



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA CIVIL

## ESTIMACIÓN DE LA ESTABILIDAD DE PAREDES COLGANTES EN FUNCIÓN DE LAS DEFORMACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL

**MARTA NOEMÍ PALMA DIAS**

PROFESORA GUÍA:  
KIMIE SUZUKI MORALES

MIEMBROS DE LA COMISIÓN:  
MARÍA ELENA VALENCIA VERA  
FELIPE OCHOA CORNEJO

Este trabajo ha sido financiado por:  
Itasca Chile Spa

SANTIAGO DE CHILE  
2024

## ESTIMACIÓN DE LA ESTABILIDAD DE PAREDES COLGANTES EN FUNCIÓN DE LAS DEFORMACIONES

El método de explotación por caserones abiertos es eficiente y seguro, permitiendo extraer grandes volúmenes de mineral mediante perforación y tronadura. Sin embargo, existen aspectos críticos que deben analizarse para garantizar su rentabilidad, como la sobre-excavación, la inestabilidad de las paredes y la dilución del material. Actualmente, los métodos más empleados para estimar la estabilidad en caserones son gráficos empíricos debido a su facilidad y rapidez de aplicación. No obstante, tienen limitaciones significativas debido a las condiciones específicas para las que fueron desarrollados y la limitada base de datos en la que se fundamentan.

El objetivo del trabajo es desarrollar una herramienta predictiva, que logre estimar la estabilidad de paredes colgantes en caserones. La idea es que su uso sea fácil y sencillo como los gráficos empíricos, pero con más condiciones y parámetros que estos métodos no permiten incorporar. Y para lograr esto, se analiza y verifica la respuesta del modelo numérico con el criterio de daño determinado.

La metodología que se emplea se divide en tres partes: la primera, comienza con una revisión bibliográfica de los parámetros que afectan la estabilidad de caserones según métodos empíricos y métodos numéricos. La segunda parte comienza con el desarrollo del modelo tridimensional de un caserón paramétrico, el que permite cambiar las condiciones geométricas y geológicas fácilmente definiendo ciertas variables en función de otras. Se utilizó un modelo numérico implementado en *FLAC3D*, teniendo como inputs los parámetros de la roca, geometría de la excavación y los esfuerzos in-situ que según la literatura son significativos al estimar la estabilidad de un caserón. La idea de este modelo numérico es que sea lo más simple pero eficiente para ahorrar tiempo y costo computacional, con tal de correr la mayor cantidad de simulaciones posibles. Con este modelo verificado, se genera una base de datos sintética con diferentes configuraciones de caserones. Finalmente, la tercera parte de la metodología corresponde al desarrollo del modelo predictivo, entrenado a partir de los resultados del modelo numérico mediante el algoritmo de random forest.

Luego de verificar el modelo numérico con el criterio de daño determinado, se alcanzan a simular 8192 casos. La base de datos se genera según una estimación de la dilución de las paredes colgantes de cada caso, y con ella se entrena el modelo predictivo. Los resultados finales que se obtuvieron con el modelo predictivo de random forest fueron: una métrica F1-score igual a 96 % para las categorías  $ELOS > 2[m]$  y  $2[m] > ELOS > 1[m]$ , 98 % para  $1[m] > ELOS > 0.5[m]$  y 86 % para  $ELOS < 0.5[m]$ . Además, se realiza una validación cruzada para verificar que el modelo predictivo pueda generalizar a datos nuevos, obteniendo un modelo robusto con una exactitud general de 94 %.

Finalmente, se cumple el objetivo del trabajo con el desarrollo de una herramienta que permite estimar la estabilidad. Además, se detalla la metodología de los pasos a seguir en cada etapa de la construcción del modelo numérico y el modelo predictivo. Para trabajos futuros, se propone implementar el cálculo de probabilidad de ocurrencia en el output junto con la cantidad de dilución. Con esto, el valor del trabajo se encuentra en su potencial uso para complementar los métodos empíricos, en el cálculo de las estimaciones de estabilidad de paredes colgantes en las primeras etapas de proyectos de ingeniería.

*Dedicado a Juana Dias Maripangue*

# Agradecimientos

Quiero agradecer a toda la gente que me acompañó en este proceso, en especial a mi familia por ser un pilar fundamental en mi vida y subirme el ánimo cuando los estudios me abrumaban.

A mis amigas de la universidad Catalina, Camila y Mariana gracias por todas las risas, los llantos, su compañía y apoyo durante estos 6 años.

Gracias a Patricio Gómez y Rodrigo Silva por darme la oportunidad de realizar mi memoria en Itasca. También quiero agradecer a María Elena Valencia y Diego Acevedo por todo el conocimiento, los consejos y tiempo entregado durante este difícil proceso. Gracias además a Felipe Acuña y Diego Ramos por los gratos momentos de distracción durante nuestros meses de memoristas.

Finalmente muchas gracias a Kimie Suzuki por todo el apoyo durante esta memoria, por ser un ejemplo a seguir y por introducirme con su entusiasmo y pasión al área de la geomecánica.

