

Anexos

Anexo a la Actividad 12: Análisis de la precipitación y su relación con los movimientos en masa

Resumen

Se presentan los hallazgos de una investigación detallada sobre cómo la precipitación influye en la ocurrencia de movimientos en masa. La lluvia incide en la inestabilidad de las laderas de dos maneras. Las lluvias intensas y de corta duración disminuyen la resistencia al corte del suelo, desencadenando frecuentemente movimientos en masa superficiales. Por otro lado, las precipitaciones de menor intensidad, pero prolongadas tienden a causar deslizamientos más profundos, generalmente a lo largo de superficies de falla preexistentes.

La complejidad de los movimientos en masa deriva no solo de estas causas directas, sino también de factores específicos del terreno y de las intervenciones humanas en las laderas. Además, la lluvia puede actuar a través de una combinación de eventos intensos de corta duración y acumulados de lluvia los días antecedentes.

Un patrón notable es el incremento en el número de emergencias por movimientos en masa durante periodos con mayor frecuencia de eventos de precipitación, especialmente en los meses de: marzo - abril - mayo y septiembre - octubre - noviembre. Esto presenta la significativa influencia de la precipitación en los movimientos en masa.

Es por esto que se vuelve relevante realizar estudios detallados de la lluvia y su influencia en los movimientos en masa. A nivel regional, se han empleado diversas metodologías para analizar y desarrollar herramientas capaces de identificar eventos de lluvia que incrementen la probabilidad de deslizamientos. Teniendo en cuenta los estudios realizados y encontrados en la literatura se plantean los análisis expuestos en este capítulo.

El proceso de investigación incluyó la recopilación de un inventario de movimientos en masa en el Valle de Aburrá y el registro histórico de lluvias asociadas a cada evento. Luego, se llevaron a cabo análisis detallados para comprender mejor la relación entre la lluvia y los movimientos en masa. Este informe presenta los métodos y resultados obtenidos, abarcando umbrales de lluvia reciente-antecedente, análisis de percentiles, modelos probabilísticos y umbrales basados en la relación Acumulado-Duración.

Datos

Inventario de movimientos en masa

La base de datos utilizada en este estudio se compila a partir de dos fuentes primarias. La primera proviene del Departamento Administrativo de Gestión del Riesgo y Desastres (DAGR), y la segunda, del Semillero de Investigación Geohazards de la Universidad Nacional.

El inventario del DAGRD se obtuvo tras una solicitud específica hecha por el proyecto SIATA. Esta base

de datos se construye a partir de reportes de emergencia recibidos a través de la línea de atención 123, los cuales informan sobre movimientos en masa. De este inventario, se recopilieron 4.153 eventos, cubriendo el período de septiembre de 2020 a marzo de 2023.

Las características de la base de datos incluyen:

- **Ubicación:** Determinada por la latitud y longitud del punto de registro del movimiento en masa.
- **Fecha y hora:** presenta la fecha y hora del evento.
- **Origen:** Causa probable del movimiento, clasificada como: *Natural*, *Antrópico*, *Natech* y *Socionatural*.
- **Tipología:** Clasificación del tipo de movimiento en masa, que incluye: *Desgarre Superficial*, *Deslizamiento*, *Movimientos complejos*, *Caída de rocas*, *Flujo*, *Reptación*, *Volcamiento*, *Subsidencia* y *Actividad Antrópica*.
- **Descripción:** Detalles del lugar en el momento de la inspección.
- **Evacuación:** Necesidad de evacuación al momento de la visita, con opciones de *Definitiva*, *Temporal* y *No requiere*.

En la Figura 1, se presentan dos gráficos. El primero muestra la distribución de registros según su origen, destacando el origen socionatural con 2.474 casos, seguido por antrópico con 688 y natural con 676. El segundo gráfico ilustra la distribución por tipo de movimiento en masa, resaltando desgarres superficiales, deslizamientos, movimientos complejos y flujos, con 1.703, 1.462, 279 y 131 registros respectivamente.

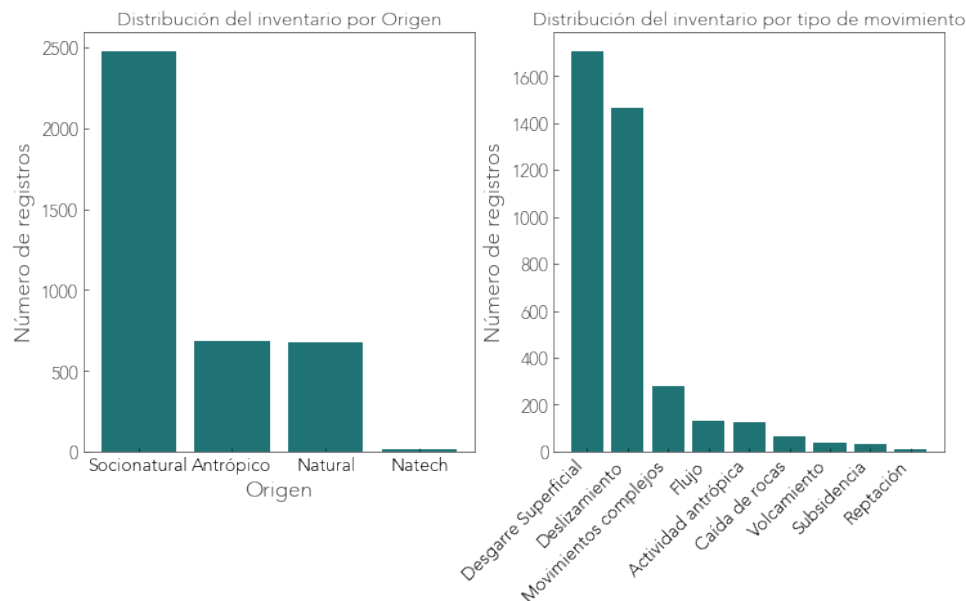


Figura 1: Distribución de los registros del inventario del DAGRD clasificados por su origen y tipología.

Por otra parte, el inventario del Semillero Geohazards ofrece información adicional sobre los otros municipios del Valle de Aburrá, complementando los datos del DAGRD, que se centran únicamente en la ciudad

de Medellín. La base de datos de Geohazards recopila registros de desastres o emergencias reportados por diferentes medios de comunicación y organismos de gestión del riesgo, abarcando desde 2012 hasta 2023. Esta selección temporal se alinea con el inicio del registro histórico de lluvias medido por pluviómetros y radar meteorológico del proyecto SIATA.

Debido a la incertidumbre en la fecha, hora y ubicación espacial de los eventos en este inventario, se realizó una depuración exhaustiva. Tras un análisis detallado, se seleccionaron 116 eventos con baja incertidumbre en la fecha y hora del evento y la ubicación espacial.

La Figura 2 muestra la distribución espacial de los eventos de ambos inventarios. Es notable que el inventario del DAGRD, que abarca toda el área de la ciudad de Medellín, representa el 97% de los eventos registrados. Este dato es crucial para análisis de lluvia a nivel regional en el Valle de Aburrá, ya que se espera en un futuro tener un mayor número de registros en los demás municipios. Además, se destaca que cada análisis implicó una rigurosa depuración del inventario, asegurando que los eventos seleccionados estuvieran principalmente relacionados con lluvias y no con otras causas, como factores antrópicos. En algunos casos, la depuración redujo los eventos de 4.268 a solo 700. Los métodos de depuración utilizados se detallan en cada capítulo del análisis correspondiente.

Datos de precipitación

Para el análisis de la precipitación, se utilizaron dos fuentes principales de registro de lluvias del SIATA: los pluviómetros y el radar meteorológico. En cuanto a los pluviómetros, se seleccionaron 35 estaciones con el registro histórico más extenso, que abarca desde 2012 hasta la actualidad. Estas estaciones que fueron las primeras instaladas por el proyecto, registran datos minuto a minuto.

La segunda fuente de datos proviene de los archivos históricos del Radar Meteorológico. Cabe mencionar que los análisis preliminares con estos datos no resultaron tan efectivos, por lo que durante este período no se realizaron más estudios con radar meteorológico, hasta lograr una mejor comprensión de estos datos.

La Figura 2 muestra la ubicación de los pluviómetros utilizados en los análisis, junto con la proximidad de los eventos del inventario de movimientos en masa. Además, se indica la posición del radar meteorológico.

En el caso de los pluviómetros para correlacionar los eventos de lluvia con los movimientos en masa del inventario, se consideraron los eventos que ocurrieron dentro de un radio de 5 km alrededor de cada pluviómetro. En casos donde un evento estuviera en el rango de dos estaciones (dentro del radio de 5 km), se asoció al pluviómetro más cercano al sitio del evento. Así, cada evento del inventario tiene asignado el pluviómetro más próximo (ver Figura 2).

El Radar Meteorológico ofrece una perspectiva espacial de la precipitación, realizando un barrido cada cinco minutos y cubriendo una distancia de aproximadamente 120 km en todas las direcciones. Este sistema proporciona una medida indirecta sobre la intensidad de la lluvia en momentos específicos.

Los datos del radar se pueden visualizar como una red de puntos o píxeles, formando una matriz en la que cada píxel representa la intensidad de la lluvia en un área determinada en un instante dado. Dado que el inventario de movimientos en masa incluye la latitud y longitud de cada evento, es posible identificar en la matriz de lluvia del radar el píxel o píxeles más cercanos al lugar del evento. Así, se extrae la serie histórica de precipitación para ese punto específico, correspondiente al sitio del movimiento en masa. Esta serie de datos proporciona información sobre la lluvia horaria en el punto desde el año 2013 hasta el año en curso.

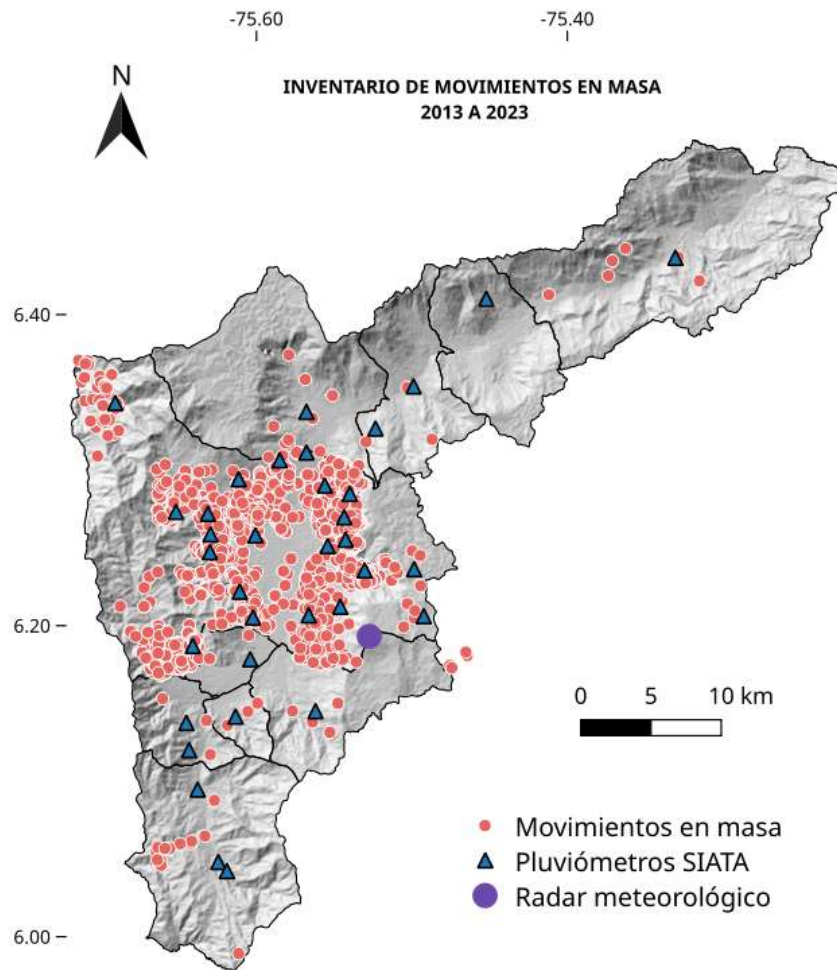


Figura 2: Mapa del Valle de Aburrá mostrando las ubicaciones de los eventos del inventario, con un total de 4.268 eventos recopilados del DAGRD y del Semillero Geohazards. Los eventos del DAGRD abarcan el período de 2020 a 2023, y los de Geohazards de 2013 a 2023. Los 35 pluviómetros seleccionados del SIATA están marcados con triángulos azules.

Limitaciones de los datos

A pesar de la rigurosa depuración realizada en los datos de movimientos en masa y precipitación, existen limitaciones significativas que deben considerarse al interpretar y valorar los resultados obtenidos. Estas limitaciones incluyen incertidumbres en la ubicación espacial y temporal de los movimientos en masa, la calidad de los datos de lluvia registrados, el área de influencia de las estaciones pluviométricas y las incertidumbres inherentes a las metodologías aplicadas.

Aunque se asume que los movimientos en masa analizados fueron desencadenados por lluvias, otros factores, como la actividad antrópica, podrían haber contribuido. La exactitud en la fecha y hora de cada evento está sujeta a la precisión del organismo o entidad que completó el registro, lo cual puede introducir un grado de subjetividad. Asimismo, la localización de los eventos no siempre refleja con exactitud el

lugar del suceso; en algunos casos, se basa en direcciones o áreas conocidas, con un margen de error que podría alcanzar decenas de metros.

La lluvia asociada a cada evento se determina por el pluviómetro más cercano, que puede estar situado a una distancia de entre 0 km y 5 km del evento. Esto, sumado a la variación en la altura de las estaciones pluviométricas, aumentando la incertidumbre sobre la lluvia específica que desencadenó el movimiento en masa. Además, la calidad de los datos de lluvia puede verse afectada por diversos factores, como daños en las estaciones, fallo de energía eléctrica del lugar o fallos en el registro y almacenamiento de los datos, lo que impide una representación precisa de las precipitaciones reales que provocaron los eventos.

Las metodologías empleadas en los análisis provienen de diversos estudios recientes. Sin embargo, muchas de estas investigaciones se diseñaron específicamente para las condiciones locales de cada región y sus características hidrometeorológicas, que pueden diferir de las del Valle de Aburrá y de cómo se recopilaban los datos en este estudio.

Finalmente, en algunas validaciones de las metodologías empleadas, no hay certeza absoluta sobre los eventos de lluvia seleccionados aleatoriamente que no desencadenaron movimientos en masa. Se asume que las fechas fuera del periodo registrado en el inventario son periodos sin movimientos en masa. No obstante, no se puede garantizar con certeza que no ocurrieran eventos de este tipo durante esas fechas seleccionadas aleatoriamente.

Metodologías de validación

Para evaluar la efectividad de los análisis realizados y su potencial aplicación en sistemas de alerta temprana para movimientos en masa detonados por lluvias, se implementaron diversas metodologías de validación. Estas incluyen enfoques empíricos, utilizando casos o eventos específicos de movimientos en masa no incluidos en la base de datos, y herramientas analíticas como tablas de contingencia (o matrices de confusión) y métricas como Accuracy y Curva AUC.

Se adoptó una clasificación binaria para diferenciar dos escenarios: aquellos donde es probable la ocurrencia de movimientos en masa y aquellos donde no lo es. Basado en esta clasificación, se desarrolló una matriz de confusión para examinar la relación entre eventos de lluvia y movimientos en masa. Las categorías utilizadas son:

- **TP (True Positive):** Eventos de lluvia que generaron movimientos en masa y están por encima del umbral establecido.
- **TN (True Negative):** Eventos de lluvia que no generaron movimientos en masa y están por debajo del umbral.
- **FN (False Negative):** Eventos por debajo del umbral que, sin embargo, resultaron en movimientos en masa.
- **FP (False Positive):** Eventos que no causaron movimientos en masa, pero se situaron por encima del umbral.

En la Figura 3, se ilustra la matriz de confusión, detallando cada uno de los valores mencionados. En sistemas de alertas tempranas, es fundamental minimizar los FN y FP. Utilizando estas categorías, se calcularon métricas clave:

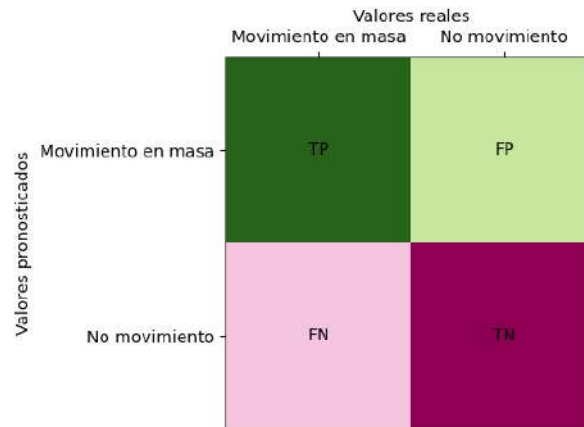


Figura 3: Esquema de la Matriz de Confusión.

- **Accuracy:** Proporción de predicciones correctas (TP + TN) respecto al total de observaciones (TP + FP + FN + TN). Indica la frecuencia con la que el clasificador acierta.
- **F1-Score:** Combina precisión y sensibilidad (Recall) en un solo valor. La precisión es la proporción de eventos correctos sobre el total predicho, mientras que el Recall refleja la capacidad del modelo para identificar correctamente eventos de lluvia que causan movimientos en masa.
- **Curva ROC - AUC:** Representación gráfica del rendimiento de un clasificador binario. La Curva ROC traza la Tasa de Verdaderos Positivos (TPR) contra la Tasa de Falsos Positivos (FPR) a diferentes umbrales. El AUC resume esta curva en un valor único entre 0 y 1.
- **Probabilidad de Detección (POD):** Tasa de Verdaderos Positivos, es decir, la proporción de eventos correctamente identificados.
- **Probabilidad de Falsa Detección (POFD):** Tasa de Falsos Positivos, refleja la proporción de predicciones incorrectas.
- **Probabilidad de Falsas Alarmas (POFA):** Relación entre falsas alarmas y el total de pronósticos correctos.
- **Hanssen y Kuipers (1965) HK:** Mide la precisión para eventos con y sin movimientos en masa. Se calcula como $HK = POD - POFD$.

Las fórmulas y valores óptimos para estas métricas se presentan en la Tabla 1.

Tabla 1: Métricas utilizadas para la validación.

Métricas	Fórmula	Rangos de variabilidad	Valor Óptimo
Accuracy	$Accuracy = \frac{TP+TN}{TP+FP+FN+TN}$	[0,1]	1
Recall	$Recall = \frac{TP}{TP+FN}$	[0,1]	1
Precision	$Precision = \frac{TP}{TP+FP}$	[0,1]	1
F1 - Score	$F1 - Score = \frac{2*(Recall*Precision)}{Recall+Precision}$	[0,1]	1
POD	$POD = \frac{TP}{TP+FN}$	[0,1]	1

Métricas	Fórmula	Rangos de variabilidad	Valor Óptimo
POFD	$POFD = \frac{FP}{FP+TN}$	[0,1]	0
POFA	$POFA = \frac{FP}{TP+FP}$	[0,1]	0
HK	$HK = POD - POFD$	[-1,1]	1

Umbral de Lluvia Reciente - Antecedente

Los umbrales de lluvia Reciente - Antecedente, a menudo referidos en la literatura como umbrales de lluvia antecedente Balducci et al. (2006), están relacionados con los tipos de umbrales de precipitación que se derivan utilizando metodologías estadísticas y empíricas. Estos umbrales se basan en el total acumulado de precipitación registrado en los días previos a un movimiento en masa. Para su elaboración, es fundamental contar con un inventario detallado de los movimientos en masa de la región y registros históricos de las precipitaciones del lugar.

Según Chleborad et al. (2006), los umbrales de precipitación acumulada antecedente se establecieron para Seattle, Washington, utilizando un inventario de 577 movimientos en masa inducidos por lluvias. Para definir estos umbrales, Chleborad et al. (2006) analizaron dos variables de precipitación antecedente, explicadas a continuación.

La primera variable, Lluvia Acumulada (LA) o precipitación antecedente a corto plazo, refleja el acumulado de lluvia registrado en los días previos al evento. Por ejemplo, un LA de 5 días evalúa la precipitación acumulada en esos 5 días previos. La segunda variable, Lluvia Acumulada Antecedente (LAA) o precipitación de largo plazo, representa el acumulado de lluvia en los días anteriores al período considerado para LA. Por ejemplo, un LAA de 15 días indicaría la precipitación acumulada durante los 15 días antes del período de 5 días de LA. La Figura 4 muestra un gráfico esquemático de una serie temporal ilustrando estas dos variables.

Finalmente, Chleborad et al. (2006) establecieron umbrales de lluvia utilizando un LA de tres días y un LAA de 15 días, integrándolos en un sistema de alerta temprana para movimientos en masa en Seattle.



Figura 4: Esquema del cálculo de la lluvia antecedente LA y LAA.

Para realizar este análisis, se realizó una depuración rigurosa del inventario de movimientos en masa. Se

estableció un filtro para seleccionar solo aquellos eventos potencialmente ocasionados o detonados por lluvias, excluyendo los movimientos con menos de 20 mm de lluvia acumulada en un día. También se descartaron eventos con más de 300 mm de lluvia acumulada en los 90 días previos al evento. Este criterio permitió filtrar y dejar solo los casos influenciados por lluvias recientes o antecedentes.

