

Susceptibilidad a deslizamientos mediante un método convencional y Machine Learning (caso de estudio Vía Aloag - Santo Domingo)

Pamela Chávez¹, Gilson Peñaloza¹

¹ *Universidad Central del Ecuador. Facultad de Ingeniería en Geología, Minas, Petróleos, y Ambiental*

Resumen

En el presente documento se analizan algunos de los diversos factores que son susceptibles de provocar situaciones de riesgo y que influyen en la generación de deslizamientos en el caso de estudio vía Aloag-Santo Domingo, se exponen las fases de las metodologías empleadas para obtener mapas de susceptibilidad en función de estos factores útiles para identificar las zonas inestables, una de ellas es la metodología convencional en la cual se emplea el software Scoops 3D, mismo que requiere del ingreso de parámetros específicos que influyen en la generación de estos fenómenos, no obstante, se aplica otra metodología mediante Machine Learning que implica el uso de algoritmos que mediante su entrenamiento y evaluación se logra obtener la predicción en cuanto a la susceptibilidad de áreas a deslizarse. Por tanto, el objetivo de este trabajo es identificar las diferencias entre las dos metodologías, y determinar aquella que presente resultados más próximos a la realidad.

Palabras clave: *Deslizamientos, Machine Learning, Scoops 3D*

1. Introducción

Los deslizamientos hoy en día constituyen un grave problema (Medina y Astudillo, 2009), y uno de los peligros geológicos con mayor frecuencia en la naturaleza (Orejuela, 2020), que desde la antigüedad han afectado a gran parte de la sociedad, debido a ello, se han desarrollado diversos estudios y a la vez implementado diversas técnicas para su evaluación, con la finalidad de valorar la susceptibilidad relacionada a estos, ante la necesidad de realizar alguna actividad antrópica. En Ecuador, la vía Aloag – Santo Domingo, fue construida en los años 60, se trata de una carretera de alta montaña, además de ser una arteria vial principal del país, la cual conecta la región Costa y Sierra, y permite el transporte de diversos productos e intercambio comercial. En los últimos años ha sufrido de proble-

mas de deslaves y desprendimientos de materiales de los taludes de corte. Las características del terreno, la actividad sísmica, así como las condiciones climáticas de la zona han potenciado la ocurrencia de estos fenómenos, sin embargo, como factor detonante principal de los mismos se considera a la intervención antrópica. Por tanto, es de mucha importancia realizar estudios que permitan zonificar las áreas más propensas para realizar una adecuada mitigación de estos peligros geológicos.

El objetivo del presente estudio es determinar la susceptibilidad de deslizamientos en la vía Aloag – Santo Domingo mediante el uso del software especializado Scoops 3D del USGS y Machine Learning, con la finalidad de reconocer las áreas de mayor susceptibilidad a deslizarse, y a la vez mediante los resultados obtenidos se pueda realizar posibles planificaciones dentro del área.

2. Estado del Arte

La susceptibilidad a deslizamientos en la vía Aloag – Santo Domingo, es un tema del cual no existe una alta disponibilidad o diversidad de estudios, se han encontrado en su mayoría publicaciones estrictamente enfocadas al ámbito civil en el mejoramiento de la vía. No obstante, de entre los diversos trabajos, dos principalmente están enfocados en la estabilidad de taludes y en la categorización de zonas susceptibles a deslizamientos, dichos trabajos son los siguientes:

2.1. Aplicaciones de métodos convencionales

Orejuela, I. (2020). Susceptibilidad a deslizamientos en la vía Aloag - Santo Domingo. Unidad de Ordenamiento Territorial, 2(17), 2-10.

- Determinó la susceptibilidad a deslizamientos mediante el uso de la técnica de Lógica Difusa, para ello utilizó diversos insumos como un DEM del cual obtuvo un mapa de pendientes, datos vectoriales como la geología, cobertura, y uso de suelo transformados a ráster, para luego utilizar el análisis de lógica difusa y posteriormente aplicar la combinación lineal ponderada, obteniendo así dos modelos de susceptibilidad, sin embargo, un ajuste a ambos modelos Fuzzy y WLC/AHP, y la combinación de los mismos representa con mayor porcentaje la realidad. Por tanto, es el modelo combinado aquel que presenta mejores resultados y a través del cual logró identificar las áreas de mayor susceptibilidad.

Pachacama, N. (2015). Caracterización de cangahuas mediante ensayos triaxiales no consolidados - no drenados (UU) y consolidado no drenado (CU), aplicación a la estabilidad de taludes. [Ingeniería].

Universidad de las Fuerzas Armadas.

- En su trabajo realizó el análisis del mecanismo de falla a partir de la resistencia al corte de la cangahua mediante ensayos triaxiales, para ello perpetró una recopilación de trabajos previos, así como la toma de muestras representativas que mediante su análisis se obtuvieron los parámetros índices-mecánicos, tales resultados fueron posteriormente utilizados en el software GeoSlope para determinar el plano de falla, como resultado las diversas muestras presentaron pesos específicos diferentes, cohesión y ángulos de fricción distintos, sin embargo, llevó a cabo una comparación con resultados previos, obteniendo similitud. Por tanto, debido al alto grado de similitud, se concluye que los datos obtenidos en este estudio son valores que se pueden adoptar para el diseño de taludes.

