

CLASIFICACIÓN DE ZONAS PROPENSAS A MOVIMIENTOS EN MASA COMO
FUENTE DE MATERIALES PARA AVENIDAS TORRENCIALES EN LA
CUENCA DEL RÍO LA PAILA (CORINTO - CAUCA) MEDIANTE
INTELIGENCIA ARTIFICIAL

EIMY DAYANA PARUMA VÁSQUEZ
JHONATAN ALEXANDER GUTIERREZ RONDÓN

Documento presentado al
Programa Académico de Ingeniería Topográfica
como requisito para optar por el título en
Ingeniería Topográfica

Director: LINA MARÍA OSPINA OSTIOS, Ph.D

Codirector: WILFREDO ALFONSO-MORALES, Ph.D.

ESCUELA DE INGENIERÍA CIVIL Y GEOMÁTICA
FACULTAD DE INGENIERÍA



Santiago de Cali, octubre 2022

Nota de aceptación

Aprobado por el comité de grado en cumplimiento de los requisitos exigidos por la Universidad del Valle para optar al título en Ingeniería Topográfica.

Jurado

Jurado

Santiago de Cali, octubre 2022

Agradecimientos

A nuestras familias, especialmente a nuestros padres gracias por su apoyo y acompañamiento sin condición, pues por ustedes hemos llegado hasta aquí con su ejemplo y trabajo, siendo el motor que nos impulsa a mejorar cada día para alcanzar las metas que nos proponíamos; los amamos.

A nuestros profesores que a lo largo de nuestra formación académica nos impartieron los conocimientos necesarios que hoy nos permiten escalar un peldaño académico más. Especialmente queremos agradecer a nuestros tutores Lina Ospina y Wilfredo Alfonso, por ser quienes nos guiaron y acompañaron durante el desarrollo de este importante proceso para nuestra formación profesional; su constante asesoramiento fue fundamental para llevar a cabo con éxito y satisfacción nuestro trabajo de grado.

Finalmente pero no menos importante, queremos agradecer a nuestros amigos, pues su apoyo fue de gran importancia en nuestro paso por la universidad, haciendo ameno este largo camino en el cual reímos, lloramos y compartimos muchos momentos que no hubieran sido lo mismo sin su compañía.

Resumen

Los fenómenos naturales (FN) son eventos que pueden afectar un territorio sin importar su ubicación o su desarrollo económico, relacionándose directamente con los factores condicionantes y detonantes inherentes a la zona donde se presentan. Colombia por su topografía, el clima y su ubicación geográfica, es un país propenso a la ocurrencia de diferentes FN como los movimientos en masa (MM). Adicionalmente, si a dichas condiciones se le suman variables políticas, económicas y sociales que dificultan el acceso a los territorios como los comprendidos por la cuenca del río La Paila (municipio de Corinto, Cauca), se hace imperativo desarrollar metodologías que permitan dotar a las comunidades de instrumentos que mejoren su capacidad de respuesta ante la posible ocurrencia de MM. El presente trabajo de grado desarrolla una metodología basada en técnicas de inteligencia artificial, las cuales permitieron predecir la ocurrencia o ausencia de un MM considerando tanto los factores condicionantes (geología, usos del suelo y pendientes) como los detonantes (precipitaciones). Se observó una buena correlación entre las variables analizadas como lo muestra la curva de aprendizaje, la cual tuvo una tendencia a estabilizarse y la precisión se mantuvo por encima del 80% en la mayoría de épocas definidas en el entrenamiento. Aunque el modelo construido arrojó una validación de entrenamiento óptima y la validación de los resultados con un AUC de 0.98 para predicciones con 10 días de anticipación; los resultados cartográficos presentan inconsistencias espaciales debido a la escala de los datos utilizados para el entrenamiento y las predicciones. Se concluye que si bien las variables usadas presentan una alta correlación para la identificación de posibles MM, se hace necesario mejorar la escala de los datos fuente buscando que la precisión cartográfica de los resultados mejore y así su comportamiento espacial muestre patrones físicos acordes al fenómeno estudiado.

Palabras clave:— Corinto, movimiento en masa (MM), inteligencia artificial (IA), gestión del riesgo (GR), factores condicionantes y detonantes, redes neuronales recurrentes (RNN).

Tabla de Contenido

	Pág.
Agradecimientos	II
Resumen	I
Glosario	II
Tabla de Contenido	II
Índice de Figuras	IV
Índice de Tablas	VI
1. Introducción.....	1
1.1 Objetivos	6
1.1.1 Objetivo general	6
1.1.2 Objetivos específicos.....	6
1.2 Organización del Documento	6
2. Marco Teórico	7
2.1 Antecedentes.....	7
2.2 Marco Conceptual	10
2.2.1 Movimiento en Masa - MM	10
2.2.2 Avenidas torrenciales	14
2.2.3 Pendientes	15
2.2.4 Geología.....	15
2.2.5 Uso del suelo.....	17
2.2.6 Precipitación	17
2.2.7 Métodos indirectos	18
2.2.7.1 Percepción remota.....	18
2.2.7.2 Machine Learning (ML).	18
2.3 Principales Hallazgos	23
3. Metodología	25
3.1 Localización de la zona de estudio	25
3.1.1 Cuenca del Río La Paila	25

3.1.2	Clima.....	26
3.2	Esquema metodológico	27
3.2.1	Fase 1.....	28
3.2.1.1	Selección de factores condicionantes y detonantes en la zona de estudio.....	28
3.2.1.2	Compilación y depuración de datos	28
3.2.2	Fase 2.....	36
3.2.2.1	Procesamiento de datos.....	36
3.2.2.2	Validación de red neuronal	43
3.2.3	Fase 3.....	43
3.2.3.1	Clasificación de zonas propensas a MM que son fuente de materiales para avenidas torrenciales	43
3.3	Síntesis.....	44
4.	Resultados y Discusiones.....	45
4.1	Fase 1.....	45
4.2	Fase 2.....	46
4.2.1	Compilación y depuración de datos	46
4.2.1.1	Geología	46
4.2.1.2	Usos de suelo.....	50
4.2.1.3	Pendiente.....	53
4.2.1.4	Precipitación	56
4.2.2	Procesamiento de datos	60
4.2.3	Validación.....	61
4.3	Fase 3.....	65
4.4	Síntesis.....	70
5.	Conclusiones y Trabajos Futuros	71
5.1	Conclusiones Generales	71
5.2	Trabajos Futuros.....	73
	Referencias.....	74
6.	Anexos	81
6.1	Clasificación Enero / 20 / 2011, Cuenca del río La Paila (Figura 4.22)	81
6.2	Clasificación Noviembre / 07 / 2017, Cuenca del río La Paila (Figura 4.23) ..	81
6.3	Clasificación Julio / 19 / 2020, Cuenca del río La Paila (Figura 4.24)	81
6.4	Zonificación propuesta para la cuenca del río La Paila (Figura 4.25)	81
6.5	Modelo de red neuronal con tratamiento de datos (Notebook Colab)	81

ÍNDICE DE FIGURAS

FIG.		Pág.
2.1	Diagrama explicativo de la metodología.....	9
2.2	Clasificación de deslizamientos Varnes 1958.....	13
2.3	Red neuronal artificial de cuatro capas.....	20
2.4	Arquitectura original de una LSTM.	21
2.5	Arquitectura red GRU.....	21
2.6	Modelos de AUC para diferentes resultados	22
3.1	Ubicación geográfica del área de estudio, Cuenca del río La Paila.	26
3.2	Metodología	27
3.3	Distribución espacial de puntos MM (rojo) y no MM (amarillo).....	32
3.4	Ventanas de tiempo.....	38
3.5	Arquitectura desarrollada para la Clasificación.....	42
4.1	Cantidad de eventos registrados (base de datos) para MM y No MM de acuerdo con las unidades cronoestratigráficas (geología).....	47
4.2	Distribución espacial de las unidades cronoestratigráficas (geología) para los puntos de entrenamiento que representan los MM	48
4.3	Cantidad de movimientos en masa registrados de acuerdo a la geología (unidades cronoestratigráficas)	49
4.4	Distribución de la ocurrencia de movimientos en masa de acuerdo a la geología (unidades cronoestratigráficas) en porcentaje.....	50
4.5	Cantidad de eventos registrados (base de datos) para MM y No MM de acuerdo con los usos del suelo.....	51
4.6	Cantidad de movimientos en masa registrados de acuerdo con los usos del suelo	51

4.7	Distribución espacial de los usos del suelo para los puntos de entrenamiento que representan los MM	52
4.8	Distribución porcentual de los usos del suelo asociados con la ocurrencia de MM	53
4.9	Cantidad de eventos registrados (base de datos) para MM y No MM de acuerdo con la clasificación de la pendiente.....	54
4.10	Cantidad de movimientos en masa registrados de acuerdo con la clasificación de la pendiente.....	54
4.11	Distribución espacial de la pendiente para los puntos que representan los MM	55
4.12	Distribución porcentual de las pendientes asociadas a la ocurrencia de MM	56
4.13	Precipitación a tres meses de la ocurrencia de un MM (datos de entrenamiento)	57
4.14	Precipitación media a tres meses de la ocurrencia de un MM (datos de entrenamiento)	58
4.15	Distribución espacial de la precipitación acumulada 4 días previos a la ocurrencia del MM registrado el 7 de noviembre del 2017 en Corinto - Cauca.....	59
4.16	Curva de aprendizaje (a) y precisión del modelo (b).....	60
4.17	Área bajo la curva ROC (AUC)	61
4.18	Matriz de confusión (Validación del modelo)	62
4.19	Matrices de confusión en diferentes escenarios de etiquetado.....	63
4.20	Matriz de confusión de los parámetros elegidos (Escenario 1)	64
4.21	Matriz de confusión de los parámetros elegidos (Escenario 2)	64
4.22	Clasificación Enero / 20 / 2011, Cuenca del río La Paila	66
4.23	Clasificación Noviembre / 07 / 2017, Cuenca del río La Paila	67
4.24	Clasificación Julio / 19 / 2020, Cuenca del río La Paila	68
4.25	Zonificación propuesta para cuenca del río La Paila	69

ÍNDICE DE TABLAS

TABLA		Pag.
2.1	Resumen de Antecedentes.....	11
2.2	Clasificación de los movimientos en masa según Sharpe 1938	12
2.3	Clasificación de avenidas torrenciales.....	14
2.4	Terminología asociada con avenidas torrenciales.....	16
2.5	Rangos de pendiente.	17
2.6	Factores condicionantes y detonantes	24
3.1	Cuenca del Río La Paila	25
3.2	Pisos térmicos - Municipio de Corinto.....	27
3.3	Unidades geológicas de Corinto - Cauca	29
3.4	Rangos de pendientes Corinto - Cauca.	29
3.5	Principales usos del Suelo de Corinto - Cauca	30
3.6	Clasificación del Movimiento.....	30
3.7	Municipios homólogos con movimientos en masa.....	31
3.8	Matriz de datos estáticos.....	39
4.1	Reporte de clasificación	62

1. Introducción

Los fenómenos naturales son procesos que se dan a nivel mundial sin tener en cuenta la zona, región o país donde se presenten. Además, son eventos físicos que no necesariamente tienen una afectación directa sobre una comunidad o población. Cuando el factor antrópico se relaciona con un fenómeno natural, se habla del término riesgo de desastre, donde el desastre genera una interrupción grave al funcionamiento de una comunidad o sociedad a cualquier escala, debido a eventos peligrosos que interactúan con las condiciones de exposición, vulnerabilidad y capacidad, que conducen a pérdidas e impactos humanos, materiales, económicos y ambientales (UNISDR, 2009). Así, Wilches-Chaux (2011) indica que “Los fenómenos naturales no son castigo divino y los desastres no son naturales sino socialmente contruidos”.

A lo largo de la historia se han documentado y estudiado diversos eventos naturales entre los cuales, uno de los más recurrentes son los Movimientos en Masa (MM). Los MM están directamente relacionados con las variables condicionantes internas e inherentes al terreno, que favorecen su ocurrencia como lo son: pendientes, aspecto, rugosidad, curvatura, índice topográfico de humedad, índice de relieve relativo, inclinación de la ladera, índice de posición topográfica, distancia relativa al cauce y la distancia a fallas. Además, se encuentran las variables detonantes como los sismos y las precipitaciones (Aristizábal et al., 2019).

A raíz de las pérdidas tanto materiales como humanas causadas ante la ocurrencia de los MM, se genera la necesidad de la gestión del riesgo ya que es imperativo conocerlo, reducirlo y manejarlo (Ministerio de Ambiente y Desarrollo Sostenible, 2020) para construir un futuro más equitativo y sostenible (Clark, 2012). Esto conlleva a que gobiernos a nivel internacional desarrollen proyectos relacionados con la gestión del riesgo. Según Mizutori (2020) “Un dólar invertido en reducir el riesgo de desastres, ahorrará entre cuatro y quince dólares en la recuperación”. El impacto que dichos eventos han tenido sobre la economía de la población entre 1990 y 2017 expresado en términos del PIB (Producto Interno Bruto) mundial es de 0.5 % (Ritchie, 2019). Para entonces, el porcentaje total de muertes durante las últimas décadas representaba un promedio del 0.1 %, aunque se estimó que se podrían presentar variaciones debido a los eventos de alto impacto que para esos años oscilaron entre el 0.01 % y el 0.4 % del total de muertes (Ritchie, 2019).

En el contexto mencionado, se conoce sobre un proyecto implementado por el Banco Mundial en Wenchuan (China) para la recuperación de los daños provocados por la ocurrencia de un sismo de magnitud momento Mw: 8.01 en el año 2008, que afectó seis de sus provincias, situándolo entre uno de los más mortíferos de los últimos tiempos. Como consecuencia de este evento sísmico se registraron 69000 muertes, 374000 heridos y aproximadamente 18000 desaparecidos causando daños físicos importantes en toda el área afectada, siendo la infraestructura, la educación y la salud los sectores más afectados (Banco Mundial, 2018). Este es un gran antecedente que evidencia los retos a los que se enfrenta una población ante la ausencia de planes y programas en gestión del riesgo y los posteriores beneficios que tendría si se le da un manejo adecuado. La experiencia y conocimiento que aportó el Banco Mundial para abordar los desafíos posteriores al desastre fue de vital importancia ya que según Pan Li, vicealcalde del condado de Yanting (uno de los lugares más afectados) el medio ambiente de la ciudad mejoró notablemente, el espacio urbano se expandió y los negocios y el turismo se incrementaron. Como consecuencia, el PIB aumentó en un 127 % y el ingreso urbano per cápita aumentó en un 180 %, lo que significó importantes beneficios ecológicos, sociales y económicos (Banco Mundial, 2018).

Como contra ejemplo (inadecuada o inexistente gestión del riesgo), se presenta el terremoto ocurrido en Puerto Príncipe (Haití) con una magnitud momento Mw: 7.0 en el año 2010, el cual según De Gouvello (2010) ha sido declarado como el desastre más severo en los últimos 200 años debido a que ocurrió en uno de los países más pobres del mundo. Desde su independencia las condiciones del país marcaron los precedentes para su extrema pobreza, la cual es consecuencia de la deuda desmedida con Francia y los préstamos solicitados a diferentes bancos a nivel mundial. Sumado a esto, se encuentra su inestabilidad política, la cual lo convierte en un país vulnerable a que eventos de esta magnitud se transformen en desastres.

Lo anterior se confirma con la información de los daños causados y las áreas afectadas producto del terremoto del 2010. Se registraron 1.3 millones de personas desplazadas, 300000 heridos, 316000 muertes, 97294 casas destruidas y 188383 casas dañadas (USGS, 2010), teniendo en cuenta las condiciones preexistentes, la poca gestión del riesgo y que las donaciones realizadas por diferentes ONG no fueron invertidas para la recuperación del país frente a este suceso, para el año 2020 (10 años después), Haití seguía en ruinas según Robert Fatton (Como se cita en BBC Mundo, 2020).

En Colombia se han registrado diversos eventos que permiten analizar el panorama relacionado con la gestión del riesgo en el país. Tres situaciones que representan claros ejemplos son los desastres ocurridos en Mocoa (Putumayo) en el 2017 y Páez (Cauca) en los años 1994 y 2008. La situación ocurrida en Mocoa a raíz de la avenida torrencial producida por fuertes lluvias que aumentaron el caudal del río Sangoyaco y otras vertientes como consecuencia de los 129.3 mm de agua precipitada en 3 horas, permite inferir que los mecanismos de prevención implementados no fueron suficientes y la falta de control en el ordenamiento territorial ocasionaron un saldo de 210 fallecidos y decenas de familias damnificadas (Corpoamazonia, 2017). Los hechos registrados en Páez muestran cómo un municipio a raíz de la avenida torrencial ocurrida en 1994, producida por la ocurrencia de más de 3000 deslizamientos que dejó 1100 víctimas (Wilches-Chaux, 2005), marcó un precedente que permitió tener un punto de partida para el desarrollo de proyectos frente a la gestión del riesgo, tal como el evento ocurrido en el 2008 causado por la erupción del Nevado del Huila, que aunque presentó daños en infraestructura importante para la movilidad del municipio, no dejó víctimas mortales debido a la evacuación oportuna de la comunidad (UNGRD, 2018).

De lo anterior se determina la importancia de la gestión del riesgo en el desarrollo de un país. Esta debe incluir a todos los actores, como los ciudadanos, los líderes políticos, el sector privado, las organizaciones de la sociedad civil, las asociaciones profesionales, los organismos científicos, técnicos y las instituciones gubernamentales (Clark, 2012). Si bien existen varios actores que deben intervenir en la gestión del riesgo, el ente regulador nacional colombiano frente a la toma de decisiones es el Ministerio de Vivienda, Ciudad y Territorio, quien mediante el Decreto 1077 de 2015 (Decreto 1807 de 2014) exigen que se establezcan las condiciones y escalas de detalle para incorporar la gestión del riesgo de desastre en la formulación o revisión de mediano y largo plazo de un POT (Plan de ordenamiento Territorial). Cuyo objetivo es la delimitación y zonificación de áreas de amenaza asociadas a fenómenos de movimientos en masa, inundaciones y avenidas torrenciales que se generan a raíz de las características topográficas e hidrometeorológicas de una zona como variables condicionantes, sumadas a las variables detonantes. Adicionalmente, una parte importante de la gestión del riesgo es el monitoreo de los fenómenos naturales, siendo el ente encargado de su sistematización el Servicio Geológico Colombiano (SGC), el cual con sus aportes posibilita implementar estudios que permiten obtener información transversal para el desarrollo de proyectos que garanticen una correcta y efectiva gestión del riesgo.

Según la información que se tiene del municipio de Corinto, los desastres se han presentado debido a sus características sociales, culturales, económicas, ambientales y especialmente geográficas. Este municipio se localiza en el sur occidente colombiano, en el extremo norte del departamento del Cauca. Esta ubicación le da un escenario paisajístico variable debido a que su territorio es montañoso y su relieve está desarrollado sobre el flanco occidental de la Cordillera Central, con alturas que superan los 4000 m.s.n.m limitando por el Este con el departamento del Tolima. Sus principales corrientes son los ríos Guengüe, Huasanó, Jagual y La Paila, entre otras corrientes secundarias que comprenden la cuenca del río La Paila. Sus temperaturas están relacionadas con los pisos térmicos templado, frío y piso bioclimático páramo, con una temperatura media de 24°C (Escuela de Administración Pública et al., 2017)

El desarrollo económico de Corinto se basa en los beneficios monetarios obtenidos de la producción primaria de caña de azúcar. Sin embargo, dichos beneficios económicos no se reflejan en el desarrollo del municipio debido a que están dirigidos al sector agroindustrial. Esto ha llevado a que se establezcan vínculos comerciales con otras ciudades aledañas basados en el cultivo de café e ilícitos; siendo este último el generador de cambios en el comportamiento social a nivel regional relacionados con la migración de sus habitantes, consecuencia del narcotráfico, que a su vez genera el conflicto armado, desplazando familias y transformando socialmente la cultura corintea (Escuela de Administración Pública et al., 2017). Con base en lo anterior, el libre acceso a muchas zonas de Corinto se imposibilita como consecuencia de la situación de orden público, puesto que la existencia de cultivos ilícitos (UNODC, 2017) y el flagelo de la guerra entre grupos armados insurgentes con el Estado Colombiano (Castiblanco, 2020) hacen que Corinto no solo sea una zona propensa a la ocurrencia de MM, sino que también, limita el acceso de personal técnico y profesional para realizar los trabajos de campo requeridos en el desarrollo de los estudios necesarios para la implementación de un Plan de Ordenamiento Territorial que cumpla con el Decreto 1077 de 2015.

Teniendo en cuenta las condiciones topográficas y geográficas de Corinto, su cercanía a una fuente hídrica como el río La Paila, las formaciones montañosas que dan lugar al abanico aluvial (Camacho, G, 2003) dentro del cual se encuentra ubicado el casco urbano y las condiciones climáticas inherentes de una zona tropical, permiten inferir según los antecedentes de desastres a nivel nacional que Corinto cumple con ciertas condiciones que lo hacen propenso a ser afectado por la ocurrencia de MM, los cuales a su vez podrían

desencadenar entre otros, las avenidas torrenciales y sus consecuencias: centenares de familias damnificadas, pérdidas económicas en el sector agropecuario, vivienda e infraestructura; tal y como se evidenció con la avenida torrencial ocurrida el 7 de noviembre del 2017 causada por un periodo intenso de lluvias (Semana, 2017).

De acuerdo a este panorama y considerando que los MM ponen en riesgo la integridad de la comunidad, es importante desarrollar una metodología con porcentajes de confiabilidad aceptable, que contribuya a la identificación y zonificación de los movimientos en masa, permitiendo tener insumos para la toma de decisiones que permitan realizar una adecuada gestión del riesgo de desastre y así proveer de información fiable al municipio de Corinto, para que en concordancia con los requisitos del gobierno nacional 'permitan la formulación de medidas de ordenamiento del territorio que respondan al conocimiento y reducción de su riesgo presente, además de evitar la generación de riesgo futuro, con lo cual se oriente la toma de decisiones por parte de las entidades territoriales para garantizar el desarrollo seguro del territorio' (UNGRD, 2015).

Conociendo la relevancia que tiene la zonificación de MM para todos los actores beneficiados en el municipio de Corinto como los entes gubernamentales (La CAR del Cauca, el gobierno departamental del Cauca, la Alcaldía municipal Corinto, UNGRD, Fondo Nacional del Riesgo de Desastre, entre otros) y su población, es necesario establecer una estrategia de trabajo adecuada para obtener los insumos esenciales que apoyan el cumplimiento del decreto 1077 de 2015.

Teniendo en cuenta lo mencionado, las condiciones topográficas presentes en la cuenca del Río La Paila, las situaciones de orden público como consecuencia de la existencia de cultivos ilícitos y grupos al margen de la ley, que dificultan el ingreso de personal técnico para el acopio de información cualitativa y cuantitativa del territorio, es pertinente la implementación de métodos indirectos (como la percepción remota y algoritmos de inteligencia artificial), que faciliten la realización de estudios espacio-geométricos de un territorio con difícil acceso como Corinto. De esta manera se disminuye la vulnerabilidad del recurso humano y material, los costos y tiempos de operación, además de ser herramientas importantes para la realización de este tipo de estudios. Bajo el escenario planteado el presente trabajo formula la siguiente pregunta de investigación:

¿Cómo clasificar zonas propensas a movimientos en masa como fuente de material para

avenidas torrenciales en la cuenca del río La Paila usando inteligencia artificial?

1.1 Objetivos

1.1.1 Objetivo general

Clasificar zonas propensas a movimientos en masa como fuente de materiales para avenidas torrenciales en la cuenca del río La Paila usando un algoritmo de inteligencia artificial.

1.1.2 Objetivos específicos

- Definir los factores condicionantes y detonantes más relevantes asociados a movimientos en masa para este estudio.
- Desarrollar un conjunto de datos actualizados con antecedentes de diferentes movimientos en masa de zonas con características morfoclimáticas similares a las de la cuenca analizada.
- Construir un modelo para la detección de movimientos en masa que pueden ser fuente de materiales para avenidas torrenciales a través de métodos indirectos.

1.2 Organización del Documento

Este documento consta de 5 capítulos. En el Capítulo 1 encontrarán una introducción en la cual se contextualiza el problema de investigación planteado y se establecen los objetivos del proyecto. El Capítulo 2 está compuesto por los trabajos que fueron considerados más relevantes según la revisión del estado del arte y la definición de ciertos temas como parte del marco teórico. En el Capítulo 3 se desarrolla la metodología utilizada de manera detallada, incluyendo la descripción de las bases de datos empleadas para la clasificación de zonas propensas a MM. Luego se presentan los resultados con su respectivo análisis en el capítulo 4. Finalmente, en el Capítulo 5 se presentan las principales conclusiones alcanzadas con el desarrollo de este proyecto y recomendaciones para trabajos futuros.

2. Marco Teórico

Este capítulo presenta los principales antecedentes usados en el desarrollo de este trabajo. Estos permitirán conocer de manera general el estado del arte relacionado proyectos que presentan diferente metodologías para dar solución a problemáticas asociadas a los MM, enfatizando en arquitecturas de aprendizaje automático, variables condicionantes y detonantes, y precisión de cada modelo implementado con el fin dar solución al primer objetivo específico. Adicionalmente, en el marco conceptual se definirán los principales conceptos relacionados con los MM, percepción remota, machine learning y los respectivos criterios de evaluación que se usarán.

2.1 Antecedentes

Los movimientos en masa (MM) son agentes geomorfológicos comunes en la región Andina, su ocurrencia y propiedades están dadas por la interacción de varios factores como los geológicos, los topográficos, los climáticos y las actividades antropogénicas (Mergili et al., 2015). Estos fenómenos han sido ampliamente estudiados a lo largo de la historia por diferentes autores dado que constituyen uno de los fenómenos que más impacto causan (víctimas y pérdidas materiales) en el mundo actual según cifras suministradas por *Our World in Data* (Ritchie & Roser, 2014).

La clasificación de zonas propensas a MM se presenta como un insumo fundamental para la correcta elaboración de planes de gestión del riesgo, por lo cual, los estudios que se implementan tienden a realizar análisis mediante la automatización o semiautomatización de procesos. Por ejemplo Goyes & Hernandez (2021), plantearon como primer paso para construir un modelo de susceptibilidad por MM la determinación óptima de las variables relacionadas con la ocurrencia de MM para el cálculo de la susceptibilidad por medio del perceptron simple y multicapa; además de redes neuronales profundas. En este trabajo obtuvieron 0.898, 0.901 y 0.902 respectivamente en su validación. Otros autores que concuerdan con esa propuesta son Ospina & Aristizábal (2021) quienes aseguran que uno de los pasos más importantes para garantizar resultados óptimos al momento de evaluar la susceptibilidad por MM se basa en la revisión, depuración y selección de variables

condicionantes y detonantes. En ese estudio usaron 11 distintos modelos de *Machine Learning* (ML) que identificaron los factores con mayor influencia en la ocurrencia de los MM, siendo los modelos SGB (0.919), XGB (0.907), RF (0.910), SVM (0.896), KNN (0.890), ADB (0.833) y DT (0.831) los que presentan mejor desempeño según su AUC.

