	$M \geq 7{,}0$/resto	Probabilidad	Odds	OR crudo	OR modelo
C. de Fuego (ref.)	295 / 597	0,331	0,494	1	1
Resto del mundo	69 / 134	0,340	0,515	1,042	1,032
C. Alpino-Himalayo	20 / 43	0,317	0,465	0,941	0,950
D. Meso-Atlántica	3 / 25	0,107	0,120	0,243	0,247

La cercanía entre las razones crudas y las del modelo muestra que el ajuste por profundidad casi no altera las comparaciones entre zonas, coherente con su contraste no significativo ($LR = 0{,}312$). Un valor de 1 indica odds iguales a la referencia; menor que 1, odds menores; mayor que 1, odds mayores.

La razón de odds se reporta porque es la medida que la regresión logística entrega de forma directa. Su lectura exige una precaución: las odds no son probabilidades, y sus porcentajes no deben confundirse.

La diferencia está en el denominador. En la Dorsal hay 28 eventos y 3 son mayores. La probabilidad divide por el total; las odds dividen por los eventos que no son mayores:

$$\text{probabilidad} = \frac{3}{28} = 0{,}107, \qquad \text{odds} = \frac{3}{25} = 0{,}120.$$

En el Cinturón de Fuego, con 892 eventos y 295 mayores, la probabilidad es $295/892 = 0{,}331$ y las odds son $295/597 = 0{,}494$.

Cada medida produce su propia comparación entre zonas. En odds: $0{,}120/0{,}494 = 0{,}243$, un 75,7 % menos. En probabilidad: $0{,}107/0{,}331 = 0{,}324$, un 67,6 % menos. Los dos porcentajes son correctos, pero miden cosas distintas y no deben intercambiarse.

Solo cuando el suceso es muy poco frecuente ambas medidas casi coinciden, porque el total y los eventos que no superan el umbral pasan a ser casi el mismo número. En este catálogo no ocurre: 32,6 % de los eventos supera el umbral. Por eso la sección de resultados dice que las odds de la Dorsal son 75,3 % menores y no que un evento mayor sea 75 % menos probable.

10.17. Intervalos de confianza y el nivel del 95 %

Un intervalo de confianza del 95 % es el resultado de un procedimiento que, aplicado a muchas muestras hipotéticas generadas bajo las mismas condiciones, contendría el valor real del parámetro en el 95 % de ellas (*Neyman 1937*). El intervalo calculado lo contiene o no lo contiene; el 95 % describe la confiabilidad del procedimiento, no una probabilidad sobre el intervalo particular. El nivel es el complemento del nivel de significación de 5 % adoptado en el informe, de modo que un intervalo del 95 % que excluye el valor nulo (el 1 en las razones de tasa o de odds) corresponde a un contraste con $p < 0{,}05$; ambos umbrales se fijaron antes de observar los resultados.

Se usaron dos métodos de cálculo. El intervalo de Wald, simétrico alrededor de la estimación:

$$\hat{\beta} \pm 1{,}96\,\widehat{EE}(\hat{\beta}),$$

en las comparaciones por pares (*véase anexo del contraste de Wald*), y el de verosimilitud perfilada, que no supone simetría y es preferible con pocos eventos (*Venables y Ripley 2002*). La comparación de ambos métodos en el modelo de eventos fuertes muestra dónde difieren:

Tabla 34. Intervalos de Wald y perfilados para las razones de odds del modelo de eventos fuertes.

Término	OR	IC 95 % de Wald	IC 95 % perfilado
Resto del mundo	1,032	[0,746; 1,427]	[0,744; 1,423]
C. Alpino-Himalayo	0,950	[0,548; 1,646]	[0,538; 1,625]
D. Meso-Atlántica	0,247	[0,074; 0,826]	[0,058; 0,712]
Profundidad (por km)	1,0002	[0,9995; 1,0009]	[0,9995; 1,0009]

En las zonas de mayor tamaño ambos métodos casi coinciden. En la Dorsal Meso-Atlántica difieren de forma visible: con 3 eventos sobre el umbral la incertidumbre no se reparte por igual a ambos lados, el intervalo perfilado lo refleja y es el que reporta el informe.

10.18. Contraste de Wald

Este anexo explica el contraste utilizado para evaluar coeficientes y comparaciones por pares a partir de sus errores estándar. El contraste de Wald determina si un coeficiente difiere de cero sin reajustar el modelo:

$$z = \frac{\hat{\beta}}{\widehat{EE}(\hat{\beta})},$$

donde $\hat{\beta}$ es el coeficiente estimado y $\widehat{EE}(\hat{\beta})$ su error estándar; bajo la hipótesis nula de coeficiente igual a cero, z se aproxima a una distribución normal estándar en muestras grandes (*Wald 1943*). Evaluado en la regresión logística de eventos fuertes para la Dorsal Meso-Atlántica:

$$z = \frac{-1{,}3986}{0{,}6159} = -2{,}271, \qquad p = 0{,}0231.$$

Para el conjunto de coeficientes del modelo:

Tabla 35. Contrastes de Wald de los coeficientes del modelo logístico de eventos fuertes.