# Tabla de Contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Objetivos	3
1.1.1. Objetivo General	3
1.1.2. Objetivos Específicos	3
1.2. Alcances	4
1.3. Organización de la Memoria	4
<b>2. Antecedentes de Estudio</b>	<b>5</b>
2.1. Estimación estabilidad de caserones	5
2.1.1. Métodos Empíricos	5
a) Gráfico de Estabilidad de Mathews	5
b) Estimación de Dilución	7
c) Ajustes y extensiones del gráfico de estabilidad de Mathews et al. (1981)	9
2.1.2. Métodos Numéricos	10
2.1.2.1. Clasificación	11
a) Modelos discontinuos	11
b) Modelos continuos	11
2.1.2.2. Modelos constitutivos	12
a) Tipos de comportamiento	12
b) Criterios de falla	14
c) Ubiquitous Joint	15
2.1.2.3. Criterios para interpretar el daño	16
a) Criterios de Velocidad	16
b) Criterios deformacionales	17
2.1.3. Elección de herramientas a usar	17
2.2. Modelos de Aprendizaje Automático	17
2.2.1. Conceptos básicos	18
2.2.2. Aplicaciones de Machine Learning en la actualidad	20
<b>3. Metodología</b>	<b>23</b>
3.1. Representación del caserón en Modelamiento Numérico	23
3.1.1. Consideraciones del modelo	24
3.1.2. Interpretación de Resultados	31
3.2. Desarrollo Modelo Predictivo	33
3.2.1. Generación Base de Datos	33
3.2.2. Preprocesamiento de datos	35
3.2.3. Entrenamiento del modelo de aprendizaje	36
3.2.3.1. Ajuste de parámetros del modelo y uso de GridSearch	38
3.2.3.2. Algoritmo Árbol de Decisión	38
3.2.3.3. Algoritmo de Random Forest	39
3.2.4. Evaluación del Modelo y Validación Cruzada	39
3.2.5. Ajuste de la base de datos para mejorar las métricas del modelo	41
<b>4. Resultados Modelo Numérico</b>	<b>42</b>
4.1. Verificación del modelo numérico	42

4.2. Distribución de las clasificaciones respecto al gráfico de Clark . . . . .	45
4.3. Correlaciones entre variables . . . . .	46
4.4. Discusiones . . . . .	52
<b>5. Modelo Predictivo</b>	<b>53</b>
5.1. Métricas de evaluación versus tamaño de la base de datos . . . . .	53
5.2. Nuevas métricas de evaluación tras balance de la base de datos . . . . .	55
5.3. Discusiones . . . . .	57
<b>6. Conclusiones y Recomendaciones a Trabajos Futuros</b>	<b>59</b>
<b>Bibliografía</b>	<b>60</b>

# Índice de Tablas

2.1.	Comparación de algunos códigos continuos disponibles en el mercado (Modificado de Lorig, 1999). . . . .	12
2.2.	Resumen herramientas predictivas que usan <i>machine learning</i> en geomecánica. . . . .	21
3.1.	Parámetros geométricos de diseño. . . . .	24
3.2.	Parámetros de esfuerzos para el diseño de diseño. . . . .	27
3.3.	Parámetros de calidad de roca para el diseño. . . . .	29
3.4.	Categorías y propiedades de las estructuras para el diseño. . . . .	29
3.5.	Orientación estructuras con respecto a inclinación pared colgante. . . . .	30
3.6.	Parámetros utilizados para modelar casos en estudio, según modo de falla. . . . .	33
3.7.	Resumen parámetros input para modelo numérico. . . . .	34
4.1.	Resumen información recolectada, base de datos Mawdesley et al. (2001). . . . .	43
4.2.	Resumen información recolectada de Hustrulid y Bullock (2001). . . . .	44
4.3.	Comparación entre la realidad, clasificación según Mathews y modelo numérico, para los tres casos de estudio. . . . .	44
4.4.	Resumen de cantidad de casos por categoría y su respectivo porcentaje de acuerdo al total de la base de datos. . . . .	45
5.1.	Nueva cantidad de casos por categoría. . . . .	55
5.2.	Mejor conjunto de hiperparámetros para el modelo árbol de decisión . . . . .	55
5.3.	Mejor conjunto de hiperparámetros para el modelo random forest . . . . .	55
5.4.	Resultados validación cruzada . . . . .	57

# Índice de Figuras

1.1.	Esquematación de la geometría de un caserón. a) Ejemplo dimensiones típicas caserones (Villaescusa, 2014) y b) Vista tridimensional de un caserón abierto (Qi, Z. et al., 2023).	2
2.1.	Factores de ajuste para determinar el número de estabilidad de Mathews (Mathews et al., 1981).	6
2.2.	Gráfico de Número de Estabilidad versus Radio Hidráulico. (Mathews et al., 1981).	7
2.3.	Guía de estimación de ELOS (Clark, 1998).	8
2.4.	Gráfico de Mathews Extendido, basado en regresión logística (Mawdesley, 2001).	9
2.5.	Línea de tiempo con la recopilación de los trabajos principales sobre modificaciones en los parámetros propuestos por Mathews et al. (1981), además de nuevos gráficos para estimar estabilidad en caserones (Elaboración propia).	10
2.6.	Métodos numéricos más utilizados en la mecánica de rocas (Elaboración propia).	11
2.7.	Representación de diferentes comportamientos de materiales mediante curvas idealizadas de esfuerzo-deformación (Sainsbury, 2012).	13
2.8.	Matriz y propiedades Ubiquitous-Joint (Modificado de Sainsbury & Sainsbury, 2017).	16
2.9.	Comparación de la resistencia a la compresión uniaxial entre el modelo Ubiquitous Joint y la solución analítica de Jaeger (1960). (Sainsbury & Sainsbury, 2017).	16
2.10.	Tipos de aprendizaje automático clásico (Elaboración propia).	18
2.11.	Matriz de confusión (Elaboración propia).	19
3.1.	Metodología del trabajo.	23
3.2.	Metodología para el dimensionamiento del modelo (Elaboración propia).	25
3.3.	Condiciones de borde del modelo (Elaboración propia).	26
3.4.	Resultados según la ecuación simplificada de Hoek & Diederichs comparada con la base de datos de China y Taiwan (Hoek & Diederichs, 2006).	27
3.5.	Escalamiento de Módulo de Young variando GSI, para UCS 100 [MPa].	28
3.6.	Escalamiento de Módulo de Young variando GSI, para UCS 25 [MPa].	28
3.7.	Representación ángulo entre estructuras y pared colgante.	29
3.8.	Criterio de Hoek-Brown y su aproximación a Mohr Coulomb (Itasca, 2023).	30
3.9.	Volumen en falla sobre la pared colgante (Elaboración propia).	31
3.10.	Volumen en falla sobre pared colgante (Elaboración propia).	32
3.11.	Comparación estado de plasticidad, con el criterio de falla para caso 1 (Elaboración propia).	32
3.12.	Comparación estado de plasticidad, con el criterio de falla para caso 2 (Elaboración propia).	33
3.13.	Comparación en la uniformidad de los datos, dependiendo el método de muestreo utilizado (Jheald, 2011).	34
3.14.	Gráfico Pairplot, para visualizar la distribución en el espacio muestral de las variables.	35
3.15.	Representación visual de algoritmo árbol de decisión con base de datos de Iris (Elaboración propia).	37
3.16.	Fronteras de decisión para algoritmo árbol de decisión con base de datos de Iris (Elaboración propia).	37
3.17.	a) Esquematación algoritmo de árbol de decisión y b) random forest (Modificado de Khan, 2024).	38
3.18.	Pasos típicos a seguir para validar un modelo predictivo (Elaboración propia).	40
3.19.	Diagrama de flujo de validación cruzada (Elaboración propia).	41
4.1.	Casos de Estudio, sobre el gráfico de estabilidad de Mathews.	43
4.2.	Resultados del modelamiento numérico de los tres casos de estudio de la mina Cannington.	44
4.3.	Histograma con frecuencias según categoría.	45
4.4.	Comparación resultados versus gráfico de estabilidad de Clark.	46