Paralelamente, de manera aleatoria se escogieron fechas desde el año 2013 hasta el 2023 donde no se registraron movimientos en masa. Estas fechas tenían más de 5 mm de lluvia acumulada en un día y más de 100 mm en 90 días, garantizando que no fueran días secos.

Este proceso de filtrado resultó en 801 eventos de movimientos en masa en el Valle de Aburrá. La Figura 5 muestra el mapa del inventario filtrado, utilizado para construir los umbrales de lluvia Reciente - Antecedente. Además, se seleccionaron 801 fechas aleatorias sin movimientos en masa pero con registros de lluvia, para equilibrar los datos y facilitar la validación de los umbrales.

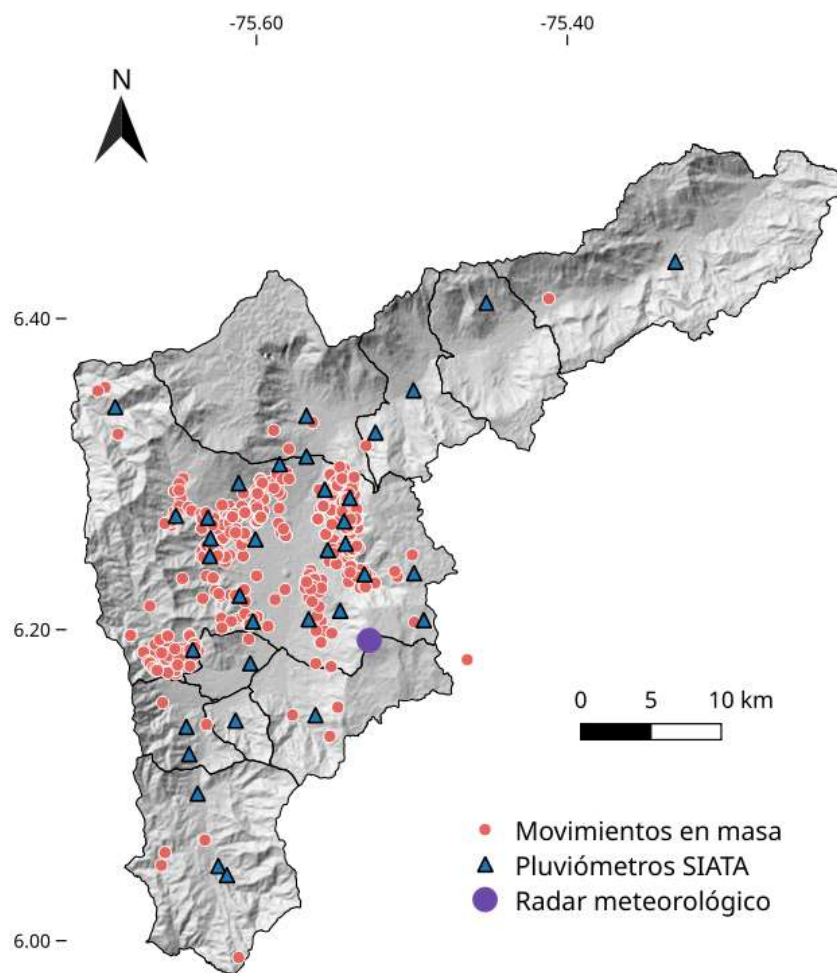


Figura 5: Esquema del calculo de la lluvia antecedente LA y LAA.

Tras compilar el inventario, se procedió a calcular la lluvia antecedente para diferentes intervalos temporales correspondientes a cada evento. Se realizaron variadas combinaciones de días para la Lluvia Acumulada

(LA) y la Lluvia Acumulada Antecedente (LAA), siguiendo parámetros reconocidos en investigaciones previas.

Para LA, los períodos estudiados fueron de 1, 3, 5 y 7 días, y para LAA, de 15, 30, 60 y 90 días. Combinando estos intervalos de LA y LAA, se buscó determinar cuáles ofrecían mayor claridad para distinguir entre los patrones de precipitación que incrementan la probabilidad de movimientos en masa y aquellos que no. El propósito era establecer un umbral claro.

La definición de este umbral se basó en el uso de técnicas de clasificación supervisada y algoritmos de aprendizaje automático. Se exploraron tres algoritmos: Regresión Logística, Máquina de Vectores de Soporte (SVM) y Análisis Discriminante Lineal (LDA), elegidos por su prevalencia en la literatura y su eficacia en clasificar linealmente entre grupos, diferenciando los eventos de lluvia que causan deslizamientos de los que no.

Para cada combinación de días considerada, se evaluó el rendimiento de los tres algoritmos de clasificación anteriormente mencionados. Con el objetivo de identificar cuál combinación de días y qué algoritmo mostraban un desempeño superior en la clasificación de eventos de lluvia que generan movimientos en masa frente a los que no, se calcularon diversas métricas de rendimiento. Estas incluyen Accuracy, F1-Score, la curva ROC y el Área Bajo la Curva ROC (AUC), cuyos detalles se explican en el capítulo correspondiente a . Para el entrenamiento y evaluación de cada algoritmo y combinación de días, se utilizó una división del 80 % de los datos para entrenamiento y el 20 % restante para evaluación.

Resultados de los Umbrales de Lluvia Reciente - Antecedente Los resultados de las métricas de **Accuracy** y **F1-Score**, expresados en porcentajes, se presentan en las Tablas 2 y 3 para cada combinación de días y algoritmo de clasificación evaluados. La Tabla 2 detalla los resultados obtenidos con los datos de los pluviómetros, mientras que la Tabla 3 muestra los resultados correspondientes a los datos recogidos por el radar meteorológico.

Tabla 2: Evaluación de algoritmos de clasificación con Accuracy y F1-Score para todas las combinaciones de días con lluvia registrada por pluviómetros.

Pluvios LA vs. LAA	Regresión Logística		SVM		LDA	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
1dvs.15d	85.67	85.67	88.16	88.16	81.93	81.93
1dvs.30d	87.23	87.23	88.47	88.47	83.18	83.18
1dvs.60d	90.03	90.03	88.47	88.47	85.05	85.05
1dvs.90d	89.1	89.1	90.03	90.03	88.47	88.47
3dvs.15d	77.26	77.26	78.19	78.19	75.39	75.39
3dvs.30d	78.19	78.19	77.57	77.57	75.08	75.08
3dvs.60d	78.5	78.5	78.82	78.82	77.57	77.57
3dvs.90d	81.0	81.0	81.93	81.93	80.06	80.06
5dvs.15d	74.77	74.77	74.77	74.77	76.32	76.32
5dvs.30d	75.39	75.39	77.57	77.57	76.01	76.01
5dvs.60d	76.01	76.01	77.57	77.57	76.32	76.32
5dvs.90d	80.69	80.69	80.69	80.69	80.06	80.06
7dvs.15d	75.08	75.08	76.01	76.01	75.7	75.7
7dvs.30d	78.19	78.19	77.57	77.57	78.5	78.5
7dvs.60d	78.82	78.82	78.5	78.5	79.44	79.44
Continúa en la siguiente página						

Tabla 2 continúa de la página anterior

Pluvios LA vs. LAA	Regresión Logística		SVM		LDA	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
7dvs.90d	80.37	80.37	82.24	82.24	80.37	80.37

Tabla 3: Evaluación de algoritmos de clasificación con Accuracy y F1-Score para todas las combinaciones de días con lluvia registrada por el radar meteorológico.

Radar LA vs. LAA	Regresión Logística		SVM		LDA	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
1dvs.15d	62.22	62.22	64.44	64.44	62.22	62.22
1dvs.30d	60.00	60.00	60.00	60.00	60.00	60.00
1dvs.60d	57.78	57.78	62.22	62.22	60.00	60.00
1dvs.90d	60.00	60.00	62.22	62.22	60.00	60.00
3dvs.15d	55.56	55.56	53.33	53.33	55.56	55.56
3dvs.30d	53.33	53.33	51.11	51.11	53.33	53.33
3dvs.60d	53.33	53.33	48.89	48.89	53.33	53.33
3dvs.90d	55.56	55.56	53.33	53.33	55.56	55.56
5dvs.15d	57.78	57.78	57.78	57.78	57.78	57.78
5dvs.30d	60.00	60.00	60.00	60.00	60.00	60.00
5dvs.60d	60.00	60.00	60.00	60.00	60.00	60.00
5dvs.90d	60.00	60.00	60.00	60.00	60.00	60.00
7dvs.15d	57.78	57.78	57.78	57.78	57.78	57.78
7dv.30d	62.22	62.22	55.56	55.56	62.22	62.22
7dvs.60d	60.00	60.00	55.56	55.56	60.00	60.00
7dvs.90d	60.00	60.00	60.00	60.00	60.00	60.00

Es notorio que los tres modelos de clasificación presentaron un rendimiento óptimo al usar datos de lluvia registrados por pluviómetros, en comparación con los datos obtenidos por radar. La mayoría de las combinaciones de días exhibieron un *Accuracy* y un *F1-Score* superiores al 75 %, lo cual sugiere que casi cualquier selección de días logra distinguir eficazmente entre las lluvias que incrementan la probabilidad de ocasionar movimientos en masa y aquellas que no.

Las combinaciones más efectivas, de acuerdo con las métricas evaluadas, son las de 1 día versus 90 días, 3 días versus 90 días, 5 días versus 90 días y 7 días versus 90 días, destacadas en negrita en la Tabla 2.

En contraste, los algoritmos de clasificación utilizados con datos de lluvia obtenidos por radar mostraron un rendimiento significativamente más bajo. Todos los valores de rendimiento fueron inferiores al 70 % en todas las combinaciones de días, como se puede observar en la Tabla 3. La mayoría de las métricas se situaron entre el 50 y el 60 %, evidenciando una limitada eficacia en la clasificación de las lluvias, independientemente de la combinación de días seleccionada. Estos resultados subrayan la importancia de explorar métodos alternativos para procesar datos de lluvia de radar y otras metodologías para evaluar los umbrales de lluvia basados en información de radar.

Además, las Figuras 6 y 7 ilustran la Curva ROC y el valor AUC de los modelos evaluados en las combinaciones de 1 día frente a 90 días y de 7 días frente a 90 días, respectivamente. En ambas figuras, se

aprecia que la Curva ROC de los modelos de Regresión Logística y Análisis Lineal Discriminante exhibe un comportamiento similar. En todos los casos, estos modelos alcanzaron valores AUC superiores a 0.90, lo que indica un rendimiento sobresaliente en la clasificación de la lluvia.

Las Figuras 8 y 9 presentan los umbrales definidos para las combinaciones de 1 día vs. 90 días y de 7 días vs. 90 días, respectivamente, que mostraron las mejores métricas con datos de pluviómetros. Además, se representan los eventos del 20 % de los datos reservados para validación, que no fueron parte del conjunto de entrenamiento. Los eventos de lluvia que no generaron movimientos en masa se indican con puntos azules, mientras que aquellos que sí lo hicieron se marcan en rojo.

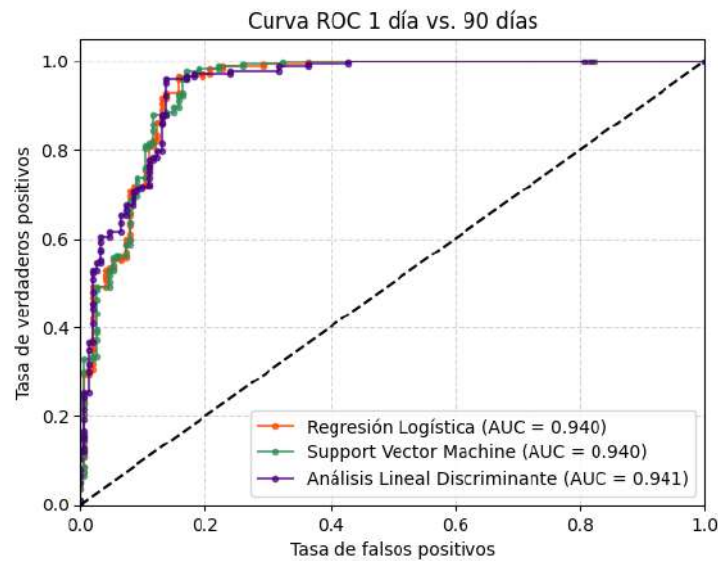


Figura 6: Curva ROC para cada uno de los tres modelos de clasificación empleados en la combinación de LA de 1 día y LAA de 90 días.

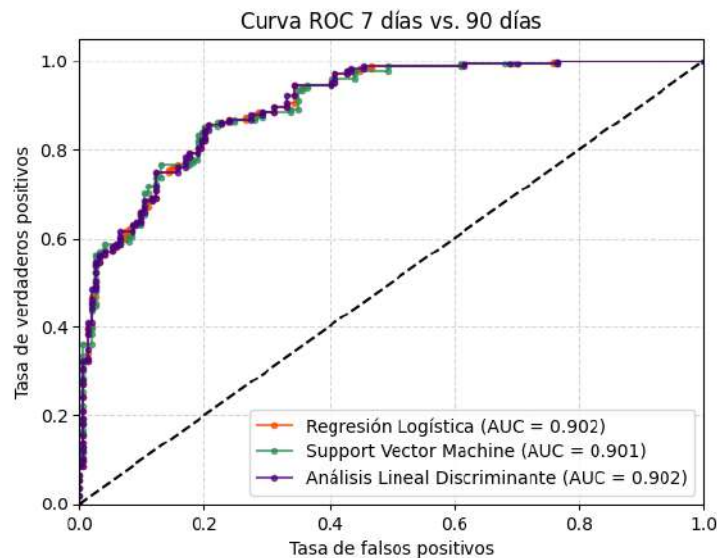


Figura 7: Curva ROC para cada uno de los tres modelos de clasificación empleados en la combinación de LA de 7 días y LAA de 90 días.

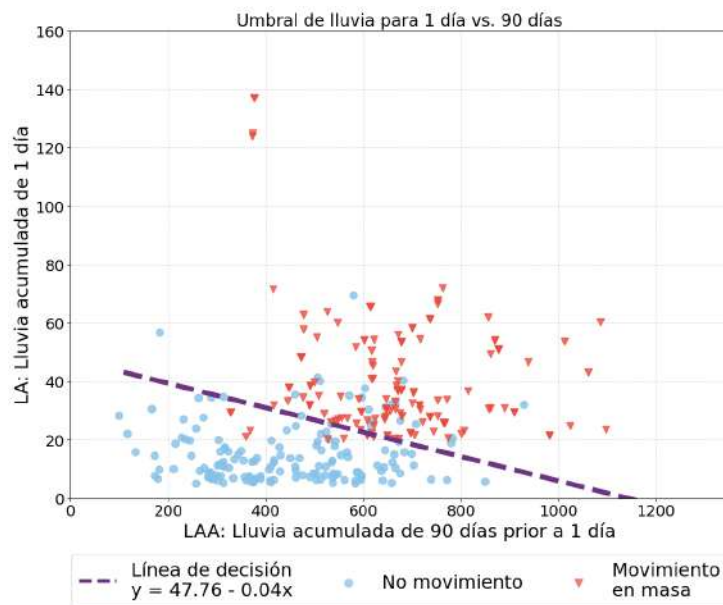


Figura 8: Combinación de LA de 1 día y LAA de 90 días para lluvia de pluviómetros y umbral establecido con cada uno de los tres modelos de clasificación evaluados.

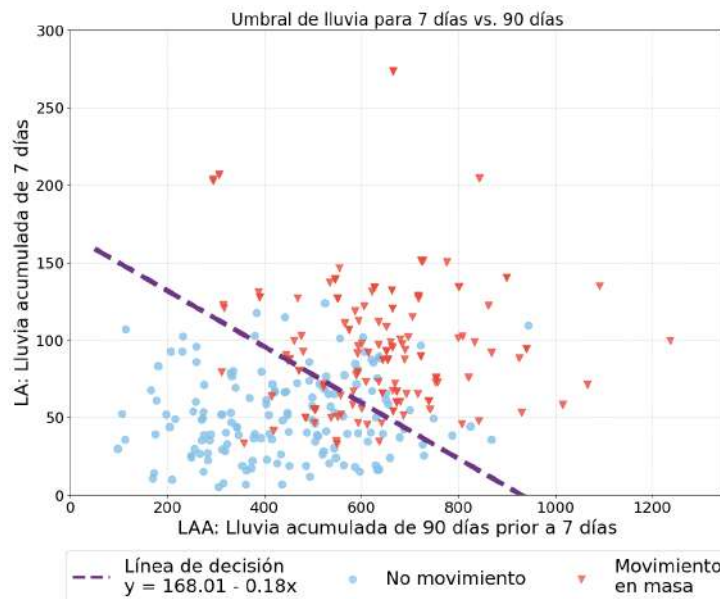


Figura 9: Combinación de LA de 7 días y LAA de 90 días para lluvia de pluviómetros y umbral establecido con cada uno de los tres modelos de clasificación evaluados.

En resumen, este análisis demuestra la eficacia de las metodologías de clasificación empleadas para diferenciar la lluvia que contribuye a la generación de movimientos en masa y para establecer un umbral crítico. Los datos de lluvia registrados por pluviómetros se clasificaron exitosamente utilizando los tres algoritmos (Regresión Logística, SVM y LDA), especialmente en las combinaciones de 1 día vs. 90 días y 7 días vs. 90 días, como se observa en las Figuras 8 y 9. Sin embargo, la misma metodología no resultó efectiva para los datos de lluvia obtenidos de radar meteorológico, lo que indica la necesidad de investigar otros métodos de extracción de datos de radar y de desarrollar enfoques alternativos para la definición de umbrales críticos de lluvia.

Análisis de percentiles

Se consideró crucial analizar y comprender los eventos de lluvia que desencadenaban los movimientos en masa en el Valle de Aburrá. Esta investigación se centró en determinar si los eventos de lluvia que desencadenaron tales movimientos correspondían a eventos extremos o atípicos en el historial de precipitaciones. Alternativamente, se evaluó si estos eventos se ajustaban a la distribución normal de lluvias, sugiriendo que las causas de los movimientos en masa podrían estar más relacionadas con las características intrínsecas del terreno que con la lluvia como factor externo. Para esto, se realizó un análisis detallado utilizando percentiles.

Los percentiles son medidas estadísticas que indican la posición relativa de un valor dentro de un conjunto de datos ordenados de menor a mayor. Un percentil se expresa en términos de porcentaje y divide el conjunto de datos en 100 partes iguales. Por ejemplo, el percentil 25 (P25) es el valor bajo el cual se encuentra el 25 % de los datos, y el percentil 75 (P75) es el valor bajo el cual se encuentra el 75 % de los datos. Los percentiles fueron utilizados para entender la distribución de los datos y para comparar valores individuales con el conjunto de datos en su totalidad.

Para llevar a cabo el análisis, se utilizaron diferentes rangos de lluvia acumulada antecedente. Los acu-

mulados de lluvia se analizaron para los siguientes rangos de tiempo: 6 y 12 horas y 1, 2, 3, 5, 7, 10, 12, 15, 20 y 30 días. Se calculó el acumulado móvil de lluvia para cada uno de estos rangos de tiempo para cada hora en cada una de las series de tiempos históricas de cada uno de los pluviómetros analizados (35 pluviómetros en total).

Con el histórico de lluvia acumulada para los rangos establecidos, se procedió a calcular los percentiles para cada rango. Los percentiles calculados fueron: 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95, 99, 99.9. Posteriormente, se identificaron cada uno de los eventos de lluvia asociados al inventario de movimientos en masa para establecer a qué percentil pertenecía el evento asociado. En total, se analizaron 2.395 eventos de movimientos en masa del inventario y, de igual forma, que para comparar con la lluvia que no detona movimientos en masa se seleccionaron de manera aleatoria el mismo número de eventos de lluvia que no generaron movimientos para posteriormente comparar la distribución de los eventos que generaron movimientos en masa con los que no lo hicieron.

Resultados de los Percentiles

Las Figuras 10 y 11 presentan dos tipos de gráficos: la Estimación de Densidad de Kernel (KDE) y un gráfico de cajas y bigotes. La KDE, que suaviza el histograma de distribución de datos en los percentiles, muestra la frecuencia de los eventos en cada percentil, con el eje horizontal representando los percentiles y el eje vertical la densidad estimada.

Por otro lado, el gráfico de cajas y bigotes resume estadísticamente la distribución de los datos, mostrando el valor mínimo, el primer cuartil (Q1), la mediana (Q2), el tercer cuartil (Q3) y el valor máximo. Los valores atípicos se indican como puntos individuales, facilitando la visualización de la dispersión y simetría de los datos.

Estos gráficos son esenciales para entender la distribución de los eventos de lluvia que causaron movimientos en masa, señalando los rangos de percentiles donde se concentran la mayoría de estos eventos y donde son menos frecuentes. La evaluación se realizó para los períodos acumulados seleccionados, que incluyen 6 y 12 horas, y 1, 2, 3, 5, 7, 10, 12, 15, 20 y 30 días.