Término	$\hat{\beta}$	$\widehat{EE}(\hat{\beta})$	z	Valor p
Resto del mundo	0,0317	0,1653	0,192	0,8481
C. Alpino-Himalayo	-0,0512	0,2804	-0,182	0,8552
D. Meso-Atlántica	-1,3986	0,6159	-2,271	0,0231
Profundidad (por km)	0,0002	0,0004	0,560	0,5751

El contraste de Wald se utilizó donde se necesitan varias comparaciones de un mismo ajuste, pues todas salen de coeficientes y errores estándar ya calculados: los intervalos por pares en frecuencia y los contrastes de cada zona en la regresión logística (las cuatro filas de la tabla anterior provienen de un único modelo). Las decisiones globales, en cambio, se tomaron con la razón de verosimilitudes, por dos razones. Primero, Wald evalúa un coeficiente a la vez: para la zona entrega tres valores p (0,8481, 0,8552 y 0,0231) y ninguno responde por la variable completa, mientras que la razón de verosimilitudes la contrasta como bloque y entrega una sola respuesta ($LR = 7{,}358$, $p = 0{,}0613$) (*véase anexo de verosimilitud y estimación de los modelos*). Segundo, Wald falla ante separación, es decir, cuando una zona presenta siempre el mismo resultado. El caso existe en el catálogo: en el umbral $M \geq 7{,}8$ la Dorsal Meso-Atlántica registra 0 eventos entre 28, de modo que sus odds son:

$$\text{odds} = \frac{0}{28} = 0,$$

y el modelo logístico, que trabaja con el logaritmo de las odds, necesitaría para esa zona un coeficiente igual a $\log(0) = -\infty$, un valor que ningún algoritmo puede estimar. En la práctica la estimación diverge hacia valores cada vez más extremos y su error estándar crece aún más rápido; como z es el cociente entre ambos, el estadístico se encoge y el valor p se acerca a 1. Wald concluiría que la zona no tiene efecto justo cuando el efecto es total (*Hauck y Donner 1977*). Por eso la categoría $M \geq 7{,}8$ se describe sin regresión, y en la clasificación de zonas, donde la separación afecta a la profundidad de la Dorsal, la inferencia se aborda con la corrección de reducción de sesgo.

10.19. Criterio de información de Akaike

Este anexo explica el criterio utilizado para comparar modelos considerando conjuntamente ajuste y complejidad. El AIC orienta la selección, pero no constituye una dócima de hipótesis.

El AIC pondera el ajuste de un modelo y la cantidad de parámetros que necesita para lograrlo (*Akaike 1974*):

$$\mathrm{AIC} = -2 \log L + 2k,$$

donde L es la verosimilitud maximizada y k el número de parámetros. La penalización $2k$ protege contra el sobreajuste. El AIC no se interpreta en valor absoluto: solo compara modelos ajustados sobre los mismos datos y se prefiere el menor; una diferencia menor que 2 indica respaldo prácticamente equivalente y se resuelve por parsimonia (*Burnham y Anderson 2004*). En la comparación entre las dos representaciones de la profundidad de los eventos fuertes, los dos modelos fueron:

Tabla 36. Comparación de los modelos logísticos según la representación de la profundidad.

Componente	Profundidad continua	Profundidad categórica
Especificación	zona + profundidad (km)	zona + categoría de profundidad
Parámetros estimados (k)	5	6
Devianza residual	1490,012	1488,261
LR de la profundidad (gl)	0,312 (1)	2,063 (2)
Valor p de la profundidad	0,5766	0,3566
LR de la zona (gl)	7,358 (3)	7,471 (3)
Valor p de la zona	0,0613	0,0583
AIC	1500,012	1500,261

La versión categórica ajusta apenas mejor (devianza 1,751 menor) al costo de un parámetro adicional, y la penalización del AIC revierte esa ventaja:

$$\Delta\mathrm{AIC} = 1500{,}261 - 1500{,}012 = 0{,}249 < 2,$$

por lo que se conservó la especificación continua, más simple y sin pérdida de información por categorización. Se usó AIC y no una prueba de significación porque las dos representaciones no son modelos anidados en el sentido requerido por la razón de verosimilitudes. El mismo criterio se aplicó en la clasificación de zonas, donde el modelo conjunto quedó a 0,899 del de solo profundidad, también por debajo de 2.

10.20. Análisis de correspondencia

Este anexo explica el análisis de correspondencia utilizado para explorar la relación entre las zonas y las categorías de magnitud y de profundidad. Es una técnica descriptiva para tablas de contingencia: resume en un plano de pocas dimensiones cuánto se apartan los perfiles observados de la independencia, la situación en que todas las zonas compartirían una misma composición (*Greenacre 2017*).

La medida de partida es la inercia total, que cuantifica la asociación global de la tabla. Equivale al estadístico chi-cuadrado de independencia (*véase anexo del contraste chi-cuadrado y frecuencias esperadas*) dividido por el número de observaciones:

$$\text{inercia total} = \frac{\chi^2}{n} = \phi^2,$$

donde n es el número de eventos. Es la misma cantidad ϕ^2 que sustenta la V de Cramér (*véase anexo de corrección por sesgo de la V de Cramér*): una inercia cercana a cero indica perfiles casi idénticos entre zonas, y una inercia mayor, perfiles más distinguibles.

La inercia se reparte en dimensiones, análogas a los ejes de un mapa. Cada dimensión tiene un autovalor λ_k, y su porcentaje sobre la inercia total indica cuánta asociación resume ese eje:

$$\text{porcentaje de la dimensión } k = \frac{\lambda_k}{\sum_j \lambda_j} \times 100.$$

Como las dos tablas tienen cuatro filas (zonas) y tres columnas (categorías), su geometría se agota en mín$(4-1,\ 3-1) = 2$ dimensiones: dos ejes representan la totalidad de la asociación. Las contribuciones señalan qué zonas y categorías definen cada eje, es decir, cuáles se apartan más del centro.

Como ejemplo con el catálogo, la tabla de zona por categoría de magnitud tiene $\chi^2 = 7{,}1222$ y $n = 1186$, de modo que su inercia total es

$$\frac{7{,}1222}{1186} = 0{,}0060,$$

el mismo valor de la sección de resultados. La tabla de profundidad alcanza una inercia de 0,0321, unas 5,3 veces mayor, y su primera dimensión concentra el 81,9 % de esa asociación: sus perfiles por zona son más distintos, sobre todo por la presencia o ausencia de eventos profundos.