4.5.	Correlación entre las variables y la cantidad de ELOS. . . . .	47
4.6.	Matriz de dispersión-Geometría del caserón. . . . .	48
4.7.	Matriz de dispersión-Parámetros relacionados al esfuerzo in situ . . . . .	49
4.8.	Matriz de dispersión-Propiedades de resistencia y de macizo rocoso. . . . .	50
4.9.	Matriz de dispersión-Orientación y propiedades de las estructuras. . . . .	50
4.10.	Matriz de dispersión-Propiedades físicas y elásticas de la roca. . . . .	51
5.1.	Comparación precisión según cantidad de datos. . . . .	54
5.2.	Comparación exhaustividad según cantidad de datos. . . . .	54
5.3.	Comparación valor-F1 según cantidad de datos . . . . .	54
5.4.	Comparación precisión para modelos árbol de decisión y random forest. . . . .	56
5.5.	Comparación exhaustividad para modelos árbol de decisión y random forest. . . . .	56
5.6.	Comparación valor-F1 para modelos árbol de decisión y random forest. . . . .	56

# Capítulo 1: Introducción

La constante explotación de las reservas minerales en superficie ha obligado a las operaciones mineras a profundizar la búsqueda de materiales con valor económico. Al ser la minería subterránea una solución a esta problemática, es que ha estado tomando un rol importante con el paso de los años. Según la revista *Minería Chilena* (2024), se anticipa un incremento del 51,6% en la producción subterránea entre 2023 y 2034, impulsado principalmente por proyectos como Chuquicamata Subterránea, Proyecto Los Bronces Integrado y el Plan de Desarrollo El Teniente. Este crecimiento permitirá que la participación relativa de la minería subterránea en la producción total de cobre pase del 11,7% en 2023 al 14,6% en 2034, reflejando la importancia creciente de esta modalidad en la industria minera chilena. Los avances tecnológicos han permitido extraer mineral en ambientes más profundos, optimizando los recursos y manteniendo la seguridad general de las operaciones. En ese sentido, las faenas mineras subterráneas han tenido que buscar soluciones para superar sus mayores desafíos mineros, tales como: estallidos de rocas, sobre-excavación del mineral y dilución entre los más relevantes.

En operaciones mineras, existen diferentes técnicas de explotación que están sujetas a las características del yacimiento, por ejemplo, la inclinación, el tamaño y la forma de la veta de mineral, como también la resistencia y la calidad del macizo rocoso alrededor de la veta. Un método de extracción atractivo debido a sus bajos costos y por su eficiencia de las operaciones sin entrada de personal, corresponde al método de *Sublevel Open Stopping* (método de caserones autosoportados). El que utiliza equipos de producción mecanizados para lograr altos índices de producción, con un nivel mínimo de personal.

Un caserón es una cavidad de gran tamaño (Figura 1.1.a) que es explotada de forma ascendente o descendente, y la estabilidad física depende de variables como la calidad del macizo rocoso y su geometría y los esfuerzos a los que se encuentra. Su geometría la compone la pared colgante (*hanging wall*), la pared adyacente (*footwall*) y el techo (*roof*). Debido a su disposición geométrica, la cara más afectada por la gravedad debido a la inclinación es la pared colgante, por esa razón es la que mayor dilución presenta en las mediciones en terreno. La pared colgante al ser la cara más propensa a la inestabilidad, variados autores han centrado sus estudios en ella.

Los métodos de extracción *Sublevel Open Stopping*, se utilizan para extraer yacimientos masivos o tabulares, a menudo de fuerte buzamiento, rodeados de rocas competentes. En general, este método tiene pocas limitaciones en cuanto a la forma, el tamaño y la continuidad de la mineralización. Su éxito depende de la estabilidad de las paredes y techos (usualmente no reforzados). En macizos rocosos de buena calidad, los caserones pueden ser excavaciones relativamente grandes, en las que la perforación y tronadura son el principal método de extracción de roca. En los límites de los caserones puede producirse una dilución del mineral consistente en roca estéril de baja calidad o materiales de relleno de mina. Además, la pérdida de mineral por insuficiencia de rompimiento del material también puede producirse dentro de los límites de los caserones (Villaescusa, 2014). Existen diferentes formas de medir la dilución, una muy utilizada corresponde al ELOS (*equivalent linear overbreak/slough*) introducida por Clark & Pakalnis (1997). El ELOS representa una profundidad promedio de sobre-excavación y tiene la ventaja de entregar el valor de la dilución de manera independiente del ancho del caserón.