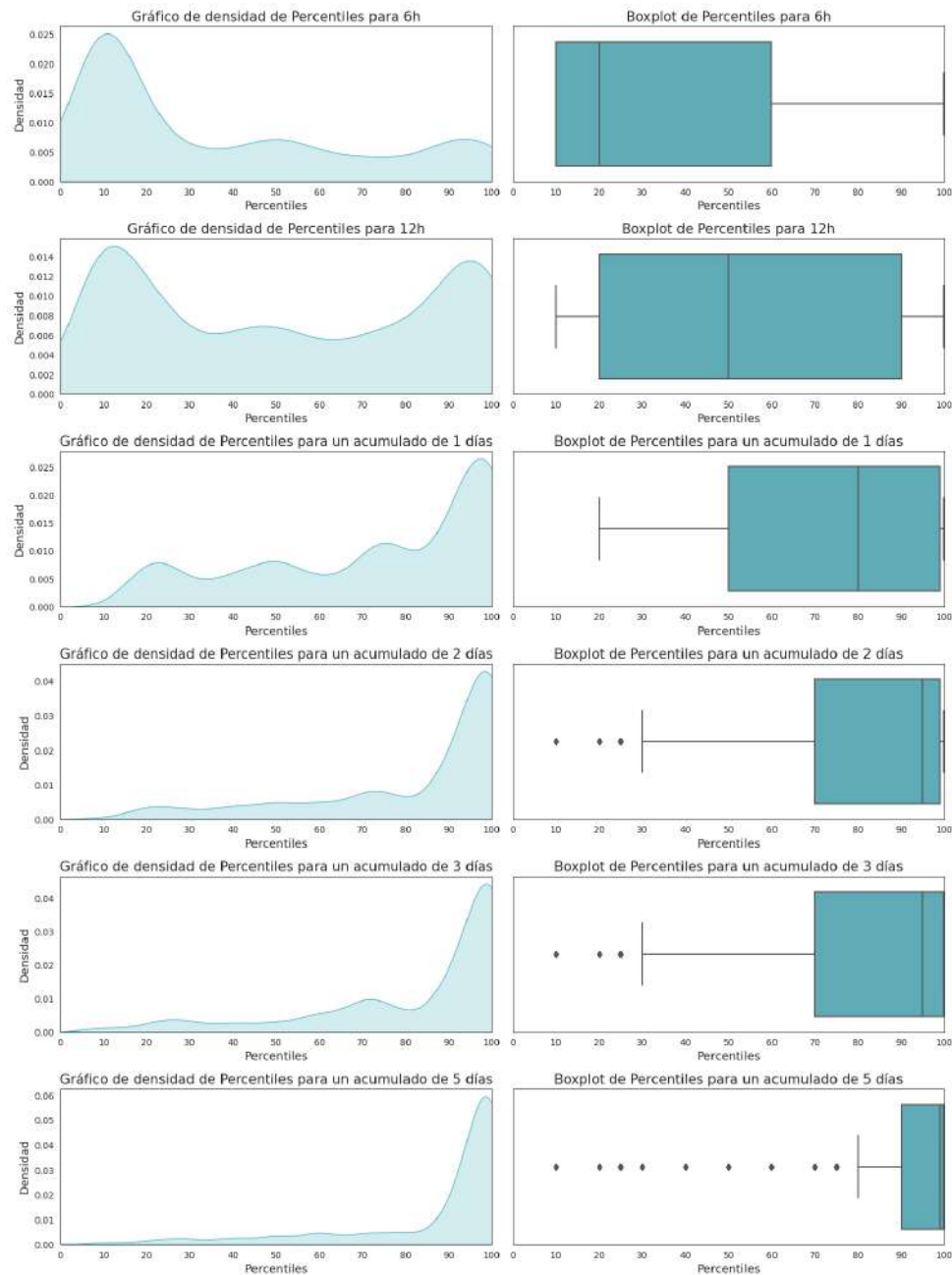


Figura 10: Gráfico de distribución de densidad de Kernel (KDE) y Gráfico de cajas y bigotes para los eventos de lluvia que generaron movimientos en masa en distintos acumulados: 6 y 12 h y 1, 2, 3 y 5 días.

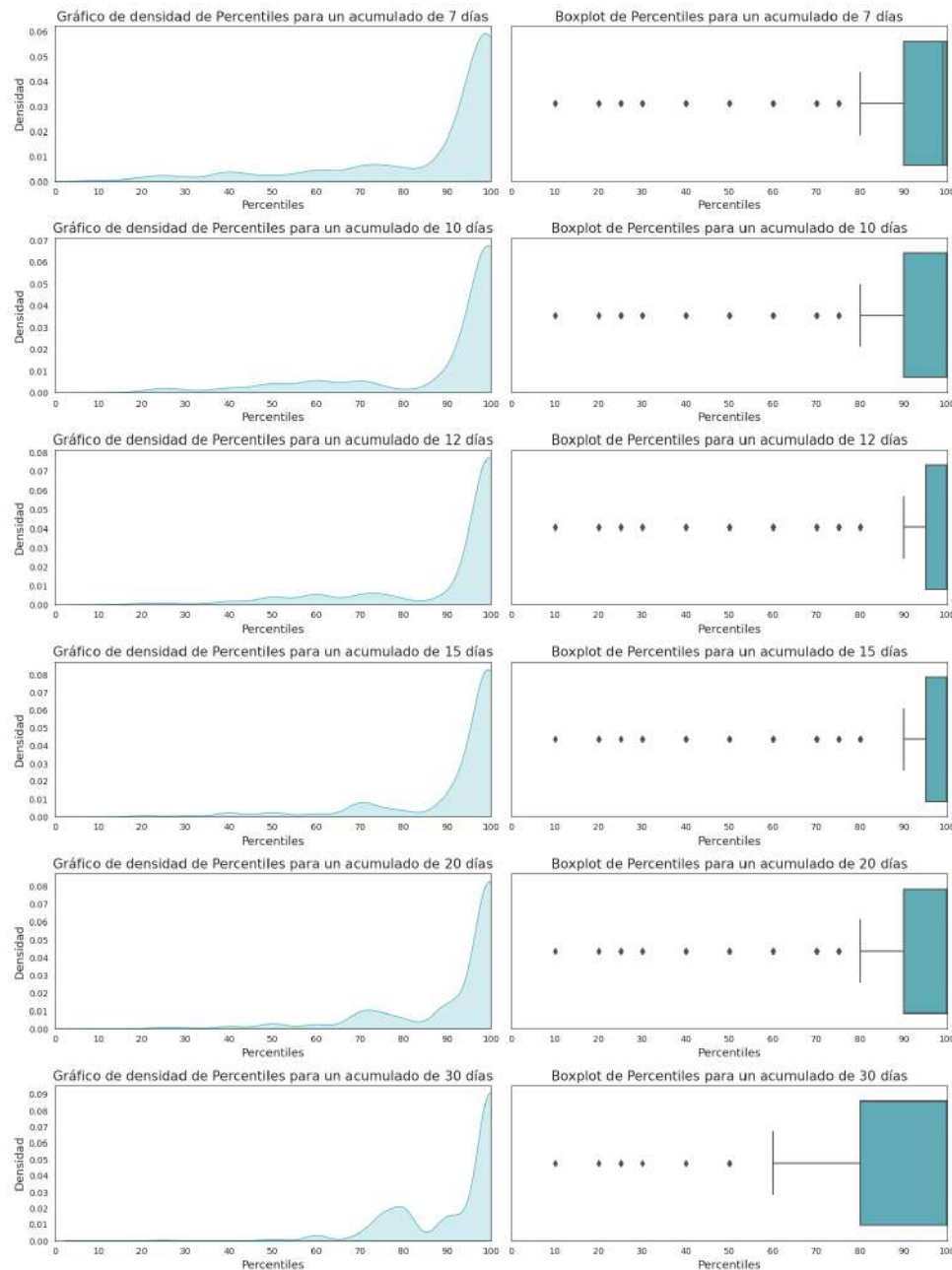


Figura 11: Gráfico de distribución de densidad de Kernel (KDE) y Gráfico de cajas y bigotes para los eventos de lluvia que generaron movimientos en masa en distintos acumulados: 7, 10, 12, 15, 20 y 30 días.

De acuerdo con las Figuras 10 y 11 se presentan algunas observaciones derivadas de los gráficos de densidad de percentiles:

- Se evidencia que, al incrementar el período de acumulación de lluvia antecedente, hay una mayor densidad de eventos en los percentiles más altos (80 a 99.9).

- Para los acumulados antecedentes de 6 y 12 horas, los eventos muestran una distribución bastante uniforme a lo largo de todos los percentiles, con una ligera predominancia en los percentiles bajos (entre 10 y 20). Esto sugiere la ausencia de un patrón definido o tendencia en el histórico de lluvias para estos períodos más cortos (véase Figura 10).
- En el caso del acumulado antecedente de 1 día, se comienza a notar una concentración de eventos en los percentiles superiores (90 a 99). Sin embargo, también hay una notable cantidad de eventos en percentiles más bajos (entre 10 a 60).
- Para los acumulados de 2, 3 y 5 días, la concentración de eventos se hace más evidente en los percentiles que van de 80 a 99.99, acompañada por un número considerable de valores atípicos, especialmente en los períodos de 2, 3 y 5 días. En los acumulados de 7, 10, 12 y 15 días, se observa una marcada concentración en los percentiles altos, particularmente para los períodos de 12 y 15 días.
- Finalmente, en los acumulados de 20 y 30 días, aunque la mayoría de los eventos continúan concentrándose en los percentiles más altos, comienza a notarse una distribución más amplia hacia los percentiles bajos, como se observa en la Figura 11.

Las Figuras 12 y 13 ofrecen un análisis comparativo más detallado entre las distribuciones de lluvia para los períodos acumulados de 6 y 12 horas, y de 12 y 15 días. Además, se incorporan las distribuciones de los eventos de lluvia que no resultaron en movimientos en masa (representados en color azul), seleccionados de manera aleatoria del registro histórico. Es interesante observar que la distribución de los eventos que desencadenaron movimientos en masa es bastante similar a la de aquellos que no lo hicieron para los acumulados de 6 y 12 horas. Específicamente, para los acumulados de 6 horas, ambas distribuciones son casi idénticas en los percentiles bajos.

En el caso de los acumulados de 12 horas, aunque algunos eventos que condujeron a movimientos en masa se encuentran en percentiles más altos, la mayoría se distribuye en percentiles bajos, presentando una distribución casi uniforme a lo largo de todo el espectro de percentiles.

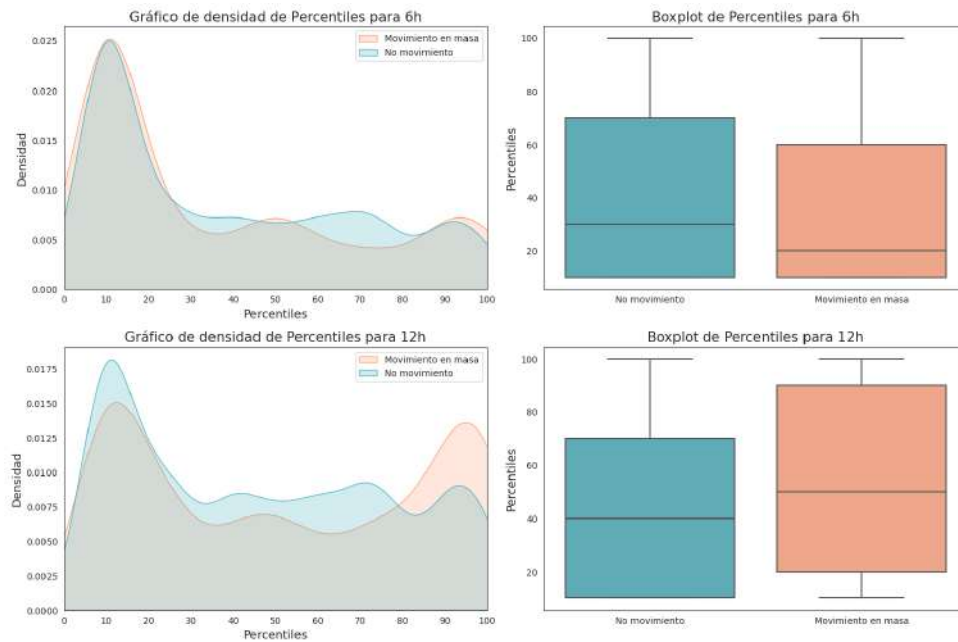


Figura 12: Gráfico de distribución de densidad de Kernel (KDE) y Gráfico de cajas y bigotes para acumulados de 6 y 12 horas, comparando eventos de lluvia que generaron movimientos en masa con aquellos que no lo hicieron.

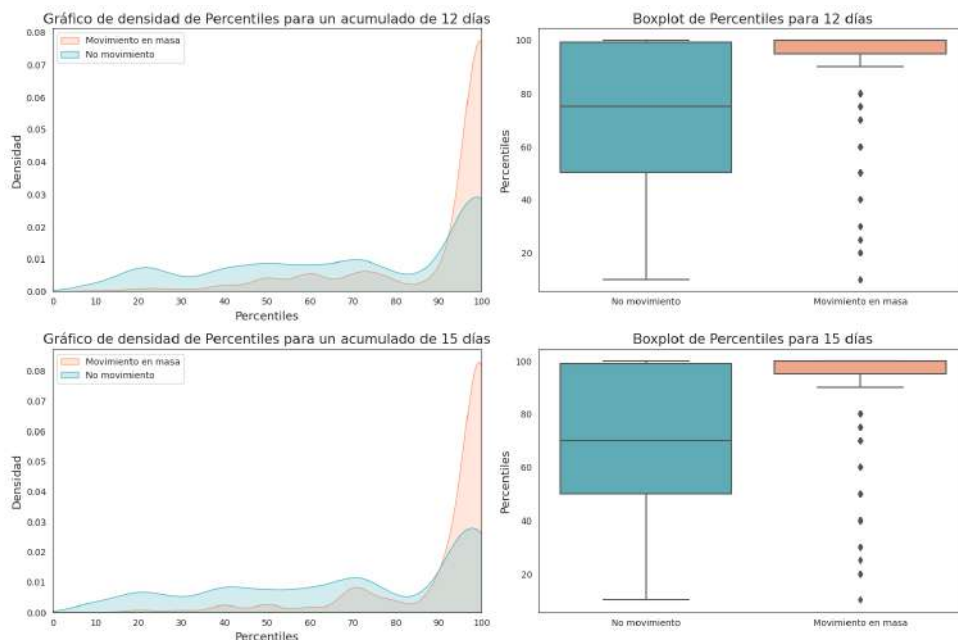


Figura 13: Gráfico de distribución de densidad de Kernel (KDE) y Gráfico de cajas y bigotes para acumulados de 12 y 15 días, comparando eventos de lluvia que generaron movimientos en masa con aquellos que no lo hicieron.

En contraste, en la Figura 13, se evidencia una clara diferencia para los acumulados de 12 y 15 días entre los eventos que generaron movimientos en masa y los que no. Los eventos que causaron movimientos en masa presentaron una concentración superior en los altos percentiles, mientras que los que no generaron movimientos en masa mostraron una distribución más uniforme en todo el rango de percentiles, oscilando en general entre los percentiles 50 y 95, y menos concentrada en los percentiles altos. En los gráficos de caja y bigotes para los acumulados de 12 y 15 días, se observaron varios eventos atípicos en los percentiles bajos, desde 10 hasta 80. Estos eventos podrían estar asociados con la incertidumbre del inventario de movimientos en masa, ya que no siempre es posible determinar si todos los eventos fueron provocados exclusivamente por lluvia.

En conclusión, al analizar los percentiles correspondientes a cada evento del inventario, se logró establecer su distribución y ubicación dentro del espectro de percentiles calculados. Esta metodología permitió identificar diferencias significativas entre los eventos que causaron movimientos en masa y aquellos que no. Se observó que la mayoría de los eventos generadores de movimientos en masa se ubicaban en los percentiles más altos, específicamente entre el 95 y 99.99. Esta distinción resultó ser más notable en los acumulados de 12 y 15 días. Por otro lado, para los periodos más cortos de 6 y 12 horas, la distribución de los eventos fue más homogénea y las diferencias menos marcadas.

Modelo probabilístico

En la continuidad del análisis de la lluvia antecedente, se plantea y examina un modelo probabilístico que define umbrales de lluvia basándose en probabilidades en lugar de valores acumulados. La metodología para desarrollar este modelo se basó en el estudio de Correa et al. (2020), que propone un sistema de alerta temprana para movimientos inducidos por lluvias en la ciudad de Manizales, operando en tiempo real y fundamentado en un enfoque probabilístico.

El modelo se construye a partir del análisis de acumulados de lluvia para distintos intervalos temporales, tanto en horas como en días. Para ello, se calcularon los acumulados de precipitación para cada evento en el inventario de movimientos en masa, abarcando desde una hora hasta 365 días (o 8760 horas) antes de cada evento. La Figura 14 ilustra ejemplos de estos acumulados calculados para varios eventos del inventario.

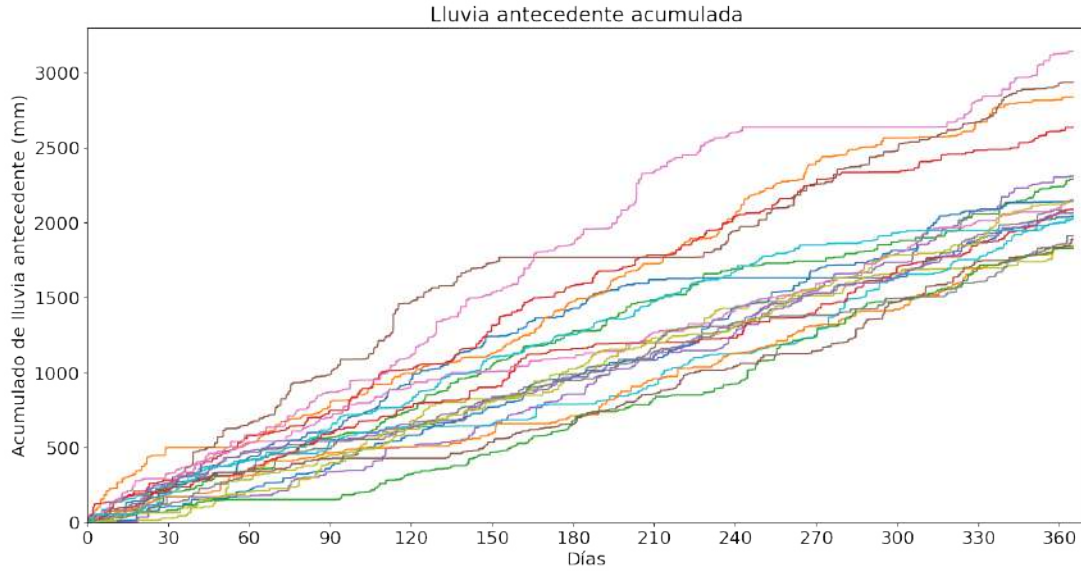


Figura 14: Acumulado de lluvia antecedente para los 365 días de 20 eventos del inventario de movimientos en masa escogidos aleatoriamente.

Después de calcular los acumulados por hora para todos los eventos del inventario, que suman un total de 2.483 eventos, se elaboró una matriz. En esta matriz, cada columna representa la hora acumulada previa, comenzando desde la primera hora y culminando en la hora 8.760 (365 días). Por otro lado, cada fila representa un evento específico del inventario. Finalmente, la matriz construida tiene las dimensiones de 2.483 x 8.760 y se presenta de la siguiente forma:

$$\mathbf{M}(\mathbf{l}, \mathbf{t}) = \begin{bmatrix} x_{t_0}^{l_1} & x_{t_1}^{l_1} & \cdots & x_{t_m}^{l_1} \\ x_{t_0}^{l_2} & x_{t_1}^{l_2} & \cdots & x_{t_m}^{l_2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t_0}^{l_n} & x_{t_1}^{l_n} & \cdots & x_{t_m}^{l_n} \end{bmatrix} \quad (1)$$

donde $x_{t_0}^{l_1}$ representa el acumulado de lluvia en la hora inmediatamente previa al primer evento registrado en el inventario. Por otro lado, $x_{t_1}^{l_1}$ se refiere al acumulado de la segunda hora anterior al mismo primer evento. Mientras que $x_{t_0}^{l_2}$ es el acumulado de la primera hora anterior al segundo evento del inventario. Este patrón continúa hasta llegar a $x_{t_m}^{l_n}$, que representa el último evento del inventario o el evento 2.483 y el último tiempo analizado que es la hora 8.760 (días 365).

Para cada columna de la matriz, se ajustó el modelo de distribución de probabilidad más adecuado. Se exploraron diversos modelos ampliamente citados en la literatura. Las distribuciones evaluadas en este estudio incluyeron: Normal, Lognormal, Gama, Weibull, Gumbel y Logística. Aunque no se abarcó una lista exhaustiva de distribuciones, se espera que con las distribuciones mencionadas cubra muchos de los comportamientos esperados de las variables utilizadas en los diferentes tiempos analizados. Para cada tiempo de acumulación se estiman los parámetros de cada distribución mediante la técnica de máxima verosimilitud.

Para determinar la distribución o modelo que mejor se adaptaba a cada intervalo de tiempo, se emplearon tres pruebas de bondad de ajuste: Chi-Cuadrado, Kolmogorov-Smirnov y Anderson-Darling. Para definir

si la variable sigue una determinada distribución, se comparó el valor P en cada una de las pruebas de bondad de ajuste con un nivel de significancia de 0.05. Esto implica que solo hay un 5 % de probabilidad de concluir erróneamente que los datos no se ajustan a la distribución propuesta. Si el valor P excede el 0.05, se sostiene que no hay razón suficiente para rechazar la hipótesis nula, indicando que los datos podrían seguir la distribución considerada. Finalmente, se optó por la distribución que superara la mayoría de las pruebas, es decir, aquella para la cual la hipótesis nula no fue rechazada en la mayor cantidad de pruebas de bondad de ajuste aplicadas, en este caso las tres pruebas mencionadas anteriormente.