2.2. Aplicaciones de ML

Ospina, J. (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, 1(38), 43-54.

- En su trabajo implementó los diferentes métodos de aprendizaje automático para la evaluación de la susceptibilidad por movimientos en masa, con el objetivo de establecer un desarrollo metodológico que considere las ventajas y limitaciones de cada uno de los métodos, para ello en un principio inició con la elaboración de un inventario de 137 movimientos en masa ocurridos en el área de estudio mediante fotointerpretación, seguido realizó el análisis exploratorio de datos y selección de variables, a continuación llevó a cabo la división de los

datos en entrenamiento y test, y a su vez la selección y aplicación de los modelos, posteriormente realizó la validación de los mismos y la optimización de los hiperparámetros mediante Grid Search y Random Search, finalmente mediante la predicción total de los datos logró obtener un mapa tipo semáforo representando la susceptibilidad en cada categoría. Por tanto, de acuerdo a los resultados concluye que los modelos no paramétricos son aquellos que arrojan los mejores resultados basados en su capacidad de adaptarse o aprender de los datos, no obstante, para ello se requiere de una base de datos extensa.

Bravo, E; Fernández, T; Sellers, C; Delgado, J. (2022). Elaboración de mapas de susceptibilidad de deslizamientos en la zona de Cuenca, Ecuador mediante la aplicación de una red neuronal artificial. Simposio Nacional sobre Taludes y Laderas Inestables, 1(2022), 1-11.

- En su trabajo implementa el algoritmo de redes neuronales para la obtención de modelos de susceptibilidad a deslizamientos rotacionales ocurridos en la zona de Cuenca, con el objetivo de obtener mapas de susceptibilidad a deslizamientos, para ello ejecutó una serie de etapas iniciando con la obtención del inventario de deslizamientos ocurridos en la zona de interés, posterior a ello llevó a cabo la selección de 10 factores condicionantes que influyen en la generación de deslizamientos, entre algunos de estos se encuentran la pendientes, aspecto, elevación, entre otros, posteriormente para el entrenamiento y evaluación del modelo realizó una división de 70 % de los datos para entrenamiento y el restante 30 % para la adecuada evaluación, subsiguientemente

realizó la configuración y optimización de los hiperparámetros, para posteriormente implementar el algoritmo de redes neuronales, una vez implementado llevó a cabo la validación del modelo mediante el método del área bajo la curva (AUC), una vez finalizada la ejecución del modelo obtuvo los valores de predicción y mediante estos logró elaborar los mapas de susceptibilidad, estableciendo los 5 niveles de susceptibilidad mediante el uso de los cuantiles en el software QGIS, junto con SAGA y GRASS. De acuerdo a los resultados obtenidos, concluye que los mismos fueron satisfactorios debido a que mostraron un buen grado de ajuste y una capacidad predictiva aceptable con valores promedio de $AUC=0.72$.

La escasez de estudios relacionados a la susceptibilidad a deslizamientos en la vía Aloag – Santo Domingo, nos lleva a reflexionar acerca de la poca importancia que se muestra a la problemática que ha sido casi constante durante muchos años en nuestra área de estudio. No obstante, mediante el presente documento se pretende identificar las zonas susceptibles a deslizamientos mediante la aplicación del software Scoops3D y Machine Learning, con la finalidad de aportar información que de una u otra manera podrá ser utilizada o tomada como base para futuras investigaciones.

3. Metodología

El área de estudio comprende los primeros 15Km de la vía Aloag-Santo Domingo, ubicada entre las provincias de Pichincha y Santo Domingo de los Tsáchilas (Figura 1).



Figura 1: Mapa de ubicación

La metodología planteada para la realización del presente proyecto se encuentra en base al siguiente diagrama de flujo (Figura 2).



Figura 2: Diagrama de flujo de la metodología usada para la identificación de zonas susceptibles a deslizamientos en la vía Aloat-Santo Domingo.

La delimitación del área fue realizada mediante Google EarthPro y ArcGis 10.5. Posteriormente la información base, en este caso el modelo de elevación digital DEM, fue obtenida del portal digital de Open Topography, el mismo que generó un Hillsha-