Muñoz et al. (2020), implementaron una red neuronal para estimar la susceptibilidad de deslizamientos, en donde identificaron los factores condicionantes a través de los valores obtenidos en los pesos sinápticos, obteniendo una precisión de 97.26 % en la red. Asimismo, Ramírez (2020) desarrolló su modelo teniendo como base fundamental variables condicionantes como la pendiente, la geología, la cobertura del suelo, el uso del suelo y las precipitaciones; con estas y el modelo de inteligencia artificial se clasificó la susceptibilidad por MM. Por su parte, Kumar & Anbalagan (2015) realizaron la clasificación de la susceptibilidad usando Redes de Enlace de Base Radial (RBFLN) que es una extensión de la Red de Función de Base Radial (RBFNN), obteniendo como resultado un AUC de 0.86 a través del uso de 116 zonas con presencia de deslizamientos y 116 con ausencia de estos. Dicho trabajo además de los factores condicionantes que consideraron pertinentes para su estudio, adicionó índices de vegetación como el NDVI (Índice de Vegetación de Diferencia Normalizada) que se extraen mediante técnicas de teledetección.

Estas técnicas de percepción remota también fueron usadas por Xu et al. (2019) mediante la interpretación de imágenes satelitales con 2m de resolución espacial para localizar 60 deslizamientos que tomaron como datos de entrada. En el análisis realizado mediante una regresión logística denominada Regresión Logística-Valor de la Información (IV-LR) el valor de la información representa la influencia de diferentes factores de control sobre la estabilidad de la amenaza de deslizamientos. En dicho estudio Usaron seis variables: (i) pendiente y elevación (consideradas como las causas más importantes de los deslizamientos); el (ii) NDVI y (iii) uso del suelo y cobertura del suelo (Land Use and Land Cover - LULC), las cuales podrían afectar los tipos de estructura de la superficie terrestre; (iv) el proceso de infiltración de precipitación; (v) las unidades geotécnicas; (vi) la proximidad a los ríos. Estas variables fueron convertidas a formato *raster* para posteriormente introducirlas como capas en el modelo, en este trabajo obtuvieron un AUC de 0.818.

El uso de métodos estadísticos es común para el análisis de la susceptibilidad por MM, por ejemplo, Aristizábal et al. (2019) utilizaron una metodología desarrollada mediante modelos

de base estadística bivariado (peso de la evidencia), el cual es uno de los recomendados por el Servicio Geológico Colombiano (SGC) y multivariado (regresión logística), siendo este último el de mayor uso a nivel mundial. Allí, los resultados de desempeño y predicción fueron satisfactorios permitiendo incorporar dicha información como insumo base para el ordenamiento del territorio. En otras palabras, los porcentajes de desempeño y predicción en resultados obtenidos mediante la implementación del modelo de regresión logística presentan un alto nivel de confiabilidad, dado que al compararlo con otros métodos como el Wi de Van Westen o el SINMAP (método de base física), presenta mejores resultados de correlación en la validación de los datos derivados (Alvarez et al., 2019). Por otra parte, Mendoza & Aristizábal (2017) desarrollaron una metodología en la que usaban distintos métodos (Figura 2.1) para el análisis de la susceptibilidad a movimientos en masa con énfasis directo en proyectos lineales (acueductos). Estos métodos son: método heurístico, método estadístico y método físico, los cuales tienen en cuenta diferentes factores condicionantes según su enfoque. En términos generales, las variables usadas fueron: geología, uso del suelo, precipitaciones, flujo acumulado, pendiente, aspecto y curvatura. Finalmente, al validar los resultados los autores obtuvieron un AUC de 0.73 para la zonificación de la susceptibilidad por MM y un AUC de 0.81 en la zonificación de la criticidad en el mapa final.

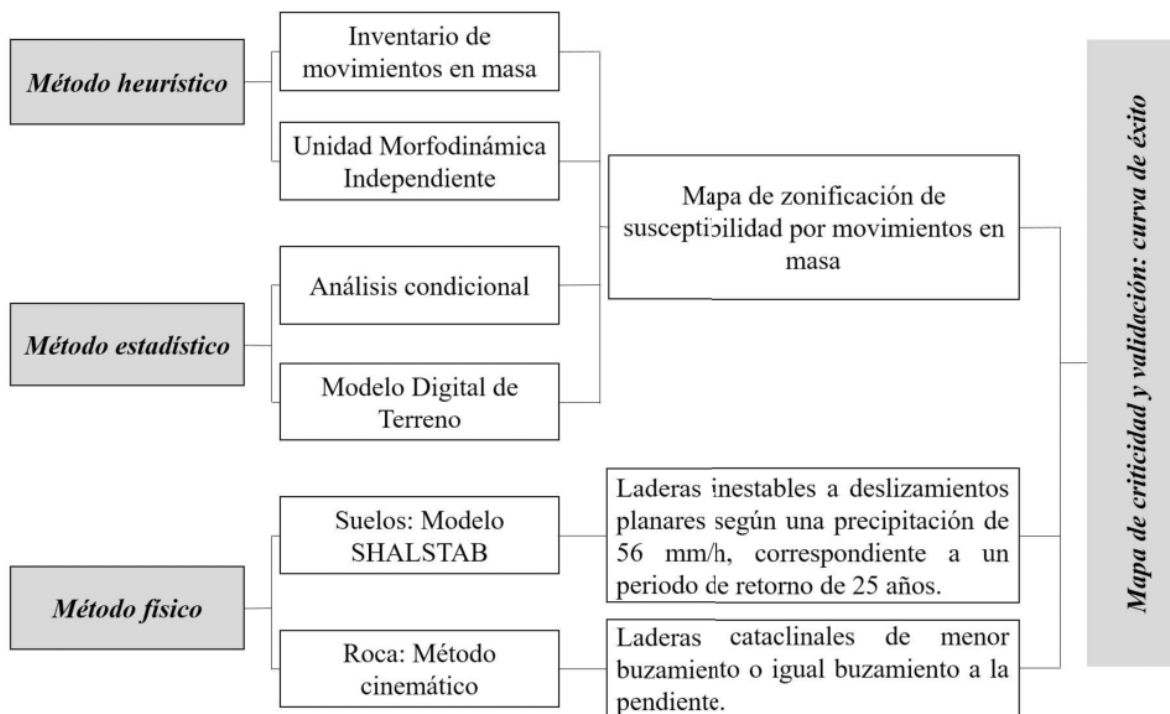


Figura 2.1: Diagrama explicativo de la metodología.

Fuente: Mendoza & Aristizábal (2017)

Finalmente, en cuanto al cálculo de la amenaza por MM, en el proyecto desarrollado por Acosta & Upegui (2020) tomaron como referencia la guía metodológica establecida por el SGC para la zonificación de amenaza por MM mencionado anteriormente. En ese estudio utilizaron los mapas obtenidos del IGAC de clase de suelo, tipo de relieve, fallas y geomorfología, vías y caminos, drenajes, cobertura de la tierra, precipitación, pendientes y el mapa de amenaza por MM. Todos estos datos los convirtieron a formato *raster* para trabajarlos en el mismo espacio. Posteriormente, realizaron un muestreo tomando puntos cada 100 metros sobre el área de estudio de cada variable para conformar el conjunto de datos para el procesamiento; luego realizaron la asignación de las 3 clases correspondientes a la capa de amenaza por MM (media, alta y muy alta) y seleccionaron el porcentaje de datos de muestreo (70 %) y validación (30 %) para los modelos *Feed Forward Network* y una red recurrente de *Elman*, dando como resultado que el modelo Elman que utiliza 2 capas ocultas es el que presenta mejor rendimiento con un error medio cuadrático (RMSE) entre 0.093 y 0.1.

La Tabla 2.1 resalta los elementos más importantes identificados en esta revisión.

2.2 Marco Conceptual

2.2.1 Movimiento en Masa - MM

Cruden & Varnes (1996) definen los MM como todos los movimientos ladera abajo de masas de roca, detritos o tierras por efecto de la gravedad. Estos movimientos se pueden clasificar de diferentes maneras teniendo cuenta el autor, escuela, país de origen o perfil profesional de quien analiza el fenómeno. Es posible diferenciarlos según el uso que se realice sobre las variables tales como la clasificación del mecanismo de falla del movimiento, el tipo de material involucrado, la actividad del movimiento y su velocidad. También se pueden clasificar de acuerdo a factores como la edad, las formas del relieve (pendiente), la ubicación geográfica, el clima y la afectación de laderas, causes o taludes (Montero, 2017).

Autor y Año	Variable Respuesta	Clasificador	Evaluación
Mendoza & Aristizábal (2017)	Zonificación de la susceptibilidad por MM (Baja, Moderada, Alta, Muy Alta)	Método estadístico, análisis condicional	AUC (0.81)
Acosta & Upegui (2020)	Clasificación de la amenaza por MM (Baja, Media, Alta)	Feed Forward Backpropagation y Elman Backpropagation	
Kumar & Anbalagan (2015)	Clasificación de la susceptibilidad por Deslizamientos (Muy Baja, Baja, Moderada, Alta, Muy Alta)	Redes de enlace de función de base radial (RBFLN)	AUC (0.86) en la predicción de la susceptibilidad
Xu, et., al (2019)	Clasificación de la susceptibilidad por Deslizamientos (Muy Baja, Baja, Media, Alta, Muy Alta)	Regresión logística	AUC (0.818)
Goyes & Hernandez (2021)	Clasificación de la susceptibilidad por MM (Muy Baja, Baja, Moderada, Alta, Muy Alta)	Perceptrón Simple, Perceptrón Multicapa y Redes Neuronales Profundas	0.898, 0.901 y 0.902 para RL, MLP y DNN
Ospina & Aristizábal, (2021)	Clasificación de la susceptibilidad por MM (Baja, Media, Alta)	PCA, RL, LDA, KNN, DT, SVM, ANN, RF, ADB, SGB, XGBoost	SGB (0.919), XGB (0.907), RF (0.910), SBM (0.896), KNN (0.890), ADB (0.833), DT (0.831).
Muñoz & Hernandez (2020)	Clasificación de la susceptibilidad por MM (Muy Baja, Baja, Moderada, Alta, Muy Alta)	Backpropagation, randomforest	
Ramírez Sánchez, (2020)	Clasificación de la susceptibilidad por MM (Baja, Moderada, alta, Muy Alta)	Inteligencia artificial no especificada	

Tabla 2.1: Resumen de Antecedentes

Montero (2017) resume las clasificaciones más conocidas en la literatura así:

■ Clasificación de Sharpe (1938)

Esta clasificación es considerada una de las más completas (Montero, 2017) dado el perfil profesional del autor. En la clasificación, el autor diferencia entre movimientos con una caída libre en sentido de inclinación de la ladera (flujos y deslizamientos de lentos a rápidos) y movimientos confinados sin cara libre en sentido descendente vertical (subsistencia) catalogados como grupos mayores; en estos identifica 16 tipos

de movimientos, además, identifica tres tipos diferentes de materiales (tierra, taluds y roca, o la mezcla de ellos) en tres ambientes geomorfológicos (glacial, fluvial, tropical o semejante) teniendo en cuenta que el desplazamiento se produce por fenómenos gravitacionales y en algunos casos interviene el hielo o el agua (Tabla 2.2).

	Movimiento		Gravedad + transporte glaciar	Gravedad	Gravedad + transporte fluvial
	Tipo	Tasa de velocidad	Tierra o roca + hielo	Tierra o roca + algo de agua hielo	Tierra o roca + agua
Con cara libre	Flujo	Imperceptible lento a rápido	1. Creep glaciar de roca	3. Creep de roca	
			2. Solifluji3n	4. Creep de talus	
			7. Avalancha de detritos	5. Creep de suelos	6. Solifluji3n
					8. Flujo de tierra
	Deslizamiento	Lento a muy rápido		11. Hundimiento	9. Flujo de lodo (Ambiente semiárido alpino volcánico)
				12. Deslizamiento de detritos	10. Avalancha de detrito
				13. Caída de detritos	
				14. Deslizamiento de roca	
				15. Caída de roca	
	Sin cara libre	Subsidencia		16. Subsistencia	

Tabla 2.2: Clasificación de los movimientos en masa según Sharpe 1938

Nota 1. En esta clasificación los movimientos con cara libre y debido a la gravedad se clasifican como deslizamientos, incluidas las caídas.

Fuente: Montero (2017)

■ Clasificación de Varnes (1958)

Esta clasificación relaciona cinco tipos de movimientos (caídas, deslizamientos, propagación lateral, flujo y avalancha) que se desarrollan en rocas, regolitos o suelos, donde se destacan características de los movimientos como su velocidad y humedad asociadas a los efectos y posibles daños, considerando adicionalmente las características cinemáticas de estos procesos. Desde otra perspectiva, hace una clasificación de los materiales desplazados como rocas o suelos, teniendo en cuenta su tamaño promedio antes de su movilización (masas más firmes o de fragmentos de diferentes tamaños), y a su vez clasifica los suelos como tierras o detritos (Figura 2.2). Finalmente, este autor es el primero que considera los movimientos complejos teniendo en cuenta la combinación de los movimientos básicos (Montero, 2017).

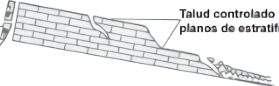
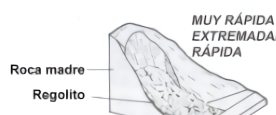
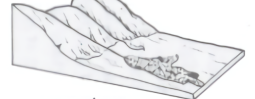
		Roca madre		Regolito (material suelto o relajado de cualquier origen que cubre la roca firme)																												
CAÍDAS		a CAÍDA DE ROCAS Diaclasas abiertas, ej.: presión hidrostática o cuñas de gelifracción  Eliminación de soporte, ej.: erosión o actividades humanas EXTREMADAMENTE RÁPIDA		e CAÍDA DE SUELO Mezcla de sedimentos  Socavación por corrientes MUY RÁPIDA																												
	DESPLAZAMIENTOS	Pequeñas deformaciones	b HUNDIMIENTO ROTACIONAL  EXTREMADAMENTE LENTA A RÁPIDA Superficie de ruptura		c DESLIZAMIENTO PLANAR DE BLOQUES  Falta a lo largo de discontinuidades MODERADA																											
Grandes deformaciones		d DESLIZAMIENTO DE ROCA Escarpe controlado por discontinuidades MUY LENTA A EXTREMADAMENTE RÁPIDA  Talud controlado por planos de estratificación		f DESLIZAMIENTOS Hundimiento Flujo de tierras Arcilla firme Arcilla blanda con material más grueso húmedo Grava arcillosa firme (i) Rotacional (ii) Planar Loess Arcilla glacial (iii) Propagación lateral Movimiento lateral general de Arcillas blandas (iv) Deslizamiento de detritos MUY RÁPIDA MUY LENTA A RÁPIDA																												
Materiales No Consolidados																																
		Principalmente fragmentos grandes de roca	Arena o limo sorteados no plásticos	Mezclas de roca y suelo	Predominantemente plástico																											
INCREMENTO EN EL CONTENIDO DE AGUA	Seco	g FLUJO DE FRAGMENTOS DE ROCA  EXTREMADAMENTE RÁPIDA	h CORRIMIENTO DE ARENA  Arena Limo firme Arena seca RÁPIDA A MUY RÁPIDA	k FLUJO LENTO DE TIERRAS  Roca lodosa meteorizada Roca lodosa	j AVALANCHA DE DETRITOS  Roca madre Regolito MUY RÁPIDA A EXTREMADAMENTE RÁPIDA																											
	Húmedo	m FLUJO DE ARENA O LIMO  Arena RÁPIDA A MUY RÁPIDA	l FLUJO RÁPIDO DE TIERRAS  MUY RÁPIDA	n FLUJO DE DETRITOS  MUY RÁPIDA	TASA APROXIMADA DE MOVIMIENTO <table><tr><td>Pie/seg</td><td></td><td></td></tr><tr><td>10³</td><td>3 m/seg</td><td>Extremadamente rápida</td></tr><tr><td>10¹</td><td></td><td></td></tr><tr><td>10⁻¹</td><td>0.3 m/min</td><td>Muy rápida</td></tr><tr><td>10⁻³</td><td></td><td>Rápida</td></tr><tr><td>10⁻⁵</td><td>1.5 m/día</td><td>Moderada rápida</td></tr><tr><td>10⁻⁷</td><td>1.5 m/mes</td><td>Muy lenta</td></tr><tr><td>10⁻⁹</td><td>1.5 m/año</td><td>Extremadamente lenta</td></tr><tr><td></td><td>0.3 m/5 años</td><td></td></tr></table>		Pie/seg			10 ³	3 m/seg	Extremadamente rápida	10 ¹			10 ⁻¹	0.3 m/min	Muy rápida	10 ⁻³		Rápida	10 ⁻⁵	1.5 m/día	Moderada rápida	10 ⁻⁷	1.5 m/mes	Muy lenta	10 ⁻⁹	1.5 m/año	Extremadamente lenta		0.3 m/5 años
Pie/seg																																
10 ³	3 m/seg	Extremadamente rápida																														
10 ¹																																
10 ⁻¹	0.3 m/min	Muy rápida																														
10 ⁻³		Rápida																														
10 ⁻⁵	1.5 m/día	Moderada rápida																														
10 ⁻⁷	1.5 m/mes	Muy lenta																														
10 ⁻⁹	1.5 m/año	Extremadamente lenta																														
	0.3 m/5 años																															

Figura 2.2: Clasificación de deslizamientos Varnes 1958.

Fuente: Montero (2017)

2.2.2 Avenidas torrenciales

Aristizábal et al. (2020) en su artículo “Definición y clasificación de las avenidas torrenciales y su impacto en los Andes colombianos” proponen el uso del término avenida torrencial (AT) para los flujos que se forman por sedimentos y agua en diferentes cantidades, los cuales se transportan a grandes velocidades a través de los cauces de la montaña que son generados por los factores detonantes como lluvias fuertes, sismos, rotura de presas o por el aporte de grandes cantidades de agua debido al deshielo. Estos autores analizan tres grandes definiciones para el término avenida torrencial, iniciando la revisión bibliográfica con la clasificación para procesos de transporte de sedimentos citando a O’Brien & Julien (1985) quienes mencionan que dicha clasificación se puede definir de acuerdo a las propiedades controladas por la concentración de sedimentos, partiendo desde inundaciones hasta movimientos en masa mediante 5 clases: crecientes, inundaciones de lodo, flujos de lodo, flujos de escombros y deslizamientos (Tabla 2.3).

Clasificación de AT	Descripción
Crecientes	Inundaciones de agua por descargas excesivas donde los sedimentos son transportados a través de suspensión y saltación por el lecho
Inundaciones de lodo	Flujos donde la concentración de sedimentos finos varía entre el 20 % y el 45 % por volumen
Flujos de lodo	Flujos donde la concentración de sedimentos finos varía del 45 % al 50 % por volumen
Flujos de escombros	Flujos de lodo con más del 50 % de sedimentos más gruesos que arena
Deslizamientos	Movimientos en masa donde la concentración por volumen de sedimentos es mayor al 50 %.

Tabla 2.3: Clasificación de avenidas torrenciales.

Fuente: O’Brien & Julien (1985)

Adicionalmente, Aristizábal et al. (2020) revisaron la clasificación de Cruden & Varnes (1996), quienes clasifican los procesos gravitacionales de vertiente en dos características: una es el tipo de movimiento continuo espacialmente (donde la superficie de cizallamiento dura poco tiempo), además el intervalo entre ellos es pequeño y por lo general no se conservan, y la otra es el tipo de material que dependiendo del tamaño de sus partículas se clasifican en flujos de tierra (si son menores a 2mm) y flujos de escombros (si son gruesas). Estos autores restringen el uso del término flujo únicamente a la fase del MM donde el material es formado por un esfuerzo continuo.

Finalmente, Aristizábal et al. (2020) citan a Borga et al. (2014), quien define las inundaciones súbitas como la consecuencia de lluvias breves de una alta intensidad que presentan un origen convectivo; estas se presentan en cuencas relativamente pequeñas por lo que se considera que son espacialmente confinadas y la transferencia de escorrentía es superficial, presentando como característica específica la morfología dado que el proceso se da en altas pendientes. Los mismos autores separan las inundaciones de los flujos de escombros y concluyen que son dos fenómenos diferentes pero que están relacionados porque son detonados por lluvias.

En definitiva, el término avenida torrencial y flujo de escombros representa un amplio abanico de procesos, desde la ocurrencia de MM en la ladera, hasta los diferentes flujos que ocurren a lo largo del sendero de montaña dependiendo de las pendientes, la concentración de agua y sedimentos. Partiendo de la revisión bibliográfica realizada por Aristizábal et al. (2020), se afirma que existen diferentes términos y condiciones relacionados con las avenidas torrenciales como se presenta en la Tabla 2.4.

2.2.3 Pendientes

La pendiente es un parámetro del terreno que mide la tasa de cambio de la elevación en la dirección descendente más empinada y describe la forma de la superficie (Wilson & Gallant, 2000). De acuerdo con el SGC (2017), el tipo de terreno puede clasificarse según el porcentaje de pendiente que se tiene o ángulo de inclinación como se observa en la Tabla 2.5.

2.2.4 Geología

La geología se define como la ciencia que estudia la Tierra y sus materiales teniendo en cuenta los fenómenos que causan su evolución. En otras palabras, la geología estudia los minerales, las rocas, la atmósfera, los sismos, la erosión y los fósiles (Oviedo, 2019). En el Mapa Geológico de Colombia (Gómez & Montes, 2020) se presentan las unidades cronoestratigráficas, estas son una subdivisión de las rocas y se definen como un conjunto de estratos caracterizados por formarse en un tiempo geológico específico (García et al., 2002).

AVENIDA TORRENCIAL	
AUTOR	DESCRIPCIÓN
Pierson (2005)	“La carga de fondo puede contener material hasta tamaño grava”
Jakob y Hungr (2005)	“Inundación de escombros se forma a través de inundaciones súbitas y el desplazamiento de grandes volúmenes de agua a lo largo del cauce, que aumenta su poder de erosión y su capacidad de transporte y remoción de los sedimentos del lecho, con la posible conjunción de deslizamientos o colapsos de la banca a pequeña escala, que al unirse al flujo aumentan volumen y concentración de sedimentos”
Hungr et ál. (2001)	“Las inundaciones de escombros son eventos torrenciales con velocidades y descargas pico similares a aquellas de las inundaciones súbitas. Debido a su alta velocidad, toman la facultad de transportar una cantidad de sedimento similar a los eventos tipo flujo de escombros canalizados, pero con bloques de tamaños menores”
Crozier (2005)	“Los movimientos en masa sobre la ladera son eventos masivos de movimientos de masa de ocurrencia regional (MORLE), los cuales corresponden a movimientos de tipo flujo que son avalanchas de escombros o flujo de escombros en ladera”
Aristizabal et ál (2020)	“Las avenidas torrenciales son un tipo flujo de escombros canalizados que se desencadenan como movimientos en masa profundos y grandes volúmenes en condiciones saturadas, lo cual les permite alcanzar el cauce y aumentar el porcentaje en volumen de agua para desplazarse a grandes velocidades y largos trayectos a lo largo del cauce.”
Ingeominas (2001)	“Fenómenos de remoción en masa donde el agua de una corriente aumenta considerablemente su volumen por el transporte de material sólido que ha caído a su cauce desde las laderas adyacentes.”
Planes de ordenación y manejo de cuencas (POMCAS)	“Inundaciones de tipo fluvial rápidas o torrenciales.”
MINAMBIENTE (2014)	“Crecientes súbitas que por las condiciones geomorfológicas de la cuenca están compuestas por un flujo de agua con alto contenido de materiales de arrastre, con un gran potencial destructivo debido a su alta velocidad.”
Fondo de Prevención y Atención de Emergencias de la ciudad de Bogotá (FOPAE, 2011)	“Sinónimo de inundaciones rápidas generadas por crecientes que ocurren de manera repentina debido a la alta pendiente del río o la quebrada y su cuenca; donde en ocasiones se produce el arrastre de una gran cantidad de material como detritos”.

Tabla 2.4: Terminología asociada con avenidas torrenciales.
Fuente: Aristizábal et al. (2020)

Inclinación (°)	Descripción
0-2	Plano a casi plano. No hay denudación apreciable.
2-4	Suavemente inclinado. Movimientos en masa de baja velocidad y procesos erosivos de diferentes tipos, especialmente bajo condiciones periglaciares (solifluxión) y fluviales (erosión laminar y surcos). Susceptible a desarrollar procesos erosivos
4-8	Inclinado. Condiciones similares a las anteriores. Alta susceptibilidad a desarrollar procesos erosivos.
8-16	Moderadamente abrupto. Movimientos en masa de todos los tipos, especialmente solifluxión periglaciara, reptación y ocasionalmente deslizamientos, también erosión de tipo laminar y en surcos. Susceptible a erosión y deslizamientos.
16-35	Abrupto. Procesos denudacionales intensos de diferentes tipos (erosión bajo cubierta forestal, reptación, deslizamientos). Alta propensión al desarrollo de procesos erosivos.
35-55	Muy abrupto. Afloramientos rocosos, procesos denudacionales intensos, depósitos granulares caóticos de poco espesor.
>55	Extremadamente abrupto. Afloramientos rocosos. Procesos denudacionales muy fuertes, especialmente “denudación de escarpe”; susceptible a rodamiento de rocas.

Tabla 2.5: Rangos de pendiente.

Fuente: SGC (2017)

2.2.5 Uso del suelo

Este término se define como el uso que se le asigna a la cobertura de manera detallada. Esta se divide según su producción y su contexto socioeconómico, además por sus características físicas (Rodríguez et al., 2018). El uso del suelo hace referencia a la cobertura y su utilidad, relacionándose con los beneficios que le puede dar a la humanidad. Como se cita en IGAC (2019), el uso está relacionado con las actividades que desarrolla el ser humano en una porción de tierra, estos usos pueden ser: reserva natural, agropecuario, industrial, entre otros.

2.2.6 Precipitación

La precipitación se define como el fenómeno meteorológico debido al cual el vapor de agua se condensa y llega al suelo en forma de lluvia, rocío, granizo o nieve, contabilizándose durante determinado periodo, en diferentes estaciones durante un tiempo específico. Este fenómeno ocurre cuando el vapor de agua condensado en el aire cae en forma de líquido o sólido sobre la superficie del suelo (IDEAM, 2013).

2.2.7 Métodos indirectos

Estos métodos se basan en observaciones indirectas, es decir que los datos provienen de fuentes secundarias como fotos (imágenes), encuestas, bases de datos, entre otras; por lo tanto, el investigador no estudia el fenómeno de forma directa sino que basa sus resultados en el análisis de las fuentes consultadas (Cajal, 2020).