En la Figura 15 y 16 se presentan algunas de las distribuciones que mejor se ajustan de acuerdo con lo mencionado anteriormente para diferentes tiempos de acumulados. Se observa que para las primeras horas previas a los eventos por ejemplo los acumulados de 6 y 24 horas el ajuste no es tan acorde con la distribución de los eventos y esto se debe a la gran cantidad de 0 presentes en los datos para las primeras horas. A partir de los 3 días de acumulados los ajustes comienzan hacer más cercanos a las distribuciones de los eventos.

Una vez los modelos de distribución de probabilidad son definidos para cada tiempo, se pueden calcular ciertas probabilidades de excedencia dado un valor de acumulado en un cierto tiempo. Una ventaja de definir un modelo probabilístico es que los umbrales o niveles de alerta se pueden definir en términos de probabilidad y no de un cierto valor de lluvia acumulada, lo que a su vez brinda una estimación indirecta de la probabilidad de movimientos en masa en una zona determinada.

Los umbrales definidos en este trabajo fueron ajustados siguiendo el trabajo realizado por Correa et al. (2020) donde plantean una metodología para definir un modelo probabilístico para un Sistema de Alerta Temprana por movimientos en masa detonados por lluvia en el área urbana de la ciudad de Manizales. Los niveles de alerta adoptados del trabajo de referencia se presentan en la Tabla 4.

Tabla 4: Niveles de alerta y probabilidad de ocurrencia.

Nivel de Alerta	Color	Probabilidad de ocurrencia
Bajo	Verde	$P < 0,32$
Medio	Amarillo	$0,32 \leq P < 0,5$
Alto	Naranja	$0,5 \leq P < 0,68$
Muy Alto	Rojo	$P \geq 0,68$

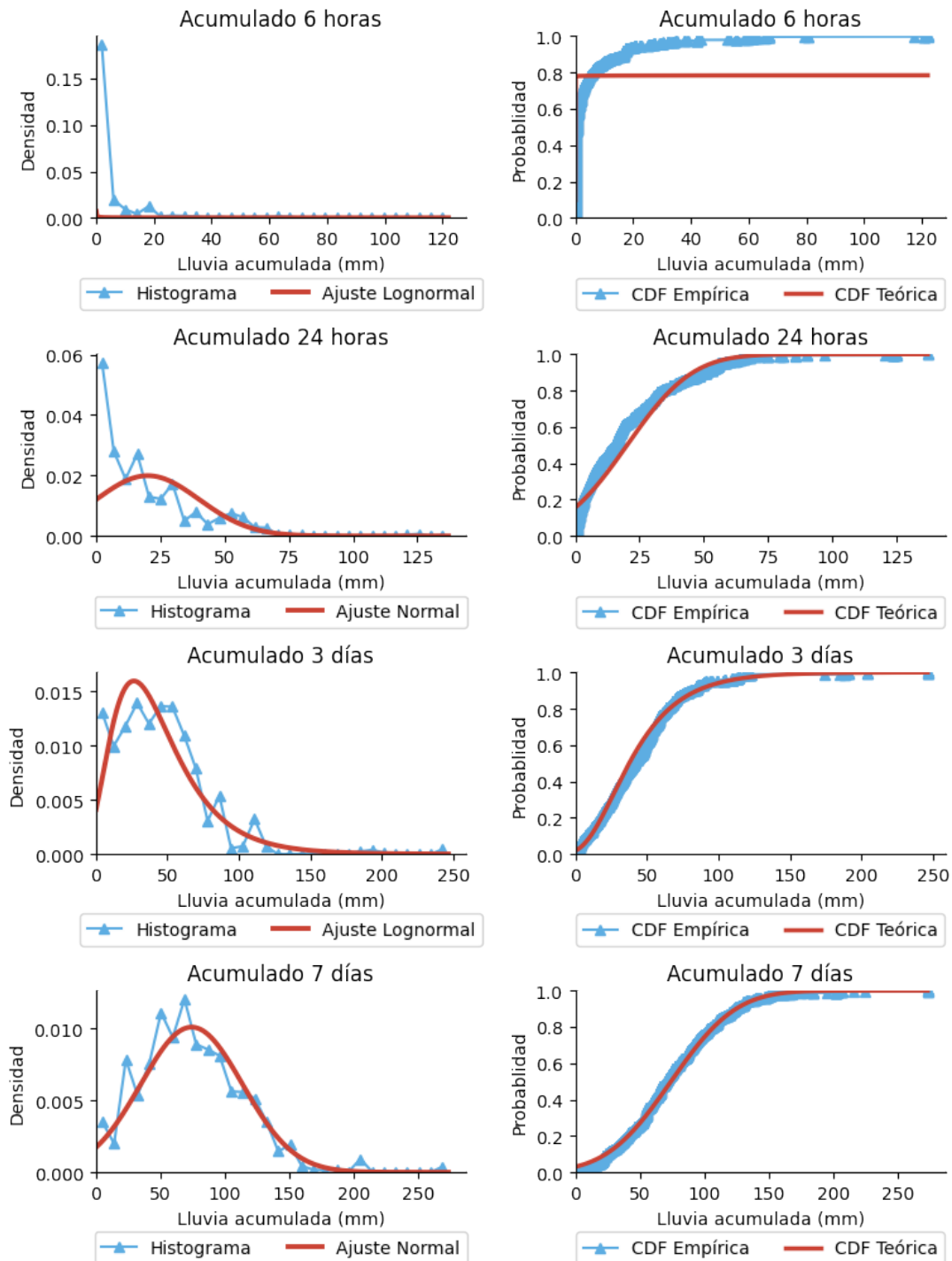


Figura 15: Algunos ejemplos de los modelos de distribución de probabilidad que mejor se ajustan para diferentes tiempos en todos los eventos del inventario.

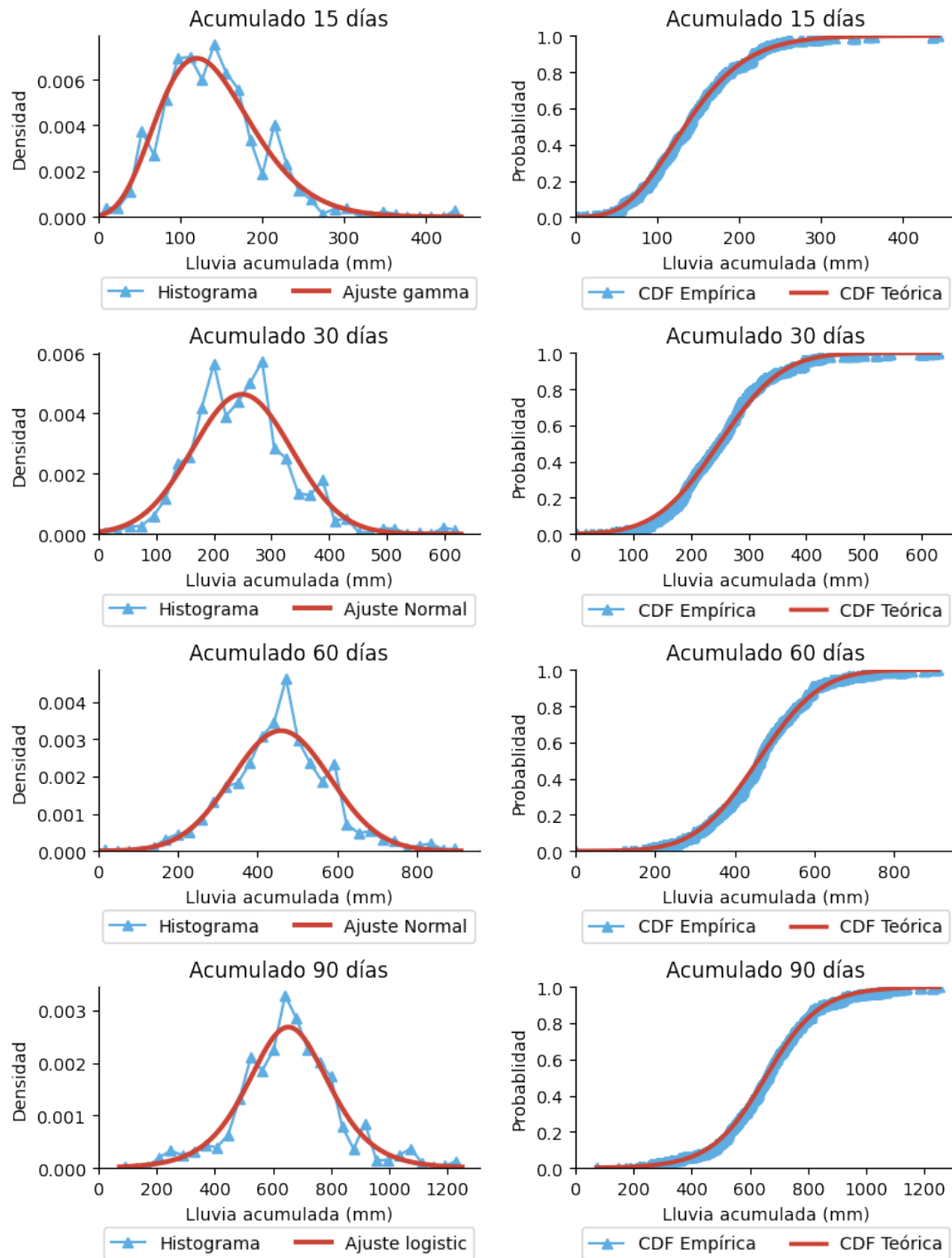


Figura 16: Algunos ejemplos de los modelos de distribución de probabilidad que mejor se ajustan para diferentes tiempos en todos los eventos del inventario.

Finalmente, se realiza una validación con el evento precipitación de alta intensidad y larga duración ocurrido en las veredas San Andrés, El Socorro y Potrerito del municipio de Girardota. El evento dio lugar a la ocurrencia de diversos movimientos en masa, desbordamientos de quebradas y avenidas torrenciales generando daños en vías, edificaciones, terrenos agrícolas, y otros componentes de la infraestructura (Ver Figura 17). El evento de precipitación se presentó entre los días 27 y 28 de mayo, con intensidades máximas de más de 120 mm/h y acumulados superiores a los 100 mm en un lapso de 24 horas.



Figura 17: Múltiples movimientos en masa capturados en imágenes tomadas con sobrevuelo de dron.

En las Figuras 18, 19 y 20 se presentan series de tiempo para diferentes días antes de la ocurrencia del evento, todas tomadas del pluviómetro 66 Instalado en la I.E. San Andrés en la vereda el Socorro, ubicado cerca de la ocurrencia de los eventos. Estas gráficas ilustran tanto las intensidades de lluvia como los acumulados previos, distinguidos por colores según los niveles de alerta descritos en la Tabla 4. En la Figura 18, que representa los 90 días previos al 28 de mayo a las 2 a.m. (marcado con una línea punteada púrpura), se destaca que durante los 30 días anteriores, la alerta fue de nivel rojo. Por otro lado, en días más distantes al evento, los niveles fluctuaron entre alertas amarillas y verdes.

En la Figura 19 se presenta la serie de tiempo para los 30 días antecedentes al evento, evidenciando que, durante todo este período, la alerta se mantuvo en rojo. Por último, la Figura 20 muestra las series de tiempo de los tres días anteriores al evento, revelando el comportamiento de la intensidad de lluvia y las alertas correspondientes. Es notable que, horas antes del evento, a pesar de presentar bajos acumulados de lluvia, el nivel de alerta se mantiene en rojo. Esta anomalía se explica debido a ajustes no óptimos en las funciones de probabilidad (ver Figura 15) para las primeras horas de los eventos del inventario, principalmente por la gran cantidad de eventos con acumulados de ceros en las primeras horas. Esto implicó que incluso acumulados bajos de lluvia reflejaran altas probabilidades de ocurrencia y, por consiguiente, altos niveles de alerta. En conclusión, el empleo de un modelo probabilístico para anticipar

eventos críticos de lluvia que puedan desencadenar movimientos en masa parece ser efectivo para alertas con tres o más días de antelación; sin embargo, su precisión disminuye con períodos más cortos.

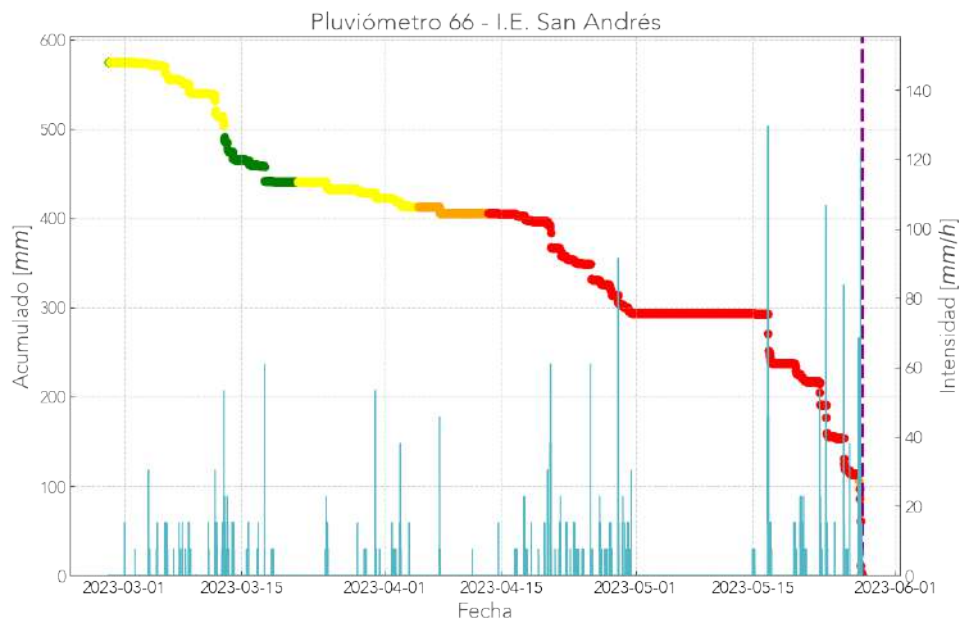


Figura 18: Serie de tiempo de lluvia para los 90 días antecedentes al evento ocurrido el día 28 de mayo en Girardota (línea puenteada púrpura). Los acumulados antecedentes se presentan con los colores de acuerdo a los niveles de alerta presentados en la Tabla 4.

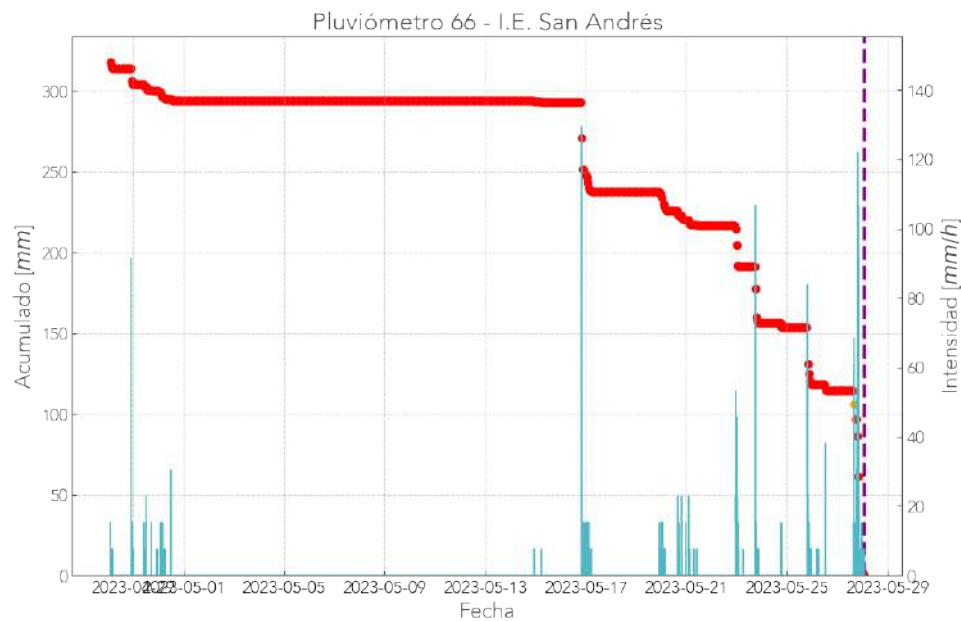


Figura 19: Serie de tiempo de lluvia para los 30 días antecedentes al evento ocurrido el día 28 de mayo en Girardota (línea puenteada púrpura). Los acumulados antecedentes se presentan con los colores de acuerdo a los niveles de alerta presentados en la Tabla 4.

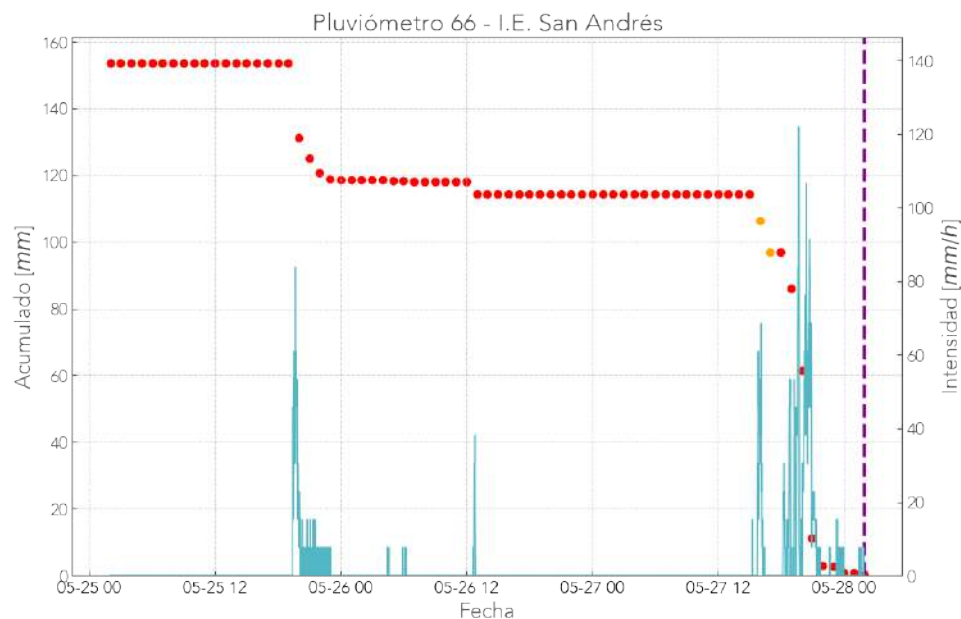


Figura 20: Serie de tiempo de lluvia para los 3 días antecedentes al evento ocurrido el día 28 de mayo en Girardota (línea puenteada púrpura). Los acumulados antecedentes se presentan con los colores de acuerdo a los niveles de alerta presentados en la Tabla 4.

Se puede concluir que se pueden construir un modelo probabilístico a partir del cálculo de los acumulados de lluvia para diferentes intervalos de tiempo antecedente. Los resultados de la metodología aplicada arrojan algunas observaciones: mientras que para períodos cortos (6 y 12 horas) la distinción entre eventos que generan y no generan movimientos en masa es menos clara, para períodos más largos (12 y 15 días) esta distinción se vuelve más evidente. Algo ya observado en el capítulo de análisis con percentiles.

La validación del modelo se realizó con un caso concreto: el evento ocurrido el 27 de mayo de 2023 en el municipio de Girardota. Esta validación evidenció una alta eficacia del modelo en predecir la probabilidad de movimientos en masa frente a lluvias antecedentes específicas, especialmente cuando se consideran periodos de 3 días o más antes del evento. No obstante, se notó una disminución en la precisión del modelo para periodos más cortos. Esto resalta la importancia de realizar ajustes y mejoras continuas en las funciones de probabilidad para intervalos de tiempo reducidos.

En conclusión, el empleo de un modelo probabilístico para calcular la lluvia antecedente en el contexto de sistemas de alerta temprana para movimientos en masa inducidos por lluvias muestra un gran potencial. Sin embargo, es esencial que este modelo sea constantemente ajustado y mejorado para garantizar su precisión y aplicabilidad efectiva en tiempo real.

Umbrales A-D

El análisis de la lluvia se direccionó hacia el estudio de otras variables relacionadas con la lluvia. Mediante el uso de métodos estadísticos, se analizó el inventario de movimientos en masa junto con las precipitaciones asociadas a cada evento, con el fin de establecer umbrales críticos de Intensidad-Duración (I-D) y de Acumulado-Duración (A-D). Tales métodos para definir umbrales han sido documentados en la literatura científica, tal como se refleja en los estudios de: Guzzetti et al. (2007), Guzzetti et al. (2008) y Peruccacci et al. (2017).

El análisis se centró en identificar las condiciones mínimas de lluvia requeridas para que ocurran movimientos en masa en el Valle de Aburrá. Estos umbrales se determinaron utilizando un método estadístico frecuentista propuesto por Brunetti et al. (2010) y aplicado en estudios como Peruccacci et al. (2012), Gariano et al. (2015), Piciullo et al. (2017) y Abraham et al. (2020).