de. Las coordenadas de estos archivos se encuentran en grados, las cuales mediante SAGA GIS 7.8.2 fueron reproyectadas a UTM WGS 84 Zona 17S. Para este caso de estudio se obtendrá el factor de seguridad en el software SCOOPS3D del Servicio Geológico de Estados Unidos (USGS), el programa utiliza un método tridimensional (3D) de análisis de equilibrio límite de columnas para evaluar la estabilidad de potenciales deslizamientos dentro de un rango de tamaño definido por el usuario. Para cada deslizamiento, Scoops3D evalúa la estabilidad de una superficie de deslizamiento esférica rotacional que abarca celdas del DEM. Proporciona áreas propensas a deslizarse para cada celda en el área, así como los volúmenes asociados. Los archivos de entrada en este programa deben estar en formato ASCII. Para esto se utilizó ArcGis 10.5, Surfer y SAGA GIS 7.8.2 y de esta manera poder seleccionar el que mejor se ajuste. A continuación se define las unidades consideradas: longitud (m), peso unitario (kN/m^3), cohesión (kPa). En el apartado de Condiciones de Superficie, para este caso de estudio se definieron 3 estratos en propiedades del material, luego completamos los parámetros de cohesión, ángulo de fricción interna, peso unitario para cada una de las capas. En análisis de estabilidad, el método utilizado fue el de Bishop simplificado, además se definió el método de caja el cual permite ingresar un criterio preliminar de volumen de deslizamientos, mismo que se determina en base a los movimientos en masa ocurridos en el sector, tanto el máximo como el mínimo y su respectivo porcentaje de tolerancia, posteriormente se define la extensión vertical, su resolución y el radio de incremento, finalmente se verifica que no se encuentre ningún error en los datos ingresados y se corre el programa. Como resultado se obtuvieron 15 archivos, cargando el DEM generado en SURFER, se realiza la clasificación de

categorías mediante el método de Natural Breaks (Jenks) en ArcGis 10.5 y se lo compara con el DEM inicial para poder definir las áreas propensas a deslizamientos a lo largo de la vía.

4. Metodología-Machine Learning

Este método, se realizó siguiendo el diagrama de flujo presentado a continuación.

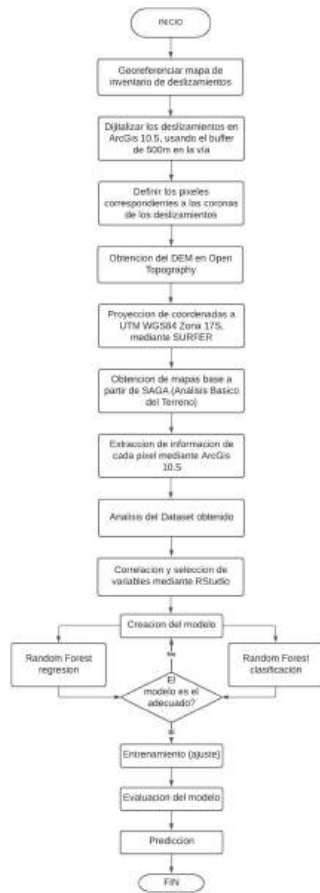


Figura 3: Diagrama de flujo de la metodología usada para la identificación de zonas susceptibles a deslizamientos en la vía Aloag-Santo Domingo.

La información base, fue obtenida a partir de un inventario de deslizamientos (Figura 3)

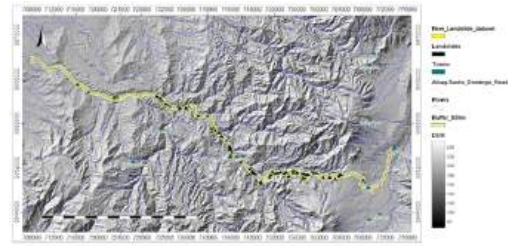


Figura 4: Inventario de deslizamientos

Mediante el software ArcGis 10.5 se digitalizó la vía Aloag- Santo Domingo y se realizó un buffer de 500m, en el mismo se obtuvo un tamaño de píxel de 92 x 92 m, posteriormente se digitalizaron los lugares en los que se ha evidenciado deslizamientos, teniendo un total de 64. Subsiguientemente se elabora un dataset con las variables tanto dependientes como la variable independiente obtenidas mediante la herramienta Basic Terrain Analysis de SAGA, el mismo que permitió generar mapas de los cuales se extrajo la información. A continuación se realizó un pre-procesamiento del Dataset y posteriormente una normalización de las variables (valores entre -1 a 1), en el caso de las variables cuantitativas, y de 0 o 1 en el caso de las variables cualitativas. Una vez armado el Dataset mediante Excel, se guarda este documento en formato .csv para posteriormente ser cargado a OneDrive y poder utilizarlo en Google Colaboratory.

1. Importación del Dataset.

2. Pre-procesamiento del Dataset.

Las variables definidas en el dataset fueron: aspecto, pendiente, curvatura planar, perfil de curvatura, distancia al canal , vegetación 1 , vegetación 2 litología 1 , litología 2 , litología 3 , las cuales son las que representan mayor incidencia en la ocurrencia o no de un deslizamiento en esta área geográfica. Para definir la correlación entre variables se utilizó RStudio, en el mismo se cargó el Dataset y mediante una serie de comando se obtuvieron gráficas de la corre-

lación entre variables, pudiendo descartar aquellas que tenían alta correlación.

3. Creación del modelo Random Forest

Es un método de aprendizaje conjunto para la clasificación, la regresión y tareas que operan mediante la construcción de una variedad de árboles de decisión en el momento del entrenamiento. Este procedimiento de arranque conduce a un mejor rendimiento del modelo porque disminuye la varianza del modelo, sin aumentar el sesgo. Esto significa que mientras que las predicciones de un solo árbol son muy sensibles al ruido en su conjunto de entrenamiento, el promedio de muchos árboles no lo es, siempre que los árboles no estén correlacionados. Simplemente entrenar muchos árboles en un solo conjunto de entrenamiento daría árboles fuertemente correlacionados.