La lectura del plano sigue tres reglas simples. Los puntos cercanos entre sí corresponden a perfiles semejantes; un punto próximo al origen tiene un perfil cercano al promedio del catálogo; y un punto alejado del origen es distintivo. Por eso, en el mapa de profundidad la Dorsal Meso-Atlántica queda apartada por ser exclusivamente superficial, mientras que el Cinturón de Fuego queda cerca del origen por cubrir todas las categorías. Esta etapa es descriptiva: sugiere qué variable separa mejor las zonas, pero no constituye prueba por sí sola, y su lectura se confirma con el modelo multinomial.

10.21. Modelo de regresión logística multinomial

Este anexo presenta el modelo utilizado para clasificar la zona de un evento a partir de su magnitud y profundidad. Es la extensión de la regresión logística a una respuesta con más de dos categorías (*véase anexo del modelo de regresión logística*): en lugar de una probabilidad de sí o no, el modelo asigna a cada evento cuatro probabilidades, una por zona, que suman 1 (*Venables y Ripley 2002*).

Con el Cinturón de Fuego como zona de referencia, el modelo estima una ecuación por cada zona alternativa:

$$\log\left(\frac{P(\text{zona}_h)}{P(\text{C. de Fuego})}\right) = \beta_{0h} + \beta_{1h}\,\text{magnitud} + \beta_{2h}\,\text{profundidad},$$

donde h recorre las tres zonas alternativas y cada ecuación compara las probabilidades relativas de una zona frente a la referencia. Si η_h denota el lado derecho de cada ecuación, las probabilidades se recuperan normalizando sus exponenciales:

$$P(\text{zona}_h) = \frac{e^{\eta_h}}{1 + \sum_j e^{\eta_j}}, \qquad P(\text{C. de Fuego}) = \frac{1}{1 + \sum_j e^{\eta_j}},$$

de modo que las cuatro probabilidades son positivas y suman 1. Se prefirió un único modelo multinomial en lugar de tres regresiones logísticas separadas porque las cuatro zonas son categorías excluyentes de una misma variable: el modelo conjunto garantiza probabilidades coherentes entre sí y permite contrastar cada variable con una sola dócima de tres grados de libertad, uno por ecuación.

Como ejemplo con el catálogo, para un evento con magnitud y profundidad iguales al promedio (puntajes Z de cero en ambas variables), los predictores lineales del modelo corregido se reducen a los interceptos: −1,5000 (Resto del mundo), −2,7017 (Cinturón Alpino-Himalayo) y −15,7648 (Dorsal Meso-Atlántica). Sus exponenciales son 0,2231, 0,0671 y 0,00000014, la suma del denominador es 1,2902 y la normalización entrega:

$$P(\text{C. de Fuego}) = \frac{1}{1{,}2902} = 0{,}775, \qquad P(\text{Resto}) = \frac{0{,}2231}{1{,}2902} = 0{,}173,$$

$$P(\text{C. Alpino-Himalayo}) = \frac{0{,}0671}{1{,}2902} = 0{,}052, \qquad P(\text{Dorsal M.-Atl.}) \approx 0.$$

Las tres primeras probabilidades quedan cerca de los pesos observados del catálogo (75,21 %, 17,12 % y 5,31 %), mientras que la de la Dorsal cae prácticamente a cero: su perfil exclusivamente superficial hace casi imposible, según el modelo, que un evento de profundidad promedio (cercana a 90 km) le pertenezca. El ejemplo muestra las dos caras del resultado de la sección de clasificación: las probabilidades se desplazan con la profundidad, pero el Cinturón de Fuego parte con una ventaja de peso muestral que ninguna zona alcanza a revertir.

10.22. Estandarización de los predictores

Este anexo explica la estandarización aplicada a la magnitud y la profundidad en el modelo de clasificación y su papel en la detección de la separación.

Estandarizar un predictor es expresarlo en desviaciones estándar respecto de su promedio:

$$z = \frac{x - \bar{x}}{s},$$

donde $\bar{x}$ es el promedio del predictor y s su desviación estándar. En el catálogo, s vale 0,417 unidades de magnitud y 164,1 km de profundidad, de modo que un puntaje $z = 1$ significa una desviación estándar sobre el promedio: 0,417 unidades más de magnitud o 164,1 km más de profundidad.

Es un cambio de escala, no de modelo. Las probabilidades ajustadas, la devianza y el AIC son idénticos en ambas versiones: con el catálogo, la devianza residual es 1728,757 y el AIC 1746,757 tanto en la escala original como en la estandarizada. Lo único que cambia es la unidad de los coeficientes, que pasan de efecto por unidad a efecto por desviación estándar, y las dos escalas se convierten multiplicando por s:

$$\beta_z = \beta_{\text{unidad}} \times s.$$

Como ejemplo con el catálogo, el coeficiente de profundidad del Resto del mundo es 0,00136 por kilómetro; multiplicado por los 164,1 km de una desviación estándar entrega $0{,}00136 \times 164{,}1 = 0{,}2235$, el mismo coeficiente del ajuste estandarizado.

La estandarización cumple dos funciones en la clasificación. Primero, hace comparables la magnitud y la profundidad: sus unidades originales no admiten comparación directa, pero una desviación estándar de cada una sí. Segundo, deja los coeficientes en una escala donde un valor extremo se reconoce a simple vista: el coeficiente de profundidad de la Dorsal Meso-Atlántica parece moderado en su unidad original ($-0{,}1685$ por km), porque por kilómetro cualquier coeficiente es numéricamente pequeño cuando la variable recorre cientos de kilómetros; reajustado con predictores estandarizados, el mismo efecto aparece como $-27{,}66$, el valor que delató la separación (*véase anexo de separación y reducción de sesgo*).

10.23. Separación y reducción de sesgo

Este anexo explica el problema de separación detectado en el modelo multinomial y la corrección con que se obtuvieron los coeficientes finales.