2.2.7.1 *Percepción remota.*

Las técnicas de la percepción remota son usadas ampliamente para recolectar información sobre facciones de la superficie terrestre. Es la técnica o la ciencia a partir de la cual se obtiene información espacial, espectral o temporal. Generalmente se entiende como la interpretación de datos de imágenes de la superficie de la Tierra, adquiridas desde un satélite o un equipo que no está en contacto físico con ella. Los datos percibidos remotamente tienen muchos usos, desde la planificación del uso de la tierra urbana e industrial, hasta el estudio del suelo agrícola y los cultivos (Bonett, 2013).

2.2.7.2 *Machine Learning (ML).*

El aprendizaje automático (ML) es un subconjunto de la inteligencia artificial, que construye modelos matemáticos basados en datos de muestra (llamados “datos de entrenamiento”) para hacer predicciones. En el aprendizaje automático, las redes neuronales, las máquinas vectoriales de soporte y la informática, generalmente proporcionan conjuntos de entrenamiento y conjuntos de prueba. El denominado *training data*, se refiere a la combinación del conjunto etiquetado y el conjunto de ejemplo sin etiquetar que pueden utilizar los algoritmos. Por otro lado, el *test data* es un conjunto de datos de prueba que consta de ejemplos que no han sido procesados por la red. El objetivo del aprendizaje automático es construir un clasificador a través del aprendizaje de un conjunto de entrenamiento y luego evaluar el rendimiento de la regresión o clasificador a través de datos de prueba (Zhang, 2020), como afirman Mitchell et al. (1997) “un programa de computadora aprende de la experiencia E con respecto a alguna clase de tareas T y la medida de desempeño P, si su desempeño en las tareas en T, es medido por P, mejora con experiencia E”.

- **Redes Neuronales Artificiales (RNA).** Según Caicedo & López (2009) una red neuronal artificial es un modelo de cómputo auto adaptativo y bioinspirado. En ella, se trata de imitar la gran cantidad de interconexiones que presentan las células nerviosas, de manera que el arreglo global de células individuales simples alcanza niveles de procesamiento muy elevados, brindándole al organismo la capacidad de enfrentarse a problemas complejos con gran eficacia y precisión. Las neuronas artificiales son las unidades básicas de procesamiento en las redes neuronales artificiales. Cada neurona artificial efectúa una evaluación de los valores que recibe como entrada a través de una función de activación, suele elegirse una función que sea de tipo no lineal y continua como la sigmoideal (ecuación 2.1) y la ReLU (ecuación 2.2).

$$\phi(x) = \frac{1}{1 + e^{-x}} \quad (2.1)$$

$$R(x) = \max(0, x) \quad (2.2)$$

Las neuronas artificiales se unen entre sí por medio de enlaces sinápticos ponderados por un conjunto de pesos. Se llama capa, a la agrupación de neuronas que se encuentran en el mismo nivel de profundidad en la red (Figura 2.3). Las señales de salida de una capa alimentan a todas las neuronas de la capa siguiente. La función de salida de la red neuronal es una composición de las funciones de activación de todas las neuronas en la red (Meneses & Alvarado, 2017). En ese sentido, la distribución de las neuronas dentro de una red neuronal artificial se realiza formando niveles de un número de neuronas determinado. En otras palabras, si un conjunto de neuronas artificiales recibe simultáneamente el mismo tipo de información, es denominada capa. En una red se pueden diferenciar tres tipos de niveles:

- Entrada: Es el conjunto de neuronas que recibe directamente la información proveniente de las fuentes externas de la red.
- Oculto: Corresponde a un conjunto de neuronas internas de la red y no tiene contacto directo con el exterior. El número de niveles ocultos puede estar entre cero y un número elevado. En general, las neuronas de cada nivel oculto comparten el mismo tipo de información, por lo que formalmente se denominan Capas Ocultas. Las neuronas de las capas ocultas pueden estar interconectadas de diferentes maneras, lo que determina junto con su número, las distintas arquitecturas de redes neuronales.

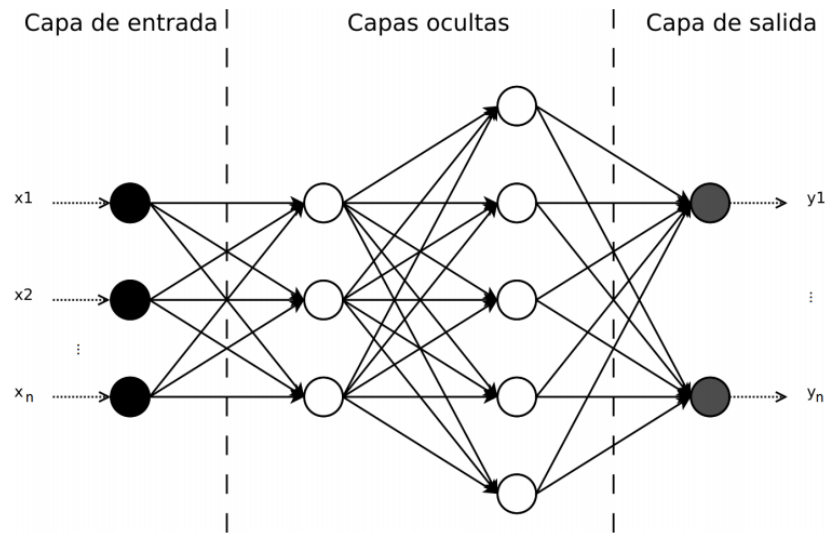


Figura 2.3: Red neuronal artificial de cuatro capas.
Fuente: Meneses & Alvarado (2017)

- Salida: Es el conjunto de neuronas que transfieren la información que la red ha procesado hacia el exterior.

- **Long-Short Term Model (LSTM).** Para el problema de la “long-term dependencies” que son un tipo especial de redes neuronales recurrentes (RNN), Hochreiter y Schmidhuber propusieron celdas LSTM como se cita en Yu et al. (2019). Estas mejoran la capacidad de recuperación de las unidades recursivas estándar mediante la introducción de “gates” o puertas en la unidad. Desde esta primera aproximación, muchos investigadores han modificado y generalizado los LSTM. Las variaciones incluyen LSTM sin puerta olvidada, LSTM con puerta olvidada y LSTM con conexiones de mirilla. Por lo general, el término celda LSTM significa LSTM con puerta olvidada. En la Figura 2.4 se puede observar su arquitectura.
- **Unidad Recurrente Cerrada (GRU).** Las unidades LSTM aprenden mejor que las RNN estándar. Sin embargo, los parámetros adicionales aumentan la carga computacional, lo cual generó la necesidad de mejorar este tipo de arquitectura con una nueva propuesta como las GRU (Yu et al., 2019), compuesta por las conexiones que se muestran en la Figura 2.5.

Para reducir el número de parámetros, la unidad GRU tiene solo 2 puertas integradas: la *forget gate* y el *input gate* del LSTM como puertas de actualización, por lo tanto

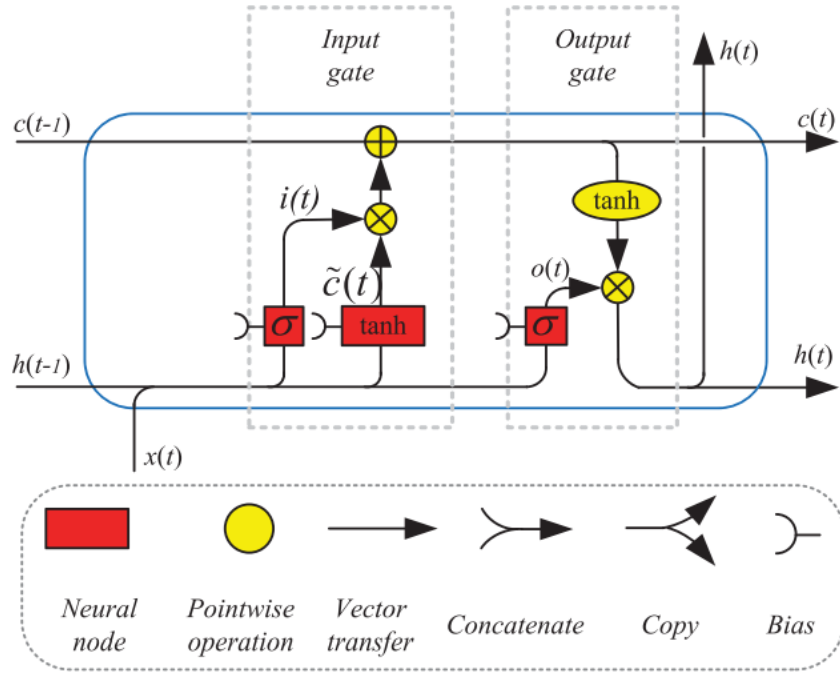


Figura 2.4: Arquitectura original de una LSTM.
Fuente: Yu et al. (2019)

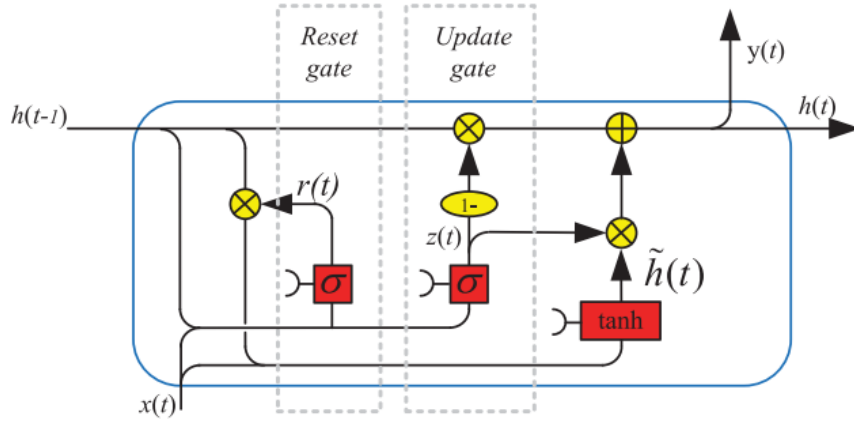


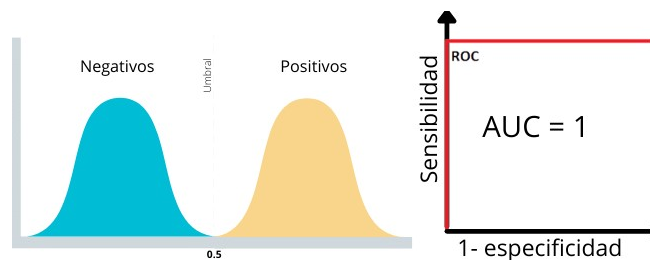
Figura 2.5: Arquitectura red GRU.
Fuente: Yu et al. (2019)

puede guardar señales de disparo y parámetros relacionados, como sonidos específicos o palabras predeterminadas a partir del análisis del espectrograma generado por la voz. Una GRU es esencialmente una variante de la LSTM normal con puertas de olvido. No obstante, debido a la falta de una puerta, la GRU de celda única no es tan poderosa

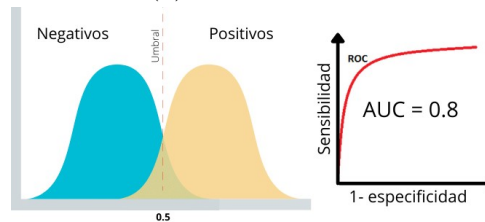
como la LSTM original, puesto que no se puede enseñar a una GRU sobre calcular y resolver lenguajes libres de contexto o problemas de traducción. Cabe aclarar que tanto la GRU como la LSTM superan a las unidades tanh-RNN tradicionales, con aproximadamente la misma cantidad de parámetros por red según lo menciona Yu et al. (2019) en su análisis bibliográfico del rendimiento de las redes LSTM, las redes GRU y la red tradicional tanh-RNN.

■ Criterios de evaluación del desempeño.

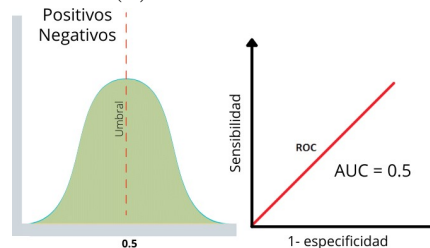
Para evaluar la capacidad de un sistema predictivo existen diferentes métricas, por ejemplo, el área bajo la curva ROC (AUC) se puede usar como métrica de evaluación puesto que cuanto mayor sea el área, mejor será la clasificación. El índice AUC varía entre 0.5 para los modelos sin capacidad predictiva y 1 para modelos con capacidad perfecta de predicción (Figura 2.6).



(a) Modelo ideal



(b) Modelo bueno



(c) Modelo aleatorio (sin capacidad de predicción)

Figura 2.6: Modelos de AUC para diferentes resultados
Fuente: Cañadas (2021)

En particular, un AUC de 0.5 o sin capacidad predictiva está definido por la diagonal que describe la ecuación 2.3 (Carrara et al., 2008).

$$y = x \quad (2.3)$$

Otro método importante para evaluar el desempeño de una red, que complementa el resultado que se obtiene del AUC, son las matrices de confusión. Esta herramienta permite visualizar mediante las columnas las predicciones de cada clase, mientras que las filas representan la clase real, es decir que este método permite conocer la cantidad y el tipo de aciertos o errores que tiene el modelo evaluado (Arce, 2019)

2.3 Principales Hallazgos

Los factores condicionantes y detonantes juegan un papel muy importante al momento de estudiar los MM. Es necesario enfatizar que las precipitaciones son el único factor detonante para este estudio que presenta un comportamiento recurrente a través del tiempo, lo cual, facilita su modelado además de ser una variable ampliamente analizada y documentada. Por otra parte, los factores condicionantes más usados según la literatura consultada son: La geología, el uso del suelo y las pendientes, como se puede observar en la Tabla 2.6. Los antecedentes consultados permitieron definir estos factores según su recurrente uso por diferentes autores en los últimos seis años, permitiendo así seleccionar para efectos de este proyecto los más influyentes en estudios relacionados con movimientos en masa teniendo en cuenta los respectivos clasificadores implementados.

Fuente	Factores	Mendoza & Aristizábal (2017)	Acosta & Upegui (2020)	Kumar & Anbalagan (2015)	Xu et al. (2019)	Goyes & Hernandez (2021)	Ospina & Aristizábal (2021)	Muñoz et al. (2020)	Ramírez (2020)
Cartografía existente	Geología	x		x			x	x	x
	Geomorfología		x		x	x			
	Relieve		x	x		x			x
	Vías y Caminos		x						
	Drenaje		x	x	x				
	Fallas		x						
	Uso de suelo	x	x			x	x	x	x
	Precipitación	x	x						x
DEM	Flujo acumulado	x				x	x	x	
	Pendientes	x	x	x	x	x	x	x	x
	Aspecto	x		x					
	Concavidad						x		
	Convexidad								
	Curvatura	x		x		x		x	
	Rugosidad					x			
	Dimension fractal local						x		
Clasificadores	Índice topográfico					x			
	Longitud del flujo					x		x	
		Método estadístico, análisis condicional	Feed Forward Back-propagation y Elman Back-propagation	Redes de enlace de función de base radial (RBFLN)	Regresión logística	Perceptrón simple, Perceptrón Multicapa y Redes Neuronales Profundas	PCA, RL, LDA, KNN, DT, SVM, ANN, RF, ADB, SGB, XGBoost	Back-propagation, randomforest.	Inteligencia artificial no especificada

Tabla 2.6: Factores condicionantes y detonantes

3. Metodología

En este capítulo se desarrolla la metodología implementada para cumplir con los objetivos planteados en este proyecto, iniciando con una breve descripción de la zona de estudio y finalizando con las fases que componen el esquema metodológico presentado en el Capítulo 2.

3.1 Localización de la zona de estudio

3.1.1 Cuenca del Río La Paila

La Cordillera Central y las estribaciones del Parque Natural Nacional Nevado del Huila permiten que se origine un sistema hidrográfico tributario a la Gran Cuenca del Río Cauca en la parte alta del territorio. La cuenca del Río La Paila es tributaria de la sub-cuenca del Río El Palo y está conformada por las cuencas Guengüe, Jagual, Quebrada seca, Quebrada La Cristalina, Zanjón Huasanó y cuatro lagunas pequeñas (Tabla 3.1). Esta cuenca comprende un área de 30160.4 ha y constituye una red abastecedora de acueductos.

MICROCUEENCA	HECTÁREAS
Guengüe	10400
Río Negro	4503
Jagual	5213
Huasanó	843.7
Cristalina	4058.2
Quebrada Seca	1881
Lagunas	3261.5
TOTAL	30160.4

Tabla 3.1: Cuenca del Río La Paila

Fuente: Adaptado de Concejo Municipal de Corinto (2016)

Aunque la cuenca comprende varios municipios como Miranda, Padilla, Puerto Tejada, Guachené, Caloto y Corinto, este último es el que mayor área abarca de la cuenca. Corinto hace parte del departamento del Cauca ubicado al suroeste de Colombia. El municipio se encuentra sobre la Cordillera Central, perteneciendo al corredor ecológico del Parque Nacional del Nevado del Huila y de la zona de influencia del Macizo Colombiano (Figura

3.1). Los rangos altitudinales varían de 900 a 4100 metros sobre el nivel del mar. El municipio limita al norte con el municipio de Miranda, al sur con Toribío y Caloto, hacia el este con del departamento del Tolima y al oeste con el municipio de Padilla. Está compuesto por los corregimientos de El Barranco, El Jagual, Media Naranja, Quebraditas y Rionegro, además de un caserío y 3 centros poblados (Concejo Municipal de Corinto, 2016).

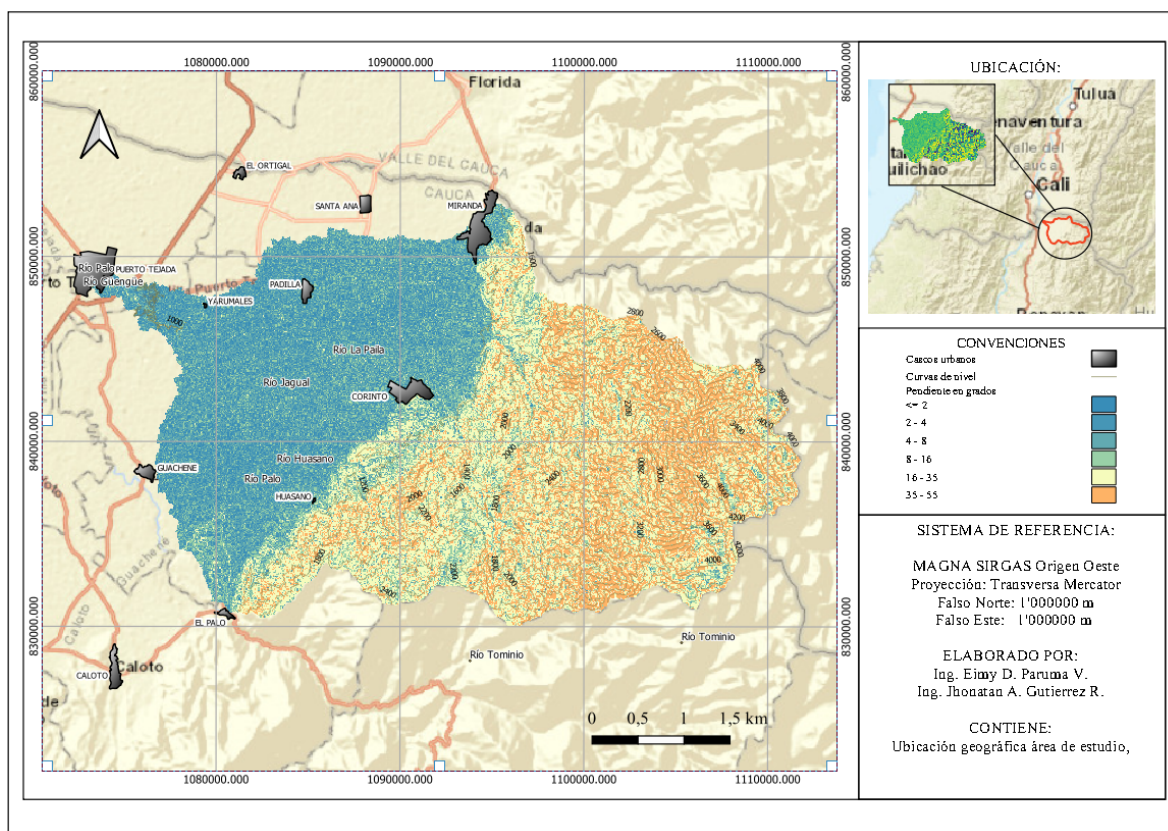


Figura 3.1: Ubicación geográfica del área de estudio, Cuenca del río La Paila.

3.1.2 Clima

El municipio de Corinto cuenta con diversas propiedades climatológicas por efectos orográficos, la circulación de vientos locales suaves y la variación de pisos térmicos que van desde templado hasta páramo bajo (Tabla 3.2), con una temperatura promedio anual de 23.6 °C (Concejo Municipal de Corinto, 2016). Debido a la ubicación geográfica de Corinto y la influencia de la Cordillera Central que determina un régimen húmedo que aumenta por el flanco occidental, se presentan dos temporadas de lluvias intensas: enero-abril y octubre-

diciembre. El valor promedio de la temperatura más alta oscila alrededor de los 35°C y la temperatura mínima alrededor de 17°C (Concejo Municipal de Corinto, 2016), con una precipitación media anual de 1.746 mm asociada a la Zona de Convergencia Intertropical (IGAC, 2016).

PISO TÉRMICO	SÍMBOLO	ALTITUD (m.s.n.m)	TEMPERATURA (°C)	ÁREA (%)
Templado	T	1000 - 2000	17 - 24	60.06
Frío	F	2000 - 3000	12 - 17	26.94
Muy frío	MF	3000 - 3600	7 - 12	6.43
Paramuno bajo	PM	3600 - 4100	4 - 7	6.57
TOTAL				100

Tabla 3.2: Pisos térmicos - Municipio de Corinto
Fuente: Adaptado de Concejo Municipal de Corinto (2016)

3.2 Esquema metodológico

La Figura (3.2) ilustra las diferentes etapas ejecutadas para la realización del presente estudio.

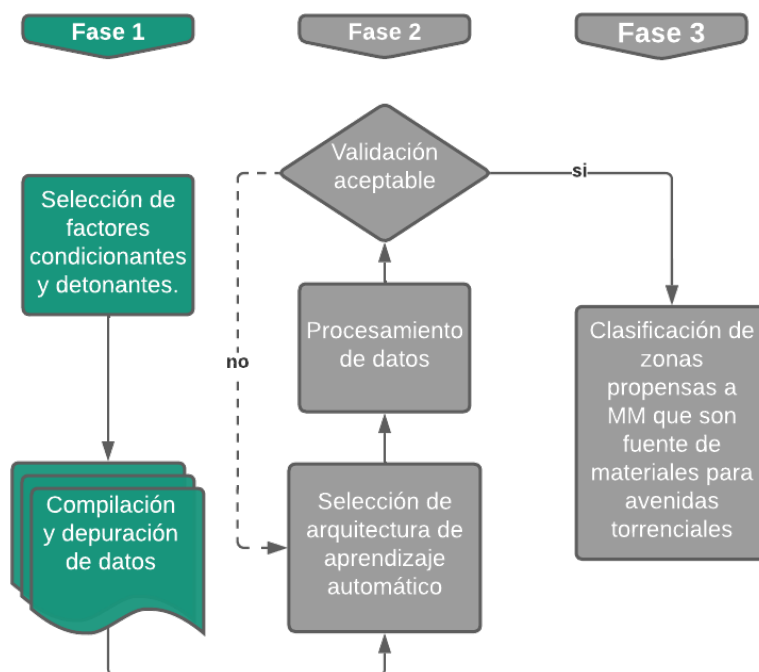


Figura 3.2: Metodología

3.2.1 Fase 1

3.2.1.1 Selección de factores condicionantes y detonantes en la zona de estudio

Para la selección de los factores condicionantes y detonantes se realizó una revisión bibliográfica partiendo de los antecedentes usados para este tipo de estudio. Con dicha revisión se identificaron los factores que cada autor usó para su respectivo estudio, analizándolos mediante una matriz comparativa (Tabla 2.6) en la que se visualizan los factores más usados. Esto permitió seleccionar los factores idóneos para la clasificación de movimientos en masa teniendo en cuenta los usados por encima del 50 % en los artículos consultados.

Identificar los factores condicionantes y detonantes posibilita definir una línea de partida para la búsqueda de las bases de datos que contengan la información espacial descrita en la siguiente fase de la metodología. Es importante aclarar que muchos de estos datos no son de acceso libre y fue necesario hacer las solicitudes respectivas a las entidades territoriales para que facilitaran el acceso a la información o generaran los usuarios de acceso requeridos para el ingreso a los geoportales con los perfiles habilitados para la descarga de los diferentes datasets.

3.2.1.2 Compilación y depuración de datos

■ Caracterización de la zona de estudio

Teniendo en cuenta los factores definidos, se realiza la búsqueda de información georreferenciada sobre la ocurrencia de MM a nivel nacional, buscando aumentar la cantidad de datos que ingresan al modelo; debido a que Corinto cuenta con muy pocos registros del fenómeno en el SIMMA del SGC (2022). Inicialmente se elabora la caracterización de los factores condicionantes definidos (Tabla 2.6), puesto que son constantes en el tiempo para así buscar municipios con condiciones similares en geología, pendientes y uso del suelo (Tabla 3.3, 3.4 y 3.5 respectivamente) que tengan información sobre la ocurrencia de MM (Tabla 3.6) y así suplir la necesidad de datos.

Con base en lo mencionado anteriormente, se generan las siguientes tablas descriptivas:

Unidad Geológica	Descripción	Edad	Clases
Q-ca	Abanicos aluviales y depósitos coluviales	Cuaternario	1
K2-Vm	Basaltos toleíticos, y doleritas, picritas, tobas básicas y brechas volcánicas.	Cretácico Superior	2
K1-Mbg	Esquistos anfibólicos, cuarzosericíticos, cloríticos y talcosos, y cuarcitas.	Cretácico Inferior	3
K1-Mmg	Anfibolitas, algunas granatíferas; metagabros; metadioritas, y serpentinitas.	Cretácico Inferior	4
T-Mmg	Gneises cuarzofeldespáticos algunos con sillimanita, cordierita y hornblenda; anfibolitas; migmatitas; esquistos, y mármoles.	Triásico	5

Tabla 3.3: Unidades geológicas de Corinto - Cauca
Fuente: Tomado de Gómez & Montes (2020)

Inclinación (°)	Descripción	Clases
0-2	Plano a casi plano. No hay denudación apreciable.	1
2-4	Suavemente inclinado. Movimientos en masa de baja velocidad y procesos erosivos de diferentes tipos, especialmente bajo condiciones periglaciares (solifluxión) y fluviales (erosión laminar y surcos). Susceptible a desarrollar procesos erosivos	2
4-8	Inclinado. Condiciones similares a las anteriores. Alta susceptibilidad a desarrollar procesos erosivos.	3
8-16	Moderadamente abrupto. Movimientos en masa de todos los tipos, especialmente solifluxión periglaciara, reptación y ocasionalmente deslizamientos, también erosión de tipo laminar y en surcos. Susceptible a erosión y deslizamientos.	4
16-35	Abrupto. Procesos denudacionales intensos de diferentes tipos (erosión bajo cubierta forestal, reptación, deslizamientos). Alta propensión al desarrollo de procesos erosivos.	5
35-55	Muy abrupto. Afloramientos rocosos, procesos denudacionales intensos, depósitos granulares caóticos de poco espesor.	6
>55	Extremadamente abrupto. Afloramientos rocosos. Procesos denudacionales muy fuertes, especialmente “denudación de escarpe”; susceptible a rodamiento de rocas.	7

Tabla 3.4: Rangos de pendientes Corinto - Cauca.
Fuente: SGC (2017)

■ Selección de municipios homólogos

De acuerdo con la caracterización geoambiental de Corinto, se realiza la búsqueda de los municipios homólogos que poseen más registros de MM en el SIMMA (SGC, 2022). Los municipios seleccionados presentan similitudes geoambientales (geología, pendientes y uso del suelo) con Corinto y fueron escogidos mediante consultas espaciales en las bases de datos utilizadas, donde el factor de búsqueda se basó en la información que se tiene de la cuenca del río La Paila. Con este filtro alfanumérico fue posible seleccionar los municipios presentados en la Tabla 3.7. Se destaca que en su mayoría se localizan sobre la Cordillera Central colombiana.