4. Entrenamiento (training) - ajuste

En esta etapa se establecieron los parámetros e hiperparámetros con los que se va a realizar el entrenamiento, el cual consiste en optimizar al modelo, para ello a medida que se va entrenando se utiliza el conjunto de datos de validación para verificar la optimización.

5. Evaluación del Modelo

En este ítem se realiza una evaluación previa, conjuntamente con los datos de entrenamiento y test para conocer el funcionamiento del modelo. Para ello se utiliza la métrica denominada como Accuracy mediante la cual se obtuvo una valoración de 100.

6. Predicción

Mediante el dataset se asigna datos a cada variable, a la vez se utilizan comandos para que los

mismos sean leídos por el modelo. Posteriormente se realiza la normalización de los datos observando que los valores estén dentro de los rangos que se establecieron. Finalmente se aplica el predict llamando al arreglo numpy, obteniendo como resultado un vector con 8429 datos con valores entre cero y uno convirtiendo todo en una matriz numérica, subsiguientemente utilizando numpy y trasladándolo a array, el resultado fue transformado a binarios para verificar la ocurrencia o no del fenómeno.

5. Resultados

La susceptibilidad a deslizamientos en determinada área se puede determinar y describir en base a la zonificación del peligro. Para ello se puede generar un mapa de susceptibilidad de deslizamientos al inicio del estudio de planificación y desarrollarlo en mayor detalle a medida que avanza el mismo. Un mapa de susceptibilidad se puede usar como herramienta para identificar áreas mejor caracterizadas para el desarrollo, examinando el riesgo potencial de los deslizamientos. Aún más, una vez que se identifique la susceptibilidad a los deslizamientos, se pueden desarrollar proyectos de inversión que eviten, prevengan o mitiguen significativamente el riesgo. El caso de estudio (vía Aloag-Santo Domingo), se realizó la evaluación de susceptibilidad en base a dos principales métodos: Convencional (uso de software Scoops 3D) y Machine Learning (usando Random Forest-Regresión). De los que se obtuvieron un total de 6 mapas. Para la representación en el mapa, se debe elegir las clases para su posterior evaluación, en nuestro caso se seleccionó 5 clases basados en la Figura 5.

SMR	Descripción	Estabilidad	FS	Susceptibilidad
≤ 30	Muy Mala	Totalmente inestable	< 1	Muy Alta
31 - 40	Mala	Inestable	1 - 1.1	Alta
41 - 60	Normal	Parcialmente estable	1.2 - 1.5	Media
61 - 80	Buena	Estable	1.5 - 2	Baja
> 80	Muy Buena	Totalmente estable	> 2	Muy Baja

Figura 5: Clases de susceptibilidad y estabilidad. Tomado de: Hernandez et al. (2008) y Minminas (2015)

Se seleccionaron 3 métodos de clasificación de datos en base al dataset y a las variables con las

que se está trabajando, y finalmente se seleccionó el método que más se ajusta al modelo.

5.1. Intervalo igual

Consiste en dividir la amplitud máxima existente entre los datos, entre el número de clases que se haya elegido $(n_f - n_i) / n$. Este método enfatiza la cantidad de un valor de atributo relativo al resto de valores. El método de clasificación de intervalos iguales se usa mejor para conjuntos de datos continuos. Entre los dos mapas obtenidos mediante las distintas técnicas se observan marcadas diferencias,

en el mapa obtenido por el método convencional, se observa que la susceptibilidad varía desde alta a muy alta en el tramo considerado, mientras que en el método no convencional, en la misma área se evidencia una susceptibilidad baja a muy baja, la cual es de manera general dominante en toda la vía, sin embargo, de forma focalizada se evidencia una susceptibilidad que va entre alta y muy alta en ciertos sectores marcados de la vía.