Hay separación cuando una variable distingue una categoría de la respuesta de manera casi perfecta. En el catálogo ocurre con la Dorsal Meso-Atlántica: sus 28 eventos son superficiales, sin ninguno intermedio ni profundo, de modo que basta una profundidad alta para descartar esa zona sin error. La separación es cuasi perfecta, y no perfecta, porque distingue a la Dorsal en un solo sentido: una profundidad alta la descarta sin error, pero una profundidad baja no la confirma, ya que los eventos superficiales son mayoría en todas las zonas (913 en el catálogo, de los cuales solo 28 pertenecen a la Dorsal). Ese solapamiento en la franja superficial es lo que impide una separación total. En esa condición la verosimilitud no alcanza su máximo en valores finitos: cada disminución adicional del coeficiente de profundidad de la Dorsal la sigue aumentando, y la estimación de máxima verosimilitud diverge (*Albert y Anderson 1984*). El algoritmo se detiene en un valor grande pero arbitrario; en el ajuste estandarizado el síntoma fue

$$\hat{\beta}_{\text{Dorsal, prof}} = -27{,}66, \qquad EE = 7{,}21,$$

un coeficiente extremo cuyo error estándar crece junto con él, lo que invalida la prueba de Wald sobre ese término (*Hauck y Donner 1977*) (*véase anexo del contraste de Wald*).

La corrección aplicada, conocida como corrección de Firth e implementada en `brglm2::brmultinom` en su extensión multinomial, ajusta las ecuaciones de estimación para reducir el sesgo medio de los estimadores; en los modelos logísticos equivale a maximizar una verosimilitud penalizada (*Kosmidis y Firth 2011*):

$$\ell^*(\beta) = \ell(\beta) + \tfrac{1}{2}\log\det I(\beta),$$

donde ℓ es la log-verosimilitud e $I(\beta)$ la matriz de información. La penalización garantiza estimaciones finitas aun con separación: cuando un coeficiente se dispara, la información asociada tiende a cero y el término de penalización cae, lo que frena la fuga hacia infinito.

Evaluada con el catálogo, la corrección dejó el coeficiente de la Dorsal en

$$\hat{\beta}^*_{\text{Dorsal, prof}} = -26{,}19, \qquad EE = 6{,}74,$$

equivalente a $-0{,}160$ por kilómetro. La corrección no busca reducir el coeficiente a un valor moderado, sino volverlo finito y con error estándar utilizable: $-26{,}19$ apenas se movió respecto del $-27{,}66$ de máxima verosimilitud y sigue siendo igual de extremo, porque es el patrón de los datos el que lo impone, no un error numérico. Por eso la corrección vuelve la estimación reportable sin eliminar la separación, y la razón de odds de la Dorsal en profundidad se informa como aproximadamente cero y sin un intervalo utilizable. En las zonas sin separación la corrección apenas mueve las estimaciones: el coeficiente de profundidad del Resto del mundo pasa de 0,2235 a 0,2246, la misma razón de odds de 1,25 reportada en los resultados. La corrección se usa solo para los coeficientes y las razones de odds; los contrastes globales de razón de verosimilitudes y el AIC se calculan sobre el modelo sin penalizar, porque comparan ajustes bajo el mismo criterio.

10.24. Métricas de clasificación

Este anexo explica las métricas con que se evaluó la capacidad clasificatoria del modelo multinomial. Todas derivan de la matriz de confusión, la tabla que cruza la zona observada de cada evento con la zona que el modelo le asigna bajo la regla de máxima probabilidad.

La exactitud global es la proporción de eventos correctamente asignados:

$$\text{exactitud global} = \frac{\text{eventos bien clasificados}}{n}.$$

Evaluada con el catálogo, el modelo asignó los 1.186 eventos al Cinturón de Fuego y acertó en sus 892:

$$\text{exactitud global} = \frac{892}{1186} = 0{,}7521.$$

Ese valor debe compararse con una vara mínima: la base trivial, la exactitud de una regla sin modelo que asigna siempre la zona mayoritaria. Como el Cinturón de Fuego reúne el 75,21 % del catálogo, la base trivial también es 0,7521: el modelo no supera lo que se lograría sin usar ninguna variable.

Con zonas tan desbalanceadas la exactitud global es engañosa por sí sola, porque acertar solo la clase mayoritaria basta para obtener un valor alto. Por eso se acompaña de la sensibilidad por zona, la proporción de eventos de cada zona que el modelo identifica correctamente:

$$\text{sensibilidad}_j = \frac{\text{eventos de la zona } j \text{ bien clasificados}}{\text{eventos de la zona } j},$$

y de la exactitud balanceada, su promedio, que pondera igual a las cuatro zonas sin importar su tamaño:

$$\text{exactitud balanceada} = \frac{1}{4}\sum_{j=1}^{4}\text{sensibilidad}_j.$$

Evaluadas con el catálogo, las sensibilidades fueron 1 (Cinturón de Fuego) y 0 en las tres zonas restantes, de modo que

$$\text{exactitud balanceada} = \frac{1+0+0+0}{4} = 0{,}25,$$

el mismo valor de la regla trivial, que también reconoce solo a la zona mayoritaria. La lectura conjunta es directa: el modelo no aporta capacidad clasificatoria adicional sobre la regla que asigna todo al Cinturón de Fuego. Las curvas ROC complementan esta evaluación: examinan si las probabilidades ajustadas contienen señal discriminante aunque la regla de máxima probabilidad no la convierta en asignaciones distintas (*véase anexo de curvas ROC y área bajo la curva*).

10.25. Curvas ROC y área bajo la curva

Este anexo explica la evaluación complementaria de la discriminación probabilística mediante curvas ROC y su resumen, el área bajo la curva (AUC), implementados con `pROC` (*Robin et al. 2011*).