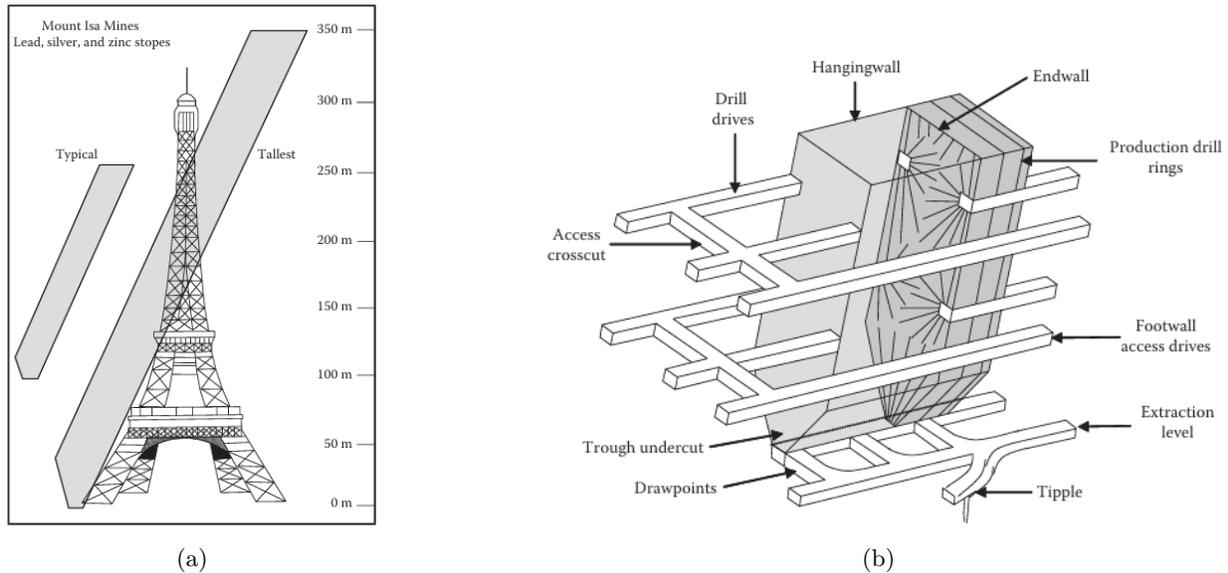


Figura 1.1: Esquematación de la geometría de un caserón. a) Ejemplo dimensiones típicas caserones (Villaescusa, 2014) y b) Vista tridimensional de un caserón abierto (Qi, Z. et al., 2023).

En gran parte de los diseños de los métodos subterráneos, un caserón es la unidad principal de explotación, es por ello que su entendimiento es de gran relevancia en minería, para lograr el equilibrio entre la estabilidad física del caserón y el negocio minero. A lo largo de los años, diversos autores han dedicado sus investigaciones a estudiar la estabilidad de los caserones mediante métodos como la determinación de la máxima luz, ábacos de hundibilidad y métodos gráficos de estabilidad, donde estos últimos destacan por ser los más utilizados en la industria minera. Los métodos gráficos de estabilidad son una herramienta empírica para diseñar caserones relacionando factores geomecánicos (calidad de macizo rocoso, nivel de esfuerzos, etc) con las dimensiones de la excavación. Estos métodos empíricos son herramientas sencillas, prácticas y muy utilizadas en etapas iniciales de proyectos de ingeniería.

Desde 1981 los gráficos de estabilidad para estudiar caserones se han desarrollado y fortalecido, agregando nuevas bases datos y reajustes de factores. Entre los métodos empíricos más conocidos se pueden mencionar a Mathews et al. (1981), Potvin (1988), Nickson (1992), Stewart & Forsyth (1995) y Mawdesley & Truman (2000). Todos estos métodos empíricos se basan en la relación entre el radio hidráulico y el concepto de número de estabilidad ( $N$ ) de Mathews et al. (1981). El radio hidráulico es un parámetro relacionado a la geometría de las paredes en estudio del caserón. Mientras que el número de estabilidad depende del índice modificado de calidad de túneles ( $Q'$ ), y tres factores: Factor A, que representa los esfuerzos que actúan en la cara libre; Factor B, que representa la influencia de las fallas geológicas en la estabilidad del caserón; y el Factor C o ajuste por gravedad, que cuantifica el efecto de la gravedad según la disposición de la cara del caserón.

Sin embargo, los métodos empíricos tienen consideraciones y limitaciones que no pueden ser obviadas. Una de ellas es que, en general, estos métodos han sido desarrollados considerando buenas prácticas de tronadura y perforación, por lo tanto, no considera el daño por explosivos, ni el efecto del desconfinamiento en las paredes. Por otro lado, las metodologías no consideran influencias de excavaciones cercanas al caserón estudiado, ni considera ningún tipo de rellenos y fallas en sectores adyacentes a la excavación (Suorineni, 2014). Otra limitación de los métodos empíricos es que se construyeron a partir de bases de datos acotadas. Este es un problema general en la geomecánica, donde la falta de información es un hecho común. Por esta razón, herramientas computacionales han

sido propuestas como solución a la deficiencia de información en el área. Por un lado, los modelos numéricos son capaces de reproducir fenómenos físicos que ocurren en la realidad, representando esfuerzos y deformaciones. Los modelos numéricos permiten simular una gran cantidad de condiciones y configuraciones de problemas, generando lo que se conoce como bases de datos sintéticas, siendo sólo limitadas por la capacidad computacional disponible. Por otro lado, se encuentran los modelos de aprendizaje automático, que permite a los computadores predecir fenómenos complejos sin programación explícita encontrando patrones en los datos. Los modelos de aprendizaje pueden aprender de datos reales extraídos de terreno, o de bases de datos sintéticas construidas con modelos numéricos.