Usos del suelo	Clases
Agrosilvopastoril con cultivos permanentes	1
Agrosilvícola con cultivos permanentes	2
Agrosilvícola con cultivos transitorios	3
Conservación de Recursos Hidrobiológicos	4
Conservación y Recuperación Erosión	5
Cuerpo de agua	6
Cultivos permanentes intensivos de clima cálido	7
Cultivos permanentes intensivos de clima medio	8
Cultivos permanentes semi intensivos de clima cálido	9
Cultivos permanentes semi intensivos de clima frío	10
Cultivos permanentes semi intensivos de clima medio	11
Cultivos transitorios intensivos de clima cálido	12
Cultivos transitorios intensivos de clima frío	13
Cultivos transitorios semi intensivos de clima cálido	14
Forestal de producción de clima cálido	15
Forestal de producción de clima medio	16
Forestal de protección	17
Pastoreo extensivo de clima cálido	18
Pastoreo semi intensivo de clima cálido	19
Protección – producción	20
Silvopastoril	21
Zonas urbanas	22

Tabla 3.5: Principales usos del Suelo de Corinto - Cauca
Fuente: Tomado de IGAC (2017)

Tipo de movimiento	Clases
Movimiento en masa	1
No movimiento en masa	0

Tabla 3.6: Clasificación del Movimiento

■ Datos de movimientos en masa

Una vez se escogen los municipios homólogos, se cuenta con 11 municipios incluyendo a Corinto. A partir de esto se obtienen 272 datos de MM a través del Sistema de Información de Movimientos en Masa (SIMMA) de la pagina web del Servicio Geológico Colombiano (SGC, 2022), del cual se descarga la información en formato shape tipo punto.

Departamento	Municipio	N° MM
Caldas	Anserma	19
	Risaralda	11
	Samaniego	38
	San José	6
Cauca	Caloto	54
	Corinto	6
Nariño	Ricaurte	20
	Nariño	4
Tolima	Santa Isabel	15
Valle del Cauca	Sevilla	37
	Tuluá	62

Tabla 3.7: Municipios homólogos con movimientos en masa
Fuente: Tomado de SGC (2022)

Teniendo los dataset por municipio, se procede a unir la información espacial a través del software libre QGIS 3.10 con el objetivo de tener en una misma base de datos los MM registrados. Es importante aclarar que la base de datos no solo debe contar con los puntos en donde ocurrieron MM, sino que también debe tener información de zonas donde no hayan ocurrido dichos movimientos (No MM). Estos últimos puntos se escogen de manera aleatoria y se almacenan en el mismo shapefile caracterizando de forma binaria los “MM” con un valor de 1 y los “No MM” con un valor de 0. Lo anterior se hace con el fin de facilitarle a la red reconocer y aprender sobre los factores condicionantes y detonantes presentes en las zonas que son o no son propensas a la ocurrencia de MM. De esta manera, se obtuvieron 238 puntos de zonas sin ocurrencia de MM y como se mencionó anteriormente 272 puntos que registran MM distribuidos en el territorio nacional entre las coordenadas: Longitud 78.076 W y Latitud 0.159 N hasta la coordenada Longitud 71.144 W y Latitud 10.328 N, como se muestra en la Figura 3.3 para un total de 510 datos.

Con los 510 puntos que se tienen georreferenciados (Latitud, Longitud) en el sistema de referencia WGS84 (4326), se continúa con la extracción de los factores detonantes (precipitaciones) y condicionantes (geología, usos de suelo y pendiente) para los puntos seleccionados. Para obtener los factores condicionantes, se utilizó la herramienta de geoprocso vectorial *intersección* del software QGIS 3.10, buscando que la información almacenada en el polígono que representa cada uno de los factores se anexe a la base de

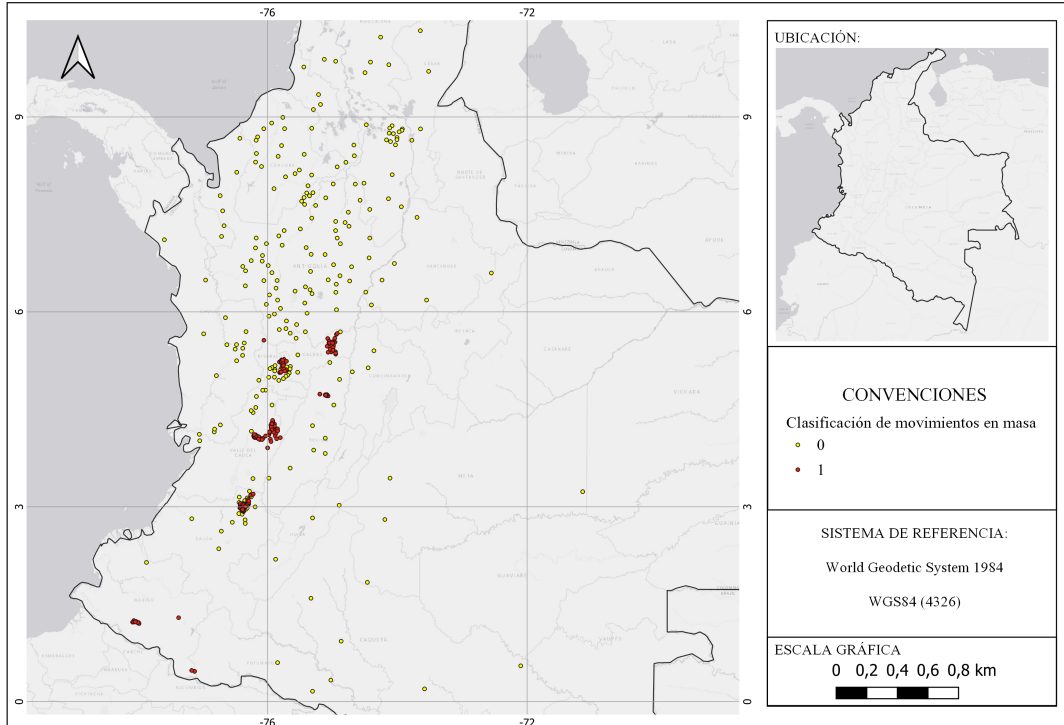


Figura 3.3: Distribución espacial de puntos MM (rojo) y no MM (amarillo)

datos (tipo punto). Por otro lado, las precipitaciones se extraen teniendo en cuenta la misma base de datos inicial, sin embargo, este factor se obtiene importando las librerías de Google Earth Engine (GEE) al entorno de desarrollo colaborativo Google Colaboratory que se ejecuta mediante el lenguaje de programación Python.

■ Pendientes

Los datos de las pendientes se obtuvieron del NASADEM Digital Elevation 30m (NASA JPL, 2020) que se encuentra alojado en la plataforma de análisis científico y visualización de conjuntos de datos geoespaciales GEE. Este conjunto de datos contiene información mejorada del reprocesamiento de los datos de elevación digital Shuttle Radar Topography Mission (SRTM) dado que integra información de los dataset ASTER GDEM, PRISM, ICESat y GLAS (GEE, 2020). Para la extracción de las pendientes de GEE es necesario cargar el shapefile con los 510 puntos de MM y no MM (Figura 3.3) en Google Colaboratory, dentro del cual mediante una codificación en lenguaje java se importa el set de datos requerido como se muestra a continuación:


```

var DEM = ee.Image("NASA/NASADEM_HGT/001");
var slope = ee.Terrain.slope (DEM);

var puntos = ee.FeatureCollection(punto);
var training = slope.sampleRegions({
  collection: puntos,
  geometries: true,
  scale: 0.1});

Export.table.toDrive({
  collection: training,
  description: "pendiente",
  selectors: ([ "OBJECTID", "TIPO_MOV", "X", "Y", "slope" ]),
  fileFormat: "CSV"
})

```

Al tener las pendientes con su respectiva coordenada X y Y (Longitud, Latitud) en formato CSV se facilita su incorporación al dataset inicial mediante el software QGIS 10.3. Con lo anterior se obtiene el primer factor condicionante de la base de datos general.

■ Geología

Por otro lado, para la geología se utiliza el shape del Mapa Geológico de Colombia a escala 1:100.000, teniendo en cuenta que es la única escala en formato digital disponible de este insumo, ya que las entidades territoriales municipales no cuentan con esta información a una mejor resolución espacial (Gómez & Montes, 2020).

Mediante el software QGIS 10.3 se realiza la corrección de las geometrías del shape con la herramienta de procesos *reparar geometrías* puesto que al ser un archivo muy extenso presenta errores topológicos. Teniendo en cuenta las correcciones necesarias, se procede con el geoproceto vectorial entre la geología y los 510 puntos usando la herramienta *intersección* con la finalidad de agregar la geología a la base de datos que ya contiene las pendientes.

■ Usos de suelo

Los usos del suelo se obtienen de la pagina de datos abiertos del IGAC, alojado en la sección de agrología. El shapefile se titula “Mapas de Clasificación de las Tierras por su Vocación de Uso a Escala 1:100.000” (IGAC, 2017). Con este insumo se debe realizar

el mismo geoproceto implementado con la geología mediante el software QGIS 10.3. Como primer paso se realiza la corrección de la geometría, y finalmente se hace la intersección entre el shape de la geología con los 510 puntos que contiene las variables necesarias para identificar los MM. De esta manera se obtiene en un mismo dataset con formato shape tipo punto los factores condicionantes planteados en la Fase 1. Las características alfanuméricas del archivo generado se presentan en la Tabla 3.8 de forma resumida teniendo en cuenta la cantidad de datos que esta contiene (tipo de movimiento y factores condicionantes estáticos).

■ Precipitaciones

Las precipitaciones se obtuvieron mediante el uso de los datos del ERA5 alojado en GEE (2022). Este dataset contiene datos a nivel global en formato raster del reanálisis atmosférico ECMWF (Centro Europeo de Previsiones Meteorológicas a Plazo Medio, por su siglas en ingles) de quinta generación, combinando datos capturados en todo el mundo con diferentes modelos. Los datos que contiene el ERA5 presenta la información de 7 variables sobre la superficie terrestre a una resolución espacial de 27,83 km. Estas variables son: La temperatura del aire, la temperatura del punto de rocío, la presión media del nivel del mar, la presión superficial, el componente U del viento, el componente V del viento y la precipitación total. Siendo esta última la variable de interés, dada como suma diaria con una frecuencia de 24h en la captura de la información.

Para extraer las precipitaciones de la colección de imágenes se usa el entorno de desarrollo colaborativo Google Colaboratory, en el cual es necesario importar la Api de GEE otorgando los permisos requeridos para la autenticación e inicialización de esta. Posteriormente, se implementa el código que define el set de datos requerido, las coordenadas del punto que se desea consultar y la ventana temporal que para este caso se define en tres meses (93 días). El código desarrollado en lenguaje Python se muestra a continuación:

```

from dateutil.relativedelta import *
from datetime import datetime
import ee

ee.Authenticate()
ee.Initialize()

rainfall_mm= pd.DataFrame()
t=len(MM[:,2])
i=0

for i in range(t):
    date=Fecha_MM[i]
    dia_mm = datetime.strptime(date, "%Y-%m-%d")
    hace_tres_meses = dia_mm + relativedelta(months=-3)
    un_dia= dia_mm + relativedelta(days=1)
    #Se llama el set de datos ERA5 que contiene las precipitaciones
    set_precipi= ee.ImageCollection("ECMWF/ERA5/DAILY")\
        .filterDate(hace_tres_meses, un_dia)\
        .select("total_precipitation");
    # Se cargan coordenadas del punto de cada MM
    punto = ee.Geometry.Point(x_MM[i],y_MM[i])
    #Se extraen los valores de los pixeles correspondientes a cada punto georeferenciado

    def poi_mean(img):
        dato_poin = img.reduceRegion(reducer=ee.Reducer.mean(), geometry=
            punto).get("total_precipitation")
        return img.set("date", img.date().format()).set("mean",dato_poin)

    dato_day = set_precipi.map(poi_mean)
    lista = dato_day.reduceColumns(ee.Reducer.toList(2), ["date","mean"]).
        values().get(0)
    df = pd.DataFrame(lista.getInfo(), columns=["date","precipitacion"])
    pre=df[["precipitacion"]]
    #Se convierten los datos extra dos a un DataFrame y se guardan
    fil= pd.DataFrame(data=pre)
    rainfall_mm= rainfall_mm.append(fil.transpose())
    i=i+1

```

Es importante aclarar que las coordenadas de cada punto se encuentran en el set de datos construido en pasos anteriores, por lo tanto, el código presentado debe recorrer las listas que contienen la información extraída del dataset general para obtener las fechas y coordenadas necesarias. Permitiendo así, tener la base de datos con toda la información requerida para el desarrollo de este proyecto.

3.2.2 Fase 2

3.2.2.1 *Procesamiento de datos*

- **Normalización de datos.**

La normalización del conjunto de datos es el proceso inicial para los estimadores de aprendizaje automático. Como la información en su estado bruto puede presentar ruido debido a redundancias, por lo tanto, es necesario asociarla con patrones similares que tengan una distribución adecuada para evitar problemas de aprendizaje en el modelo que se desea desarrollar. Teniendo en cuenta lo anterior, para las variables categóricas (pendiente, uso del suelo y geología) se implementa el método *get_dummies()* de la librería *Pandas* (The pandas development team, 2020), permitiendo convertir cada dato en una variable ficticia o indicadora a partir de columnas compuestas por valores de 1 y 0 según la cantidad de clases que existan en el dataframe indicado (Kumari & Kashish, 2020).

Por otra parte, para la normalización de la variable temporal (precipitaciones) se debe implementar el método *MinMaxScaler()* de la librería *Sklearn* (Pedregosa et al., 2011) debido a que para este proyecto es importante mantener la temporalidad definida por el factor condicionante. El método *MinMaxScaler* es un estimador de escala que transforma individualmente las características de los datos conservando la proporcionalidad según el rango que se defina (Scikit-learn developers, 2020), en este caso se definió un rango entre 0 y 1 con la misma finalidad que en la variables categóricas: reducir los problemas de aprendizaje del algoritmo.

- **Construcción de la Red Neuronal.**

Para iniciar con la construcción del modelo, se hace la división del dataset entre datos de entrenamiento y datos de testeo, los cuales se definen con un 80 % y 20 %

respectivamente, donde los datos de entrenamiento se subdividen en un 80 % para entrenar modelo y un 20 % para validarlo en el *model.fit* (Anexo 6.5).

A partir de la división del dataset general y el análisis de los hiperparámetros, se puede identificar qué tipo de arquitectura es la adecuada para el procesamiento de los datos. Conociendo el comportamiento y tipo de información que se va a procesar, se puede observar que la precipitación debido a la frecuencia en la captura de sus datos en un punto (granularidad), registra información temporal o secuencial, por lo tanto se puede inferir que los datos se procesan mejor en una RNN, no obstante, para este proyecto se implementa la GRU, puesto que son usadas para captar señales de disparo, es decir que pueden identificar las variaciones inusuales en la precipitación, además son una mejora de la LSTM, que a su vez son mas óptimas que las RNN, considerando que las RNN presentan una deficiencia en el manejo del gradiente descendente para la memoria a largo plazo (Shewalkar, 2019).

En cuanto a las tres variables estáticas (pendiente, geología y usos del suelo) se ingresan al modelo mediante una capa densa de forma individual, para luego ser procesadas y finalmente concatenadas con el resultado del procesamiento de las precipitaciones. A continuación se explica en detalle el uso de las variables.

- Variables de entrada:

- Precipitación diaria
- Pendiente
- Geología
- Uso del suelo

- Matriz de datos:

Teniendo en cuenta la información que se extrae de cada punto que representa la presencia o ausencia de MM, se generan dos matrices.

- Matriz de datos temporales:

Es una matriz de tres dimensiones (None, 93, 3) que contiene la información de las precipitaciones extraídas con cada punto del set de datos. Contiene 93 columnas que representan cada uno de los días en los 3 meses analizados y tres ventanas temporales por punto. Las ventanas de tiempo se definen con el fin de ingresar diferentes temporalidades a la red

y mejorar su aprendizaje sobre una misma línea de tiempo. Estas presentan un rezago de 10 días entre si a partir de la fecha en que se registra el MM (Figura. 3.4).

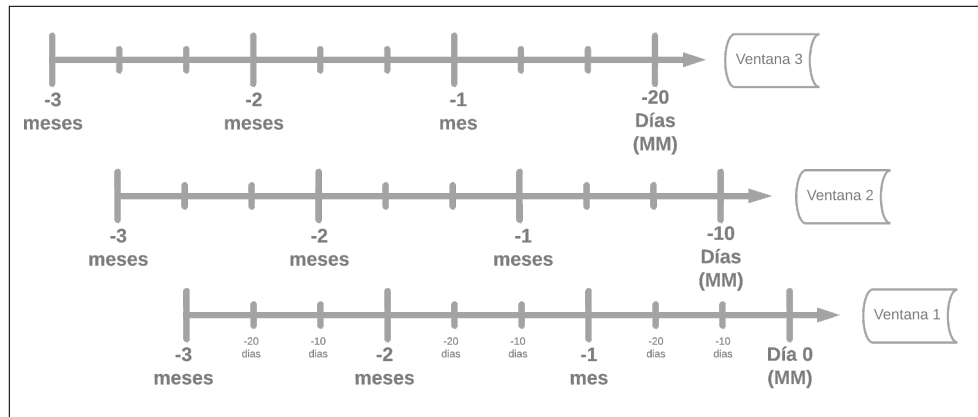


Figura 3.4: Ventanas de tiempo

o Matriz de datos estáticos:

Es la matriz compuesta por la información correspondiente a los factores condicionantes, es decir que está compuesta por las pendientes, la geología y el uso del suelo, además del tipo de movimiento y coordenada geográfica asociada a cada punto del dataset (Tabla 3.8).

TIPO DE MOVIMIENTO	X	Y	PENDIENTE	GEOLOGÍA	USO DE SUELO
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...
0	-76,389	3,012	18,386	K2-Vm	Conservación y Recuperación de erosión
0	-76,400	2,905	22,975	K1-Mbg	Protección y producción
0	-76,322	3,068	17,644	K2-Vm	Protección y producción
0	-76,333	2,960	27,427	K1-Mbg	Protección y producción
1	-77,122	0,466	6,658	E3N1-Sct	Protección y producción
1	-77,122	0,466	6,658	E3N1-Sct	Protección y producción
1	-77,167	0,480	3,818	Q-ca	Cultivos permanentes semi intensivos de clima cálido
1	-77,368	1,292	20,263	N2-p	Pastoreo extensivo de clima cálido
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...

Tabla 3.8: Matriz de datos estáticos

Es importante aclarar que la anterior Tabla solo muestra el tipo de movimiento con valores de 0 y 1 (No MM y MM respectivamente) ya que la codificación generada con la normalización del resto de datos genera una matriz compuesta por unos y ceros difícil de presentar en este documento. Sin embargo, las matrices que ingresan al modelo son producto de la normalización de las pendientes, la geología y el uso del suelo. Estas tienen tres dimensiones y están definidas por (None, 7, 1), (None, 62, 1) y (None, 22, 1) respectivamente.

- Hiperparámetros relevantes:

La descripción de los hiperparámetros se debe hacer en dos bloques al igual que las matrices de entrada, puesto que el procesamiento que se hace con cada

tipo de dato es diferente. No obstante, el modelo definido concatena los resultados de cada submodelo, permitiendo así definir los hiperparámetros que se usan en el modelo general.

- Hiperparámetros para los datos temporales:

El modelo implementado para las precipitaciones está compuesto por una capa convolucional 1D, tres capas de normalización por lotes, una capa de activación, dos capas GRU, dos dropout y finaliza con una capa densa definidas de la siguiente manera:

```
X = Conv1D(filters=196, kernel_size=1, strides=1)(
    X_input)
X = BatchNormalization()(X)
X = Activation("relu")(X)
X = Dropout(rate=0.25)(X)
X = GRU(200, return_sequences=True)(X)
X = BatchNormalization()(X)
X = Dropout(rate=0.25)(X)
X = GRU(200, return_sequences=True)(X)
X = Flatten()(X)
X = BatchNormalization()(X)
X = Dropout(rate=0.25)(X)
X = Dense(128, activation="relu")(X)
```

- Hiperparámetros para los datos estáticos:

Para las tres variables estáticas analizadas se implementa la misma arquitectura en cada submodelo. La diferencia está en los datos que ingresan a cada uno teniendo en cuenta la cantidad de categorías resultantes en las diferentes variables. Cada submodelo está compuesto por dos capas densas y una capa de activación, así:


```

#Pendientes

X2= Dense(196, activation="relu")(X_input2)
X2= Activation("relu")(X2)
X2= Dense(128, activation="relu")(X2)

# Geologia

X3= Dense(196, activation="relu")(X_input3)
X3= Activation("relu")(X3)
X3= Dense(128, activation="relu")(X3)

# Uso del suelo

X4= Dense(196, activation="relu")(X_input4)
X4= Activation("relu")(X4)
X4= Dense(128, activation="relu")(X4)

```

Con los hiperparámetros definidos, se concatenan los 4 submodelos en un modelo general, buscando relacionar los resultados de cada proceso. Por lo tanto, el modelo concatenado está compuesto por una capa de concatenación, tres capas densas y dos dropout definidos así:

```

Y = Concatenate(axis=1)([X,X2,X3,X4])
Y = Dense(200, activation="relu")(Y)
Y = Dropout(0.25)(Y)
Y = Dense(128, activation="relu")(Y)
Y = Dropout(0.2)(Y)
Y = Dense(1, activation="sigmoid")(Y)

```

Finalmente, el modelo generado como producto de los procesos anteriores se representa gráficamente de la siguiente manera en la Figura 3.5:

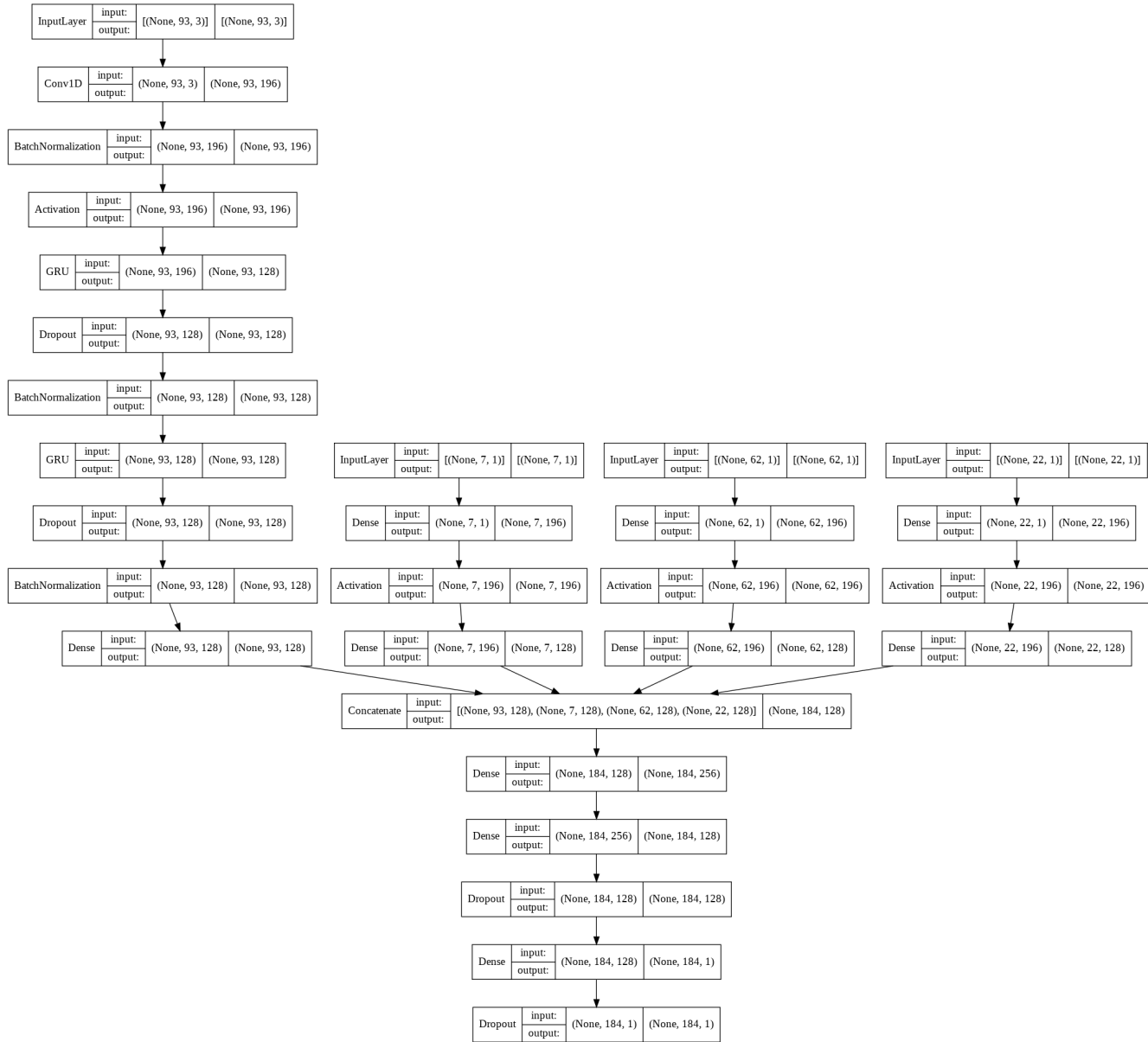


Figura 3.5: Arquitectura desarrollada para la Clasificación

3.2.2.2 Validación de red neuronal

Para evaluar el desempeño de la red construida, se utilizaron 2 tipos de validaciones que se complementan. Primero se realiza el área bajo la curva “ROC” para evaluar la capacidad predictiva del modelo y después la matriz de confusión que permite evaluar el desempeño de cada clase dentro de la red.