Como la respuesta es multinomial, la evaluación se realiza una zona a la vez bajo el esquema uno contra el resto: los eventos de la zona evaluada son la clase positiva y todos los demás, la negativa. Para un umbral de probabilidad t, se definen:

$$\text{sensibilidad}(t) = P(\hat{p} \geq t \mid \text{zona}), \qquad \text{especificidad}(t) = P(\hat{p} < t \mid \text{otra zona}),$$

donde $\hat{p}$ es la probabilidad que el modelo asigna a la zona evaluada. La curva ROC dibuja la sensibilidad contra $1 -$ especificidad (los falsos positivos) al recorrer todos los umbrales posibles; su lectura no depende, por tanto, de fijar un umbral particular.

El AUC resume la curva en un número con una lectura de ordenamiento: es la probabilidad de que, tomados al azar un evento de la zona y otro de fuera de ella, el modelo asigne mayor probabilidad al primero:

$$\text{AUC} = P(\hat{p}_{\text{evento de la zona}} > \hat{p}_{\text{evento de otra zona}}).$$

Un AUC de 0,5 equivale a ordenar al azar y un AUC de 1, a un ordenamiento perfecto. Evaluado con el catálogo, el AUC de la Dorsal Meso-Atlántica fue 0,877: si se toma al azar un evento de la Dorsal y otro de fuera, en el 87,7 % de los pares el modelo asigna mayor probabilidad de Dorsal al evento que efectivamente le pertenece. Los AUC restantes fueron 0,689 (Cinturón de Fuego), 0,528 (Cinturón Alpino-Himalayo) y 0,470 (Resto del mundo), este último bajo 0,5 y por tanto sin señal discriminante útil, coherente con su carácter residual y heterogéneo.

La curva ROC complementa a la matriz de confusión porque responde otra pregunta (*véase anexo de métricas de clasificación*). La matriz juzga la etiqueta final: bajo la regla de máxima probabilidad la Dorsal nunca gana y su sensibilidad es 0. El AUC juzga el ordenamiento de las probabilidades: el 0,877 muestra que el modelo sí eleva la probabilidad de Dorsal en los eventos correctos, aunque nunca lo suficiente para superar al Cinturón de Fuego. Esa es la diferencia entre contener señal y convertirla en clasificación. Con solo 28 eventos en la clase positiva, el AUC de la Dorsal se interpreta además con cautela, como se declara en los resultados.

10.26. Ajuste de valores p por multiplicidad

En una prueba individual con $\alpha = 0{,}05$, el riesgo aceptado de error tipo I es 5 %. Cuando se realizan cinco pruebas independientes, la probabilidad de obtener al menos un falso positivo puede aproximarse mediante:

$$1 - (1 - 0{,}05)^5 = 0{,}226.$$

Por tanto, bajo ese supuesto ilustrativo, el riesgo de encontrar al menos una diferencia por azar puede alcanzar aproximadamente 22,6 %. Para reducir este problema se utilizó el procedimiento de Benjamini-Hochberg, que controla la proporción esperada de errores tipo I entre los resultados declarados como significativos (*Benjamini y Hochberg 1995*).

La reducción del error tipo I tiene como contrapartida un posible aumento del error tipo II. Un ajuste demasiado estricto puede hacer que una diferencia real deje de ser detectada. Benjamini-Hochberg se escogió porque controla la tasa de falsos descubrimientos y es menos restrictivo que los procedimientos orientados a evitar cualquier error dentro de la familia de pruebas. La documentación oficial de R señala que esta condición es menos estricta que el control del error familiar y que, por ello, suele conservar mayor potencia estadística (*R Core Team, s. f.-c*).

El procedimiento ordena los $m = 5$ valores p originales de menor a mayor:

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(5)}.$$

Para cada posición i, el valor ajustado se obtiene mediante:

$$\tilde{p}_{(i)} = \min_{j \geq i} \left(\frac{m}{j} p_{(j)}, 1 \right).$$

La implementación se realizó mediante `p.adjust(..., method = "BH")` del paquete `stats` de R. Los valores obtenidos fueron:

Contraste	Valor p original	Valor p ajustado
RMS por período	$< 0{,}0001$	$< 0{,}0001$
Período por tipo de magnitud	0,00010	0,00025
RMS por zona	0,01583	0,02638
Zona por fuente de magnitud	0,28047	0,35059
Zona por tipo de magnitud	0,67073	0,67073

Los valores ajustados se compararon con el nivel de 0,05. El RMS por zona, el RMS por período y la asociación entre período y tipo de magnitud permanecieron estadísticamente detectables después del ajuste. En consecuencia, la corrección redujo el riesgo de falsos descubrimientos sin modificar las decisiones obtenidas en este bloque.

Mientras que Benjamini-Hochberg se aplicó al bloque exploratorio de comparabilidad, las comparaciones por pares de carácter confirmatorio, es decir, las seis razones entre zonas de la sección de frecuencia y las seis comparaciones de Dunn por variable en magnitud y profundidad, se ajustaron mediante el procedimiento de Holm (*Holm 1979*). A diferencia de Benjamini-Hochberg, que controla la proporción esperada de falsos positivos, Holm controla la probabilidad de cometer al menos un falso positivo en toda la familia, criterio más estricto y apropiado para conclusiones confirmatorias.

El procedimiento ordena los m valores p de menor a mayor, $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(m)}$, y calcula el valor ajustado mediante:

$$\tilde{p}_{(i)} = \max_{j \leq i} \ \min \left((m - j + 1)\, p_{(j)},\, 1 \right).$$

El valor p más pequeño se multiplica por m; el siguiente, por $m - 1$, y así sucesivamente, de modo que solo el primero enfrenta el umbral más estricto. Por eso Holm es uniformemente más potente que la corrección de Bonferroni, que multiplica todos los valores por m, mientras controla el mismo error por familia. La implementación se realizó mediante `p.adjust(..., method = "holm")` del paquete `stats` de R (*R Core Team, s. f.-c*).