La determinación de estos umbrales se realizó aplicando un enfoque estadístico frecuentista, siguiendo la metodología propuesta por Brunetti et al. (2010), y que ha sido implementada en diversas investigaciones Peruccacci et al. (2012), Gariano et al. (2015), Piciullo et al. (2017), Abraham et al. (2020). Este método frecuentista se descompone en los siguientes pasos, basados y adaptados de Brunetti et al. (2010):

1. Se aplica la ecuación propuesta por Caine (1980), ampliamente reconocida en la literatura para establecer umbrales de Intensidad y Duración:

$$I = \alpha D^{\beta} \quad (2)$$

Donde I es la intensidad media del evento en mm/h, D es la duración total del evento de lluvia, y α y β son parámetros constantes, α representa el intercepto con el eje vertical, mientras que β es la pendiente de la recta. Para determinar estos parámetros, se utilizó el método de mínimos cuadrados. La Figura 21 muestra un ejemplo de los eventos de precipitación vinculados a cada evento del inventario, junto con la línea que mejor se ajusta a la ecuación 2.

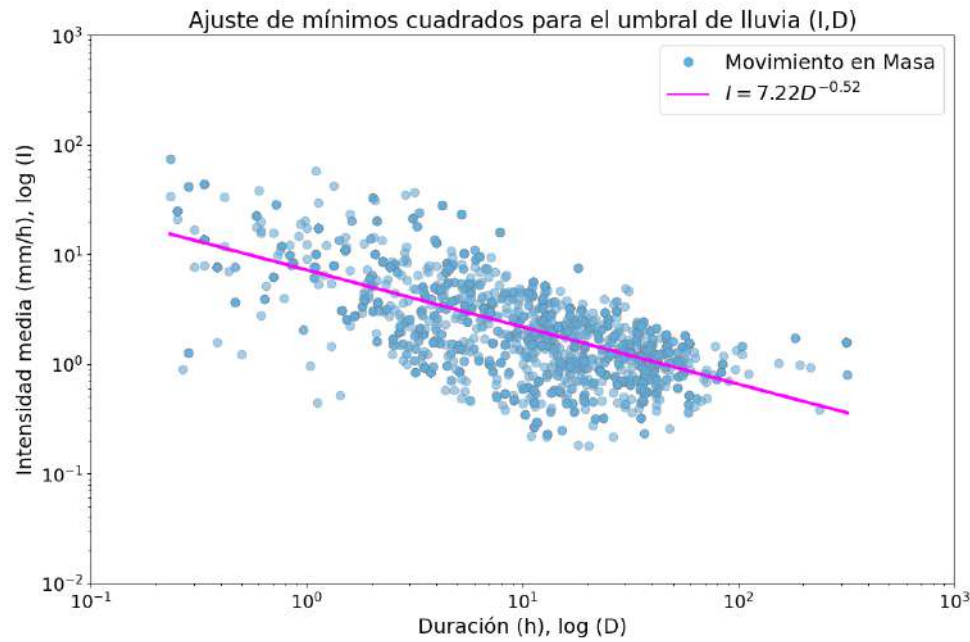


Figura 21: Duración (D) e Intensidad media (I) de cada uno de los eventos de lluvia asociados al inventario de movimientos en masa. La recta de color rosa describe el ajuste por mínimos cuadrados de las condiciones empíricas de I-D.

- II. A partir de la línea ajustada del paso anterior, se calcularon las diferencias $\delta(D)$ entre el logaritmo de las intensidades reales de los eventos $\log[I(D)]$ y las intensidades esperadas a partir del ajuste $\log[Ia(D)]$:

$$\delta(D) = \log[I(D)] - \log[Ia(D)] \quad (3)$$

Una vez calculadas estas diferencias, se establece la función de distribución de probabilidad (PDF) de $\delta(D)$ usando la Estimación de Densidad de Kernel. Posteriormente, se realiza un ajuste de esta función con una distribución Gaussiana. La Figura 22 ilustra esta función PDF (línea punteada azul) junto con el ajuste Gaussiano (línea sólida negra).

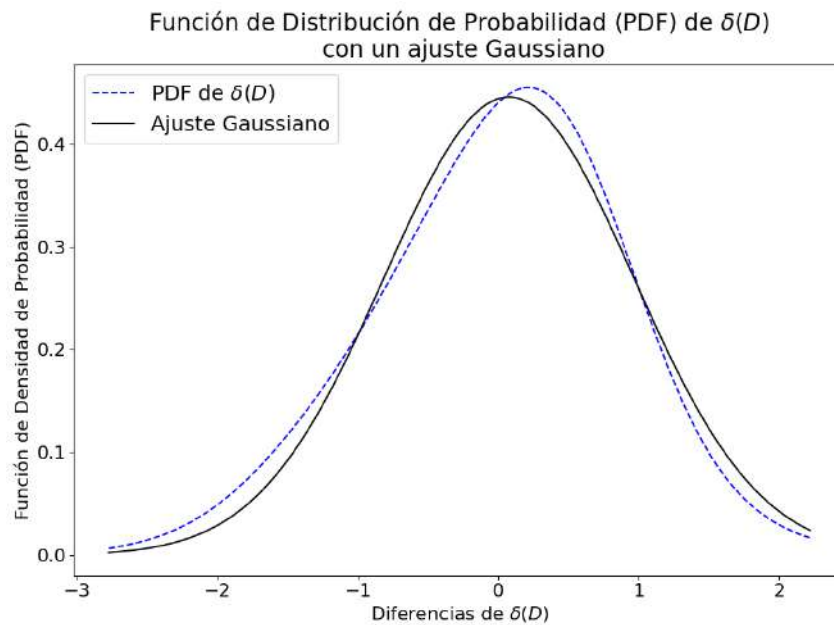


Figura 22: Estimación de la Densidad de Probabilidad de Kernel de las diferencias $\delta(D)$ (en línea punteada azul) ajustada con una función Gaussiana (línea negra) para la distribución de los datos empíricos I-D.

- III. Por último, se establecieron umbrales de lluvia correspondientes a distintas probabilidades de excedencia, con base en la distribución modelada en el paso 2. Por ejemplo, en la Figura 23 se muestra el valor de probabilidad del 5 %, representado con una línea punteada roja. La línea punteada gris indica el valor medio de la distribución de datos. La diferencia δ^* entre la media de los datos y la probabilidad del 5 % se emplea para determinar el intercepto o valor α del umbral de excedencia del 5 %. En la Figura 24 se presenta el umbral del 5 % de probabilidad de excedencia denominado como T5. Una probabilidad de excedencia del 5 % indica que la probabilidad de ocurrencia de un movimiento en masa con una lluvia que esté por debajo de ese umbral es inferior al 5 %.

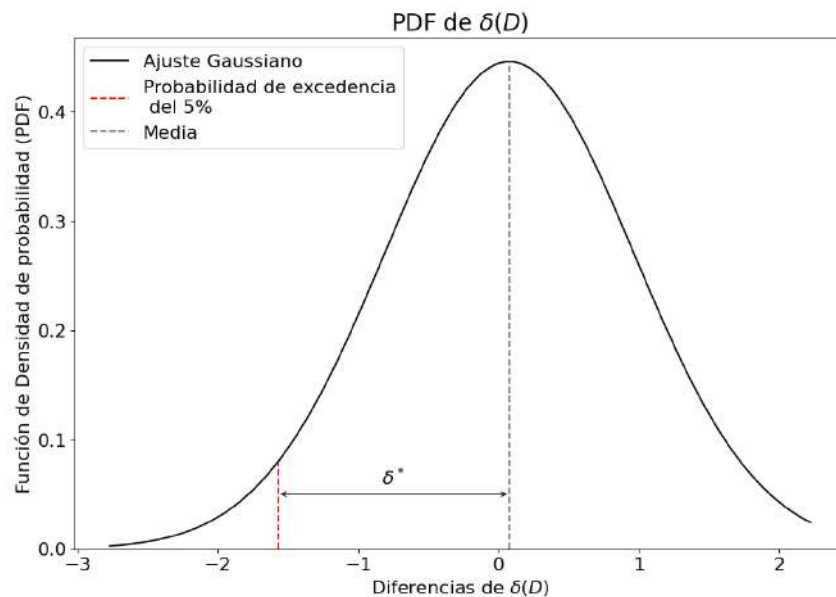


Figura 23: Probabilidad del 5 % en la función de distribución Gaussiana para la diferencia de datos $\delta(D)$. δ^* es la diferencia entre la media de los datos y el valor de la probabilidad del 5 %, representado como T5.

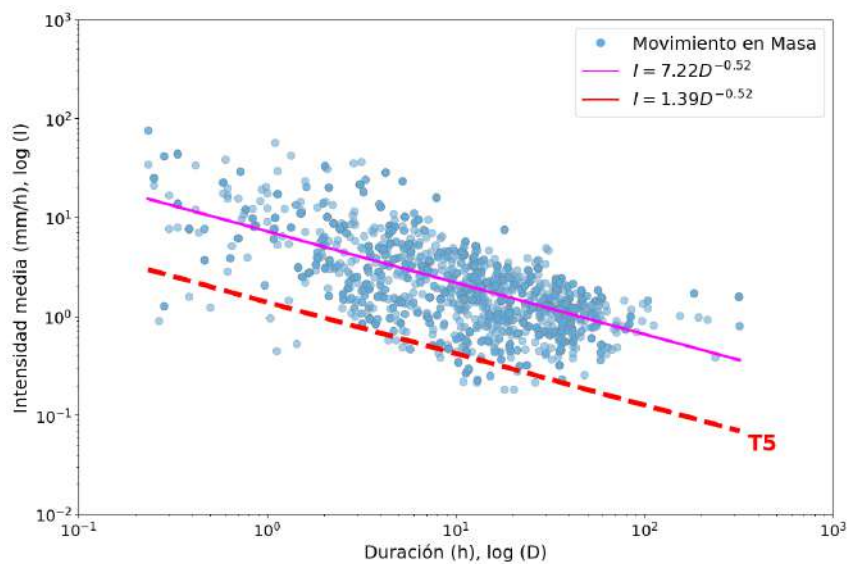


Figura 24: Umbral crítico de lluvia para una probabilidad de excedencia del 5 % representado por la línea punteada roja como T5.

Según Brunetti et al. (2010), una de las principales ventajas del método Frecuentista radica en la posibilidad de definir múltiples umbrales basados en distintos niveles de probabilidad de excedencia. Esta característica facilita establecer diferentes rangos para emitir alertas en respuesta a eventos de lluvia. En concreto, al contar con diferentes niveles se pueden diseñar gráficos que analicen la precipitación en tiem-

po real que exceden ciertos niveles de umbral. Brunetti et al. (2010) propone 4 niveles de probabilidad de excedencia acompañados de sus respectivas alertas:

- **Umbral T0,05 %:** Corresponde a una probabilidad de excedencia del 0,005 % del área sobre la curva con el ajuste Gaussiano (ver Figura 22). Representa el nivel de alerta más bajo, color verde.
- **Umbral T0,5 %:** Se relaciona con una probabilidad de excedencia del 0,5 %. Su nivel de alerta es bajo, color amarillo.
- **Umbral T1,5 %:** Alineado con una probabilidad de excedencia del 1,5 %. Es un nivel de alerta medio, color naranja.
- **Umbral T5 %:** Representa una probabilidad de excedencia del 5 %. Este es el nivel de alerta alto, color rojo.

La Figura 25 presenta cada uno de estos umbrales, coloreados de acuerdo a su alerta, junto con la ecuación de la recta asociada. Es notable que, para el umbral T0,05 % (el más bajo), ninguno de los eventos del inventario se ubica por debajo, mientras que para el nivel más alto, T5 %, el 5 % de los datos del inventario se encuentra por debajo de este umbral.

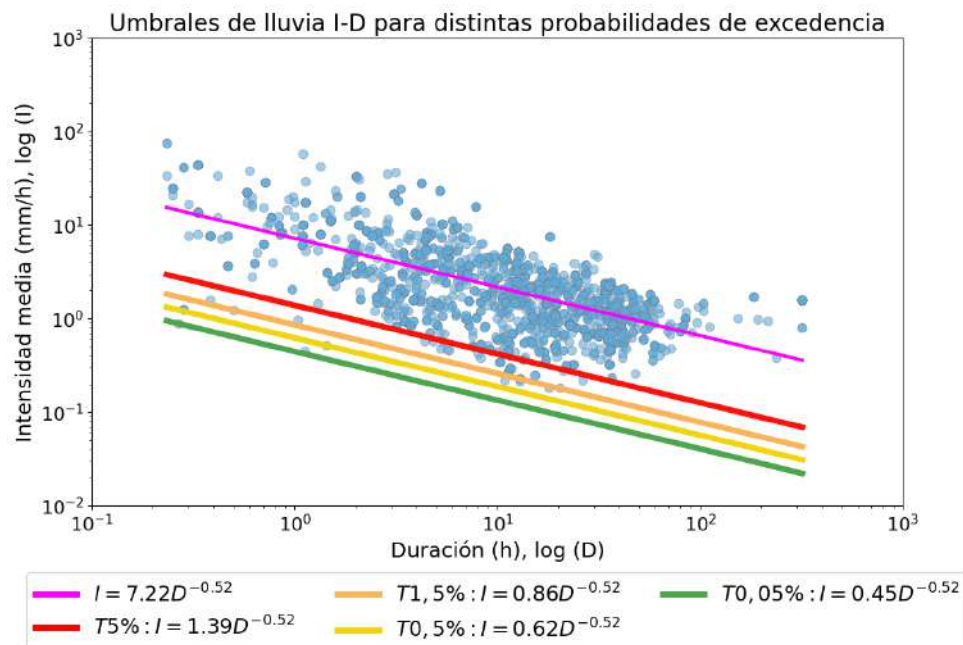


Figura 25: Umbrales críticos de lluvia I-D para realizados con distintas probabilidades de excedencia: T5 %, T1,5 %, T0,5 % y T0,05 %.

Realizando una revisión detallada de la literatura diversos autores argumentan que los umbrales A-D (Acumulado - Duración) podrían brindar más información que los umbrales I-D (Intensidad - Duración).

Peruccacci et al. (2017) sostienen que los umbrales I-D y A-D pueden considerarse funcionalmente equivalentes dado que la intensidad de la lluvia se calcula dividiendo el acumulado total por la duración del evento. Al construir umbrales con dos variables se supone que las variables expresan o miden cantidades

independientes. Sin embargo, al buscar una relación estadística entre la duración de la lluvia D y la intensidad media de la lluvia I , se viola el supuesto porque la intensidad media de la lluvia depende de la duración de la lluvia, a través de la lluvia acumulada.

En contraste, las variables de Acumulado (en mm) y Duración (en horas) derivan casi directamente de las lecturas de pluviómetros, siendo medidas independientes entre sí. Además, para la determinación de umbrales en tiempo real, resulta más práctico y eficiente computar el acumulado y la duración que la intensidad, ya que esto simplifica y agiliza los cálculos. Por lo tanto, en el análisis correspondiente a este mes, se priorizó el examen de los umbrales A-D (Acumulado - Duración). Es por esto que durante este estudio se trabajó con los umbrales de A - D (Acumulado - Duración).

Por otro lado, una de las limitaciones inherentes a los umbrales de lluvia definidos mediante metodologías estadísticas es que su efectividad está directamente vinculada al inventario utilizado. Tanto la cantidad como la calidad de los datos influyen en la incertidumbre asociada a dichos umbrales: un inventario con una colección más extensa y representativa de eventos tiende a generar resultados más precisos, como señalan Peruccacci et al. (2017).

Para cuantificar la incertidumbre en los umbrales, resulta crucial evaluar las incertidumbres asociadas a los parámetros α y β de la ecuación 2. El cálculo de estas incertidumbres no solo proporciona una medida de confianza en los valores de los parámetros y, por ende, en los umbrales establecidos, sino que también facilita la comparación estadística entre diferentes umbrales.

Determinación de las incertidumbres de los parámetros α y β :

Para una evaluación precisa de las incertidumbres, se pueden emplear métodos como el Bootstrap, el análisis de regresión o técnicas bayesianas, que permiten estimar intervalos de confianza para α y β . Estos intervalos proporcionan una base cuantitativa para la interpretación de los umbrales y su variabilidad potencial ante distintas condiciones y conjuntos de datos.

Para este análisis se utilizó la técnica Bootstrap que es una herramienta utilizada para estimar la distribución de una estadística (como la media o la mediana) a partir de un conjunto de datos, al generar múltiples muestras a través del remuestreo con reemplazo. En el contexto de nuestro análisis para los parámetros α y β de la ecuación de Caine (Ec. 2), esta técnica implica seleccionar aleatoriamente subconjuntos de los datos originales (con la posibilidad de repetir datos) para crear “nuevas” muestras. Para cada una de estas muestras generadas, se calcula la media y la desviación estándar de α y β .

En esta aplicación específica, se han realizado $k = 5,000$ iteraciones del proceso de remuestreo, lo que proporciona una sólida base para la estimación de los parámetros y sus incertidumbres. Con cada iteración, se ha obtenido un conjunto de valores para α y β , y a partir de estos, se ha calculado la media y la desviación estándar correspondiente. Posteriormente, las medias y las desviaciones estándar se promedian para obtener una estimación final más precisa de los parámetros y sus incertidumbres.

Con las incertidumbres estimadas, se puede proceder a identificar las curvas de umbral mínimo y máximo que enmarcan la curva del umbral central, basadas en la variabilidad de los parámetros obtenidos. Estas curvas delimitan el rango dentro del cual es más probable que se sitúen los verdaderos valores de α y β , proporcionando así una representación visual y cuantitativa de la confianza que se puede tener en los umbrales calculados. Finalmente, se identifican las curvas mínimas y máximas alrededor de la curva del umbral de acuerdo con las incertidumbres definidas de la siguiente forma:

$$A_{\min, \Delta\alpha} = (\alpha - \Delta\alpha)D^\beta \quad (4)$$

$$A_{\min, \Delta\beta} = \alpha D^{(\beta - \Delta\beta)} \quad (5)$$

$$A_{\max, \Delta\alpha} = (\alpha + \Delta\alpha)D^\beta \quad (6)$$

$$A_{\max, \Delta\beta} = \alpha D^{(\beta + \Delta\beta)} \quad (7)$$

donde $\Delta\alpha$ y $\Delta\beta$ son las incertidumbres asociadas a los parámetros α y β .

Las curvas definidas por las ecuaciones 4, 5, 6 y 7 varían con el incremento de la duración de la lluvia. Para evaluar el efecto combinado de las incertidumbres calculadas con los parámetros α y β por cada duración de lluvia, se compara los valores con las dos incertidumbres de los parámetros y se elige el valor más alto o más bajo de las incertidumbres de los dos parámetros de para obtener las curvas mínimas o máximas de la siguiente forma:

$$\begin{cases} A_{\min} = \min[A_{\min, \Delta\alpha}, E_{\min, \Delta\beta}] \\ A_{\max} = \max[A_{\max, \Delta\alpha}, E_{\max, \Delta\beta}] \end{cases} \quad (8)$$

lo anterior define la región sombreada que es la incertidumbre asociada al umbral A-D.

Mediante la técnica de Bootstrap y utilizando un valor de k igual a 5000, se evaluó el número de desviaciones estándar para α y β utilizando distintas cantidades de eventos en el dataset para umbrales con una probabilidad de excedencia del 5 %. Este análisis se llevó a cabo seleccionando al azar distintas cantidades de eventos, desde 50 hasta el total de 991 eventos del inventario.

Las Figuras 26 y 27 presentan el análisis de la media y desviación estándar para los parámetros α y β evaluados en relación con un umbral de probabilidad de excedencia del 5 % para diferentes tamaños de muestra del inventario. Se constata que con un inventario de 500 eventos o más (indicado por las líneas punteadas negras), las desviaciones estándar para los parámetros α y β se reducen significativamente, siendo menores a 0.8 y 0.04, respectivamente. Esto implica que para el inventario actual, compuesto por 991 eventos, es viable establecer un umbral con incertidumbres relativamente bajas. Según Peruccacci (2017), un umbral se considera aceptable en un sistema de alerta temprana si la incertidumbre relativa ($\Delta\alpha/\alpha$) es menor o igual al 10 %. En este caso, para un inventario de 991 eventos, las incertidumbres relativas en los parámetros α y β son del 10.2 % y 8.8 %, respectivamente, lo que permite concluir que las incertidumbres asociadas a un umbral de excedencia del 5 % son lo suficientemente bajas para ser utilizadas en futuros sistemas de alerta temprana.