3.2.3 Fase 3

3.2.3.1 Clasificación de zonas propensas a MM que son fuente de materiales para avenidas torrenciales

Teniendo en cuenta los resultados de la validación y los análisis que se realizaron para definir los rangos que indican la existencia o no de un MM según las predicciones que calcula algoritmo, se procede a clasificar las zonas propensas a la ocurrencia de movimientos en masa presentes en la cuenca del Río La Paila de acuerdo con los porcentajes de correlación que arroje el modelo entrenado, facilitando así la mapificación de los MM que son fuente de material para avenidas torrenciales. Finalmente, se representan las zonas propensas a este tipo de movimiento por medio de herramientas GIS.

Es importante aclarar que este modelo computacional predice la posible ocurrencia de un MM con 10 días de anticipación si la zona analizada presenta las condiciones ideales para el desarrollo del fenómeno. De acuerdo al modelo, cuando la probabilidad de ocurrencia es mínima o nula clasifica el área con ausencia de MM y hace una tercera clasificación “incierto” cuando no puede decidir si existe o no la posibilidad de un MM; sin embargo se genera la alerta. En otras palabras, el clasificador desarrollado no discrimina entre los tipos de movimiento en masa existentes debido a que fue entrenado para decidir sobre la posible ocurrencia o ausencia del fenómeno analizado.

Para obtener las salidas cartográficas finales, la red neuronal recibe la información espacial de la zona de estudio. Sin embargo, dichos datos requieren un geoproceso previo en el software QGIS (intersección), debido a que es necesario tener los factores condicionantes y detonantes asociados a un identificador (id) que será usado al final del

procesamiento para conocer la predicción realizada para cada combinación de datos existente en el dataset de polígonos entregado. En otras palabras, el entrenamiento de la red se hizo mediante datos puntuales como se muestra en la figura 3.3 buscando que el modelo aprenda sobre las diferentes combinaciones posibles entre los factores condicionantes y detonantes, sin embargo los datos de la zona de estudio se entregan en forma de polígono con el fin de reducir la carga computacional que requiere procesar dicha información tipo punto.

3.3 Síntesis

El desarrollo metodológico permitió plantear una perspectiva para la identificación de zonas propensas a la ocurrencia de MM que pueden ser fuente de materiales para avenidas torrenciales. Se tuvo en cuenta la construcción de una red neuronal que parte de una arquitectura de red basada en redes neuronales de aprendizaje profundo y en la experiencia adquirida durante su desarrollo, permitiendo así procesar las variables condicionantes y detonantes ampliamente analizadas por diferentes autores (Capítulo 2), quienes desarrollan estudios relacionados con el fenómeno que se plantea clasificar en este trabajo de grado a través de un método indirecto.

4. Resultados y Discusiones

Este Capítulo presenta los resultados obtenidos a partir de la metodología planteada en el Capítulo 3 con el fin de desarrollar las discusiones pertinentes. Aquí, se expone el desempeño obtenido a partir del algoritmo generado para la clasificación de los movimientos en masa (GRU y RNA). La sección 4.1. presenta los resultados obtenidos en la revisión bibliográfica de los antecedentes para definir los factores condicionantes y detonantes que se usaron. La sección 4.2. hace un análisis descriptivo de la base de datos construida, además del análisis respectivo del desempeño que obtuvo la red neuronal desarrollada. Finalmente, la sección 4.3. presenta los resultados del procesamiento aplicado a los datos de precipitación, pendiente, geología y uso del suelo correspondientes a la cuenca del río La Paila.

4.1 Fase 1

En la primera fase se realizó una revisión bibliográfica con la cual se generó la Tabla 2.6 de los antecedentes relacionados con el presente trabajo. Este ejercicio permitió definir varios aspectos en el desarrollo del proyecto. Principalmente se definieron los factores condicionantes y detonantes mas usados para la solución de la problemática planteada según estudios desarrollados por otros autores. Además, fue posible identificar los diferentes esquemas metodológicos teniendo en cuenta clasificadores implementados, fuentes de datos y variables a las que se dan respuesta con su correspondiente validación.

Como resultado final se escogieron 3 factores condicionantes y un factor detonante. Los factores condicionantes escogidos fueron: La pendiente, la geología y el uso del suelo. En cuanto al factor detonante, se seleccionó la precipitación debido a que esta presenta patrones de comportamiento que facilitan su análisis a través del tiempo basados en series temporales óptimas para una RNN (GRU). Por otra parte, aunque los eventos sísmicos pueden ser identificados a través de la caracterización sismotectónica de una región, estos no son predecibles con precisión (Prado, 2022). Es decir que no presentan un comportamiento temporal recurrente que pueda ser modelado a lo largo del tiempo y aunque son un factor detonante para la ocurrencia de MM, ninguno de los antecedentes consultados usa los sismos como variable de estudio, por lo tanto no se tuvieron en cuenta en el presente trabajo.

4.2 Fase 2

4.2.1 Compilación y depuración de datos

En la compilación y depuración de datos se utilizaron los insumos mencionados en el capítulo anterior: El catálogo histórico de los movimientos en masa se obtuvo del SIMMA (SGC, 2022), la precipitación obtenida del ERA5 (GEE, 2022) y la pendiente extraída del NASADEM alojadas GEE (NASA JPL, 2020), el shapefile del Mapa Geológico de Colombia 2020 (Gómez & Montes, 2020) y el shapefile del Mapa de Clasificación de las Tierras por su Vocación de Uso a Escala 1:100.000 del IGAC (IGAC, 2017).

Con la base de datos estructurada y normalizada en el notebook de Google Colaboratory, se realizó un análisis descriptivo de los datos compilados, para así observar su comportamiento haciendo énfasis en los datos donde ocurrió un MM. La Tabla 3.8 describe el dataset general de las variables estáticas.

4.2.1.1 Geología

El factor geológico presenta variedad en las unidades geológicas. En general, la base de datos se divide en 62 clases como se muestra en la Figura 4.1 compuesta por “MM” y “No MM”, donde se evidencia que las clases predominantes son K2-Vm y Q-al, con 150 y 49 datos respectivamente. Con lo anterior, se puede inferir que la geología está mayormente compuesta por rocas ígneas y volcánicas de composición máfica: basalto toleíticos; doleritas, picritas, tobas básicas y brechas volcánicas (Gómez & Montes, 2020).

En las Figuras 4.2 y 4.3 se presentan las unidades geológicas presentes en las zonas seleccionadas para el entrenamiento de este factor en zonas con MM. La Figura 4.4 exhibe que el tipo de geología predominante en zonas que históricamente han presentado MM. De acuerdo con esta gráfica, la unidad “Formación Volcánica” (K2-Vm), es la unidad geológica que se presenta en la mayoría de zonas donde ocurrió un MM según la base de datos utilizada; puesto que el 48.5 % de los MM están relacionado con esta unidad geológica. Por lo cual, se puede inferir que los terrenos mas propensos a la ocurrencia de MM son los compuestos geológicamente por rocas ígneas y volcánicas de composición máfica.

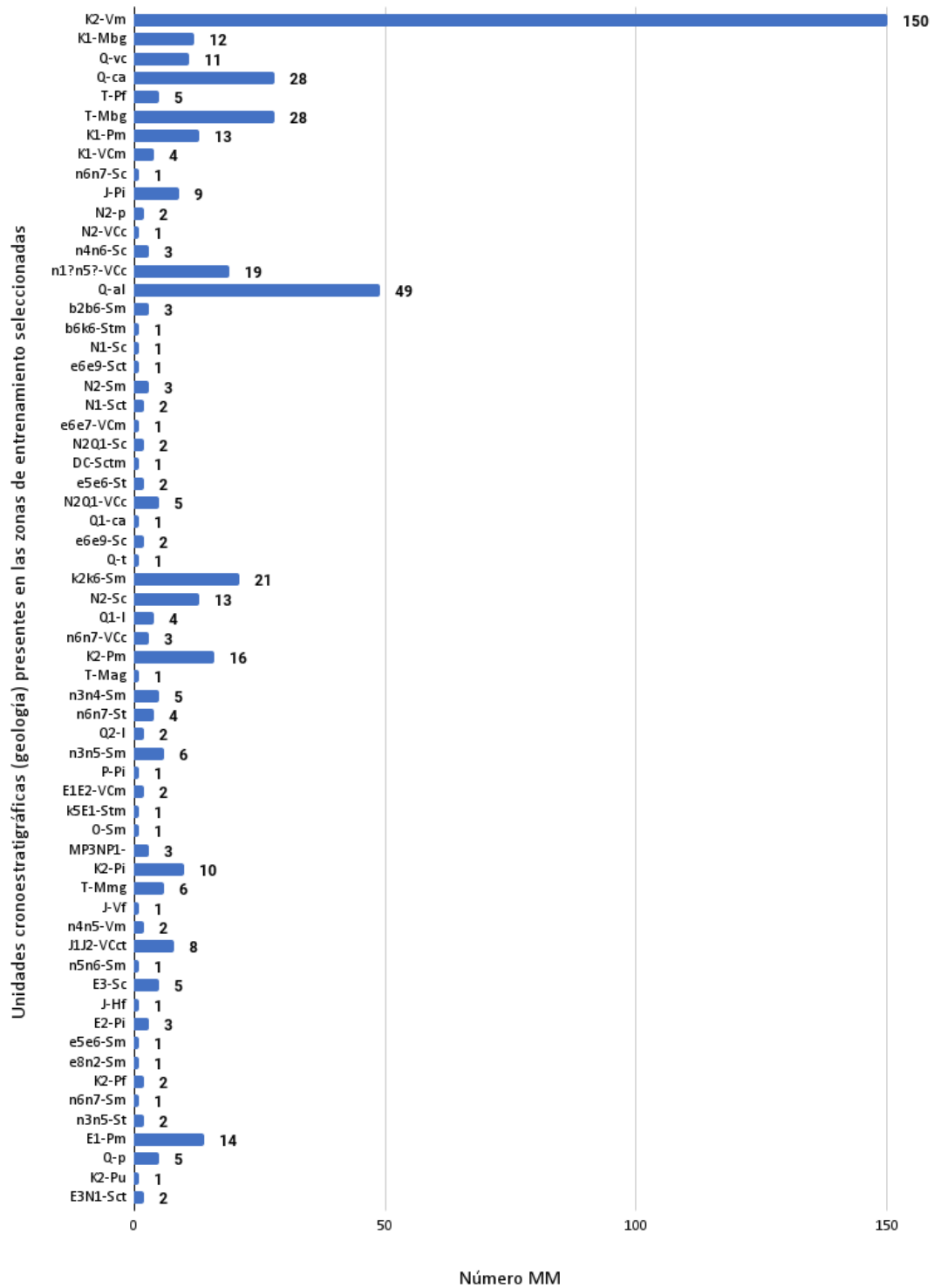


Figura 4.1: Cantidad de eventos registrados (base de datos) para MM y No MM de acuerdo con las unidades cronoestratigráficas (geología)

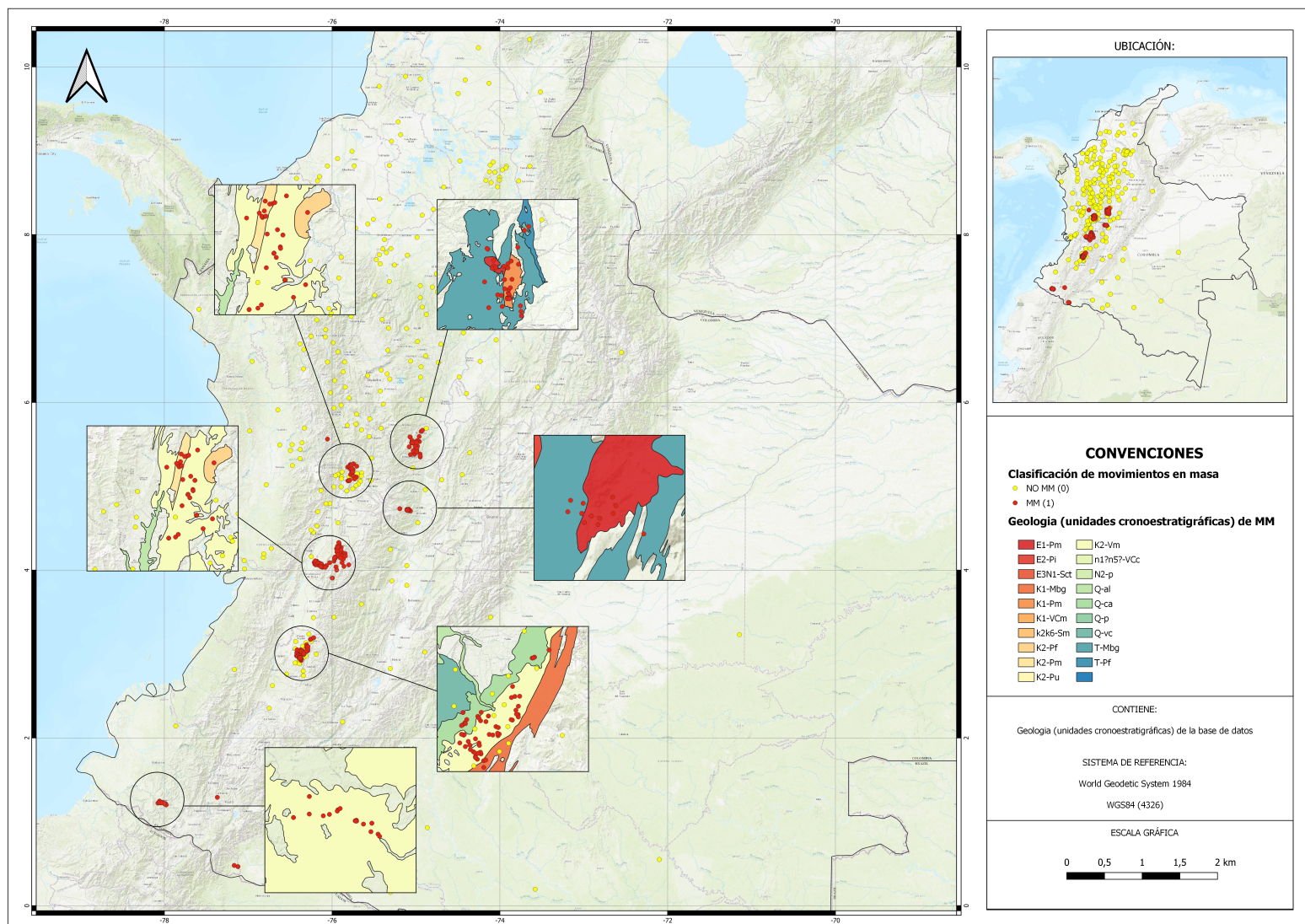


Figura 4.2: Distribución espacial de las unidades cronoestratigráficas (geología) para los puntos de entrenamiento que representan los MM

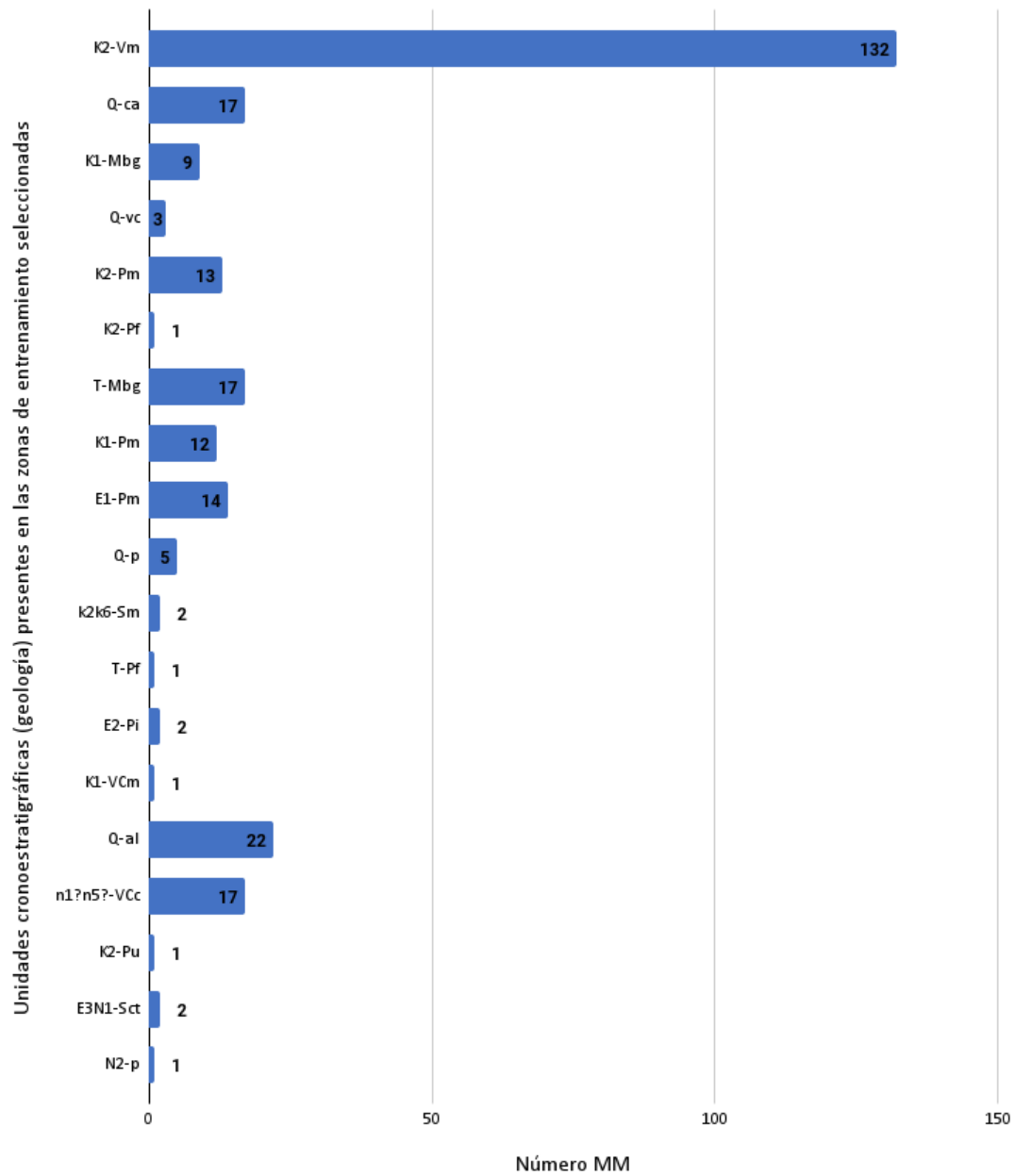


Figura 4.3: Cantidad de movimientos en masa registrados de acuerdo a la geología (unidades cronoestratigráficas)

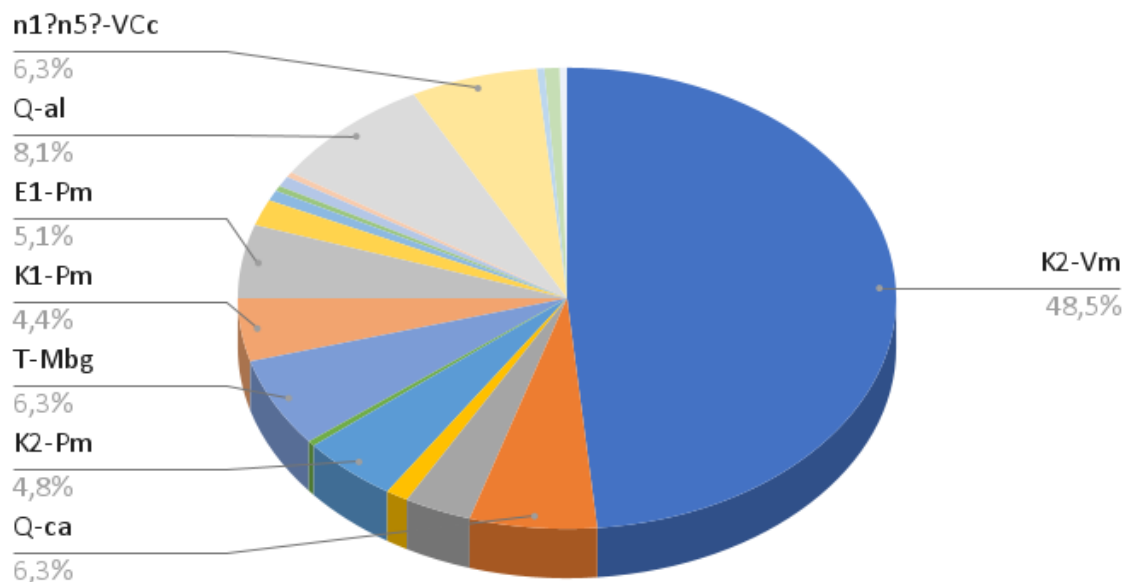


Figura 4.4: Distribución de la ocurrencia de movimientos en masa de acuerdo a la geología (unidades cronoestratigráficas) en porcentaje

4.2.1.2 Usos de suelo

Los usos de suelo en la base de datos usados para el entrenamiento del modelo definieron 23 clases. El uso principal presente en la mayoría de puntos seleccionados (MM y No MM) es el de protección y producción como se muestra en la Figura 4.5. Este uso cubre aproximadamente el 38 % del territorio colombiano (UPRA, 2014) y están formados por territorios de protección forestal, cultivos agrícolas, ganadería entre otros.

En la Figura 4.6 se observa que el uso de suelo predominante en zonas que han presentado movimientos en masa es el de protección y producción seguido por conservación y recuperación de erosión con 152 y 28 datos respectivamente, lo cual se evidencia espacialmente en la figura 4.7.

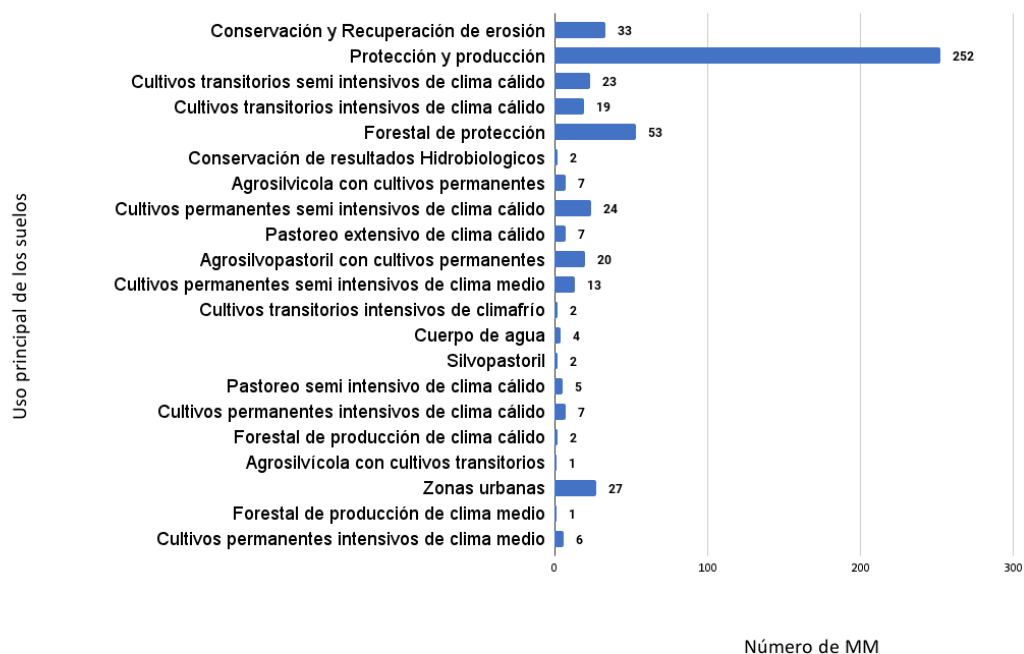


Figura 4.5: Cantidad de eventos registrados (base de datos) para MM y No MM de acuerdo con los usos del suelo

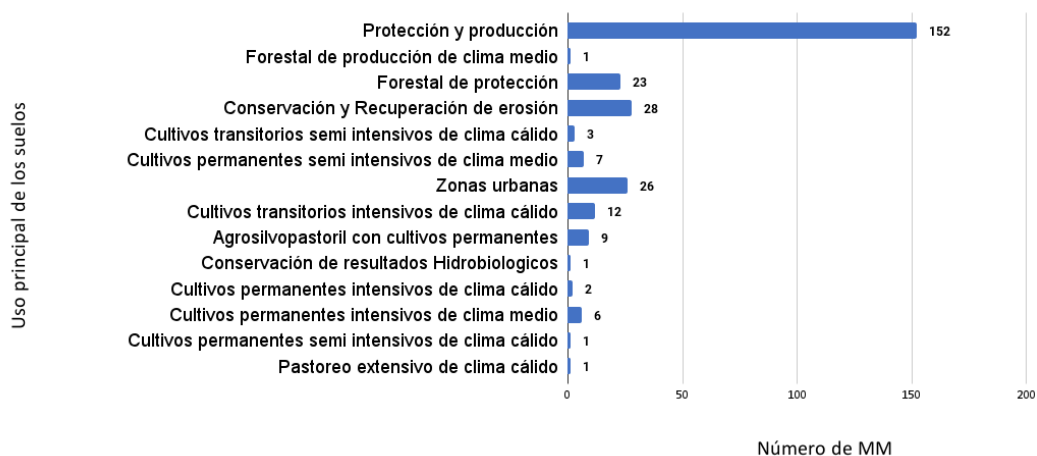


Figura 4.6: Cantidad de movimientos en masa registrados de acuerdo con los usos del suelo

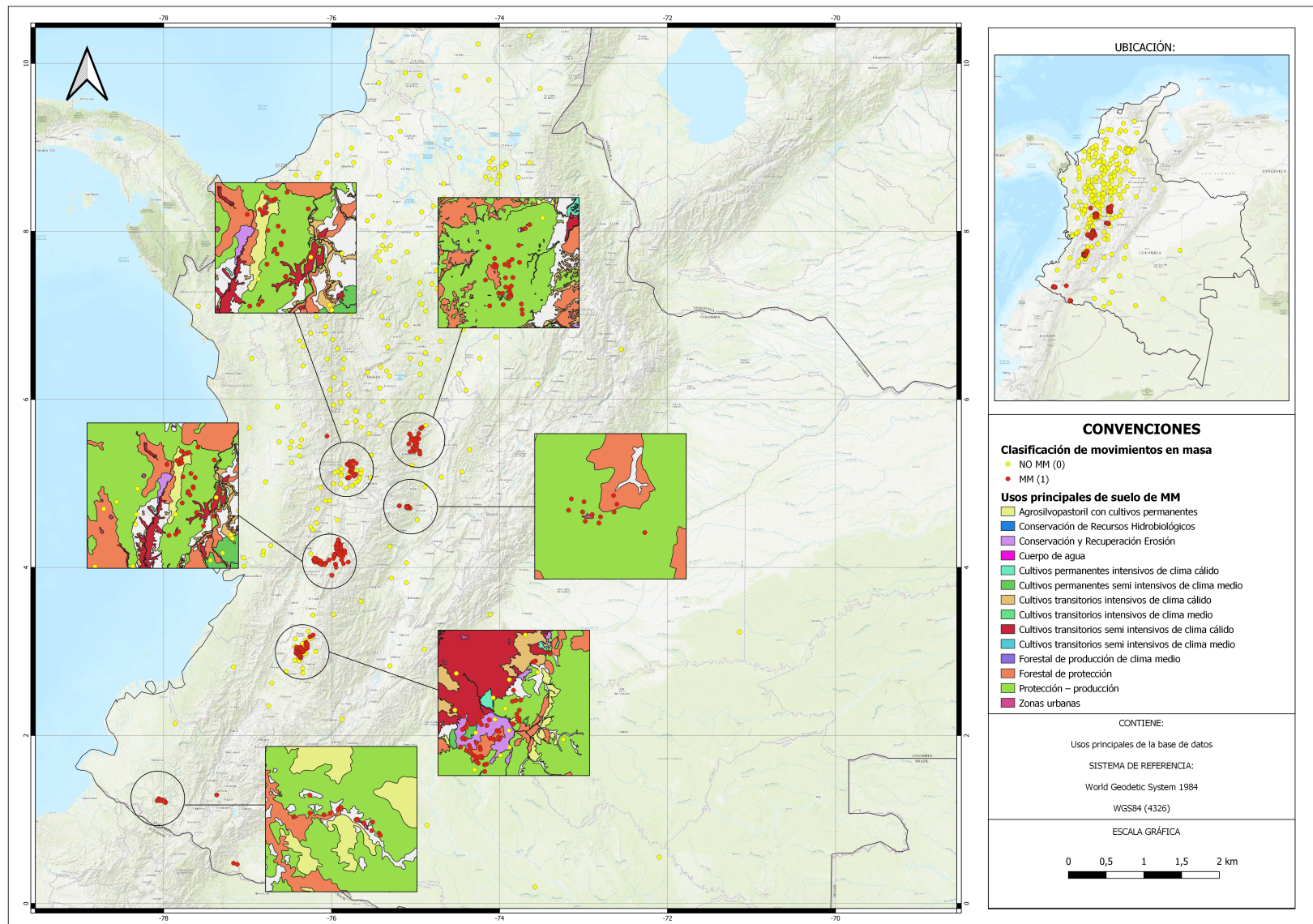


Figura 4.7: Distribución espacial de los usos del suelo para los puntos de entrenamiento que representan los MM

Al analizar porcentualmente la información del Uso del Suelo, se observa que el 55.9 % (Figura 4.8) de los puntos en donde se registraron MM presentan un uso de protección y producción, lo que permite inferir que dicho uso es el correspondiente a las zonas más propensas para la ocurrencia de MM.