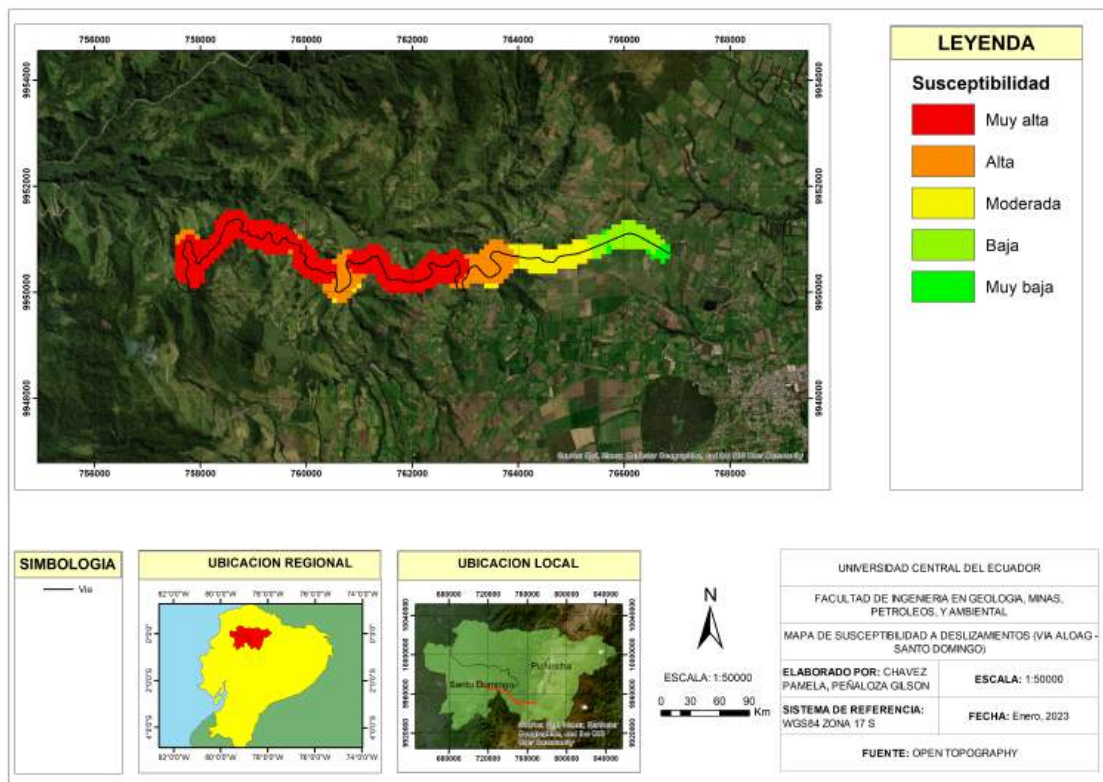


Figura 6: Mapa de susceptibilidad a deslizamientos mediante clasificación Intervalo Igual

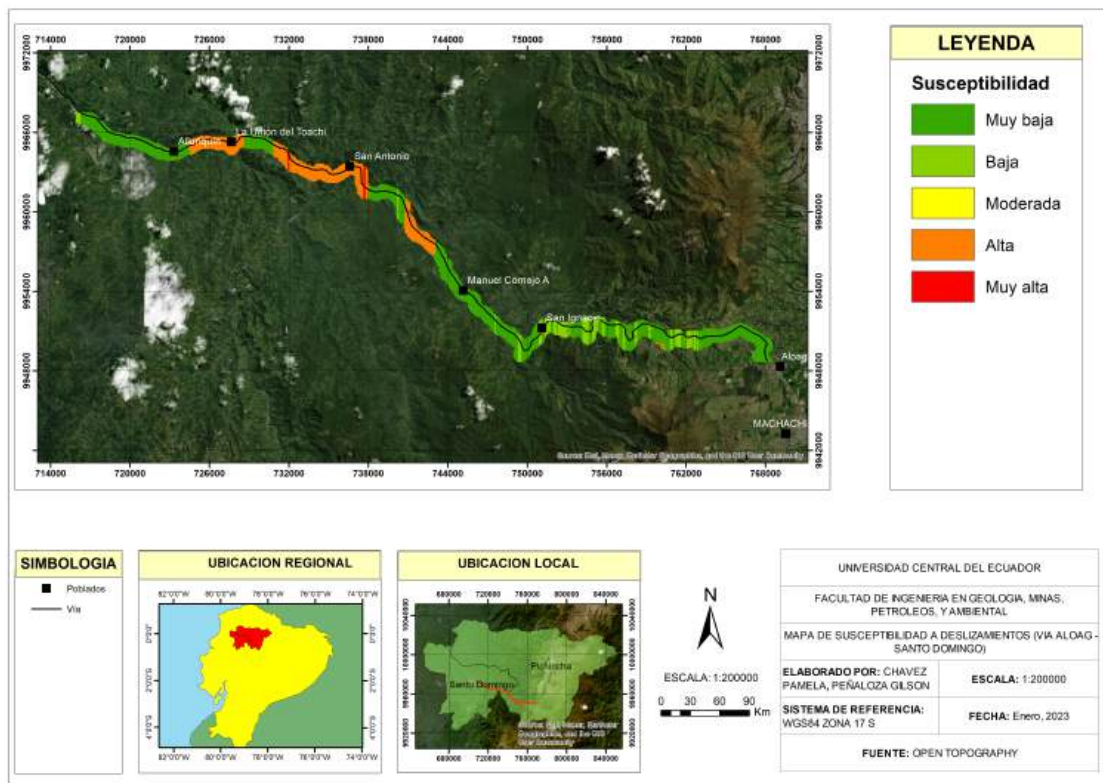


Figura 7: Mapa de susceptibilidad a deslizamientos mediante clasificación Intervalo Igual

5.2. Clasificación cuantiles

En una clasificación de cuantiles, cada clase contiene el mismo número de entidades. La clasificación de cuantiles es adecuada para los datos distribuidos linealmente. El cuantil asigna el mismo número de valores de datos a cada clase. Puesto que las entidades se agrupan en igual número en cada clase mediante la clasificación de cuantiles, el mapa

resultante puede ser erróneo en algunas ocasiones: Entidades similares se pueden situar en clases adyacentes, o entidades con valores muy diferentes se pueden poner en la misma clase. En los mapas generados se observa que los grados de susceptibilidad al no encontrarse en una distribución lineal, se podría tener un grado de error alto debido al dataset con el que se está trabajando, por la definición misma del método de cuantiles.