El uso de herramientas del tipo *machine learning* se ha incrementado de forma exponencial en los últimos años en diferentes disciplinas. En el área de geomecánica se ha estado introduciendo su aplicación en variadas problemáticas. Por ejemplo, se han empleado redes neuronales para determinar la luz de caserones (Adoko et al., 2022; Wang et al., 2002), clasificadores para determinar la estabilidad de excavaciones (García-Gonzalo et al., 2016) y árboles de decisión para predecir la estabilidad de caserones (Qi et al., 2018b, 2018c), entre otros. La principal dificultad de este enfoque es similar al presentado en los ábacos empíricos. Los modelos predictivos están limitados a la base de datos con cual fueron entrenados. Si bien su capacidad predictiva puede ser alta, no es posible generalizar el modelo a otros tipos de roca, condiciones de esfuerzos y geometrías. Existe gran potencial para estas herramientas, donde los mejores resultados pueden obtenerse a partir de modelos que consideren un gran volumen de datos de entrada en un rango amplio de condiciones. Lo anterior es posible, cuando se utilizan modelos numéricos y bases de datos sintéticas, generando modelos predictivos robustos y generalizables.

Este trabajo presenta una metodología para estimar la estabilidad de caserones basada en “meta-modelos” (Gramacy, 2021). Estos se definen como modelos predictivos construidos a partir de bases de datos sintéticas, las cuales pueden ser generadas a partir de resultados obtenidos de modelos numéricos. Los metamodelos tienen la capacidad de entregar respuesta rápida utilizando técnicas de regresión u otras técnicas de aprendizaje automático. La primera parte del trabajo consiste en una revisión bibliográfica de métodos de estimación de estabilidad de caserones. De esta revisión, se identifican las variables que deben ser incluidas en un modelo numérico simplificado de la realidad. La segunda parte explica la metodología aplicada para construir el modelo numérico. La tercera parte presenta en detalle la construcción del modelo sustituto. Finalmente, en la última parte del trabajo se presentan las conclusiones respecto a los resultados y una propuesta de posibles siguientes pasos en el desarrollo de este tipo de soluciones.

## **1.1. Objetivos**

### **1.1.1. Objetivo General**

El objetivo general del trabajo corresponde a generar una herramienta capaz de predecir la estabilidad de paredes colgantes en caserones, mediante un criterio deformacional.

### **1.1.2. Objetivos Específicos**

Los objetivos específicos son los presentados a continuación:

- Identificar los parámetros que influyen en la estimación de la estabilidad de caserones a partir de métodos empíricos y de modelos numéricos.
- Representar de forma simple y eficiente las condiciones de un caserón en profundidad, con un modelo numérico.

- Desarrollar un modelo predictivo para analizar la estabilidad de paredes colgantes en caserones, a partir de los resultados obtenidos de los modelos numéricos.

## 1.2. Alcances

El trabajo se centra en el estudio de la dilución no planificada de paredes colgantes en caserones aislados en *Sublevel Open Stopping*. Para esto, se utiliza el software *FLAC3D* desarrollado por ITASCA, el que corresponde a un programa de modelamiento numérico continuo. Se utiliza además, el modelo constitutivo elasto-plástico Ubiquitous Joint, que sólo considera el set más desfavorable de estructuras de un macizo rocoso.

Un alcance importante del trabajo, fue la definición del rango de valores mínimos y máximos en los parámetros de entrada del modelo numérico, ya que los resultados del modelo predictivo solo permiten estimar resultados que se encuentren dentro del intervalo de valores simulados en la base de datos. Otro alcance del estudio realizado son los factores que no se consideraron y pueden afectar a la estabilidad, tal como el nivel freático, tener más de una estructura en el macizo rocoso o la consideración de tener más excavaciones cercanas que afecten en los esfuerzos inducidos.

Otro alcance es la limitada información sobre casos de estudio para validar los modelos, sin embargo se revisa el comportamiento de 3 casos de estudios con el criterio de daño escogido. También, se debe mencionar que se determinará la estabilidad a partir de la cantidad de dilución según el parámetro ELOS.

## 1.3. Organización de la Memoria

La organización de este trabajo consta de 6 capítulos, los cuales se separarán en los siguientes temas:

- Capítulo 2: Éste capítulo sintetiza los métodos más utilizados para la estimación de estabilidad de caserones, los cuales se dividen en métodos empíricos y métodos numéricos. Además se presenta un resumen sobre los conceptos básicos del *Machine Learning* y sus potenciales usos.
- Capítulo 3: Se presenta la metodología utilizada en la realización del trabajo, explicada detalladamente. Se separa en el desarrollo del modelo numérico y el desarrollo del modelo predictivo.
- Capítulo 4: Éste capítulo presenta los principales resultados y discusiones de los resultados de los modelos numéricos. Se muestra de forma resumida la cantidad de casos y la dilución obtenida para cada una de ellos.
- Capítulo 5: Se muestran los principales resultados y discusiones del modelo predictivo. Aquí se presenta una comparación entre las métricas de rendimiento de dos algoritmos y elige el mejor modelo predictivo para la problemática analizada.
- Capítulo 6: En este capítulo se concluye sobre los resultados encontrados, además de recomendaciones para posibles trabajos futuros.

# Capítulo 2: Antecedentes de Estudio

En este capítulo se analizan los diversos métodos disponibles para el diseño de caserones, así como las técnicas para estimar su estabilidad. Además, se discuten de manera general el uso y las oportunidades que ofrece el aprendizaje automático, con un enfoque específico en el campo de la geomecánica.