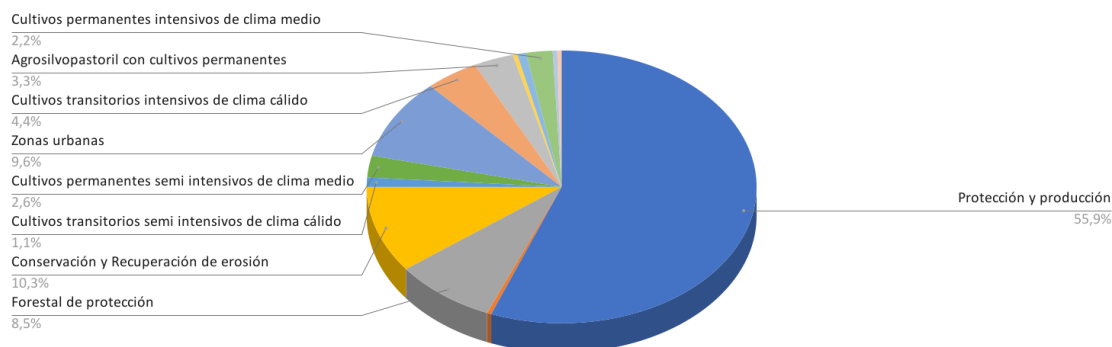


Figura 4.8: Distribución porcentual de los usos del suelo asociados con la ocurrencia de MM

4.2.1.3 Pendiente

Como se mencionó en el Capítulo 3 (Tabla 3.4), las pendientes se dividieron en 7 clases. En la base de datos general (MM y No MM) se observaron pendientes en su mayoría de clase 5 (abruptas) (Figura 4.9), las cuales se generan por procesos denudacionales intensos de diferentes tipos (erosión bajo cubierta forestal, reptación y deslizamientos), con alta propensión al desarrollo de procesos erosivos. La otra clase con alta presencia en este set de datos es la clase 4, correspondiente con pendientes moderadamente abruptas. Este tipo de pendientes presenta movimientos en masa de todos los tipos, especialmente solifluxión periglaciaria, reptación y ocasionalmente deslizamientos, además de ser propensa a erosión laminar y en surcos (SGC, 2017).

Se observó que para la mayoría de puntos seleccionados con ocurrencia de MM también, las pendientes son abruptas (Figura 4.10), seguida de las pendientes de clase 4, que son moderadamente abruptas. Además, esta figura ilustra un comportamiento acorde a la descripción de diferentes autores, puesto que los datos de MM usados para el entrenamiento confirma que la ocurrencia de este fenómeno es directamente proporcional

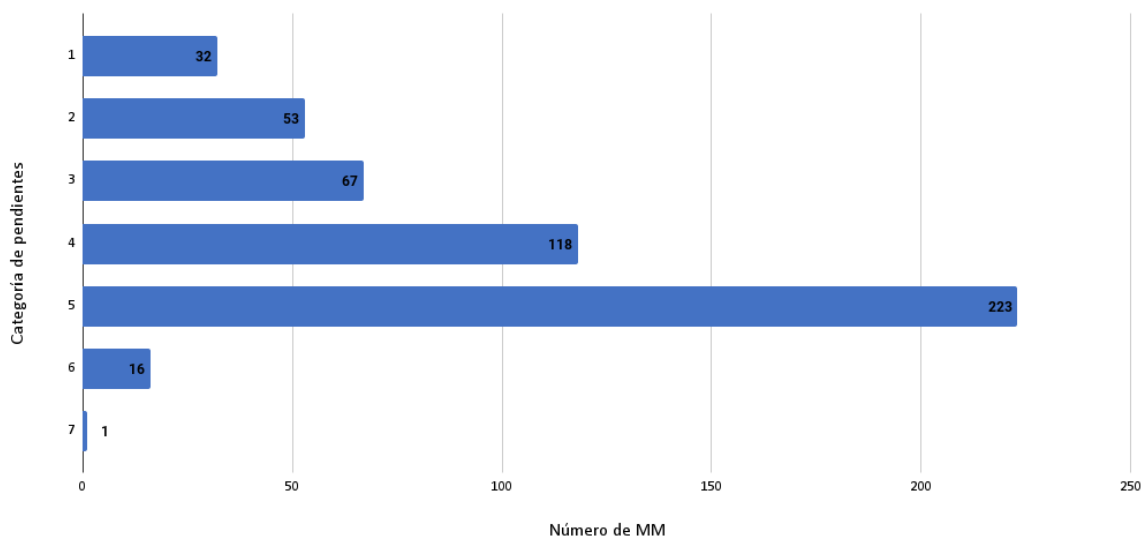


Figura 4.9: Cantidad de eventos registrados (base de datos) para MM y No MM de acuerdo con la clasificación de la pendiente

al aumento de la pendiente; cabe aclarar que el número de MM registrados con pendiente de clase 6 y 7 disminuyeron, debido a que en la base de datos consultada (SGC, 2022) existen pocos datos de MM con dichas pendientes para las zonas homólogas escogidas como se observa espacialmente en la figura 4.11.

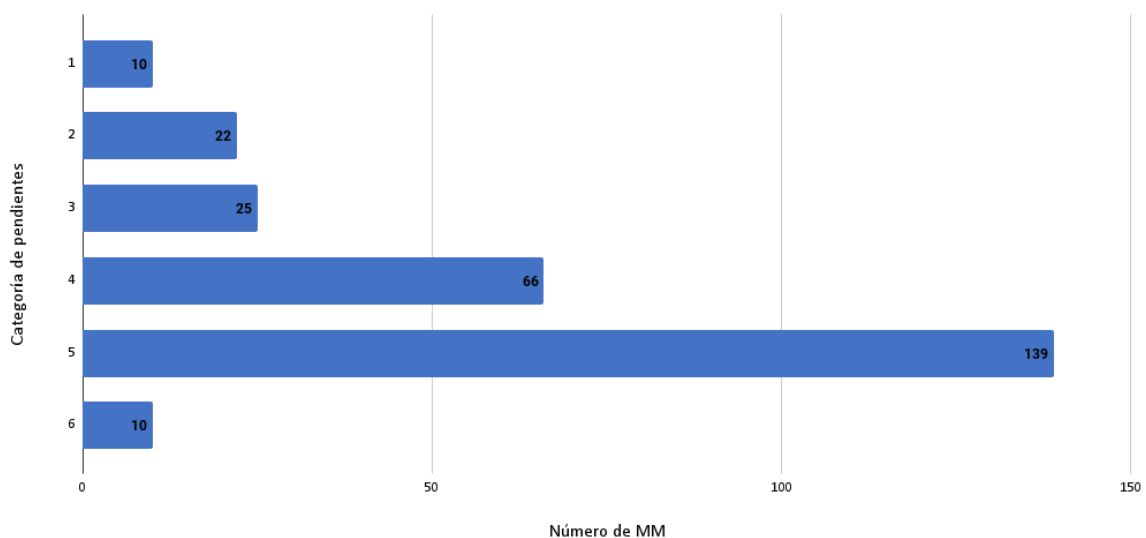


Figura 4.10: Cantidad de movimientos en masa registrados de acuerdo con la clasificación de la pendiente

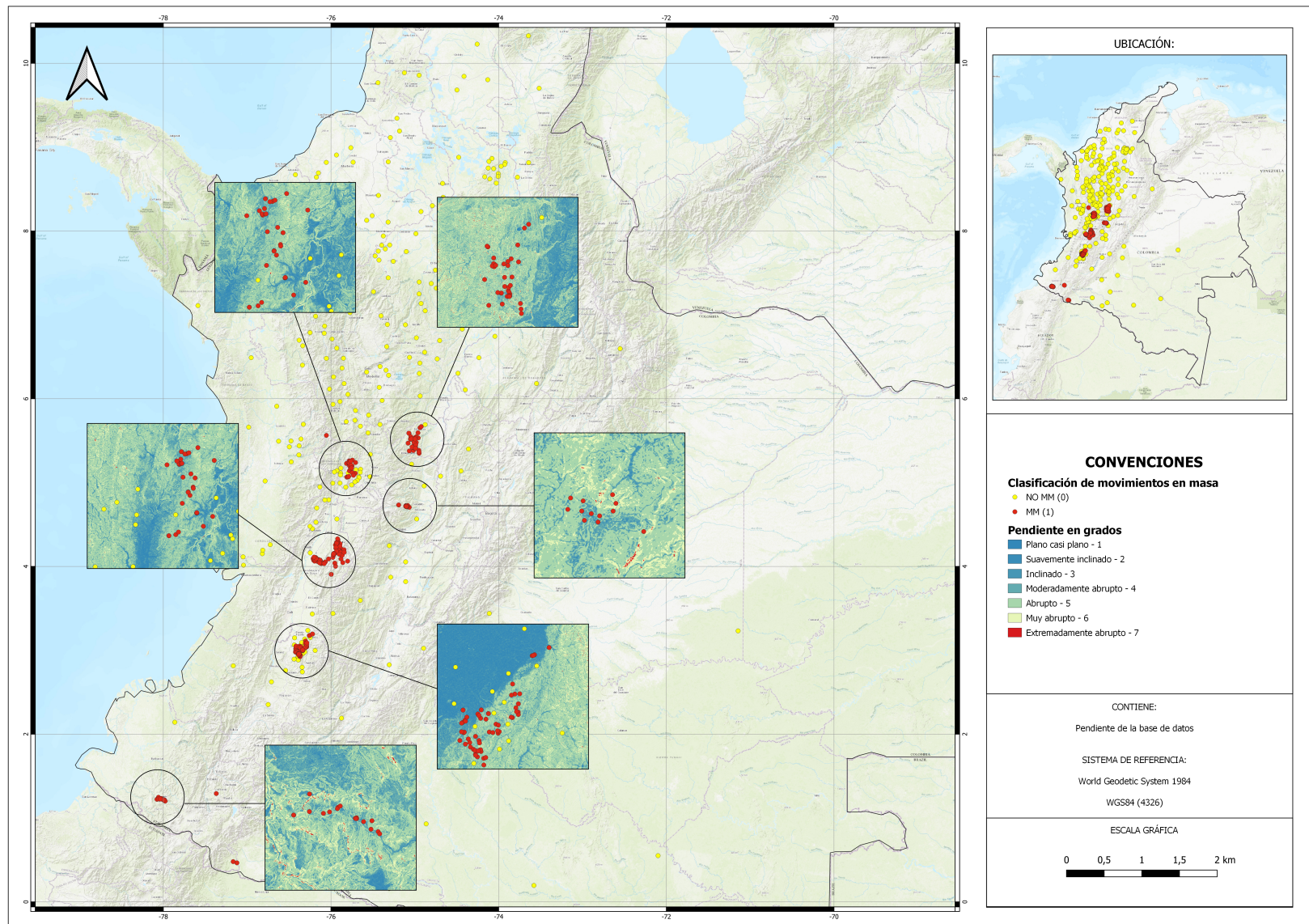


Figura 4.11: Distribución espacial de la pendiente para los puntos que representan los MM

Al analizar el comportamiento de la pendiente en zonas que históricamente han presentado algún tipo de MM, se puede observar según la Figura 4.12 que el 30.9 % se ha desarrollado en pendiente de clase 5. Por lo tanto, se puede inferir que una zona propensa a MM por lo general tiene una pendiente abrupta y es propensa a erosión bajo cubierta forestal, reptación y deslizamientos, como se mencionó anteriormente. Cabe aclarar que las pendientes suavemente inclinadas, las moderadamente inclinadas y las inclinadas también se presentan de forma recurrente en áreas que han presentado MM. Lo anterior permite deducir que aunque la mayoría de MM se presentan en zonas con pendientes de 16° a 35° (abrupta), las zonas con pendientes con rangos entre 2° a 16° también tienen una probabilidad de presentar MM en la zona de estudio.

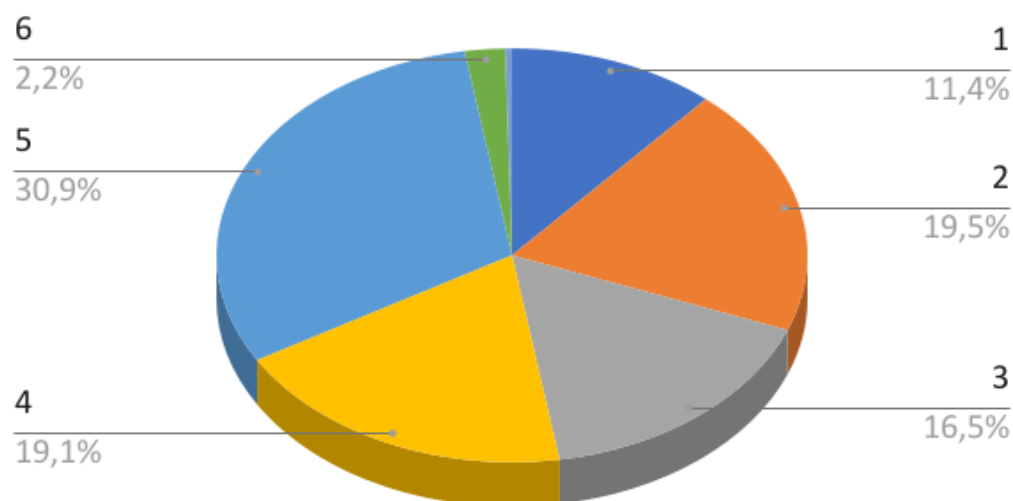


Figura 4.12: Distribución porcentual de las pendientes asociadas a la ocurrencia de MM

4.2.1.4 Precipitación

La precipitación es un dato que presenta diferentes comportamientos a través del tiempo en cada punto seleccionado. Para efectos de un mejor análisis, solo se muestra la información correspondiente a la ventana de tiempo en la que se encuentra registrado el día cero o día en el que ocurrió el MM para cada punto seleccionado (Figura 3.4). La Figura 4.13 muestra el comportamiento de la precipitación durante tres meses en cada zona donde existe un

registro de un MM. En esta, se aprecian todas las series temporales que representan el comportamiento de la precipitación, donde en el día 93 para cada serie de tiempo, indica la precipitación del día en el que ocurrió el MM.

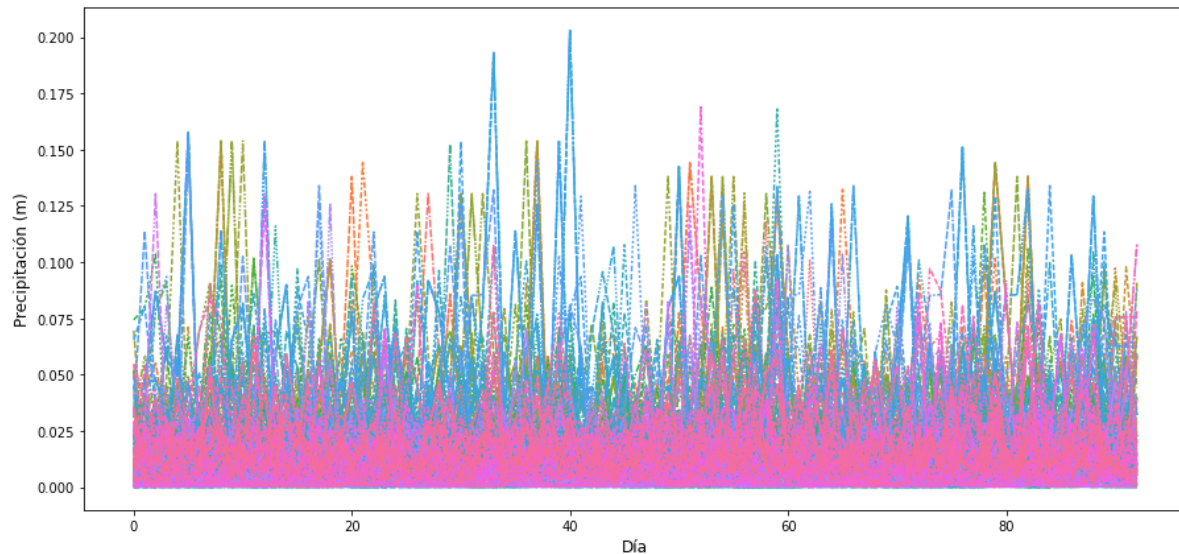


Figura 4.13: Precipitación a tres meses de la ocurrencia de un MM (datos de entrenamiento)

Nota 1. La figura se observa saturada de información debido a la cantidad de datos de precipitación usados y su respectiva granularidad temporal.

En la anterior figura se puede apreciar que si bien, existen picos altos de precipitación previos a la ocurrencia de un MM, estos no ocurren necesariamente un día con fuertes lluvias. La Figura 4.14 permite observar la media de las precipitaciones por día, evidenciando el comportamiento analizado en la Figura 4.13. No obstante, conociendo la media de la precipitación diaria, se puede inferir que el día en que ocurrieron los MM no es necesariamente cuando mayor precipitación hubo. A partir del día 65 (28 días antes de la ocurrencia del fenómeno) se visualizan picos de lluvias que evidencian incrementos importantes en el volumen de las precipitaciones. Sin embargo, al observar los picos presentados en los días previos al evento, se sugiere que la acumulación de agua satura el terreno, originando inestabilidad en el suelo, lo que genera la condición detonante idónea para la ocurrencia de un MM.

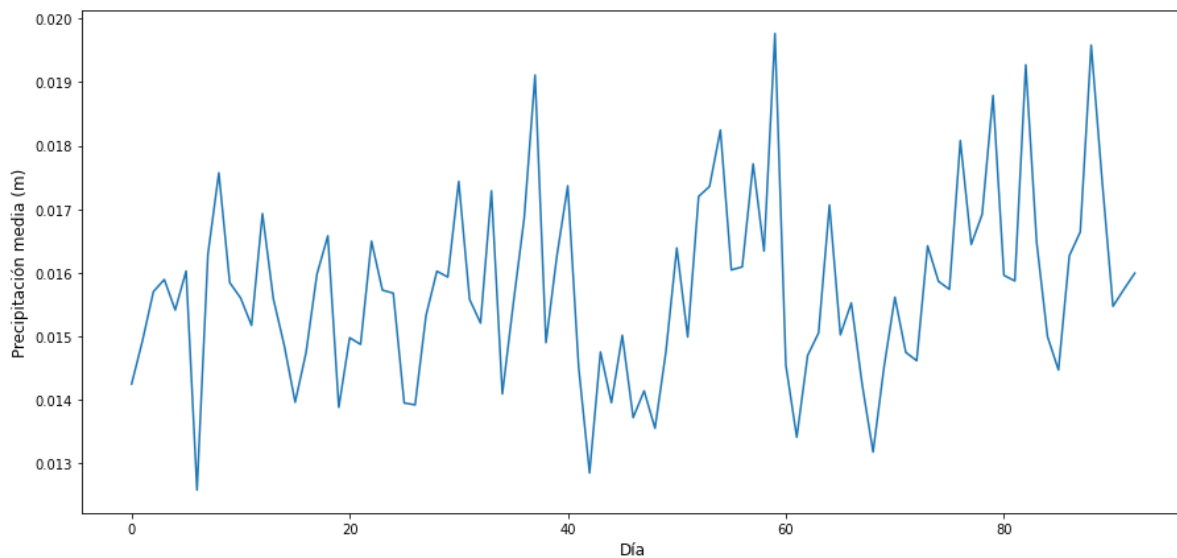


Figura 4.14: Precipitación media a tres meses de la ocurrencia de un MM (datos de entrenamiento)

Finalmente se presenta de manera gráfica (figura 4.15) la distribución espacial de la precipitación para un solo punto de entrenamiento debido a que esta variable por ser temporal no permite la visualización de todos los puntos de entrenamiento con el total de fechas utilizadas. Sin embargo, esta gráfica muestra la precipitación acumulada durante los 4 días anteriores al MM que se presentó el 7 de noviembre del 2017 en Corinto - Cauca, mostrando que aunque existe una pequeña variabilidad acumulativa, esta es casi imperceptible de manera visual debido a que la granularidad de la variable es diaria.

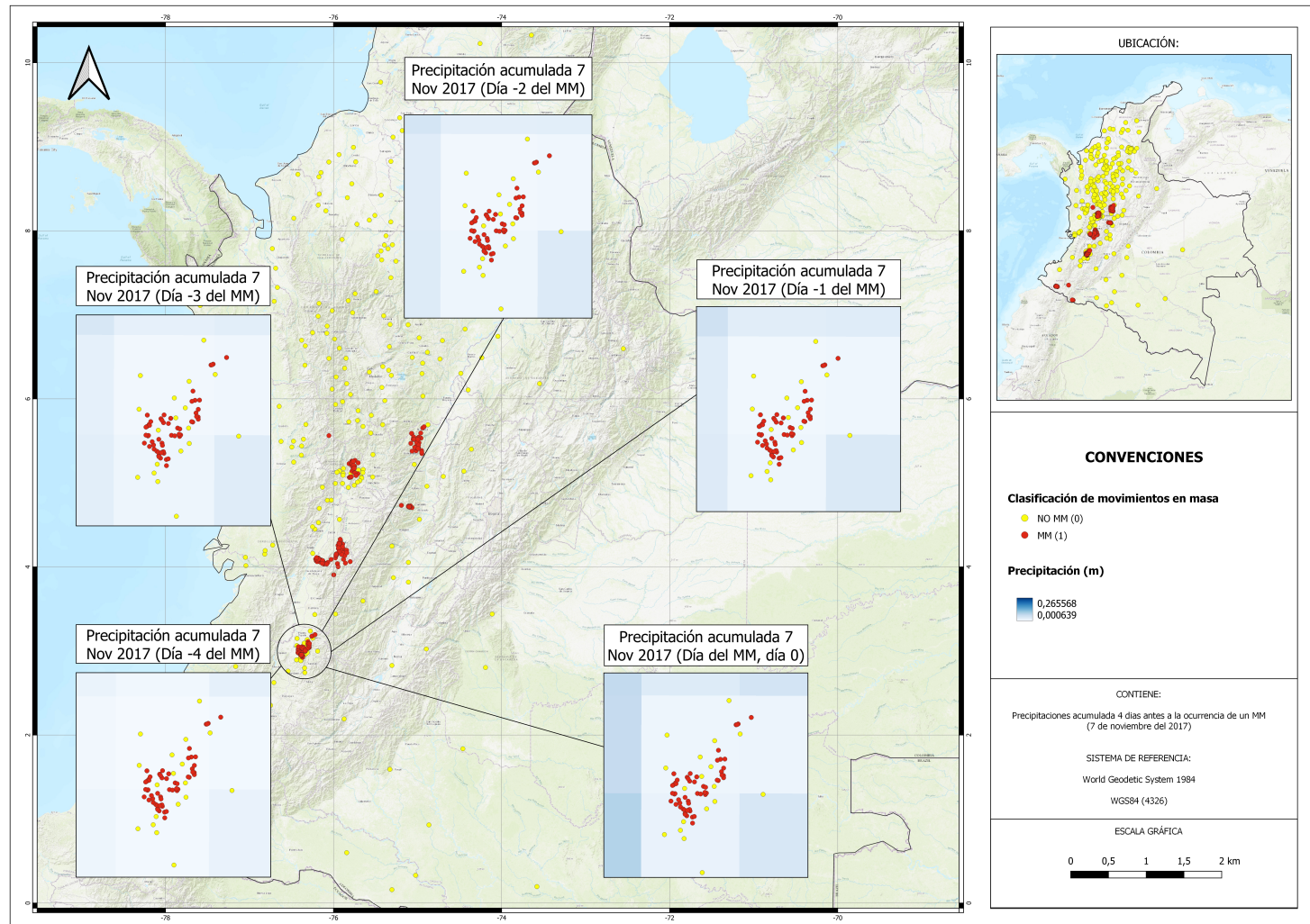
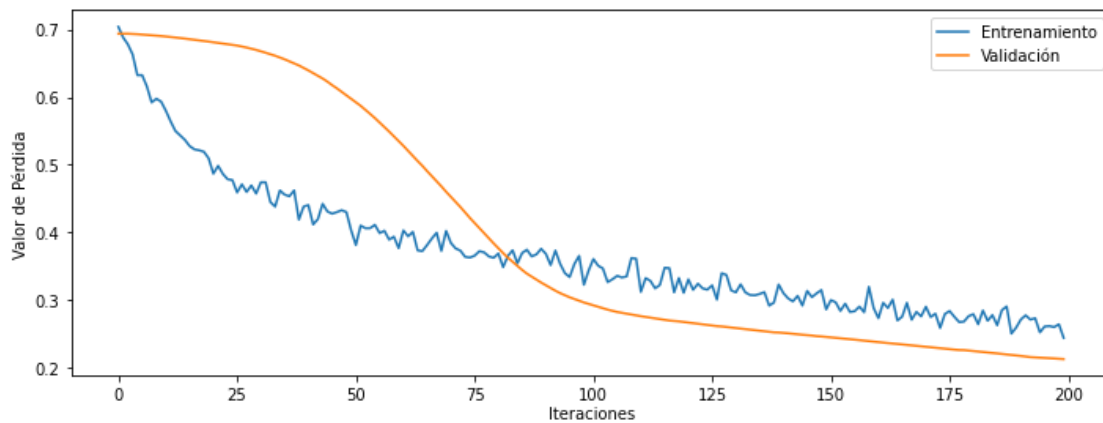


Figura 4.15: Distribución espacial de la precipitación acumulada 4 días previos a la ocurrencia del MM registrado el 7 de noviembre del 2017 en Corinto - Cauca

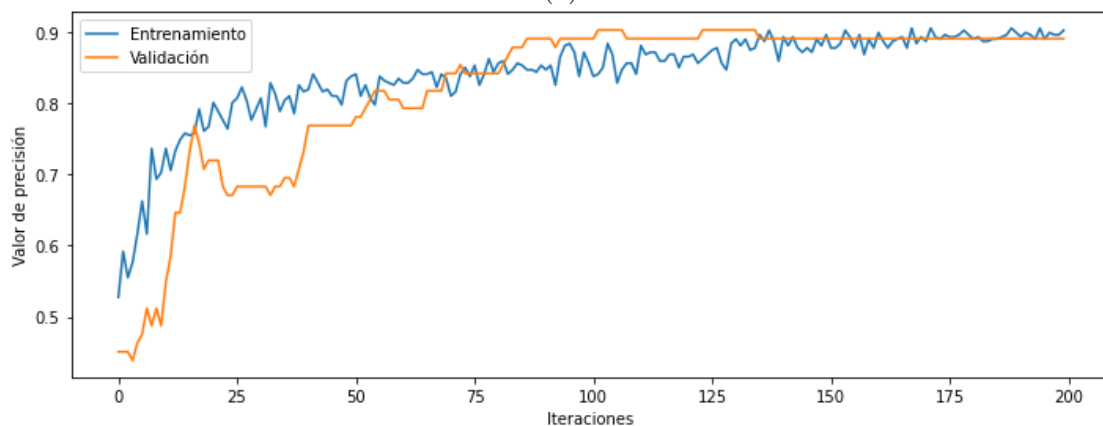
Nota 1: Se presentan todos los puntos de entrenamiento con el fin de mostrar la distribución espacial que tienen estos sobre los píxeles de la precipitación

4.2.2 Procesamiento de datos

La red desarrollada en el Capítulo 3 presentó un comportamiento óptimo tanto en su aprendizaje como en su validación. Como se observa en la Figura 4.16 el modelo tuvo un buen ajuste teniendo en cuenta que la función de costo disminuye en cada iteración y aunque se aprecian pequeños sobreajustes por el ruido que se genera en la curva de aprendizaje (Figura 4.16a), esto se debe a los batches (tramos de datos) en los que se procesa la información. Además, el comportamiento que presenta la curva que grafica la precisión del modelo es ascendente en la mayoría de iteraciones, lo cual permite deducir que el modelo se entrenó con un porcentaje de precisión por encima del 80 % (Figura 4.16b).