En la familia de seis comparaciones por pares de la sección de frecuencia, el mayor valor p ajustado fue 0,00045, de modo que las seis diferencias se mantuvieron significativas tras el ajuste.

10.27. Comparaciones completas Dunn-Holm

Este anexo presenta los resultados completos de las comparaciones por pares posteriores a Kruskal-Wallis. La justificación del ajuste por multiplicidad se desarrolla en su anexo específico (*véase anexo de ajuste de valores p por multiplicidad*). Los valores p ajustados mediante Holm son los que se utilizan para la decisión inferencial.

Tabla 37. Comparaciones completas de Dunn-Holm para magnitud y profundidad.

Variable	Comparación	Z	p sin ajustar	p ajustado
Magnitud	C. Fuego - Resto	-0,761	0,4466	1,0000
Magnitud	C. Fuego - C. Alp.-Him.	-0,141	0,8880	0,8880
Magnitud	Resto - C. Alp.-Him.	0,283	0,7771	1,0000
Magnitud	C. Fuego - Dorsal M.-Atl.	2,713	0,0067	0,0333
Magnitud	Resto - Dorsal M.-Atl.	2,877	0,0040	0,0241
Magnitud	C. Alp.-Him. - Dorsal M.-Atl.	2,374	0,0176	0,0705
Profundidad	C. Fuego - Resto	3,965	0,00007	0,00029
Profundidad	C. Fuego - C. Alp.-Him.	3,552	0,00038	0,00076
Profundidad	Resto - C. Alp.-Him.	1,073	0,2833	0,2833
Profundidad	C. Fuego - Dorsal M.-Atl.	7,020	$<$ 0,0001	$<$ 0,0001
Profundidad	Resto - Dorsal M.-Atl.	5,154	$<$ 0,0001	$<$ 0,0001
Profundidad	C. Alp.-Him. - Dorsal M.-Atl.	3,893	0,00010	0,00030

10.28. Diagnósticos de los modelos

Este anexo documenta los diagnósticos utilizados para decidir entre los modelos Poisson y binomial negativo en la comparación de frecuencia entre zonas. La formulación, justificación e interpretación del modelo binomial negativo se presentan en el anexo correspondiente (*véase anexo del modelo binomial negativo para conteos con sobredispersión*). Aquí se presentan únicamente las comprobaciones aplicadas al catálogo y sus resultados.

La sobredispersión se evaluó mediante tres vías concordantes.

La dispersión de Pearson es un índice descriptivo: la suma de los residuos de Pearson al cuadrado dividida por los grados de libertad residuales,

$$\hat{\phi} = \frac{1}{n-p} \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i},$$

con y_i los conteos observados, $\hat{\mu}_i$ los valores ajustados, $n = 104$ observaciones y $p = 4$ parámetros. Un valor cercano a 1 indica ausencia de sobredispersión; aquí $\hat{\phi} = 1{,}410$ (*McCullagh y Nelder 1989*).

La prueba de Cameron-Trivedi contrasta $H_0 : \mathrm{Var}(Y) = \mu$ frente a $H_1 : \mathrm{Var}(Y) = \mu + \alpha\,\mu$. Mediante una regresión auxiliar estima el parámetro de sobredispersión α y contrasta $\alpha = 0$ con un estadístico asintóticamente normal estándar; su forma exacta se encuentra en la fuente original (*Cameron y Trivedi 1990*).

La comparación de verosimilitudes entre Poisson y binomial negativa evalúa si el parámetro de dispersión mejora el ajuste, mediante

$$LR = 2\left(\hat{\ell}_{\mathrm{BN}} - \hat{\ell}_{\mathrm{P}}\right),$$

donde $\hat{\ell}_{\mathrm{BN}}$ y $\hat{\ell}_{\mathrm{P}}$ son las log-verosimilitudes maximizadas de la binomial negativa y de la Poisson.

Índice o prueba	Valor	Lectura
Dispersión de Pearson	1,410	Levemente por encima de 1
Prueba de Cameron-Trivedi	$z = 2{,}149$; $p = 0{,}0158$	Rechaza ausencia de sobredispersión
Razón de verosimilitudes Poisson vs binomial negativa	$LR = 5{,}628$; $p = 0{,}0088$	Favorece la binomial negativa
Parámetro de dispersión θ (binomial negativa)	42,35	Sobredispersión leve pero consistente

Tabla 38. Diagnóstico de sobredispersión del modelo de frecuencia.

Las tres comprobaciones apuntaron a una sobredispersión leve, pero consistente ($\theta = 42{,}35$), por lo que se utilizó la binomial negativa. La comparación de verosimilitudes contrasta si hace falta variabilidad adicional respecto de la Poisson. Conviene expresarlo con el inverso del parámetro de dispersión, $\alpha = 1/\theta$, que mide la cantidad de varianza extra: la Poisson corresponde a $\alpha = 0$ (es decir, θ infinitamente grande) y la binomial negativa, a $\alpha > 0$. Este contraste requiere un ajuste, porque α no puede ser negativo; bajo la hipótesis nula, $\alpha = 0$ queda situado en el extremo de su rango posible. En esa situación, el estadístico no sigue la distribución χ^2_1 habitual, que supone que el parámetro puede moverse hacia ambos lados del valor nulo. La distribución de referencia correcta asigna la mitad de la probabilidad al valor cero y la otra mitad a una χ^2_1. Como resultado, el valor p correcto es la mitad del que entrega la prueba ordinaria: 0,0088 en lugar de 0,0177 (*Self y Liang 1987*). La prueba χ^2_1 sin corregir es demasiado conservadora cuando el parámetro está en el extremo de su rango, de modo que la corrección ajusta el valor p a la baja y refuerza la evidencia de sobredispersión, sin modificar la conclusión.