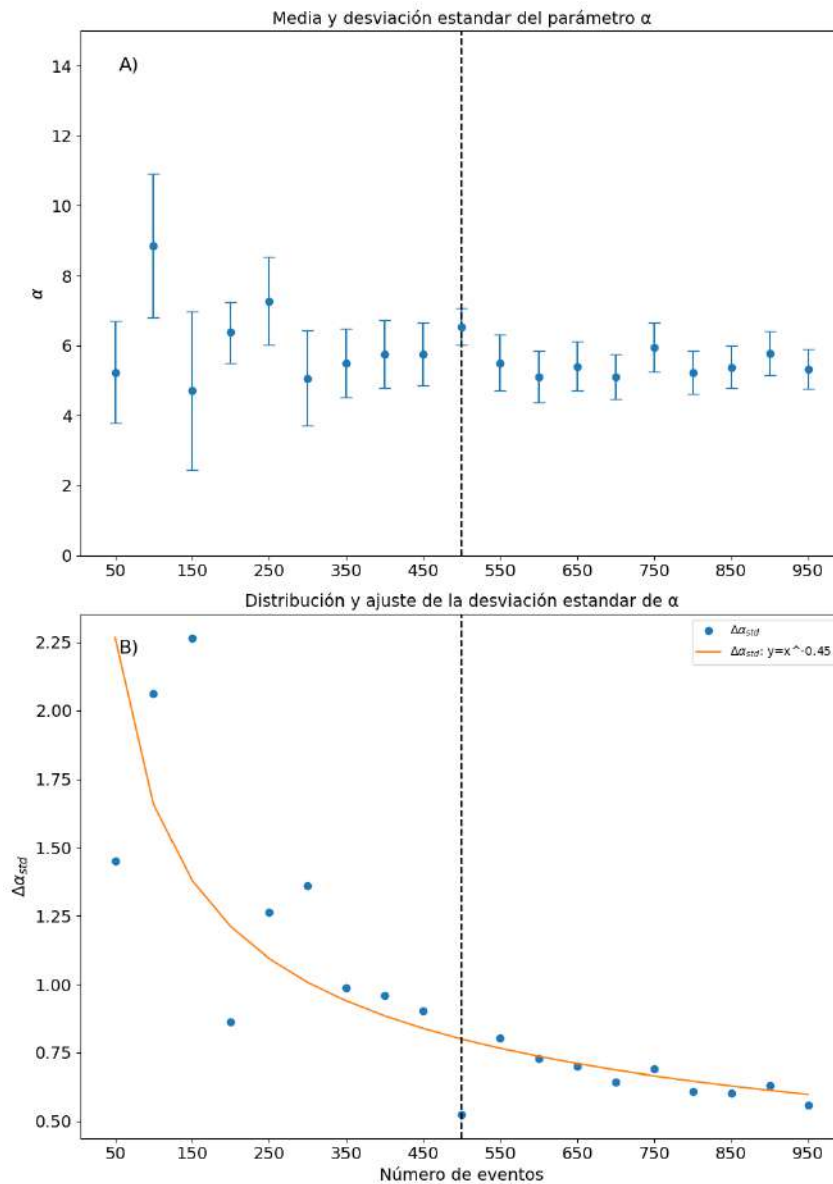


Figura 26: (A) Media y desviación estándar del parámetro α para diferentes cantidades de eventos del inventario. (B) Tendencia de la desviación estándar de α en función del número de eventos analizados.

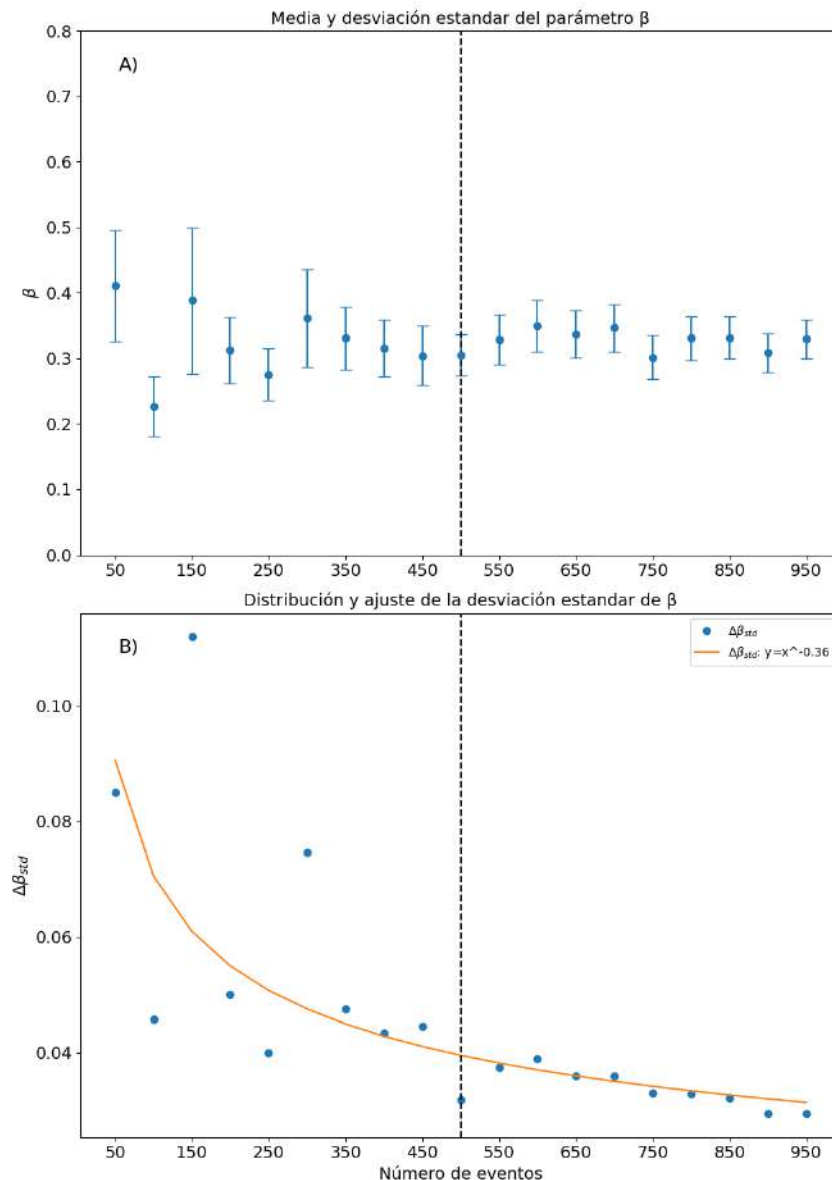


Figura 27: A) Media y desviación estándar del parámetro β en función del número de eventos analizados del inventario. (B) Comportamiento de la desviación estándar de β con variaciones en el tamaño del conjunto de datos.

Diferentes umbrales A-D de acuerdo a distintas variables del terreno:

Se establecieron umbrales críticos de precipitación A-D para todo el Valle de Aburrá. No obstante, resulta evidente que las condiciones del terreno varían considerablemente en toda el área estudiada, lo que implica que diferentes zonas pueden mostrar respuestas diversas a los eventos de lluvia. Por esta razón, se clasificaron los eventos de lluvia asociados al inventario de movimientos en masa según diversas variables del terreno, como la geología, cobertura vegetal, usos del suelo, pendientes de las laderas y distintas zonas geográficas. Este enfoque estratificado tiene como objetivo discernir si las características

específicas del terreno pueden influir de manera significativa en la correlación entre los umbrales de precipitación y la ocurrencia de movimientos en masa.

Umbrales de lluvia asociados a tipos de unidades geológicas:

Se llevó a cabo un análisis exploratorio de datos para determinar si existe alguna correlación entre la geología y los umbrales de precipitación. Para ello, se dividió el Valle de Aburrá en cinco principales unidades geológicas: Rocas Volcanosedimentarias, Rocas Ultramáficas, Rocas Metamórficas, Rocas Ígneas Intrusivas y Depósitos Cuaternarios. A cada registro del inventario de deslizamientos se le asignó una categoría geológica basándose en su localización geográfica.

La Figura 28 exhibe el mapa de las unidades geológicas predominantes en el Valle de Aburrá y la distribución espacial de los eventos de movimientos en masa documentados. Se destaca que la mayoría de los sucesos se localizan en las laderas orientales y occidentales de Medellín. La Figura 29A se presenta un diagrama de torta que representa el porcentaje del área total que ocupa cada unidad geológica dentro del Valle. Las Rocas Ígneas Intrusivas son las que predominan, cubriendo el 32.2% del área, seguidas por las Rocas Metamórficas con un 29% y los Depósitos Cuaternarios con un 23.6%. Las Rocas Volcanosedimentarias y las Rocas Ultramáficas representan un 8% y un 7.3% del área, respectivamente.

En la Figura 28B, se clasifican los movimientos en masa registrados por tipo geológico. El 40% de los eventos se ha documentado en Rocas Ultramáficas, a pesar de que esta unidad comprende solo el 7.3% del área del Valle. Esto podría atribuirse a que la mayoría de los registros provienen de las laderas de Medellín, donde esta unidad geológica es predominante. Sin embargo, también puede deberse a las características de los suelos derivados de esta unidad que aumentan la susceptibilidad a la ocurrencia de movimientos en masa ante eventos de lluvias. Los eventos en Rocas Ígneas Intrusivas constituyen un 26.9% y los que ocurrieron en Depósitos Cuaternarios un 21.4%.

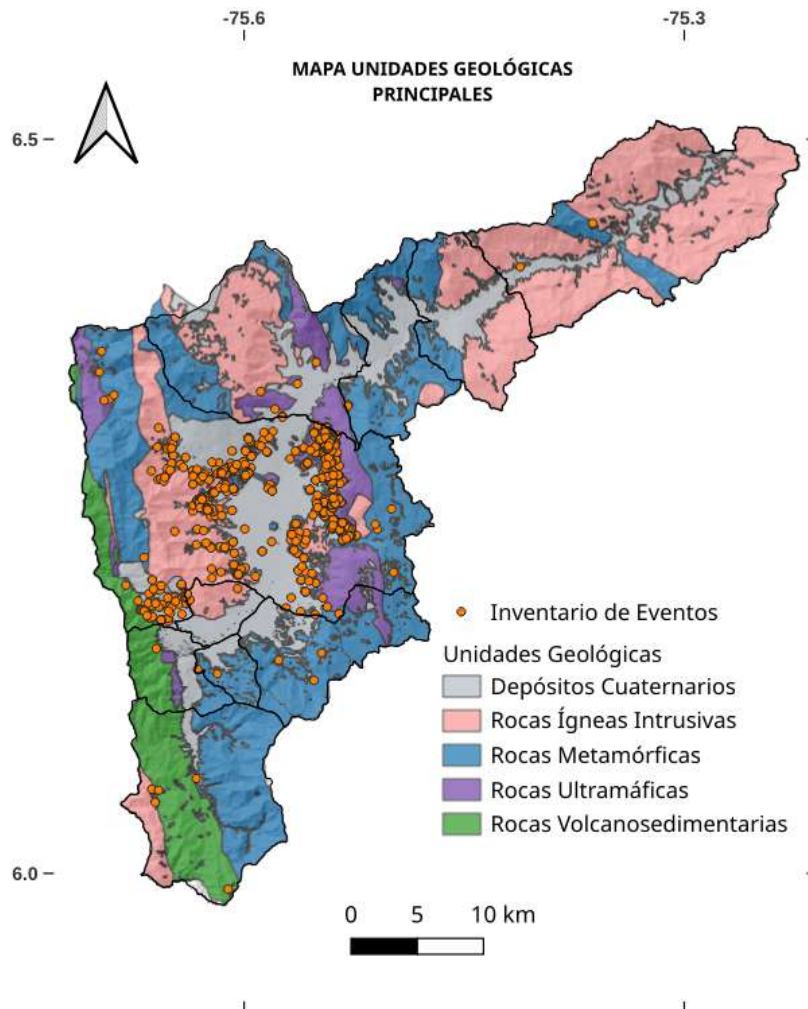


Figura 28: Mapa de unidades geológicas principales en el Valle de Aburrá y la distribución espacial de los eventos de movimientos en masa.

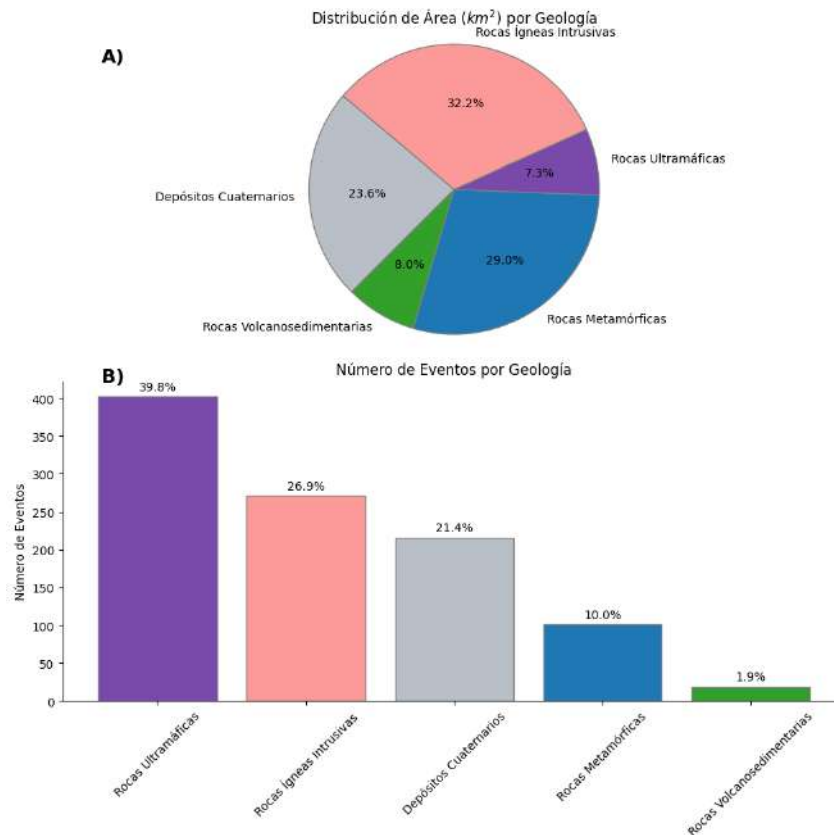


Figura 29: A) Distribución porcentual del área correspondiente a cada unidad geológica en el Valle de Aburrá. (B) Clasificación de los eventos del inventario por unidad geológica, en función de su ubicación geográfica.

En la Figura 30 se analizan las condiciones de lluvia que desencadenaron los movimientos en masa, diferenciadas por unidad geológica (distintos colores). Asimismo, se establecen los umbrales de excedencia del 5 % para cada categoría de roca. Se evidencia que los umbrales más elevados corresponden a las Rocas Metamórficas y a las Rocas Volcanosedimentarias, siendo sus valores muy parecidos entre sí. Por otro lado, los umbrales para las Rocas Ígneas Intrusivas y los Depósitos Cuaternarios presentan similitudes, mientras que para las Rocas Ultramáficas se observa la pendiente más pronunciada en comparación con los demás umbrales.

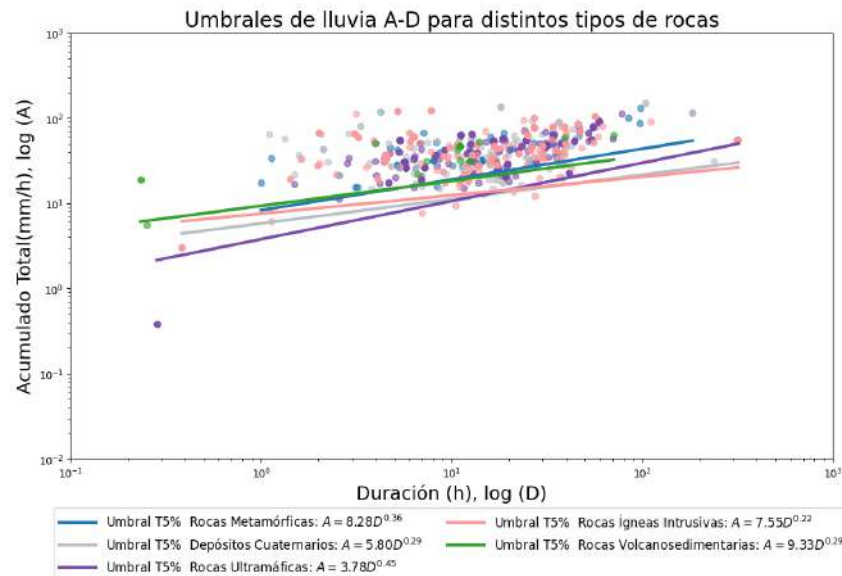


Figura 30: Umbrales críticos de lluvia establecidos para cada una de las unidades geológicas principales del Valle de Aburrá.

Umbrales de precipitación para diversos tipos de cobertura del suelo:

Se efectuó un análisis paralelo al estudio de las unidades geológicas, enfocándose esta vez en las coberturas vegetales y los usos del suelo del Valle de Aburrá. El mapa fue categorizado en seis clases principales: Aguas Continentales, Áreas Agrícolas Heterogéneas, Bosques, Pastos, Zonas de Extracción Minera y Escombreras, y Zonas Urbanizadas. La Figura 31 muestra el mapa de coberturas y usos del suelo, junto con la distribución del inventario de movimientos en masa del Valle.

La Figura 32A ilustra la distribución porcentual del área (en km²) de cada tipo de cobertura en el Valle. Los Bosques constituyen el 47.2 %, seguidos por los Pastos con el 27.5 %, y las Zonas Urbanizadas con el 21 %. Las Áreas Agrícolas Heterogéneas ocupan un 2.9 %, mientras que las Zonas de Extracción Minera, Escombreras y Aguas Continentales comprenden menos del 1 % del área total. En la Figura 32B se muestra la cantidad de eventos de movimientos en masa por tipo de cobertura. Es notable que el 87.1 % de los eventos se registra en Zonas Urbanizadas, lo cual es coherente dado que el inventario recoge eventos que han impactado infraestructura o viviendas. Aproximadamente el 10 % de los eventos suceden en Pastos y Bosques, y menos del 3 % se ubican en Zonas de Extracción Mineras y Áreas Agrícolas Heterogéneas.

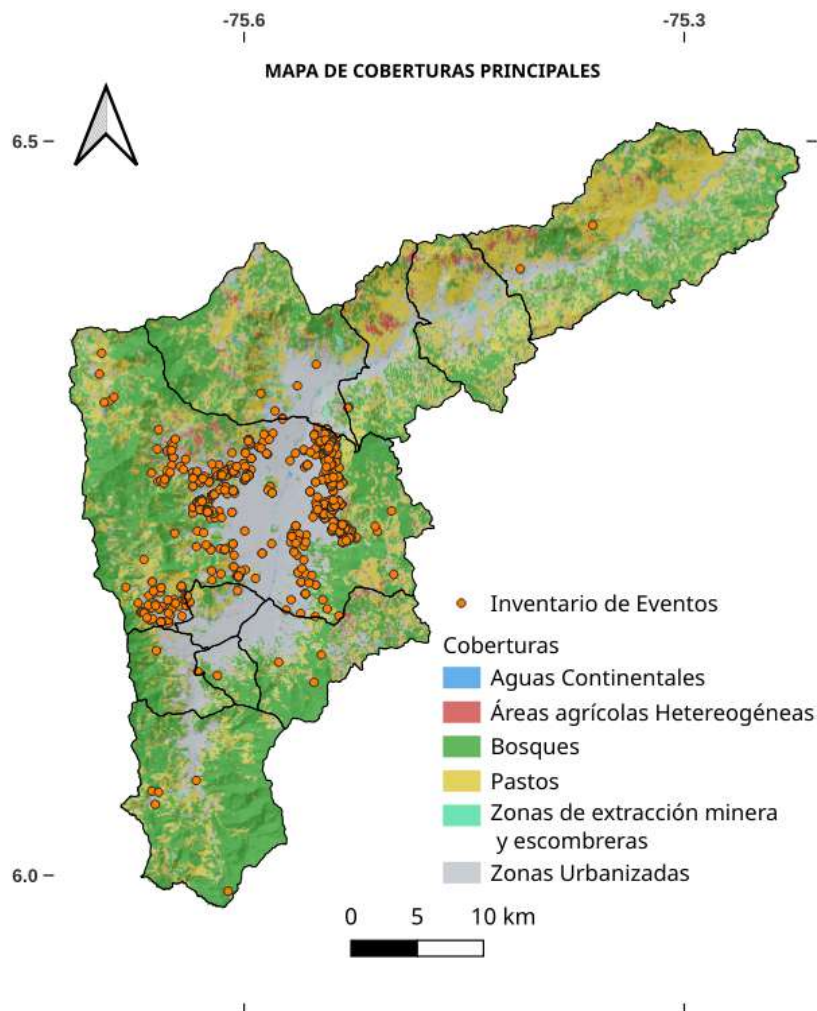


Figura 31: Mapa de coberturas y usos del suelo en el Valle de Aburrá con la distribución espacial de los eventos de movimientos en masa.

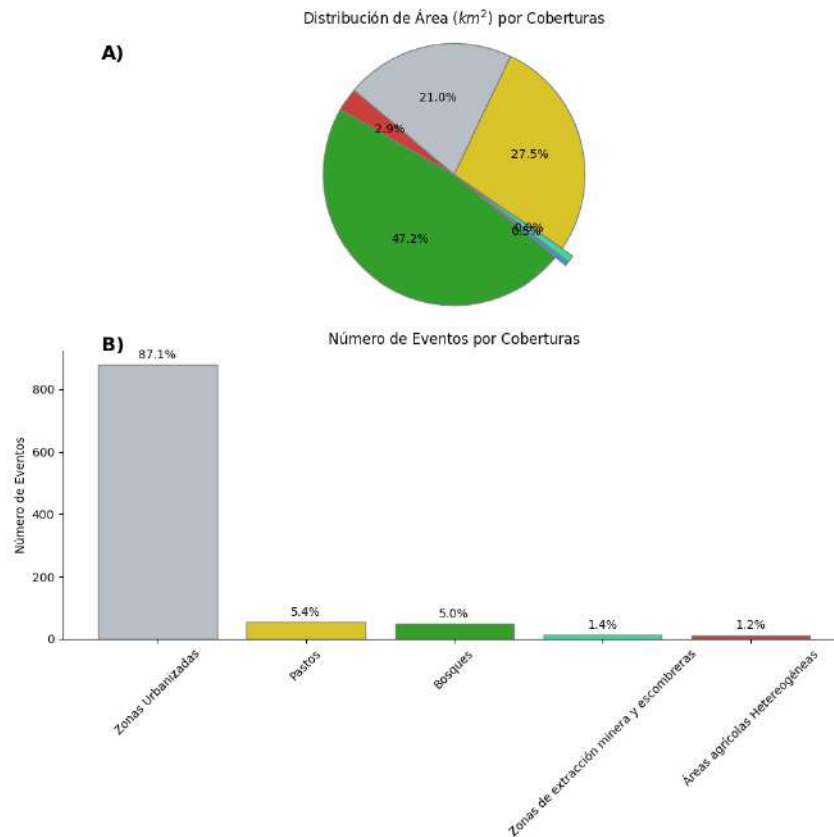


Figura 32: (A) Porcentaje del área que ocupa cada tipo de cobertura en el Valle de Aburrá. (B) Clasificación de los eventos del inventario por tipo de cobertura, basada en su localización geográfica.