(a)



(b)

Figura 4.16: Curva de aprendizaje (a) y precisión del modelo (b)

4.2.3 Validación

El análisis estadístico realizado al desempeño de la red con los datos de validación arroja un AUC igual a 0.98 (Figura 4.17). Esto nos permite inferir que existe una buena correlación entre la sensibilidad y la especificidad de acuerdo con la curva que se grafica entre la tasa de verdaderos positivos y falsos positivos respectivamente.

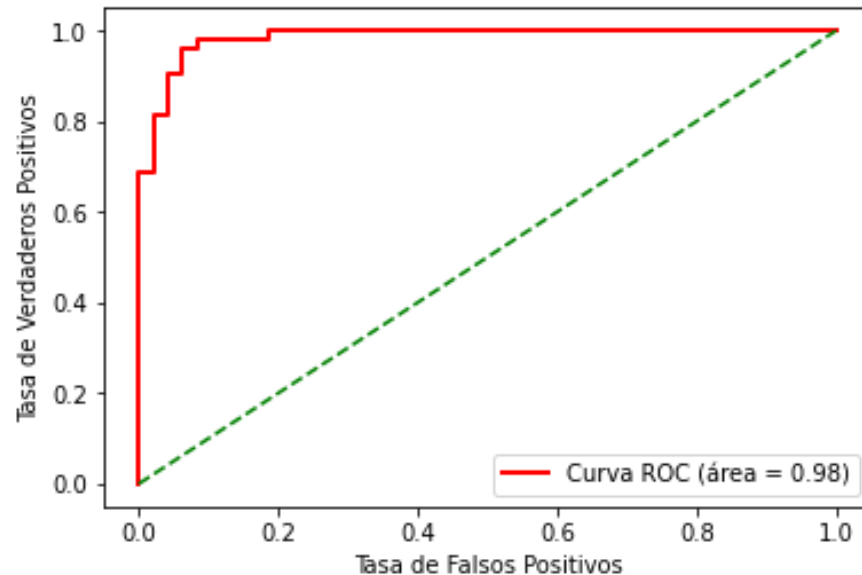


Figura 4.17: Área bajo la curva ROC (AUC)

A partir de la anterior figura surge la necesidad de conocer la cantidad de datos clasificados como verdaderos positivos o falsos positivos. Para dar solución a dicha incógnita se genera la matriz de confusión que se muestra en la Figura 4.18 con su respectivo reporte de clasificación (Tabla 4.1). Esto permite identificar varias cosas importantes: inicialmente se puede afirmar que la precisión o capacidad del clasificador para etiquetar correctamente la información es del 90 % para “No MM” y del 96 % para “MM” con una media armónica ponderada de la precisión y la exhaustividad (f1-score) del 93 %. Además, se puede observar que existen 5 falsos negativos y 2 falsos positivos en las predicciones, es decir que se debe prestar mucha atención a las 5 zonas clasificadas como falsos negativos puesto que no se generarán las alertas pertinentes frente a la ocurrencia de un MM en zonas que tengan variables iguales a los casos clasificados como “No MM” que debían ser “MM”.

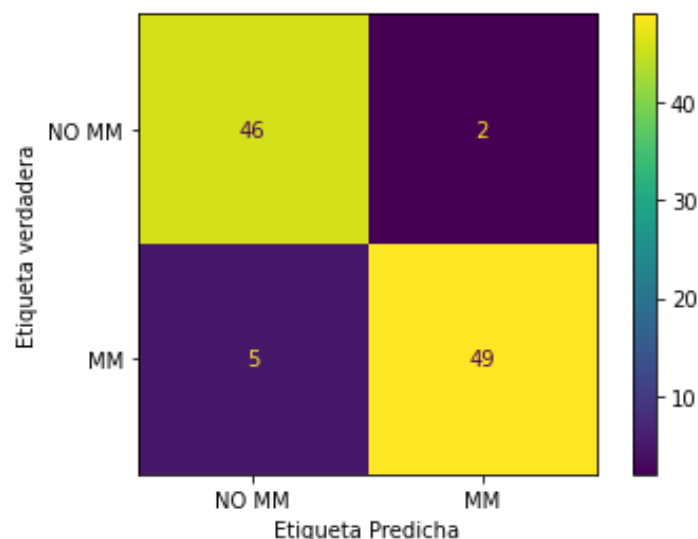


Figura 4.18: Matriz de confusión (Validación del modelo)

	precision	recall	f1-score	support
No MM	0.90	0.96	0.93	48
MM	0.96	0.91	0.93	54
accuracy			0.93	102
macro avg	0.93	0.93	0.93	102
weighted avg	0.93	0.93	0.93	102

Tabla 4.1: Reporte de clasificación

Teniendo en cuenta lo mencionado, es pertinente hacer un análisis adicional frente al etiquetado de los resultados que arroja el modelo como consecuencia de los 5 datos clasificados como falsos positivos por la matriz de confusión. Es importante aclarar que la matriz de confusión presentada en la Figura 4.18 asocia todos los valores inferiores a 5 como “No MM” y los valores mayores o iguales a 5 como un MM. En este sentido la Figura 4.19a se generó modificando los rangos de etiquetado. En esta los MM son todos los valores superiores o iguales a 0.43 y los inferiores a dicho valor son los que no presentan MM. Por su parte, la Figura 4.19b representa el escenario 2, donde los valores superiores o iguales a 0.24 son MM.

La modificación de los rangos de etiquetado con base a los parámetros propuestos anteriormente, permite mejorar el problema relacionado con los falsos positivos que identifica la matriz de confusión como se observa en la Figura 4.19. No obstante, es necesario realizar un ajuste que tenga en cuenta tanto los falsos positivos como los falsos

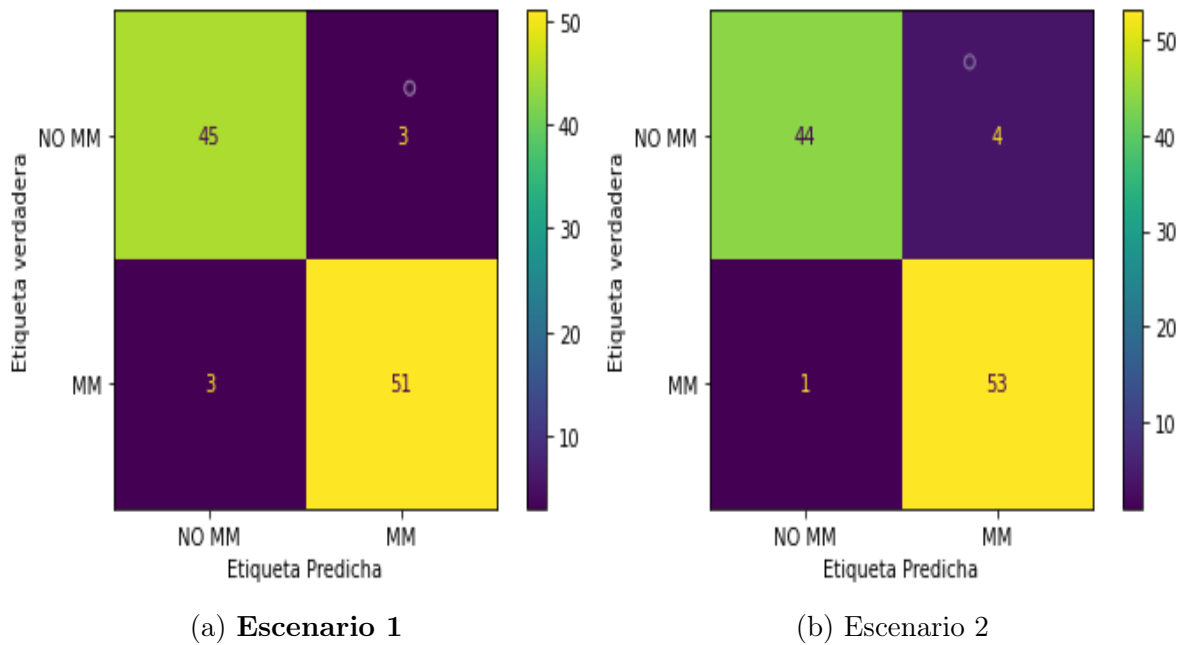


Figura 4.19: Matrices de confusión en diferentes escenarios de etiquetado

negativos buscando un balance en los resultados del modelo a partir de la curva AUC (Figura 4.17). Los rangos para asignar las etiquetas de “MM” y “No MM” se deben definir según la experiencia de un experto en gestión del riesgo. Para efectos de este trabajo se definen los siguientes escenarios:

■ Escenario 1

En este escenario los “No MM” están representados por los valores de predicción inferiores a 0.24. Los valores que están entre 0.24 y 0.64 se definen como valores inciertos, es decir que no se puede determinar la ocurrencia o no de un “MM”. Sin embargo, este genera una alerta y finalmente, todos los valores superiores o iguales a 0.64 se clasifican como “MM” (Figura 4.20).

■ Escenario 2

Para el segundo escenario, se configuró una escala similar en donde los “No MM” están definidos por los valores de predicción inferiores a 0.4. Los valores inciertos son aquellos iguales o mayores a 0.4 y menores a 0.5. Finalmente para identificar los “MM” se eligen los valores iguales o superiores a 0.5 (Figura 4.21).

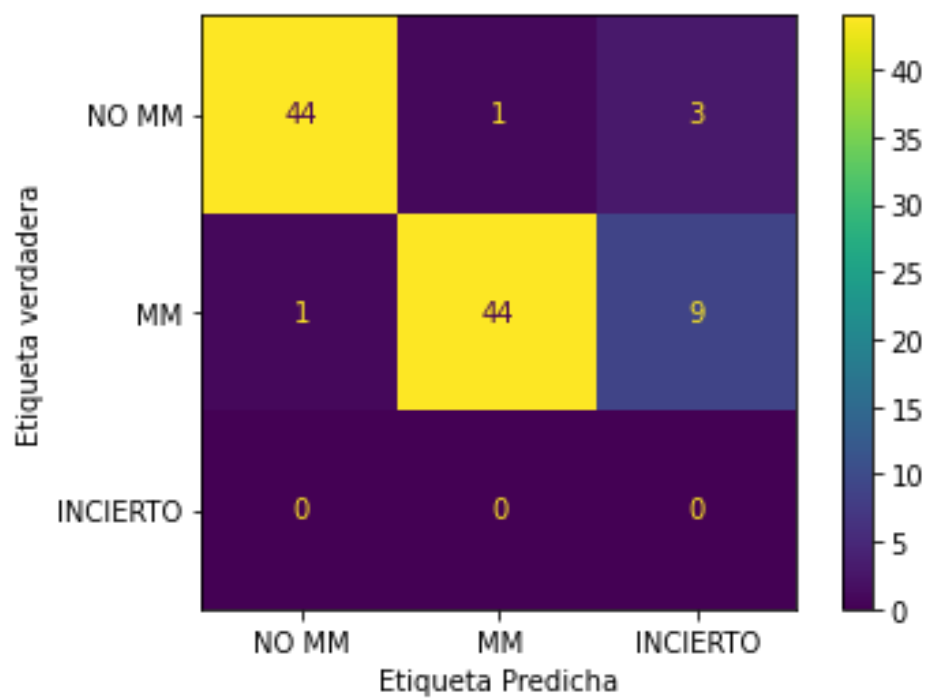


Figura 4.20: Matriz de confusión de los parámetros elegidos (Escenario 1)

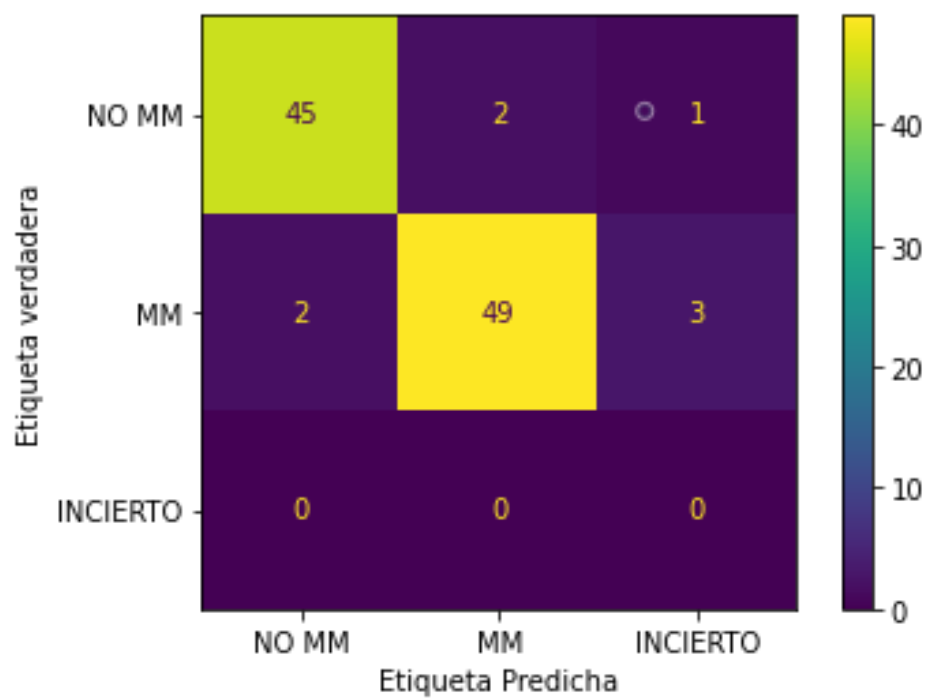


Figura 4.21: Matriz de confusión de los parámetros elegidos (Escenario 2)

Con base en los escenarios descritos anteriormente se infiere que el escenario 1 (Figura 4.20) presenta un buen ajuste estadístico frente a los falsos negativos y falsos positivos. Pero al analizar la cantidad de datos inciertos se observa que se clasificaron 12 puntos en dicha categoría. Si bien reducir los falsos negativos es de suma importancia, tener una decisión ambigua en aproximadamente el 11.76 % de los datos, incrementaría la necesidad de analizar dichos resultados después del procesamiento con el fin de generar las alarmas pertinentes dado que la red no los generó automáticamente.

Por otra parte, aunque en el escenario 2 (Figura 4.21) se identifica una mayor cantidad de falsos negativos, la red tiene una mejor capacidad de decisión frente a la ocurrencia o no de un MM dado que los datos inciertos pasaron del 11.76 % al 3.92 %. Puesto que este escenario reduce la cantidad de resultados ambiguos, se define como el adecuado para la clasificación de las zonas propensas a MM. Sin embargo, es necesario analizar los resultados cartográficos que se obtengan, con el fin de tomar decisiones frente a la adecuada gestión del riesgo desde el punto de vista de los expertos en este campo.

4.3 Fase 3

Bajo el análisis realizado a los resultados de los escenarios del modelo propuesto anteriormente, se realiza la clasificación de 3 fechas diferentes obteniendo sus respectivas predicciones con las 3 clases definidas (MM, No MM e Incierto).

La Figura 4.22 presenta la predicción realizada para un MM registrado en la base de datos del SIMMA. Según el resultado obtenido, se tiene que toda la cuenca es propensa a MM según la clasificación definida. Esta predicción nos permite inferir que la red generaliza sus resultados, dado que los datos de entrenamiento y predicción del modelo tienen una escala espacial que no permiten un pronóstico a detalle. Por esta razón, se observa que hacia el noroeste de la cuenca, es decir la parte baja de esta, donde se ubican los cascos urbanos de Puerto Tejada, Yarumales, Padilla, Huasano, Miranda y Corinto, se clasificaron como propensas a MM. No obstante, basados en el marco teórico (Capítulo 2) es poco probable que en zonas con pendientes bajas ocurra un MM, aunque son propensas a ser afectadas por avenidas torrenciales que contengan material de MM ocurridos en la parte alta de la cuenca.

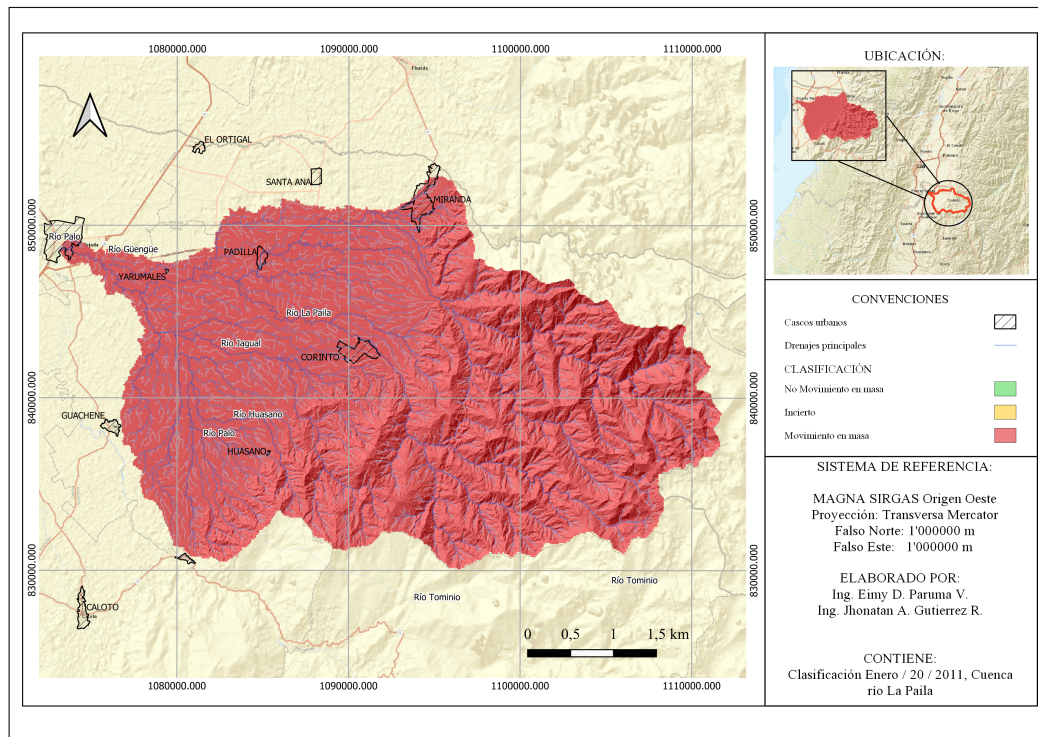


Figura 4.22: Clasificación Enero / 20 / 2011, Cuenca del río La Paila

Nota: Esta salida cartográfica se presenta en el anexo 6.1 para una mejor visualización

Otro escenario de predicción generado es el que se muestra en la Figura 4.23. En esta se presenta el pronóstico calculado desde el 29 de octubre del 2017 para la avenida torrencial que ocurrió el 7 de noviembre del 2017 mencionada en el capítulo 1. En esta figura se observa un escenario diferente al anterior en donde el modelo divide la cuenca en 2 partes debido al tamaño del píxel de la precipitación, obteniendo que la parte alta de la cuenca es propensa a la ocurrencia de un movimiento en masa que puede generar la ocurrencia de una avenida torrencial si además de la pendiente se suman altas precipitaciones, pero en la otra mitad, es decir en la zona plana donde se encuentran los cascos urbanos no presenta propensión a un MM debido a la disminución de la pendiente; no obstante, la zona con pendiente baja puede sufrir afecciones por avenidas torrenciales, como lo ocurrido el 7 de noviembre en Corinto a raíz de fuertes lluvias en la parte alta de la cuenca. Con lo anterior se puede inferir que si el día en el que ocurre un MM hay fuertes precipitaciones o estas ocurren días posteriores antes de la compactación del material particulado, la ocurrencia de una avenida torrencial es muy probable, afectando poblaciones que se encuentren en las zonas bajas, en este caso se podrían afectar cascos urbanos como Puerto Tejada, Yarumales, Padilla, Huasano, Miranda y Corinto.

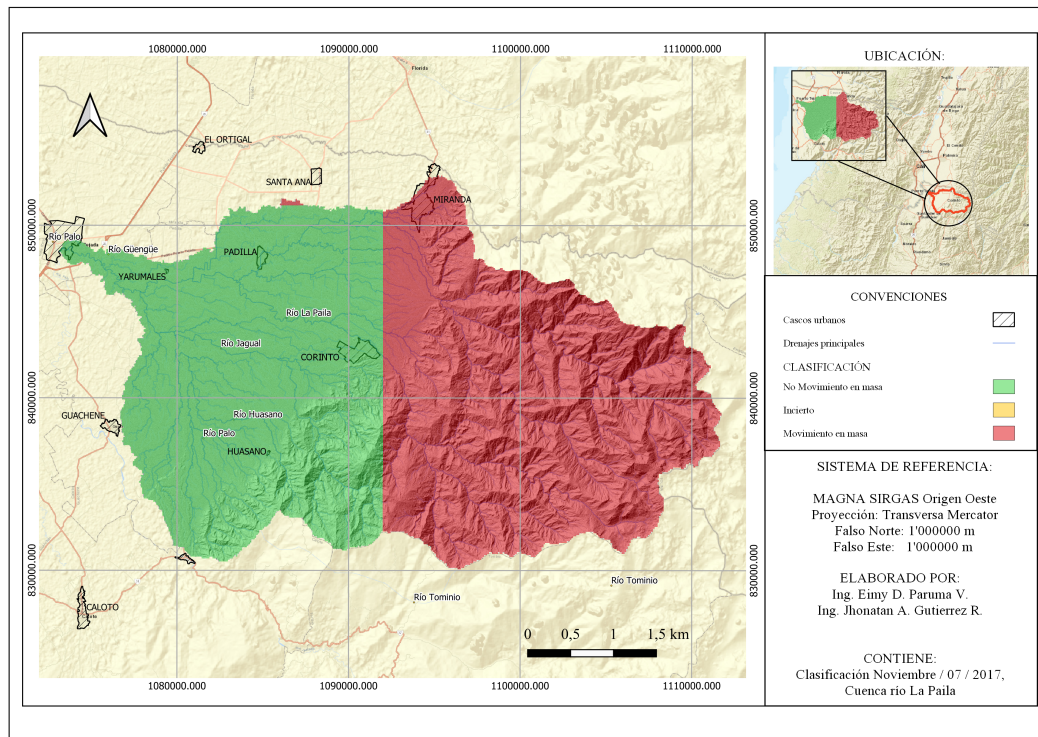


Figura 4.23: Clasificación Noviembre / 07 / 2017, Cuenca del río La Paila
 Nota: Esta salida cartográfica se presenta en el anexo 6.2 para una mejor visualización

Por otro lado, se presenta la Figura 4.24 que predice el comportamiento de la cuenca para el 19 de julio del 2020 partiendo de la ultima fecha que el ERA5 registra información en GEE (9 de julio del 2020). Esta presenta un comportamiento diferente a los escenarios anteriores al dividir la cuenca en 2 partes (zona este y zona oeste), siendo un poco más acorde a la realidad. Debido a las pendientes, el tipo de geología, los usos de suelo y las precipitaciones propias de una zona montañosa (zona Este de la cuenca), se infiere que hay una mayor tendencia a la ocurrencia de un MM desde el punto de vista geológico.

De acuerdo con lo anterior, la predicción del modelo va medianamente acorde al punto de vista geológico, puesto que predice la zona montañosa como propensa a presentar MM. La división marcada entre la zona Este, Oeste y una pequeña parte hacia el Norte se genera de acuerdo al tamaño del píxel o resolución que tienen las precipitaciones. Esto permite deducir que la precipitación como factor detonante tiene un gran peso según la red, el cual se produce al momento de realizar el respectivo ajuste que la red neuronal hace a los pesos sinápticos en las neuronas mediante el gradiente descendente que se ajusta con la función de pérdida. Por lo tanto, la suma de pendientes altas y condiciones específicas de

la precipitación, el uso del suelo y la geología marcan las clases definidas por la red para el 19 de julio del 2020.