El contraste global del efecto de zona, en cambio, no requiere esa corrección. A diferencia del parámetro de dispersión, los coeficientes de zona pueden tomar valores positivos o negativos, porque una zona puede presentar mayor o menor frecuencia que la referencia. Bajo la hipótesis nula no quedan, entonces, en el extremo de su rango, sino en su interior, y en esa situación el estadístico sí sigue una distribución χ^2 ordinaria, con tres grados de libertad, uno por cada zona comparada con el Cinturón de Fuego.

Las seis comparaciones por pares se obtuvieron mediante contrastes de Wald, es decir, a partir de los coeficientes ya estimados del modelo binomial negativo y de sus errores estándar, sin volver a ajustar el modelo. El cálculo se realizó sobre las dos versiones: el modelo de tasa anual, sin offset, y el de densidad, con el logaritmo de la superficie como offset. Para cada par, la razón entre dos zonas proviene de la diferencia de sus coeficientes, y el error estándar de esa diferencia, de la matriz de covarianzas del modelo. Los valores p se ajustaron mediante Holm dentro de cada familia de seis comparaciones.

Conviene una precisión de lectura: el ajuste de Holm se aplica solo a los valores p. Los intervalos de confianza del 95 % que acompañan a cada razón son individuales, es decir, no se ensancharon para considerar las seis comparaciones de forma simultánea.

Los intervalos de confianza de las razones frente a la zona de referencia, en la *Tabla 4*, se obtuvieron por verosimilitud perfilada. La verosimilitud mide qué tan bien el modelo reproduce los datos para cada valor posible del parámetro; su máximo es la estimación. El intervalo perfilado conserva todos los valores de la razón cuya verosimilitud, tras reajustar el resto de los parámetros, permanece dentro de un umbral basado en la χ^2. A diferencia del intervalo de Wald, que toma la estimación más o menos un múltiplo de su error estándar y supone que la incertidumbre es simétrica, la verosimilitud perfilada no impone esa simetría, por lo que es preferible en las zonas con pocos eventos, donde la razón se estima con incertidumbre asimétrica. En este catálogo ambos métodos difieren solo en la tercera cifra, de modo que ninguna conclusión depende de la elección.

10.29. Comparaciones completas por pares de frecuencia

Este anexo presenta los resultados completos de las comparaciones de frecuencia entre zonas en ambas métricas. El procedimiento de cálculo se explica en el anexo del contraste de Wald (*véase anexo del contraste de Wald*). Todos los valores p ajustados mediante Holm quedan por debajo de 0,001; el mayor es 0,00045, correspondiente al par Cinturón Alpino-Himalayo frente a Dorsal Meso-Atlántica en tasa anual.

Métrica	Comparación (zona 1 / zona 2)	$\|z\|$ (Wald)	p ajustado (Holm)
Tasa anual	C. Fuego / C. Alp.-Him.	19,32	$< 0{,}0001$
Tasa anual	C. Fuego / Dorsal M.-Atl.	17,61	$< 0{,}0001$
Tasa anual	C. Fuego / Resto	16,69	$< 0{,}0001$
Tasa anual	C. Alp.-Him. / Dorsal M.-Atl.	3,51	0,00045
Tasa anual	C. Alp.-Him. / Resto	7,78	$< 0{,}0001$
Tasa anual	Dorsal M.-Atl. / Resto	9,61	$< 0{,}0001$
Densidad	C. Fuego / C. Alp.-Him.	10,64	$< 0{,}0001$
Densidad	C. Fuego / Dorsal M.-Atl.	12,90	$< 0{,}0001$
Densidad	C. Fuego / Resto	41,48	$< 0{,}0001$
Densidad	C. Alp.-Him. / Dorsal M.-Atl.	4,66	$< 0{,}0001$
Densidad	C. Alp.-Him. / Resto	14,76	$< 0{,}0001$
Densidad	Dorsal M.-Atl. / Resto	5,54	$< 0{,}0001$

Tabla 39. Comparaciones por pares entre zonas en frecuencia: estadístico de Wald y valor p ajustado por Holm.

10.30. Declaración ampliada de uso de inteligencia artificial

☒ El equipo declara que sí utilizó herramientas de inteligencia artificial generativa durante el desarrollo de esta etapa de la asesoría, según se detalla a continuación:

Herramienta utilizada	Fecha aproximada de uso	Propósito del uso	Sección o producto asociado	Tipo de apoyo recibido	Verificación realizada por el equipo
Codex de OpenAI	julio de 2026	Apoyar la búsqueda de alternativas estadísticas frente a supuestos incumplidos o características detectadas por el equipo.	Modelos de conteo, análisis temporal, comparación de distribuciones y clasificación.	Búsqueda dirigida en la documentación de R, sugerencia de funciones y paquetes estadísticos e identificación de publicaciones originales asociadas.	Revisión de la ayuda de R y de la documentación de cada paquete, descarga y lectura de las publicaciones originales y apoyo de NotebookLM para precisar fuentes en inglés.
NotebookLM	julio de 2026	Apoyar la lectura y revisión de bibliografía publicada en inglés.	Antecedentes, metodología y anexos estadísticos.	Traducción de fragmentos, síntesis de documentos incorporados por el equipo y localización de contenidos dentro de las fuentes.	Comparación de las respuestas con los documentos originales y revisión de su correspondencia con las afirmaciones del informe.
Zotero	julio de 2026	Organizar y comprobar las fuentes utilizadas durante el análisis.	Biblioteca del proyecto, citas y archivo `referencias.bib`.	Gestión bibliográfica de autores, títulos, fechas, DOI, URL y documentos consultados.	Contraste de los registros con las publicaciones originales y comprobación de la información bibliográfica.
QGIS	julio de 2026	Calcular y verificar la superficie de las zonas utilizadas en la comparación espacial.	Variable `Area`, capa `zonas_inf2.geojson` y modelos de frecuencia por superficie.	Cálculo geométrico de las áreas, revisión de los polígonos y preparación de la exposición incorporada como offset.	Comprobación del sistema de referencia, inspección visual de las zonas y contraste de las superficies utilizadas en los modelos.
Codex de OpenAI	julio de 2026	Apoyar la organización documental y la revisión formal del informe.	Archivo QMD, anexos, citas, referencias cruzadas y presentación final.	Sugerencias de claridad, coherencia y sintaxis Quarto y LaTeX.	Revisión, modificación y aprobación manual de cada cambio por el equipo.