La Figura 33 muestra la relación entre las condiciones de precipitación que desencadenaron los deslizamientos del inventario y los tipos de coberturas del suelo. Además, establece los umbrales del 5% de excedencia para cada categoría de cobertura. El umbral más alto corresponde a las Zonas de Extracción Minera y Escombreras, aunque su fiabilidad es limitada debido a que se construyó a partir de solo el 1.2% de los datos del inventario, equivalente a 15 eventos. Los umbrales de Bosques, Zonas Urbanizadas y Áreas Agrícolas Heterogéneas muestran similitudes en su pendiente e intersección. Por otro lado, el umbral correspondiente a los Pastos se interpone entre los umbrales de Bosques y Zonas Urbanizadas.

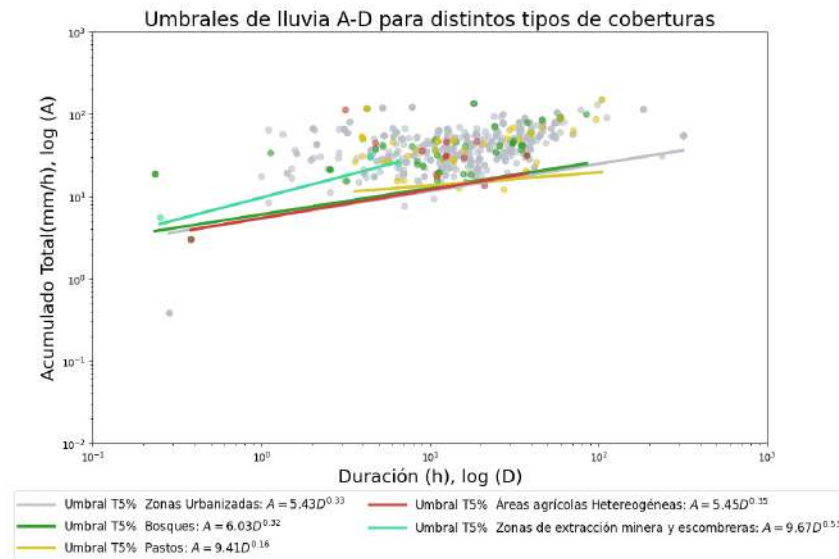


Figura 33: Umbrales críticos de precipitación determinados para cada tipo de cobertura y uso del suelo en el Valle de Aburrá.

Umbrales para diferentes tipos de pendientes de ladera:

La clasificación de las pendientes del terreno en el valle se realizó conforme a los intervalos definidos por el Servicio Geológico Colombiano Colombiano (2017). La Tabla 5 muestra esta clasificación basada en los grados de inclinación. Asimismo, la Figura 34 exhibe el mapa de pendientes del terreno para todo el Valle de Aburrá, junto con la distribución de los eventos del inventario de movimientos en masa.

Tabla 5: Calificación de Pendientes con base en la Susceptibilidad a los Movimientos en Masa, tomado de Colombiano (2017)

INCLINACIÓN (GRADOS)	DESCRIPCIÓN	DESCRIPCIÓN DEL COM-PORTAMIENTO
< 7	PLANA A SUAVEMENTE INCLINADA	Muy baja tendencia a desarrollar movimientos en masa
7 – 11	INCLINADA	Baja tendencia a generar movimientos en masa
11 – 19	MUY INCLINADA	Media tendencia a generar movimientos en masa
19 – 40	ABRUPTA	Alta tendencia a generar movimientos en masa
> 40	ESCARPADA	Muy alta tendencia a generar movimientos en masa

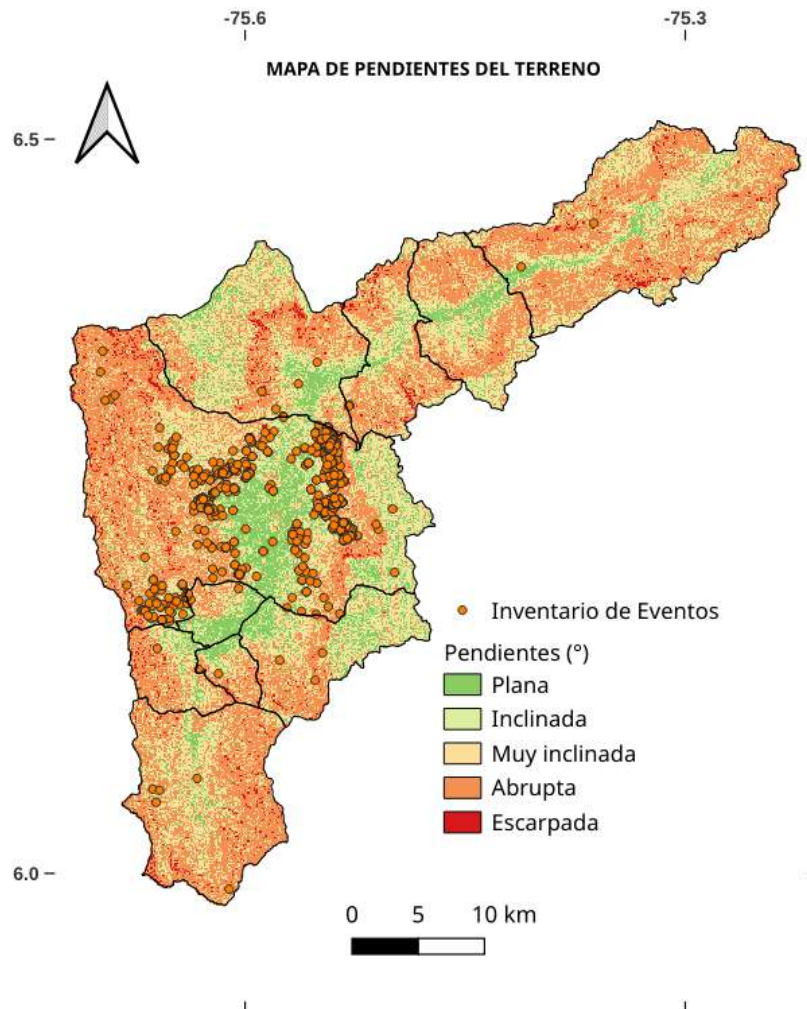


Figura 34: Mapa de clasificación de pendientes en el Valle de Aburrá y la distribución espacial de los eventos de movimientos en masa.

La Figura 35A ilustra el porcentaje de distribución areal (en km²) de las distintas categorías de pendiente en el Valle de Aburrá. Predominan las pendientes categorizadas como Abruptas (19-40 °), las cuales comprenden el mayor porcentaje de superficie, seguidas de las pendientes Muy inclinadas (11-19 °), que ocupan el 28 % del área total del valle. Las clasificaciones de pendientes Inclinada, Plana a Suavemente Inclinada y Escarpada ocupan menos del 15 % de la superficie total del Valle de Aburrá, con las Escarpadas representando solo el 3 %. La Figura 35B muestra la cantidad de eventos del inventario correspondientes a cada tipo de pendiente, destacando las Abruptas con más de 500 eventos, equivalentes al 52 % del total, y las Muy inclinadas con cerca de 300 eventos, o el 30 % del total. Las categorías de pendientes Plana, Inclinada y Escarpada acumulan menos del 10 % de los eventos del inventario.

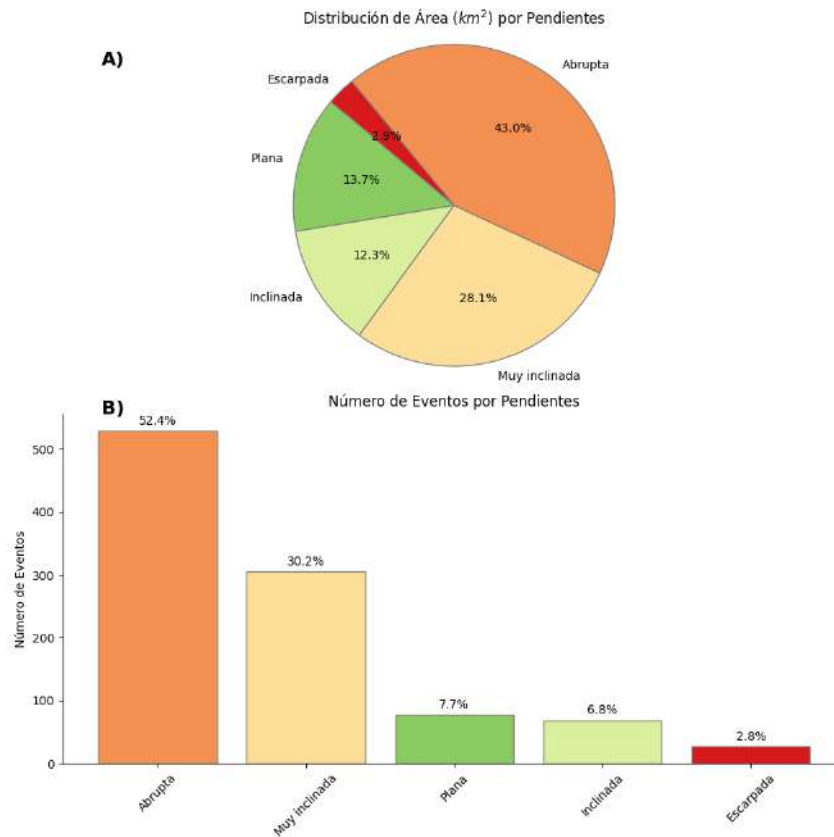


Figura 35: (A) Distribución porcentual del área correspondiente a cada categoría de pendientes en el Valle de Aburrá. (B) Distribución de los eventos del inventario según el tipo de pendiente, basada en su ubicación geográfica.

En la Figura 36 se muestra la distribución de los eventos de lluvia que han provocado movimientos en masa, organizados por la inclinación de las pendientes. Además, se definen los umbrales de excedencia del 5 % para cada tipo de pendiente. Se observan diferencias entre los umbrales; el más elevado corresponde a las pendientes Abruptas, aunque este se basa en aproximadamente 30 eventos del inventario. Los umbrales de las pendientes Abruptas y Muy inclinadas son similares, pero se intersectan en duraciones superiores a las 20 horas. Los umbrales para pendientes Inclinadas y Planas son más altos, aunque se construyeron con menos de 100 eventos.

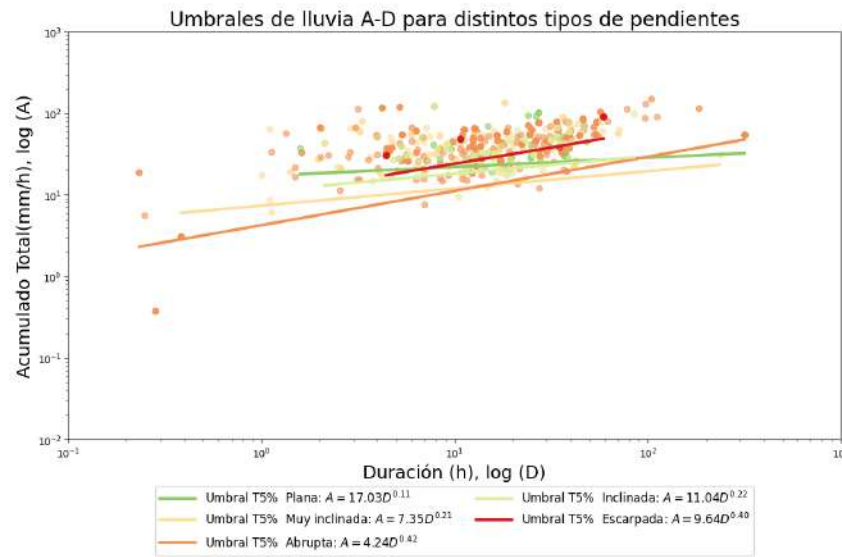


Figura 36: Distribución de los umbrales críticos de precipitación para diferentes tipos de pendientes del terreno.

Umbral de precipitación por zonas geográficas en Medellín:

La categorización de umbrales de precipitación se efectuó por zonas geográficas. Para este estudio, se centró la atención en el municipio de Medellín, que concentra el 90 % de los datos del inventario de eventos. El municipio se dividió en seis zonas, denominadas: Noroccidental, Nororiental, Suroriental, Suroccidental, San Antonio de Prado y Palmitas. La Figura 37 muestra el mapa de Medellín dividido en estas zonas, junto con la distribución del inventario de eventos de movimientos en masa. Se destaca que todas las zonas, excepto Palmitas, muestran una abundante concentración de eventos.

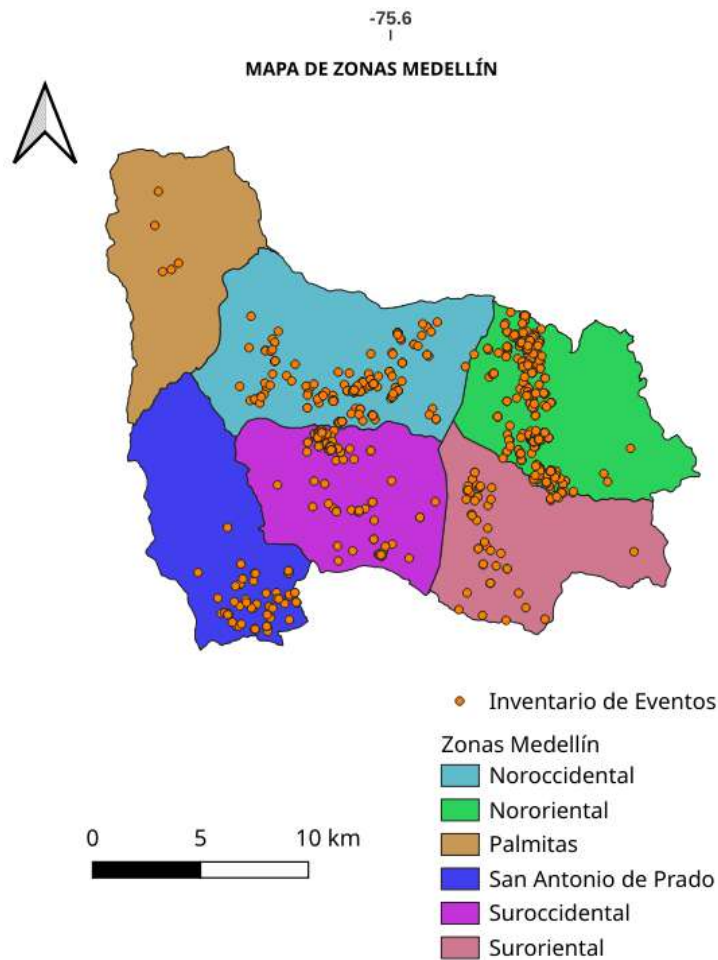


Figura 37: Mapa del municipio de Medellín dividido por zonas geográficas y la distribución espacial de los eventos de movimientos en masa registrados en el inventario.

La Figura 38A ilustra la distribución porcentual del área que comprende cada zona en relación con el área total de Medellín. Se observa una distribución bastante equitativa entre las zonas, variando del 15 al 22 %. Las zonas Nororiental y Noroccidental son las que ocupan una mayor proporción del área, con un 21 % y 17 % respectivamente. La Figura 38B muestra la cantidad de eventos del inventario distribuidos por zonas. La zona Nororiental destaca con casi la mitad de los eventos, representando el 47 % del total. Le sigue la zona Noroccidental con el 27 %. En contraste, Palmitas registra menos del 0.5 % de los eventos.

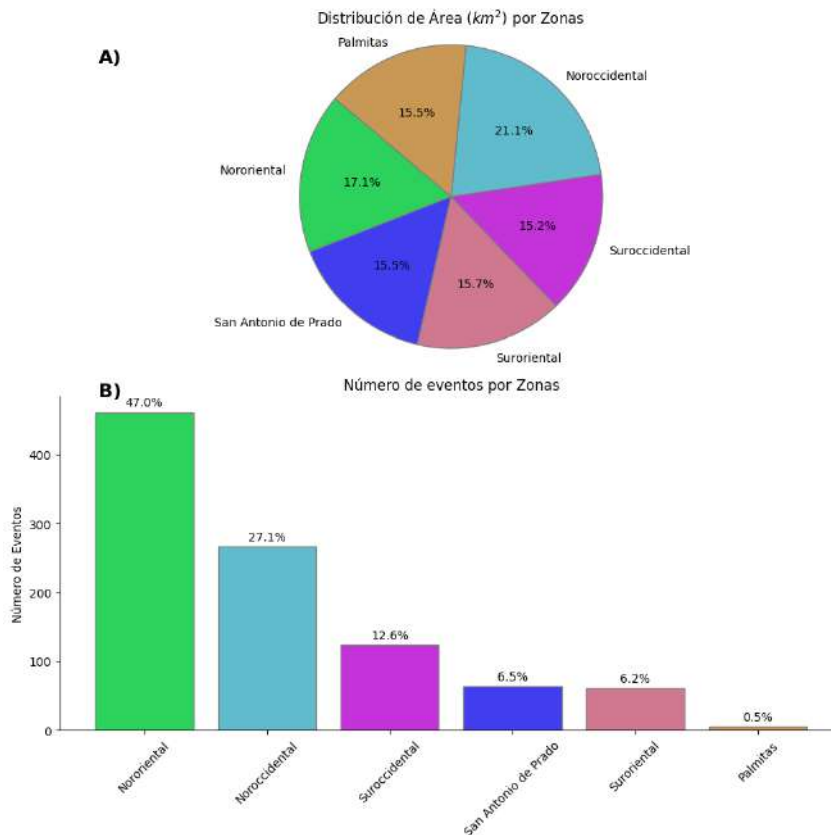


Figura 38: ((A) Porcentaje del área total de Medellín correspondiente a cada zona geográfica. (B) Número de eventos del inventario clasificados por zona.

La Figura 39 presenta la distribución de las condiciones de precipitación asociadas con los movimientos en masa, clasificadas por zonas de Medellín. También se determinan los umbrales del 5 % de excedencia para cada zona. El umbral para Palmitas es el más alto; sin embargo, su validez se ve limitada por la escasa cantidad de datos. Los umbrales para San Antonio de Prado y la zona Suroccidental son similares, con una pendiente de línea cercana, distinta a la del umbral de la zona Nororiental. La zona Noroccidental presenta el umbral más bajo, aunque sus valores se solapan con los de la zona Suroccidental para duraciones superiores a 20 horas.

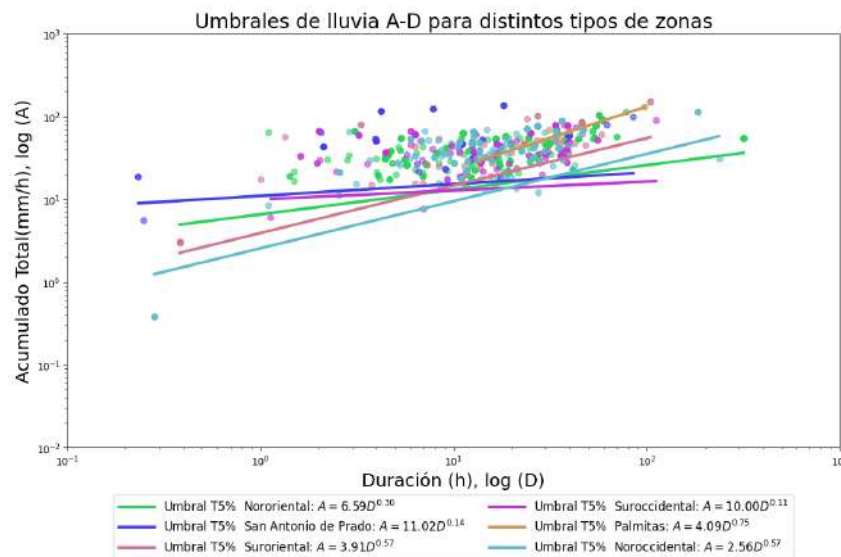


Figura 39: Umbrales de precipitación críticos asociados a movimientos en masa en diferentes zonas de Medellín.

En conclusión, tras el análisis de la lluvia clasificado según diversas variables del terreno, no se encontraron diferencias significativas para las distintas clases evaluadas en cada una de las variables estudiadas. Esto podría atribuirse principalmente a dos factores: en primer lugar, al sesgo del inventario de eventos analizados, dado que muchos de estos eventos están registrados en zonas urbanas, y es posible que muchos de ellos no estén directamente asociados a la lluvia como factor desencadenante. En segundo lugar, el comportamiento de la lluvia en relación con los movimientos en masa podría, de alguna manera, aplicarse de manera general a todo el Valle de Aburrá. Es decir, por el momento, se podría trabajar con un umbral crítico para todo el Valle de Aburrá. No obstante, este umbral tendría una incertidumbre asociada al inventario de eventos analizado.