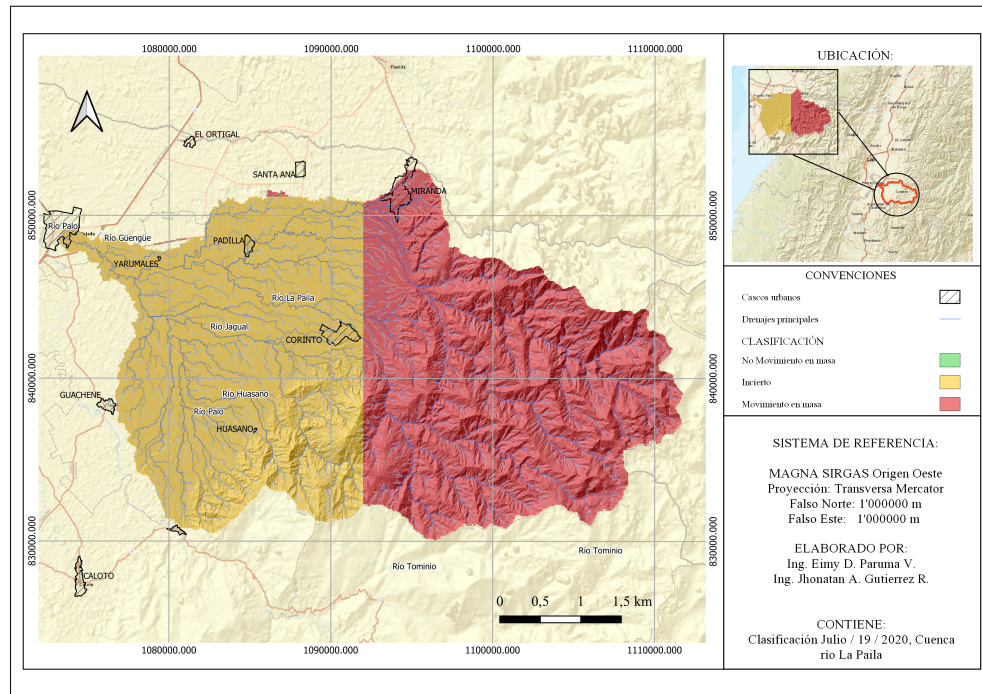


Figura 4.24: Clasificación Julio / 19 / 2020, Cuenca del río La Paila

Nota: Esta salida cartográfica se presenta en el anexo 6.3 para una mejor visualización

Con base al resultado cartográfico obtenido para esta fecha y partiendo de los resultados estudiados en los anteriores escenarios (figuras 4.22 y 4.23), se deduce que aunque la red obtuvo buenos resultados tanto en el entrenamiento como en su validación, la escala cartográfica de los datos llevan a que esta generalice sus predicciones. Esto se evidencia con las predicciones clasificadas como inciertas dado que geológicamente hablando, la zona noreste y parte de la sureste de la cuenca cuentan con pendientes bajas. Por lo cual, con base a los resultados cartográficos obtenidos y la literatura consultada, es poco probable la ocurrencia de un MM en dichas zonas. Por otra parte, la zona sur de las cabeceras urbanas de Corinto y Huasano, posiblemente exista una incertidumbre frente a la ocurrencia o no de un MM puesto que es un terreno con pendientes más pronunciadas. Finalmente, revisando los resultados calculados para la zona Este de la cuenca, se observa que si bien las pendientes tomaron gran relevancia, la línea que marca el píxel de la precipitación posiblemente influyó en que la zona con pendientes bajas ubicadas al sur de Miranda fueran catalogadas como propensas a MM a pesar de que estas no cuentan con las características típicas de una zona propensa a MM.

Es importante mencionar que aunque el modelo tuvo una confiabilidad del 98 % según el AUC y un *f1-score* del 93 %, la precisión espacial de los resultados es baja debido a la escala de los datos de entrenamiento. Por esto y de acuerdo a lo observado para los eventos seleccionados en las fechas analizadas, se presenta la Figura 4.25 como propuesta cartográfica producida a partir del análisis realizado al mapa que contiene la clasificación del 19 de julio del 2020 (Figura 4.24) y la pendiente del terreno. En este orden de ideas, se definieron las zonas con pendiente baja como “No MM”, las que presenta cambios considerables en la pendiente pero fueron clasificadas como “No MM” se reclasificaron como “Incierto” y las zonas que fueron clasificadas con posibilidad de “MM” se mantuvieron en dicha clase, excepto las zonas que presentan poca variación en su pendiente ya que se reclasificaron como “No MM”.

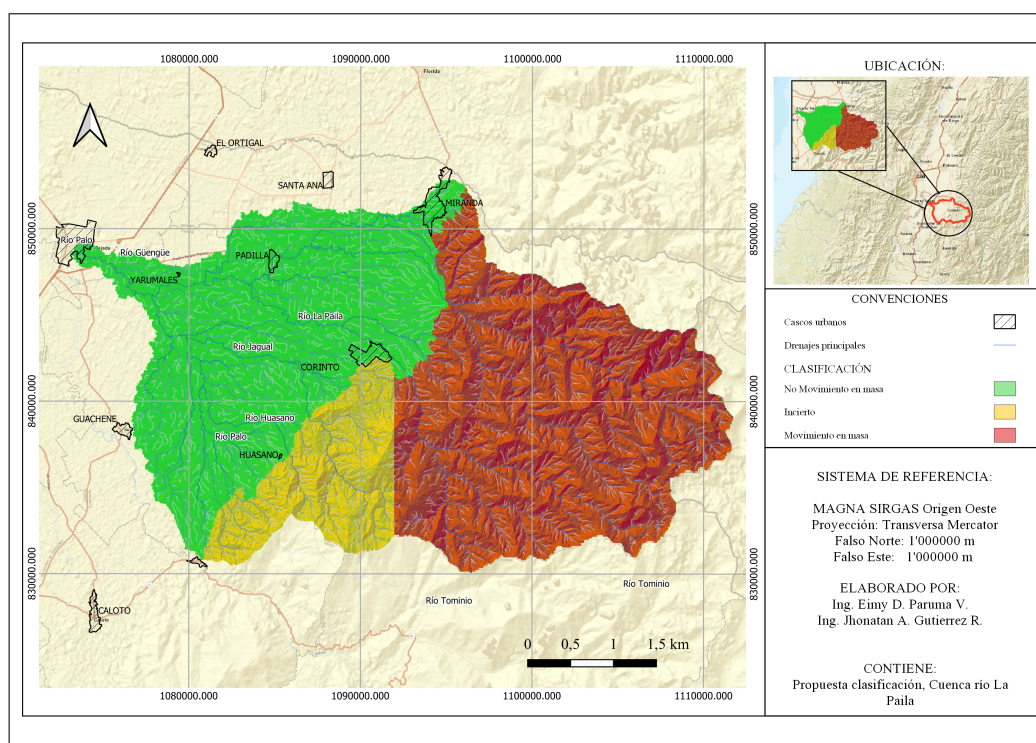


Figura 4.25: Zonificación propuesta para cuenca del río La Paila

Nota: Esta salida cartográfica se presenta en el anexo 6.4 para una mejor visualización

4.4 Síntesis

En términos generales, la metodología planteada mostró a través de la cartografía generada la clasificación de zonas propensas a MM que pueden ser fuente de materiales para avenidas torrenciales con diez días de anticipación. Es importante resaltar que dichos resultados presentan una generalización del fenómeno dada la escala espacial de los datos con la cual fue entrenado el modelo de red neuronal. No obstante, es posible identificar la coherencia de los resultados de cada clasificación presentada teniendo en cuenta la cantidad de datos utilizados y la correlación entre la variables implementadas; esta última se evidencia con las curvas de aprendizaje y precisión del modelo (Figura 4.16) construido para zonas con condiciones morfoclimáticas similares a las de la cuenca del río La Paila. Permitiendo así esbozar un método generador de alertas tempranas del fenómeno analizado que puede mejorar su precisión usando datos de entrenamiento que presenten una mejor resolución espacial.

5. Conclusiones y Trabajos Futuros

En este capítulo se presentan las principales conclusiones del presente trabajo de grado basados en los resultados obtenidos y su respectiva comparación con la literatura consultada. Además, se presentan los trabajos futuros que se pueden desarrollar a partir de los resultados obtenidos y la experiencia adquirida.

5.1 Conclusiones Generales

La red neuronal desarrollada en el presente trabajo de grado concatena 4 modelos (GRU y ANN) de aprendizaje automático. Esta procesa información captada por sensores remotos e información espacial georeferenciada para la clasificación de zonas propensas a movimientos en masa en la cuenca del río La Paila. La red se entrenó teniendo en cuenta los puntos que representan la ocurrencia histórica de MM registrados en la base de datos del SIMMA. Es importante resaltar que para seleccionar los puntos de entrenamiento se garantizó que estuvieran en municipios con condiciones similares a las del área de estudio.

La arquitectura del modelo desarrollado no tiene precedentes a nivel nacional en estudios enfocados para el análisis de fenómenos que presentan variables importantes con un componente temporal como la precipitación para los MM. La GRU es una arquitectura que permite procesar óptimamente el comportamiento de los datos de la precipitación, puesto que presentan patrones de comportamiento similares cada año y facilita su estudio a través de series de tiempo. En una GRU las series temporales de la precipitación se procesan de manera eficiente debido a que el algoritmo logra identificar variaciones inusuales en dicha variable, registrándolas como señales de disparo. En otras palabras, el modelo puede identificar las variaciones inusuales que existen en la precipitación días previos a la ocurrencia de un MM según la tendencia que esta presente para la fecha analizada.

Los resultados obtenidos a partir del área bajo la curva AUC (0.98) calculada y el *f1-score* medido por la matriz de confusión (0.93), permiten concluir que el modelo puede predecir de manera eficiente las zonas propensas a presentar MM. No obstante, los

resultados cartográficos muestran que la resolución espacial de la precipitación y la escala cartográfica de las variables estáticas que actualmente se encuentran disponibles (en escala nacional), generan errores asociados a la precisión espacial con la cual se obtienen los resultados, llevando a la generalización de patrones de comportamiento en el área de estudio. Es decir, que aunque la red desarrollada muestra una alta correlación entre las variables analizadas, es importante tener insumos con **mayor resolución espacial** para que los resultados cartográficos permitan observar con mayor precisión las zonas propensas a MM que pueden ser fuente de materiales para avenidas torrenciales.

Cabe aclarar que aunque existe información sobre la precipitación con mayor resolución espacial de libre acceso, estas carecen de resolución temporal, lo cual en términos de gestión del riesgo generaría incertidumbre temprana en la población expuesta. Por el contrario, si la información cuenta con una resolución temporal diaria, se convierte en información oportuna que disminuye la dificultad para el planteamiento de los planes prevención que pueden desarrollar las entidades territoriales, dotando de capacidad de decisión a la comunidad ante la ocurrencia de un MM. Por tal razón, para el presente trabajo de grado se priorizó la resolución temporal, puesto que esta permite generar las alertas tempranas requeridas en una temporalidad específica. Para este caso la temporalidad se definió con 10 días de anticipación a la ocurrencia del evento.

El análisis cartográfico realizado en la fase 3 del Capítulo 4 permite concluir que tanto la pendiente como la precipitación son los factores que tienen mayor importancia para la red puesto que el modelo les asignó un mayor peso en el cálculo de las probabilidades, el cual se hizo evidente por dos características fundamentales en las clasificaciones: (i) la división marcada por los límites del píxel de la precipitación en la cuenca y (ii) las zonas que fueron clasificadas como propensas a MM mayormente cuentan con terrenos montañosos.

El análisis estadístico realizado para las variables estáticas y la variable temporal, permite deducir que las zonas que cuentan con ciertos factores condicionantes y detonantes similares a la cuenca del río La Paila, tienden a ser propensas a la ocurrencia de un MM. Por ejemplo, los terrenos con la Formación Volcánica (K2-Vm) y usos del suelo de tipo “Protección y Producción” se encuentran comúnmente en zonas propensas a presentar MM. Además, los terrenos montañosos con pendientes mayormente abruptas, son propensos a MM según su tendencia a presentar erosión bajo cubierta vegetal, reptación y deslizamientos. Finalmente, la precipitación como variable temporal no necesariamente detona un MM por fuertes lluvias

el mismo día en el que se presenta el fenómeno, sino que este se puede generar a raíz de **la saturación de agua en el suelo** producida por precipitaciones previas que contribuyen a la inestabilidad del terreno. Es importante resaltar que si un MM ocurre el mismo día en el cual se presentan fuertes lluvias, es altamente probable que se genere una avenida torrencial, debido a la acción conjunta de las altas precipitaciones que saturan el suelo de agua, el material particulado producto del MM y la pendiente.

De manera general, el modelo desarrollado en el presente trabajo de grado sirve como un método para identificar zonas propensas a la ocurrencia de MM y por ende ser fuente de material para las avenidas torrenciales. No obstante, se debe aclarar que el producto cartográfico presenta sesgos debido a la escala espacial de los datos con los cuales fue entrenada la red. Por lo tanto, estos resultados deben ser analizados e interpretados por un profesional en el área para tomar decisiones frente al producto cartográfico, con el fin de disminuir la generalización de datos en la que puede incurrir el modelo al calcular las predicciones y así generar planes de acción adecuados frente a la gestión del riesgo.

5.2 Trabajos Futuros

Con base en el trabajo realizado y las limitaciones encontradas, se presentan una serie de trabajos futuros que permitirían mejorar los resultados en investigaciones posteriores:

- Utilizar datos con una mejor resolución espacial para el entrenamiento del modelo que permitan obtener resultados detallados sobre el área de estudio. Esta mejora en el set de datos permitirá generar un insumo con mayor precisión para la adecuada gestión del riesgo.
- Agregar el clasificador AdaBoost al modelo ya entrenado, con el fin ajustar los resultados de este y así mejorar las predicciones finales.

Referencias

- Acosta, J. N., & Upegui, E. S. (2020). Determinación de procesos de remoción en masa utilizando técnicas de inteligencia computacional. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 10(E28), 344–359.
- Alvarez, M. A., Mastrantonio, L., & Moreiras, S. M. (2019). Análisis de susceptibilidad de flujos de detritos en el parque provincial aconcagua, mendoza, argentina. *Revista de la Facultad de Ciencias Agrarias UNCuyo*, 51(2), 177–191.
- Arce, J. I. B. (2019). La matriz de confusión y sus métricas. <https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>. Accessed: 2022-7-3.
- Aristizábal, E., Carmona, M. I. A., & López, I. K. G. (2020). Definición y clasificación de las avenidas torrenciales y su impacto en los andes colombianos. *Cuadernos de Geografía: Revista Colombiana de Geografía*, 29(1), 242–258.
- Aristizábal, E., Vásquez, M., & Ruíz, D. (2019). Métodos estadísticos para la evaluación de la susceptibilidad por movimientos en masa. *TecnoLógicas*, 22(46), 43–64.
- Banco Mundial, B. (2018). China: Reconstruir de manera más firme. <https://www.bancomundial.org/es/results/2018/05/11/building-back-better-plus>.
- BBC Mundo, B. (2020). 10 años del terremoto de haití: 5 cosas que devastaron al país caribeño antes del sismo de 2010. <https://www.bbc.com/mundo/noticias-america-latina-51067830>.
- Bonett, J. J. P. (2013). Uso de la espectroscopía de reflectancia difusa (mir) para la determinación de las propiedades químicas en suelos agrícolas de colombia. *Departamento de Ingeniería Civil y Agrícola*.
- Borga, M., Stoffel, M., Marchi, L., Marra, F., & Jakob, M. (2014). Hydrogeomorphic response to extreme rainfall in headwater systems: flash floods and debris flows. *Journal of Hydrology*, 518, 194–205.

- Caicedo, E. F., Bravo, & López, J. A., Sotelo (2009). *Una aproximación práctica a las redes neuronales artificiales*. Programa Editorial UNIVALLE.
- Cajal, A. (2020). Observación indirecta: características, ventajas, desventajas, ejemplo. <https://www.lifeder.com/observacion-indirecta/>. Accessed: 2022-NA-NA.
- Camacho, G (2003). Memoria explicativa mapa geológico del departamento del Cauca en revisión-01. <http://recordcenter.sgc.gov.co/B4/13010040019624/documento/pdf/0101196241101000.pdf>. Accessed: 2020-9-3.
- Carrara, A., Crosta, G., & Frattini, P. (2008). Comparing models of debris-flow susceptibility in the alpine environment. *Geomorphology*, 94(3), 353–378. GIS technology and models for assessing landslide hazard and risk.
- Castiblanco, C. A. (2020). Efectos del desplazamiento forzado sobre el acceso a la educación en colombia. *Revista de investigación, desarrollo e innovación*, 10(2).
- Cañadas, R. (2021). Curvas ROC. <https://abdatum.com/machine-learning/curvas-roc>. Accessed: 2022-7-3.
- Clark, H. (2012). La importancia de reducir el riesgo de desastres para fortalecer las naciones. <https://www.undp.org/content/undp/es/home/ourperspective/ourperspectivearticles/2012/08/15/building-resilience-the-importance-of-disaster-risk-reduction.html>.
- Concejo Municipal de Corinto, C. (2016). Plan de desarrollo municipal 2016-2019: "nuevas ideas para la paz". *Alcaldía Municipal de Corinto*.
- Corpoamazonia (2017). Evaluación ambiental preliminar de la emergencia generada por avenidas torrenciales de los ríos sangoyaco y mulato y las quebradas taruca, conejo y almorzadero. <https://onx.la/e1417>. Accessed: 2020-3-9.
- Cruden, D. M., & Varnes, D. J. (1996). Landslides: investigation and mitigation. chapter 3-landslide types and processes. *Transportation research board special report*.
- De Gouvello, C. (2010). Emergency program for solar power generation and lighting for haiti, as a consequence of the earthquake in port au prince. *The World Bank*.

- Escuela de Administración Pública, nacional de gestión de riesgo de desastre, S., & Amunorca (2017). Estrategia municipal de respuesta a emergencias del municipio de corinto cauca. *UNGRD*.
- García, P., Montellano, M., A Quiroz, S., Sour, F., Ceballos, S., & Chávez, L. (2002). *Paleobiología. Lecturas seleccionadas*. Universidad Nacional Autónoma de México.
- GEE (2020). Nasadem: Nasa nasadem digital elevation 30m nbsp;|nbsp; earth engine data catalog nbsp;|nbsp; google developers. https://developers.google.com/earth-engine/datasets/catalog/NASA_NASADEM_HGT_001#description.
- GEE (2022). ERA5 daily aggregates - latest climate reanalysis produced by ECMWF / copernicus climate change service. https://developers.google.com/earth-engine/datasets/catalog/ECMWF_ERA5_DAILY.
- Gómez, J., & Montes, N. (2020). Mapa geológico de colombia en relieve. escala 1:1 000 000. Tech. rep., Servicio Geológico Colombiano.
- Goyes, P., Peñafiel, & Hernandez, A., Rojas (2021). Doble evaluación de la susceptibilidad por movimientos en masa basada en redes neuronales artificiales y pesos de evidencia. *Boletín de Geología*, 43(1), 173–191.
- Gómez, J., & Montes, N. (2020). Mapa geológico de colmbia. https://www2.sgc.gov.co/MGC/Paginas/mgc_1M2020.aspx.
- IDEAM (2013). *Formato Común de Hoja Metodológica de Indicadores Ambientales*.
- IGAC (2016). Corinto, uno de los municipios del cauca en los que renacerá la paz. <https://igac.gov.co/noticias/corinto-uno-de-los-municipios-del-cauca-en-los-que-renacera-la-paz>.
- IGAC (2017). Datos abiertos agrología. <https://geoportal.igac.gov.co/contenido/datos-abiertos-agrologia>. Accessed: 2022-1-9.
- IGAC (2019). Lineamientos para el uso de información geográfica en el desarrollo del componente rural de los planes de ordenamiento territorial - POT. Tech. rep., IGAC.

- Kumar, R., & Anbalagan, R. (2015). Remote sensing and gis based artificial neural network system for landslide suceptibility mapping. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, (pp. 4696–4699). IEEE.
- Kumari, R., & Kashish (2020). Python Pandas – método get_dummies(). https://www.geeksforgeeks.org/python-pandas-get_dummies-method/.
- Mendoza, R. J., & Aristizábal, E. (2017). Metodología para la zonificación de la susceptibilidad por movimientos en masa en proyectos lineales. estudio de caso en el acueducto del municipio de fredonia, antioquia. *Ingeniería y ciencia*, 13(26).
- Meneses, F. D., & Alvarado, M. (2017). Pronóstico del tipo de cambio usd/mxn con redes neuronales de retropropagación. *Res. Comput. Sci.*, 139, 97–110.
- Mergili, M., Marchant, S. C., & Moreiras, S. M. (2015). Causes, characteristics and impacts of landslide processes in contrasting areas of the andean region. *Cuadernos de Geografía: Revista Colombiana de Geografía*, 24(2), 113–131.
- Ministerio de Ambiente y Desarrollo Sostenible, M. (2020). Gestión del riesgo de desastres. <https://www.minambiente.gov.co/index.php/gestion-del-riesgo-de-desastres>.
- Mitchell, T. M., et al. (1997). Machine learning. 1997. *Burr Ridge, IL: McGraw Hill*, 45(37), 870–877.
- Mizutori, M. (2020). Lecciones de la pandemia para la cooperación internacional. [urlhttps://www.undrr.org/es/news/lecciones-de-la-pandemia-para-la-cooperacion-internacional](https://www.undrr.org/es/news/lecciones-de-la-pandemia-para-la-cooperacion-internacional).
- Montero, O. J. (2017). *Clasificación de movimiento en masa y su distribución en terrenos geológicos de Colombia*. Guías técnicas y métodos de trabajo en geociencias y asuntos nucleares. Libros del Servicio Geológico Colombiano.
- Muñoz, E., Sánchez, O., & Hemandez, F. (2020). Methodology for estimating landslides susceptibility using artificial neural networks. In *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)*, (pp. 104–109). IEEE.
- NASA JPL (2020). Nasadem merged dem global 1 arc second v001 [data set]. https://lpdaac.usgs.gov/products/nasadem_hgtv001/.

- Ospina, G. J. P., & Aristizábal, E. (2021). Aplicación de inteligencia artificial y técnicas de aprendizaje automático para la evaluación de la susceptibilidad por movimientos en masa. *Revista Mexicana De Ciencias Geológicas*, 38(1), 43–54.
- Oviedo, L. G. (2019). *Geología, ciencia y cultura*. Servicio Geológico Colombiano.
- O'brien, J., & Julien, P. (1985). Physical properties and mechanics of hyperconcentrated sediment flows. *Proc. ASCE HD Delineation of landslides, flash flood and debris flow Hazards*.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Prado, G. V. (2022). *Caracterización sismotectónica de la superficie lunar asociada a la falla Lee-Lincoln*. Tesis pregrado. Universidad del Valle.
- Ramírez, C. C., Sánchez (2020). Construcción de un sistema de alertas tempranas de deslizamientos basado en un modelo de analítica predictiva e inteligencia artificial, aplicado a las líneas de transmisión operadas por intercolombia sa esp. *Universidad de Antioquia*.
- Ritchie, H. (2019). Natural disasters. *Our World in Data*.
- Ritchie, H., & Roser, M. (2014). Natural disasters. *Our World in Data*.
- Rodríguez, V. A. F., Flórez, Z. A. N., & Siachoque, B. R. F. (2018). Evaluación de tierras para la zonificación con fines agropecuarios a nivel nacional metodología a escala general 1: 100.000. *Unidad de planificación Rural Agropecuaria*.
- Scikit-learn developers (2020). sklearn.preprocessing.MinMaxScaler. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>.
- Semana (2017). Avalancha que aterrorizó a los habitantes de corinto deja cuatro muertos. <https://www.semana.com/nacion/articulo/corinto-cuaca-en-emergencia-por-desbordamiento-del-rio-la-paila/546370/>.

- SGC (2017). *Guía metodológica para la zonificación de amenaza por movimientos en masa ESCALA 1: 25.000*.
- SGC (2022). Sistema de información de movimientos en masa. <https://simma.sgc.gov.co/#/public/basic/>.
- Shewalkar, A. (2019). Performance evaluation of deep neural networks applied to speech recognition: Rnn, lstm and gru. *Journal of Artificial Intelligence and Soft Computing Research*, 9(4), 235–245.
- The pandas development team (2020). Pandas-dev/pandas: Pandas. <https://doi.org/10.5281/zenodo.3509134>.
- UNGRD (2015). Guía de integración de la gestión del riesgo de desastres y el ordenamiento territorial municipal. <http://portal.gestiondelriesgo.gov.co/Documents/Guia-Integracion-Gestion-Riesgo-Ordenamiento-Territorioal-Octubre2015.pdf>.
- UNGRD (2018). Hoy 20 de noviembre: diez años de la erupción del nevado del huila. <http://portal.gestiondelriesgo.gov.co/Paginas/Noticias/2018/Hoy-20-de-noviembre-diez-anos-de-la-erupcion-del-Nevado-del-Huila.aspx>.
- UNISDR, E. (2009). Terminología sobre reducción del riesgo de desastres. *Ginebra, Ginebra, Suiza*.
- UNODC, G. d. C. (2017). Colombia: Monitoreo de territorios afectados por cultivos ilícitos 2016. *Recuperado de https://www.unodc.org/documents/colombia/2017/Julio/Censo_cultivos_oca_2016.pdf*.
- UPRA (2014). https://www.minagricultura.gov.co/Documents/UPRA_Oferta_Institucional.pdf. Accessed: 2022-6-30.
- USGS (2010). M 7.0 - 10 km al SE de léogâne, haití. <https://earthquake.usgs.gov/earthquakes/eventpage/usp000h60h/impact>. Accessed: 2020-3-9.
- Wilches-Chaux, G. (2005). 10. el terremoto, la avalancha y los deslizamientos de la cuenca del río páez, 1994. *DESASTRES*, (p. 121).

- Wilches-Chaux, G. (2011). Los desastres no “naturales”: una propuesta para la recuperación con enfoque de derechos. <https://onx.la/c76d0>. Accessed: 2020-3-9.
- Wilson, J. P., & Gallant, J. C. (2000). *Terrain analysis: principles and applications*. John Wiley & Sons.
- Xu, C., Zhang, W., Yi, Y., & Xu, Q. (2019). Landslide susceptibility mapping using logistic regression model based on information value for the region along china-thailand railway from saraburi to sikhio, thailand. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, (pp. 9650–9653). IEEE.
- Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, *31*(7), 1235–1270.
- Zhang, X.-D. (2020). Machine learning. In *A Matrix Algebra Approach to Artificial Intelligence*, (pp. 223–440). Springer.

6. Anexos

- 6.1 Clasificación Enero / 20 / 2011, Cuenca del río La Paila (Figura 4.22)**
- 6.2 Clasificación Noviembre / 07 / 2017, Cuenca del río La Paila (Figura 4.23)**
- 6.3 Clasificación Julio / 19 / 2020, Cuenca del río La Paila (Figura 4.24)**
- 6.4 Zonificación propuesta para la cuenca del río La Paila (Figura 4.25)**
- 6.5 Modelo de red neuronal con tratamiento de datos (Notebook Colab)**