## 2.1. Estimación estabilidad de caserones

### 2.1.1. Métodos Empíricos

#### a) Gráfico de Estabilidad de Mathews

Es un método empírico para estimar la estabilidad de las paredes de un caserón. Fue creado mediante la recopilación de 26 casos de 2 minas canadienses, una australiana y otros casos obtenidos de literatura. Con enfoque en profundidades hasta los 1000 [m], este método reconoce fallas controladas por esfuerzos, fallas estructurales y por combinación de las dos antes mencionadas. Mathews et al. (1981) sugiere que la estabilidad de cada pared en el caserón, debe analizarse por separado, y se deben ir ajustando las dimensiones hasta que todas las caras sean estables.

Para el análisis, se consideraron dos parámetros; el primero es el Número de Estabilidad ( $N'$ ) (Ecuación 2.1), utilizado para representar la calidad del macizo rocoso, el estado de esfuerzos y la orientación de las superficies expuestas. El segundo parámetro corresponde al Radio Hidráulico ( $S$ ) (Ecuación 2.2), el cual da cuenta de la forma y tamaño de la excavación.

$$N' = Q' * A * B * C \quad (2.1)$$

$$S = \left( \frac{\text{Área Pared}}{\text{Perímetro Pared}} \right) \quad (2.2)$$

Donde:

- **Q': Modified NGI Rock Mass Rating**

El índice  $Q'$  fue modificado por Mathews en 1981, a partir del factor  $Q$  propuesto por Barton, Lien y Lunde (1974) del Instituto Geotécnico Noruego (NGI), el cual se observa en la Ecuación 2.3. A diferencia del índice original, se estableció el factor de reducción de esfuerzos (SRF) igual 1. Manteniendo todos los demás factores, por lo que el índice modificado  $Q'$ , representa solamente la resistencia y estructuras del macizo rocoso.

$$Q = \frac{RQD}{(Jn)} * \frac{(Jr)}{(Ja)} * \frac{(Jw)}{SRF} \quad (2.3)$$

Donde:

- RQD: Rock Quality Designation.
- Jn: Joint set number (número de sets de fracturas).
- Jr: Joint roughness number (forma de las superficies de fracturas).

- Ja: Joint alteration number (ángulo de fricción aproximado de las superficies de las estructuras).
- Jw: Joint water reduction factor.
- SRF: Stress reduction factor (condiciones de esfuerzos y del desconfinamiento del macizo rocoso).

- **A: Factor de esfuerzos**

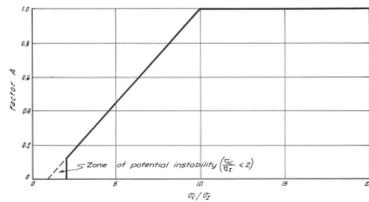
Para la construcción del gráfico empírico, se propone el factor A para reemplazar al factor SRF, en la caracterización de esfuerzos en los que se encuentra un macizo rocoso. Este factor es la relación entre la resistencia a la compresión uniaxial de la roca, y el esfuerzo inducido actuando paralelo a la cara expuesta bajo análisis del caserón. Su valor se puede estimar según el gráfico de la Figura 2.1.a.

- **B: Factor de orientación de discontinuidades**

Este factor da cuenta sobre la presencia de estructuras persistentes, es decir, planos de debilidad en cierta dirección que están presentes de forma continua a través del macizo rocoso. Para seleccionar el factor B, primero se debe determinar la orientación del set de fracturas más desfavorable, basado en el espaciamiento y continuidad relativa. Luego, se determina el ángulo de intersección entre las estructuras, con la superficie expuesta en consideración. Finalmente, se elige el valor del factor B según la Figura 2.1.b.

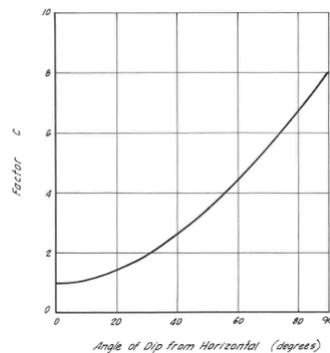
- **C: Factor de ajuste por gravedad**

Este factor refleja el efecto de la gravedad y debe aplicarse a techos, paredes colgantes y posiblemente a footwalls empinadas, cuando una estructura adversa ha sido identificada, es decir, una estructura con orientación y propiedades desfavorables para la estabilidad de la pared. La forma de obtener el factor C se muestra en la Figura 2.1.c.

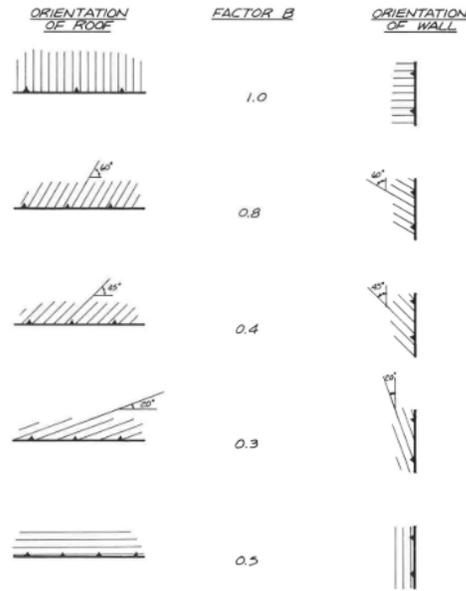


$\sigma_3$  = Uniaxial compressive strength of intact rock.  
 $\sigma_1$  = Induced compressive stress.

(a) Factor de esfuerzos A



(c) Factor C de diseño de superficie



(b) Factor B de orientación de defectos en la roca

Figura 2.1: Factores de ajuste para determinar el número de estabilidad de Mathews (Mathews et al., 1981).

Utilizando los factores mencionados anteriormente, se graficó el número de estabilidad en función del radio hidráulico para una recopilación de casos de estudio, además de otros extraídos de literatura. En la Figura 2.2 se muestran los resultados, divididos en tres zonas, que se describen a continuación:

- Zona estable: La excavación se mantiene en buen estado sin utilización de soporte, o sólo con necesidad local de soporte.
- Zona potencialmente inestable: La excavación presenta falla localizada, pero tiende a formar arcos mecánicos estables. Sumando la utilización de soporte puede disminuir el riesgo de falla.
- Zona de potencial hundimiento: La excavación falla y no logra estabilizarse antes del relleno de toda la cavidad.