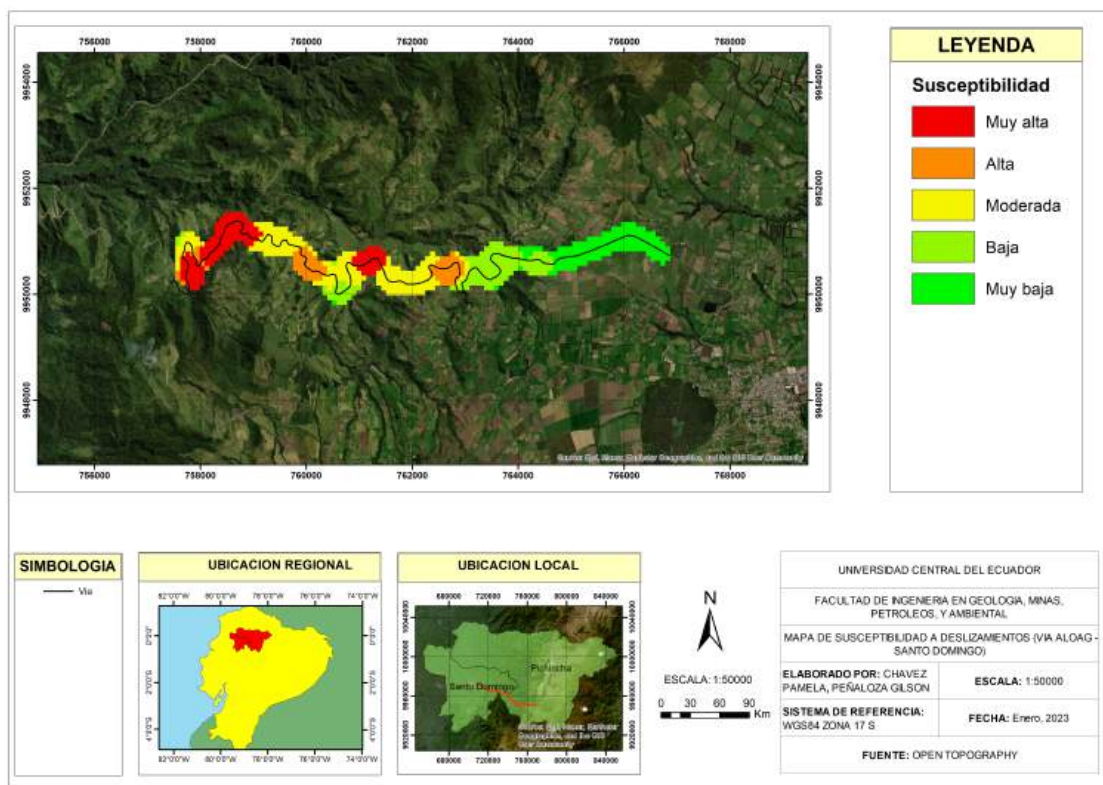


Figura 8: Mapa de susceptibilidad a deslizamientos mediante clasificación Cuantiles