Tabla 40. Declaración detallada de las herramientas utilizadas durante el Informe Asesoría 2.

10.30.1. Detalle aclaratorio del uso

El procedimiento comenzaba con un diagnóstico realizado por el equipo. Cuando los resultados mostraban un supuesto incumplido o una característica que requería otro tratamiento, se consultaba a Codex si existían alternativas documentadas en R o en sus paquetes, de manera similar a una búsqueda mediante `help()`, `?función` o `??término`. En algunos casos, Codex sugirió directamente funciones o paquetes que podían implementar el procedimiento requerido. Estas sugerencias se consideraron puntos de partida y no decisiones metodológicas definitivas.

El equipo revisaba la ayuda de R, la documentación del paquete, los argumentos y ejemplos de las funciones y las referencias bibliográficas indicadas por sus autores. Cuando la documentación remitía a un artículo metodológico, el equipo localizaba y descargaba la publicación original para revisar sus supuestos, fundamentos y condiciones de aplicación. Si el texto estaba en inglés o requería precisar conceptos técnicos, se incorporaba a NotebookLM como apoyo para la lectura, la traducción de fragmentos y el contraste de la interpretación. La decisión final se basaba en la

documentación y en el artículo original, no únicamente en la respuesta de Codex o NotebookLM. Solo después de esta revisión se incorporaba la alternativa a los scripts y se contrastaban sus resultados con los diagnósticos iniciales.

Este procedimiento se aplicó frente a la sobredispersión de los conteos, en la revisión de estacionariedad e independencia serial, en la elección de una ruta no paramétrica ante el incumplimiento de los supuestos multivariantes y en las dificultades de estimación del modelo multinomial.

La verificación documental se apoyó además en Zotero, utilizado para organizar las publicaciones, comprobar sus metadatos y conservar la correspondencia entre autores, títulos, DOI, URL y archivos consultados. Zotero no corresponde a una herramienta de inteligencia artificial, pero se declara por su función dentro del proceso de trazabilidad y verificación. La selección de los métodos, su implementación, la ejecución de los análisis y la interpretación de los resultados permanecieron bajo responsabilidad de los integrantes.

QGIS se utilizó para calcular la superficie de los polígonos que representan las zonas de comparación y registrar esos valores en la variable `Area` de la capa `zonas_inf2.geojson`. Las superficies obtenidas se incorporaron posteriormente como exposición en los modelos de densidad. El equipo verificó el sistema de referencia, la delimitación de los polígonos y la correspondencia entre los valores calculados y los utilizados en el análisis. QGIS tampoco corresponde a una herramienta de inteligencia artificial, pero se incluye por su participación directa en la construcción de una variable utilizada por los modelos.

De manera complementaria, Codex se utilizó para organizar referencias bibliográficas y revisar aspectos formales del archivo QMD, como claridad, coherencia, referencias cruzadas y sintaxis Quarto y LaTeX. Las propuestas fueron revisadas y modificadas por el equipo antes de incorporarse al documento.

La inteligencia artificial no se utilizó como fuente de resultados estadísticos ni como sustituto de la evaluación metodológica. Los resultados incluidos en el informe proceden de los scripts reproducibles y fueron comprobados por el equipo mediante sus tablas, gráficos, diagnósticos y fuentes de respaldo. El equipo asume la responsabilidad por la selección metodológica, la exactitud de los resultados, su interpretación y la versión final del informe.

Este informe evalúa inferencialmente los principales patrones descritos en el Informe Asesoría 1 sobre 1.186 terremotos con $M \geq 6{,}5$ registrados por USGS entre 2000 y 2025. Se confirma que la principal diferencia entre zonas corresponde a la frecuencia. El Cinturón de Fuego concentra 892 eventos y una tasa de 34,3 por año, muy por encima de las demás zonas, y mantiene la mayor concentración después de considerar la superficie. Las fuentes y los métodos de reporte no condicionan materialmente esta comparación.

También se mantiene la similitud práctica de las magnitudes, cuyas distribuciones permanecen ampliamente superpuestas. La profundidad diferencia mejor a las zonas, especialmente por el carácter superficial de la Dorsal Meso-Atlántica y el rango más amplio presente en el Cinturón de Fuego. En cambio, las fluctuaciones observadas entre períodos no constituyen evidencia de un cambio sostenido en la tasa de ocurrencia. Tampoco se identifica estacionalidad mensual ni una composición de severidad diferente entre zonas.

La magnitud y la profundidad, consideradas conjuntamente, no permiten clasificar de manera confiable la zona de un evento. El modelo asigna todos los terremotos al Cinturón de Fuego y obtiene una exactitud de 75,21%, igual a clasificar siempre en la zona mayoritaria. En consecuencia, los patrones descriptivos de frecuencia y profundidad quedan respaldados, mientras que las diferencias temporales y de severidad requieren una interpretación más acotada.

Los resultados se limitan al catálogo analizado y no permiten estimar amenaza, riesgo sísmico ni probabilidades futuras de ocurrencia.

Análisis global
Distribución geográfica de terremotos extremos a nivel mundial.

Evolución temporal
Estudio de la variación histórica de eventos extremos.

Zonas más activas
Identificación de regiones sísmicas con mayor frecuencia y profundidad.

Modelamiento estadístico
Evaluación de patrones y diferencias significativas entre zonas.

www.tetrametric.cl

tetrametricssa@gmail.com

Santiago, Chile

USACH | UNIVERSIDAD DE SANTIAGO DE CHILE