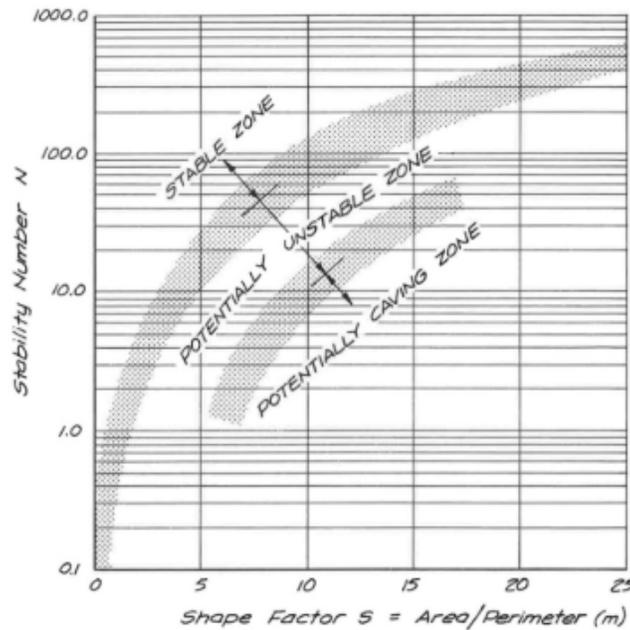


Figura 2.2: Gráfico de Número de Estabilidad versus Radio Hidráulico. (Mathews et al., 1981).

En este sentido, un número de estabilidad grande y un radio hidráulico pequeño, debería representar una cara estable, mientras que la configuración contraria, un número de estabilidad bajo y un radio hidráulico alto, debería representar mayor inestabilidad. Como conclusión del estudio realizado por Mathews et al. (1981), se llegó a que las zonas de estabilidad definidas en la Figura 2.2, se consideran conceptualmente validas, pero no se obtuvieron datos suficientes para confirmarlas. Por lo que se plantea la necesidad de visitas adicionales al sitio, para obtener datos confirmatorios.

## b) Estimación de Dilución

La dilución se define como el material de baja ley, que se extrae y es procesado junto con el flujo de mineral, reduciendo así su valor. La dilución se define y cuantifica con respecto a un límite del caserón idealizado. Según lo anterior, se pueden clasificar dos tipos de dilución, la planificada y la no planificada (Villaescusa, 2014).

Clark (1998), desarrolló un nuevo enfoque de diseño empírico para estimar la dilución no planificada a partir de las paredes colgantes y yacentes de caserones abiertos. Los gráficos de diseño resultantes

se basaron en mediciones cuantificables de sobreexcavación/desprendimiento realizadas con el Sistema de Monitoreo de Cavidades (CMS), y se desarrollaron a partir de una base de datos exhaustiva recompilada de seis operaciones subterráneas de caserones abiertos en Canadá. Se introdujo un nuevo parámetro denominado ELOS (desprendimiento/sobreexcavación equivalente lineal), mostrado en la Ecuación 2.4, y se incorporó a los gráficos de diseño como una medida de la dilución no planificada.

$$ELOS = \frac{\text{Volumen en falla [m}^3\text{]}}{\text{Área de la Pared [m}^2\text{]}} \quad (2.4)$$

La justificación teórica para la metodología de diseño del gráfico empírico de Clark, se demostró mediante un estudio de modelamiento numérico que examinó la zona de relajación alrededor de los caserones. Además utilizó métodos estadísticos, redes neuronales y estudios adicionales de casos para validar las zonas de diseño propuestas. Este nuevo enfoque para el diseño de caserones fue una mejora sobre los métodos existentes, ya que permitió que las dimensiones de los caserones se determinaran basándose en un nivel 'aceptable' de dilución, en lugar de descripciones cualitativas de estabilidad tales como: 'estable', 'zona de transición' o 'potencialmente inestable'. Las zonas se describen a continuación y se muestran en la Figura 2.3.

- **Zona:  $ELOS \leq 0.5$  [m]:** Se asocia principalmente a daño por tronadura.
- **Zona:  $0.5$  [m]  $\leq ELOS \leq 1.0$  [m]:** Considerado como dilución menor. Es provocada por fenómenos de deslizamiento de bloques o lajamiento de las paredes del caserón.
- **Zona:  $1.0$  [m]  $\leq ELOS \leq 2.0$  [m]:** Considerado como dilución moderada. Está asociada al desprendimiento y caída de bloques de roca desde la pared colgante y techo, los que deslizan entre sí o caen por gravedad.
- **Zona:  $2.0 \leq ELOS$  [m]:** Considerado como dilución severa. Esta zona corresponde a una sobre excavación mayor producto del desconfinamiento de la pared asociado a la zona de relajación de esfuerzos que ha concluido en un arco mecánico estable mucho más atrás de la delimitación de la pared colgante.