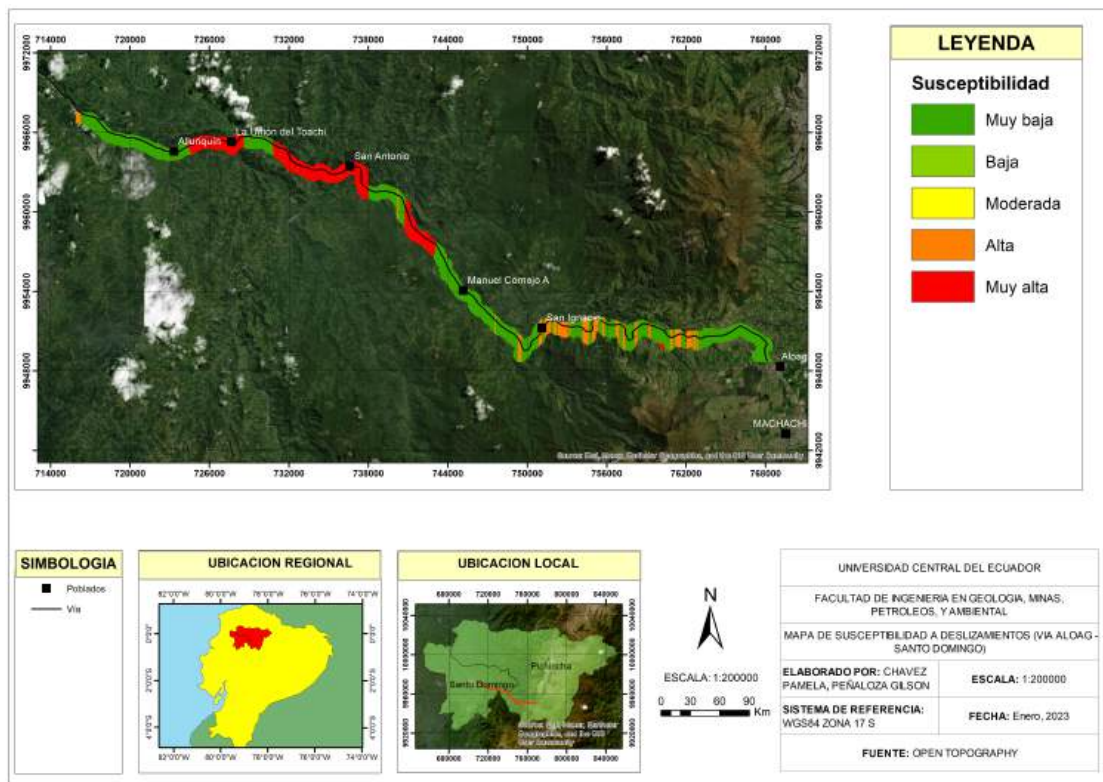


Figura 9: Mapa de susceptibilidad a deslizamientos mediante clasificación Cuantiles

5.3. Natural breaks (Jenks)

Con una clasificación de cortes naturales (Jenks), las clases se basan en las agrupaciones naturales inherentes a los datos. Las rupturas de clase se crean de manera que los valores similares se agrupen mejor y se maximizan las diferencias entre clases. Las entidades se dividen en clases cuyos límites quedan establecidos dónde hay diferencias considerables entre los valores de los datos. Este método se usa mejor con datos que están distribuidos de manera desigual pero no sesgados hacia ninguno de

los extremos de la distribución. En otras palabras, el método busca reducir la varianza dentro de las clases y maximizar la varianza entre clases. Para el caso de estudio, tanto el mapa generado con el método convencional como el generado a partir de Machine Learning se encuentran mejor representados por la clasificación de cortes naturales de Jenks, debido a que trabaja mejor con las variables del dataset las mismas que fueron seleccionadas en base a la información disponible y características principales que influyen en los deslizamientos de la zona.

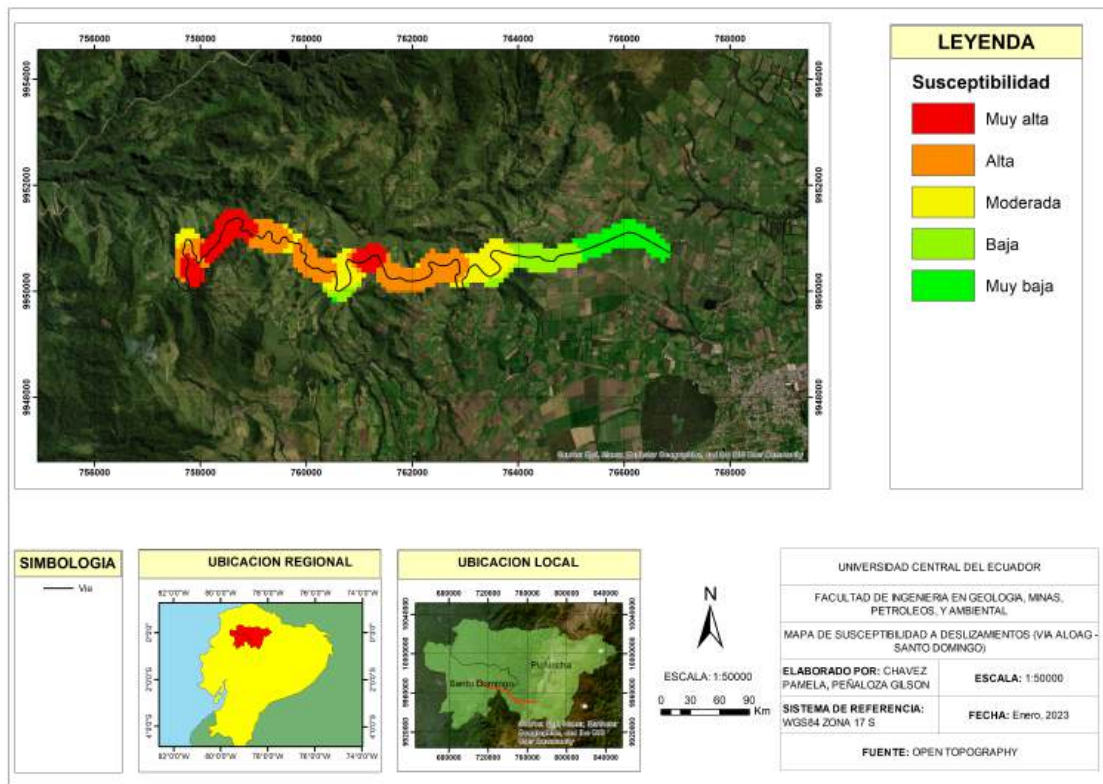


Figura 10: Mapa de susceptibilidad a deslizamientos mediante clasificación Natural Breaks (Jenks)