Definición de eventos de lluvia para la construcción de umbrales de A-D (Acumulado - Duración):

Es importante identificar con precisión los eventos de lluvia que detonaron los movimientos en masa. Para ello, es necesario establecer criterios que diferencien un evento de precipitación de otro. Hay dos criterios principales: el primero se refiere a que el pluviómetro registre un valor superior a 0, indicando el comienzo del evento. El segundo implica definir el número de horas sin lluvia que se requiere para considerar que un evento ha concluido y comienza otro. Según la investigación de Rosi et al. (2015), la definición de este intervalo sin lluvia es clave, ya que influye significativamente en la sensibilidad de los umbrales de lluvia establecidos.

Por esta razón, se realizó un análisis minucioso de los umbrales de precipitación Acumulado-Duración (A-D), centrándose especialmente en la variación del segundo criterio que define un evento de lluvia. Este enfoque particular se orientó a evaluar cómo los distintos intervalos de tiempo sin lluvia afectan la determinación de un evento de lluvia y su respectivo umbral. Para ello, se examinaron los umbrales de precipitación para una amplia gama de intervalos sin lluvia, que incluyen 10, 15, 30 y 45 minutos, así como 1, 2, 3, 4, 5, 6, 7, 8... hasta 48 horas sin lluvia. En cada uno de estos intervalos, se identificaron los umbrales de precipitación asociados a varias probabilidades de excedencia, incluyendo 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 20 % y 50 %.

Finalmente, para cada intervalo de tiempo sin lluvia, se seleccionó la probabilidad de excedencia que ofrecía los resultados más confiables y adecuados de acuerdo con las métricas de validación. Este proceso permitió afinar los umbrales y mejorar significativamente la precisión en la identificación de eventos de lluvia potencialmente causantes de movimientos en masa.

Para el cálculo y evaluación de los umbrales, se partió de la base de 991 eventos de lluvias que ocasionaron movimientos en masa. De estos, el 80 % (aproximadamente 793 eventos) se utilizó para la construcción de los umbrales, mientras que el 20 % restante (aproximadamente 198 eventos) se destinó a la validación de los mismos. Adicionalmente, para una validación más robusta, se incorporaron datos de eventos de lluvia históricos tomados de manera aleatoria registrados por los pluviómetros que no resultaron en movimientos en masa. A partir de esta información, se evaluaron cada uno umbrales para cada intervalo de lluvia definido y cada una de las probabilidades de excedencia mencionadas anteriormente.

Los resultados de estas métricas, evaluados para distintos intervalos de tiempo sin lluvia y con la mejor probabilidad de excedencia evaluada se presentan en la tabla 6. La tabla inicia desde el minuto 10, luego 15,30 y 45 para luego pasar a horas desde la hora 1 hasta 48. Cada intervalo de tiempo presenta las métricas para la probabilidad de excedencia que mejor obtuvo resultados.

Tabla 6: Resultados métricas evaluadas para los umbrales de A-D para cada intervalo de tiempo.

Intervalo sin lluvia	Probabilidad de excedencia	TP	FP	FN	TN	POD	POFD	POFA	HK
10 min	0.08	22	1	5	5	0.815	0.1670	0.043	0.648
15 min	0.04	43	5	2	20	0.956	0.200	0.104	0.756
30 min	0.07	72	16	4	33	0.947	0.327	0.182	0.621
45 min	0.08	81	21	9	51	0.900	0.292	0.206	0.608
1 h	0.05	92	22	5	47	0.948	0.319	0.193	0.630
2 h	0.09	106	25	10	67	0.914	0.272	0.191	0.642
3 h	0.10	114	22	13	71	0.898	0.237	0.162	0.661
4 h	0.10	122	18	14	80	0.897	0.184	0.129	0.713
5 h	0.08	134	21	8	89	0.944	0.191	0.135	0.7530
6 h	0.10	138	18	8	98	0.945	0.155	0.115	0.790
7 h	0.10	145	21	4	100	0.973	0.174	0.127	0.800
8 h	0.10	147	24	3	98	0.980	0.197	0.140	0.783
9 h	0.08	149	23	4	105	0.9740	0.180	0.134	0.794
10 h	0.09	152	28	3	102	0.981	0.215	0.156	0.765
11 h	0.08	152	25	4	108	0.974	0.188	0.141	0.786
12 h	0.09	154	25	4	111	0.975	0.184	0.140	0.791
13 h	0.09	152	23	6	113	0.962	0.169	0.131	0.793
14 h	0.09	153	24	5	112	0.968	0.176	0.136	0.792
15 h	0.09	153	25	5	111	0.968	0.184	0.140	0.785
16 h	0.07	154	30	5	111	0.969	0.213	0.163	0.756
17 h	0.09	152	30	7	111	0.956	0.213	0.165	0.743
18 h	0.10	154	27	5	108	0.969	0.200	0.149	0.769
19 h	0.10	157	34	4	112	0.975	0.233	0.178	0.742
20 h	0.09	157	35	4	111	0.975	0.240	0.182	0.735
21 h	0.10	155	37	6	113	0.963	0.247	0.193	0.716
22 h	0.10	155	38	6	112	0.963	0.253	0.197	0.709
23 h	0.09	157	43	4	107	0.975	0.287	0.215	0.688
24 h	0.08	157	43	4	107	0.975	0.287	0.215	0.688
25 h	0.08	157	43	4	109	0.975	0.283	0.215	0.692

Intervalo sin lluvia	Probabilidad de excedencia	TP	FP	FN	TN	POD	POFD	POFA	HK
26 h	0.08	157	43	4	111	0.975	0.279	0.215	0.696
27 h	0.10	157	42	4	112	0.975	0.273	0.211	0.702
28 h	0.08	157	44	4	110	0.975	0.286	0.219	0.689
29 h	0.08	157	44	4	110	0.975	0.286	0.219	0.689
30 h	0.08	156	44	5	110	0.969	0.286	0.220	0.683
31 h	0.08	156	44	5	110	0.969	0.286	0.220	0.683
32 h	0.08	155	44	6	110	0.963	0.286	0.221	0.677
33 h	0.10	155	42	6	112	0.963	0.273	0.213	0.690
34 h	0.10	155	43	6	111	0.963	0.279	0.217	0.684
35 h	0.10	155	44	6	111	0.963	0.284	0.221	0.679
36 h	0.10	155	45	6	110	0.963	0.290	0.225	0.672
37 h	0.10	155	45	6	110	0.963	0.290	0.225	0.672
38 h	0.10	155	45	6	110	0.963	0.290	0.225	0.672
39 h	0.10	154	45	7	110	0.957	0.290	0.226	0.666
40 h	0.09	154	45	7	110	0.957	0.290	0.226	0.666
41 h	0.09	154	45	7	110	0.957	0.290	0.226	0.666
42 h	0.10	154	45	7	110	0.957	0.290	0.226	0.666
43 h	0.10	154	45	7	110	0.957	0.290	0.226	0.666
44 h	0.10	154	44	7	111	0.957	0.284	0.222	0.673
45 h	0.10	154	45	7	110	0.957	0.290	0.226	0.666
46 h	0.10	154	46	7	109	0.957	0.297	0.230	0.660
47 h	0.10	154	46	7	110	0.957	0.295	0.230	0.662
48 h	0.10	155	46	6	110	0.963	0.295	0.229	0.668

En las Figuras 40 y 41 se presenta el resumen de los valores dados por las métricas para todos los intervalos sin lluvia definidos. En la Figura 40 se observan los valores de la tabla de contingencia, los cuales se explican a continuación:

- **Figura 40A:** Ilustra la cantidad de TP (Verdaderos Positivos) en función del intervalo sin lluvia. La línea azul indica que el número de TP aumenta rápidamente en las primeras horas y luego, a partir de la hora 10 sin lluvia, la cantidad de Verdaderos Positivos se estabiliza.
- **Figura 40B:** Muestra la cantidad de TN (Verdaderos Negativos) en función del intervalo sin lluvia. Similar a la Figura 40A, la línea verde exhibe un incremento inicial seguido de una estabilización alrededor de la hora 10 sin lluvia, lo que sugiere un cambio menor en la cantidad de TN a partir de esta hora.
- **Figura 40C:** Representa la cantidad de FP (Falsos Positivos) en función del intervalo sin lluvia. A diferencia de las Figuras 40A y 40B, aquí se muestra un incremento inicial seguido de un descenso gradual entre las 5 y las 9 horas, para luego aumentar gradualmente entre las 10 y 20 horas y estabilizarse a partir de las 23 horas. Se observa que entre las 5 y 15 horas, el número de Falsos Positivos es menor a 30.
- **Figura 40D:** Presenta la cantidad de FN (Falsos Negativos) en función del intervalo sin lluvia. Idealmente, se busca tener una cantidad baja de FN, ya que representan los eventos no detectados por estar por debajo del umbral. Se observa un incremento en las primeras horas, pero entre las 8 y 11 horas se presentan los valores mínimos. Posteriormente, la cantidad de FN aumenta levemente comparado con las primeras horas.

Por otro lado, la Figura 41 muestra las métricas evaluadas para cada uno de los intervalos sin lluvia, detalladas en la Tabla 6. El comportamiento de cada una de las métricas se explica a continuación:

- **Figura 41A:** Exhibe los valores de la métrica POD (Probabilidad de Detección) en función del intervalo sin lluvia. Se observa un aumento significativo en las primeras horas seguido de una disminución hasta la hora 5. Los mejores valores se presentan entre las horas 8 y 11, después de lo cual los valores se mantienen relativamente constantes.
- **Figura 41B:** Muestra los valores de la métrica HK (Hanssen y Kuipers) en función del intervalo sin lluvia. La gráfica exhibe variaciones abruptas en las primeras horas, pero a partir de la hora 5 se observa un aumento gradual, alcanzando los mejores valores entre las horas 8 y 15. Después de la hora 15, los valores disminuyen gradualmente.
- **Figura 41C:** Ilustra los valores de la métrica POFD (Probabilidad de Falsas Alarmas) en función del intervalo sin lluvia. Los valores óptimos de esta métrica es cuando se aproximan a 0. La gráfica muestra aumentos significativos en las primeras horas seguidos de una disminución desde la hora 5 hasta la 15. Posteriormente, se observa un aumento en los valores de esta métrica.
- **Figura 41D:** Presenta los valores del índice POFA (Porcentaje de Falsas Alarmas) en función del intervalo sin lluvia. Al igual que la métrica POFD, los valores óptimos se acercan a 0. Se observa un incremento inicial seguido de una disminución gradual hasta la hora 15, momento a partir del cual los valores aumentan gradualmente.

Tabla de contingencia

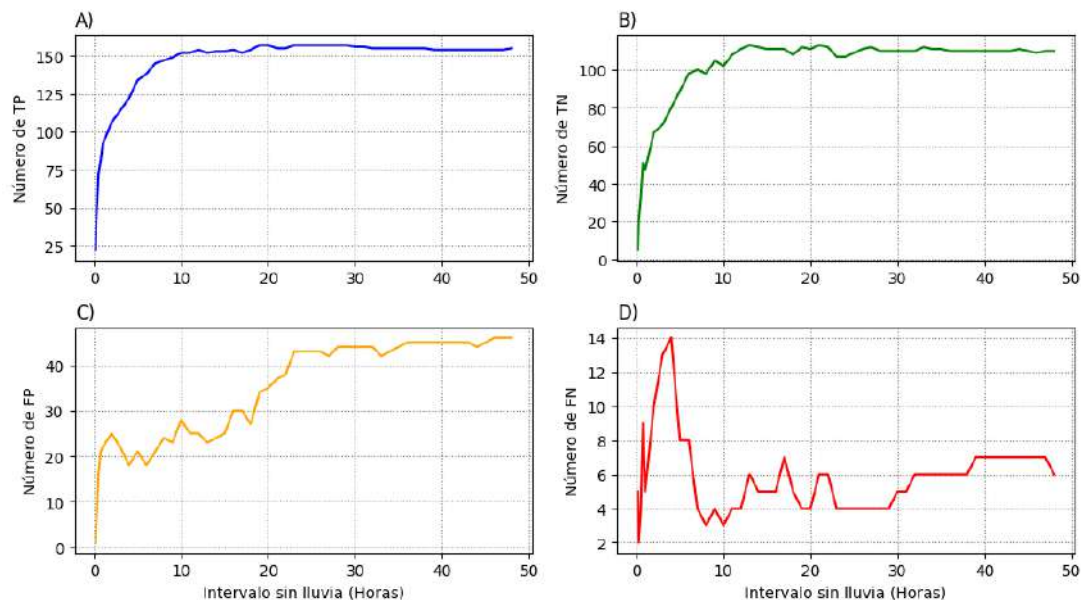


Figura 40: Representación gráfica de los valores de la tabla de contingencia en función del tiempo transcurrido sin lluvia. A) Cantidad de Verdaderos Positivos (TP) vs. horas sin lluvia. B) Cantidad de Verdaderos Negativos (TN) vs. horas sin lluvia. C) Número de Falsos Positivos (FP) vs. horas sin lluvia, indicando un incremento inicial. D) Número de Falsos Negativos (FN) vs. horas sin lluvia.

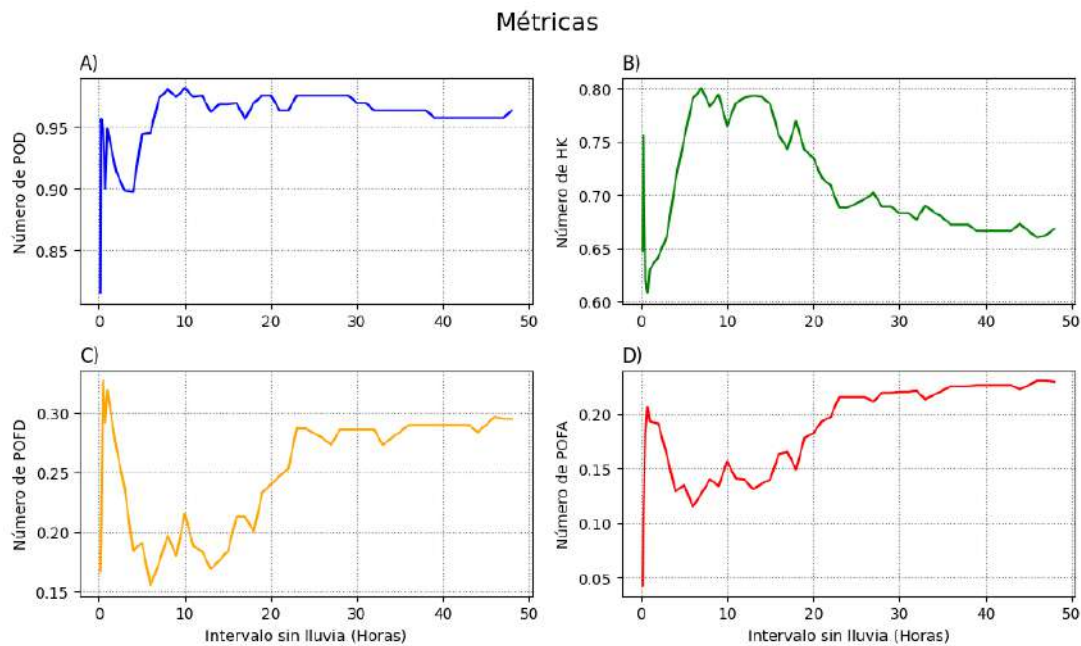


Figura 41: Análisis gráfico de métricas de rendimiento en función de las horas sin lluvia. A) Probabilidad de Detección (POD) vs. horas sin lluvia. B) Puntuación de Hanssen y Kuipers (HK) vs. horas sin lluvia. C) Tasa de Falsas Alarmas (POFD) vs. horas sin lluvia. D) Índice de Falsas Alarmas (POFA) vs. horas sin lluvia.

En conclusión, los gráficos previamente analizados revelan una tendencia consistente entre las 6 y las 15 horas, donde se registran los valores más óptimos según las métricas evaluadas y la tabla de contingencia. Esta observación sugiere que el intervalo ideal sin lluvia para definir eventos de precipitación se sitúa en este rango horario. Finalmente, en la Figura 42 ilustra el umbral para eventos de lluvia determinados por intervalos sin lluvia de 8 horas y con una probabilidad de excedencia del 10%. Es notable la reducida cantidad de Falsos Negativos (3) y Falsos Positivos (23), indicando que este umbral podría ser óptimo para su evaluación en tiempo real a través de la herramienta SIAMMO (Sistema Integrado de Alertas y Monitoreo de Movimientos en Masa).

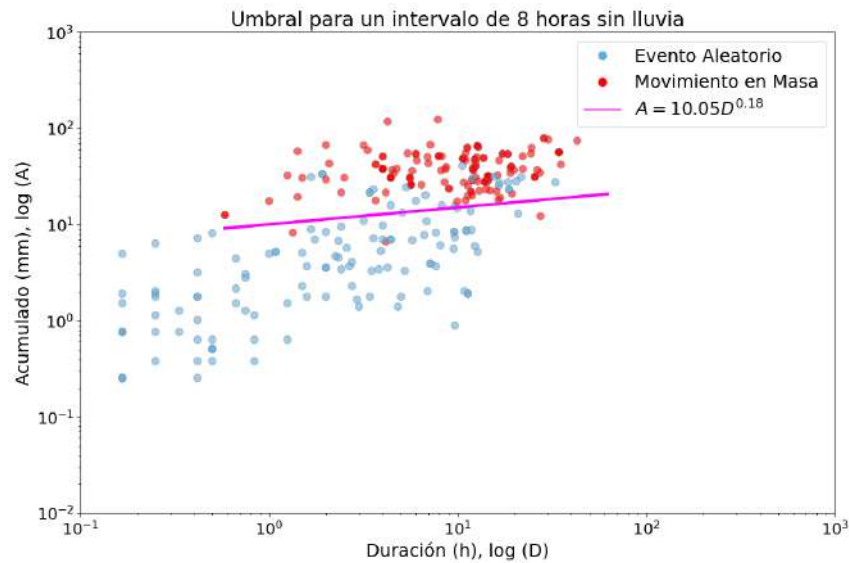


Figura 42: Umbral de lluvia evaluado para eventos definidos con un intervalo sin lluvia de 8 horas y una probabilidad de excedencia del 10 %. Los puntos rojos y azules representan eventos de validación, es decir, eventos de lluvia con y sin deslizamiento no considerados en la construcción del umbral.

Referencias

- Abraham, M. T., Satyam, N., Rosi, A., Pradhan, B., and Segoni, S. (2020). The selection of rain gauges and rainfall parameters in estimating intensity-duration thresholds for landslide occurrence: Case study from wayanad (india). *Water (Switzerland)*, 12.
- Balducci, V., Guzzetti, F., Peruccacci, S., and Rossi, M. (2006). Rainfall thresholds for the initiation of landslides.
- Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F. (2010). Natural hazards and earth system sciences rainfall thresholds for the possible occurrence of landslides in italy.
- Caine, N. (1980). The rainfall intensity: Duration control of shallow landslides and debris flows. *Geografiska Annaler. Series A, Physical Geography*, 62(1/2):23–27.
- Chleborad, A. F., Baum, R. L., Godt, J. W., and Leahy, P. P. (2006). Rainfall thresholds for forecasting landslides in the seattle, washington, area-exceedance and probability.
- Colombiano, S. G. (2017). *Guía metodológica para la zonificación de amenaza por movimientos en masa escala 1: 25.000*.
- Correa, O., García, F., Bernal, G., Cardona, O. D., and Rodriguez, C. (2020). Early warning system for rainfall-triggered landslides based on real-time probabilistic hazard assessment. *Natural Hazards*, 100:345–361.
- Gariano, S. L., Brunetti, M. T., Iovine, G., Melillo, M., Peruccacci, S., Terranova, O., Vennari, C., and Guzzetti, F. (2015). Calibration and validation of rainfall thresholds for shallow landslide forecasting in sicily, southern italy. *Geomorphology*, 228:653–665.
- Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P. (2007). Rainfall thresholds for the initiation of landslides in central and southern europe. *Meteorology and Atmospheric Physics*, 98:239–267.
- Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P. (2008). The rainfall intensity-duration control of shallow landslides and debris flows: An update. *Landslides*, 5:3–17.
- Peruccacci, S., Brunetti, M. T., Gariano, S. L., Melillo, M., Rossi, M., and Guzzetti, F. (2017). Rainfall thresholds for possible landslide occurrence in italy. *Geomorphology*, 290:39–57.
- Peruccacci, S., Brunetti, M. T., Luciani, S., Vennari, C., and Guzzetti, F. (2012). Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central italy. *Geomorphology*, 139-140:79–90.
- Piciullo, L., Gariano, S. L., Melillo, M., Brunetti, M. T., Peruccacci, S., Guzzetti, F., and Calvello, M. (2017). Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides. *Landslides*, 14:995–1008.
- Rosi, A., Lagomarsino, D., Rossi, G., Segoni, S., Battistini, A., and Casagli, N. (2015). Updating ews rainfall thresholds for the triggering of landslides. *Natural Hazards*, 78:297–308.