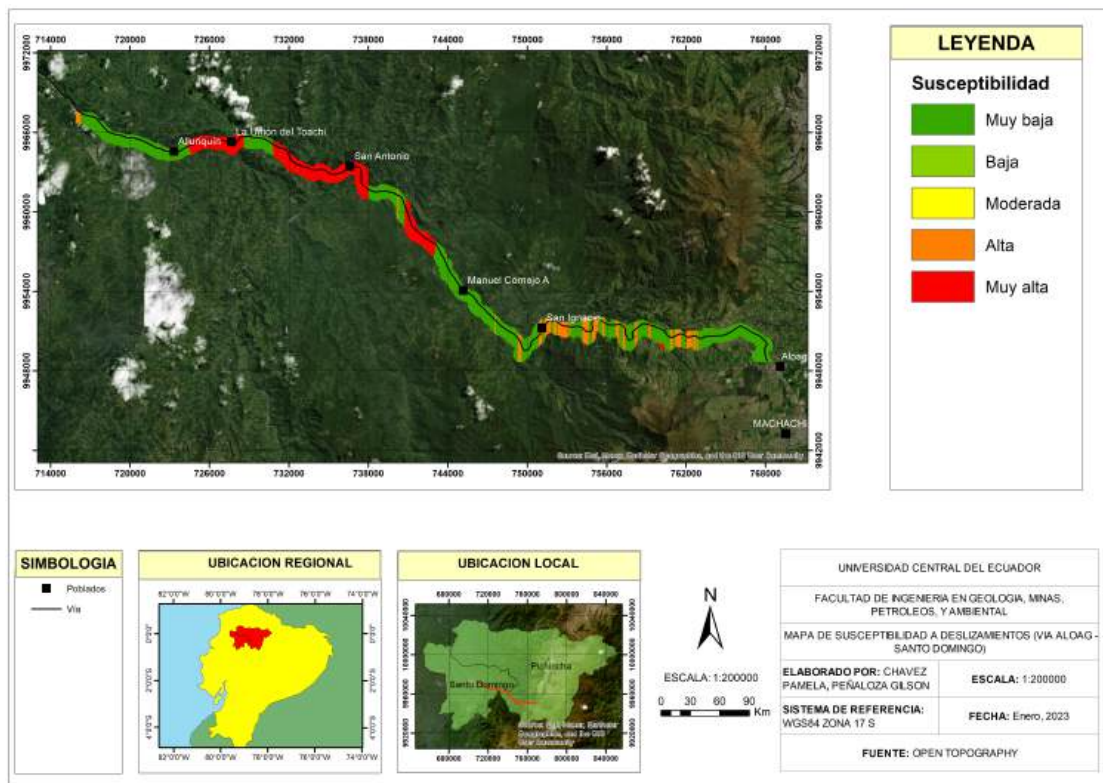


Figura 11: Mapa de susceptibilidad a deslizamientos mediante clasificación Natural Breaks (Jenks)

6. Conclusiones

Ambas técnicas muestran una clasificación en cuanto a la susceptibilidad a lo largo de la vía Aloag-Santo Domingo, la técnica convencional al seguir una serie de instrucciones está limitada a representar de manera más aproximada la realidad, mientras que la técnica no convencional al tener un aprendizaje de datos o ejemplos de la realidad tiende a realizar de manera más aproximada la discriminación.

A pesar de existir una similitud entre el mapa de natural breaks y el mapa de cuantiles generados mediante Machine Learning, se considera a la clasificación por natural breaks, como la más adecuada y aproximada a la realidad, puesto que su principio es generar grupos de datos similares estableciendo los límites entre estos en función de la varianza de datos, además de que existe similitud en los resul-

tados obtenidos tanto por el método convencional así como por el no convencional.

7. Recomendaciones

La existencia de información sobre la aplicación de metodologías no convencionales en cuanto al tratamiento de estos fenómenos en el área es escasa, por tanto es recomendable la implementación de estudios en los que se apliquen distintos algoritmos con la finalidad de obtener resultados cercanos a la realidad, y mediante los cuales poder realizar un adecuado ordenamiento territorial.

Las distintas clasificaciones muestran resultados que en su mayoría divergen, sin embargo, entre las clasificaciones por Natural breaks y cuantiles en la metodología no convencional coinciden en sus resultados, aquello representa un problema puesto que se asume que ambas clasificaciones muestran resulta-

dos cercanos a la realidad, por tanto, se recomienda trabajo, con el objetivo de evitar estas similitudes que en un próximo trabajo, se realice una o varias y así poder obtener una mejor discriminación. clasificaciones diferentes de las ya aplicadas en este

Referencias

- Medina, C., y Astudillo, J. (2009). *Ubicación e identificación de los principales deslizamientos de la vía macará-sabiano de la provincia de Loja*. Maraca: TESIS AEIRNNR.
- Orejuela, I. F. P. (2020). Susceptibilidad a deslizamientos en la vía alóag-santo domingo, mediante lógica difusa. *Revista Geoespacial*, 17(2), 01–12.