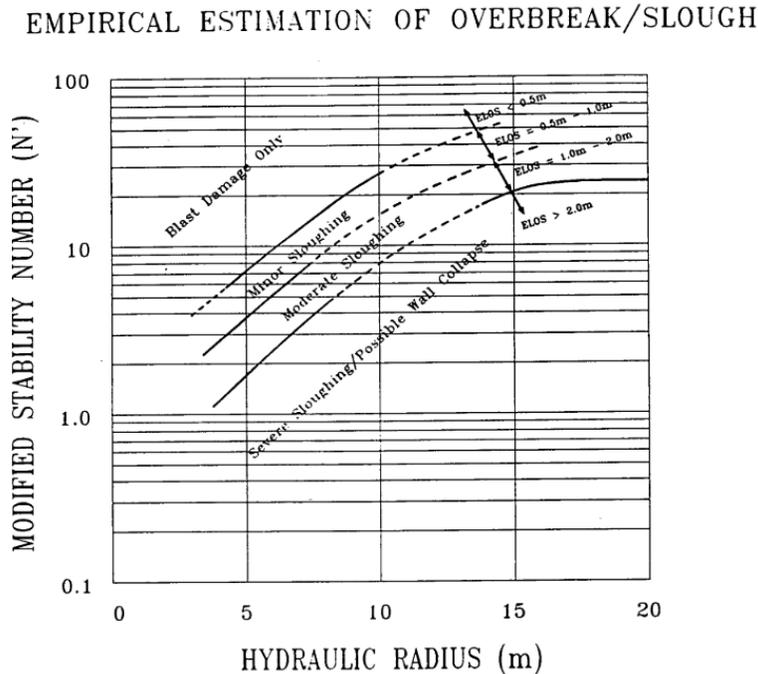


Figura 2.3: Guía de estimación de ELOS (Clark, 1998).

Al igual que el Gráfico Empírico de Mathews et al. (1981), las principales limitaciones del gráfico de Clark(1998) se centran en el tamaño de la base de datos que utilizó. Clark propone al final de su estudio que la base de datos necesita caserones de mayor tamaño y en calidades de roca de peor calidad, para refinar el método. Además, el método esta restringido a paredes colgantes y yacentes en un estado de esfuerzos bajos o relajación, con estructuras paralelas críticas para la estabilidad de las paredes. En su estudio, Clark identifica otros parámetros que afectan la estabilidad, pero solo se dan pautas generales con respecto a su influencia en la estimación de ELOS, por ejemplo geometría irregular de las paredes o tipo el tipo de tronadura utilizada.

### c) Ajustes y extensiones del gráfico de estabilidad de Mathews et al. (1981)

Una serie de investigaciones y recopilación de datos en diferentes minas del mundo se realizó, con el objetivo de ampliar la base de datos de estos gráficos. Una de las actualizaciones más destacadas fue realizada por Mawdesley et al. (2001), quienes lograron aumentar de 26 casos que componían el gráfico original, a un total de 483, como se observa en la Figura 2.4. Además, a lo largo de los años, se han introducido diversos ajustes en los parámetros para mejorar la estimación de la estabilidad, así como también nuevas propuestas de gráficos empíricos. Estos últimos comparten la característica de basarse en el gráfico de Mathews et al. (1981), como fundamento metodológico. En la Figura 2.5, se presenta una línea de tiempo explicativa que ilustra estos desarrollos a través de los años.

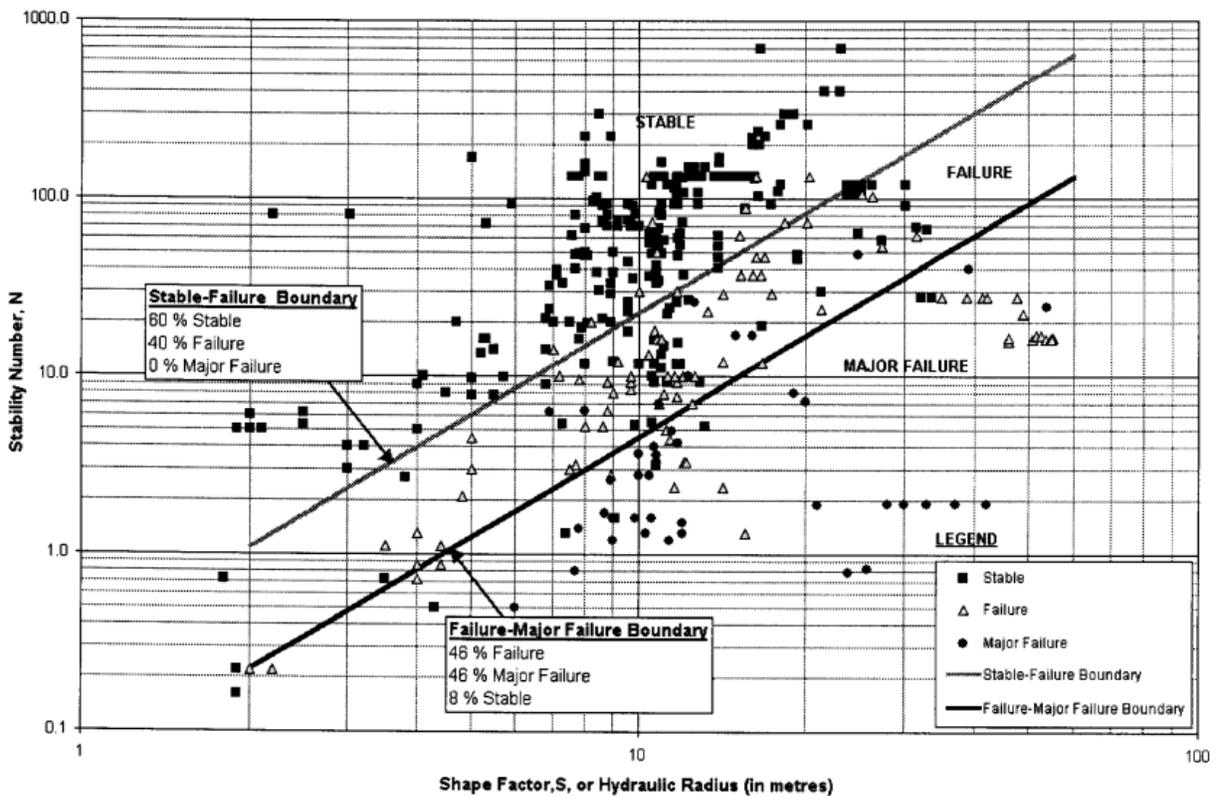


Figura 2.4: Gráfico de Mathews Extendido, basado en regresión logística (Mawdesley, 2001).

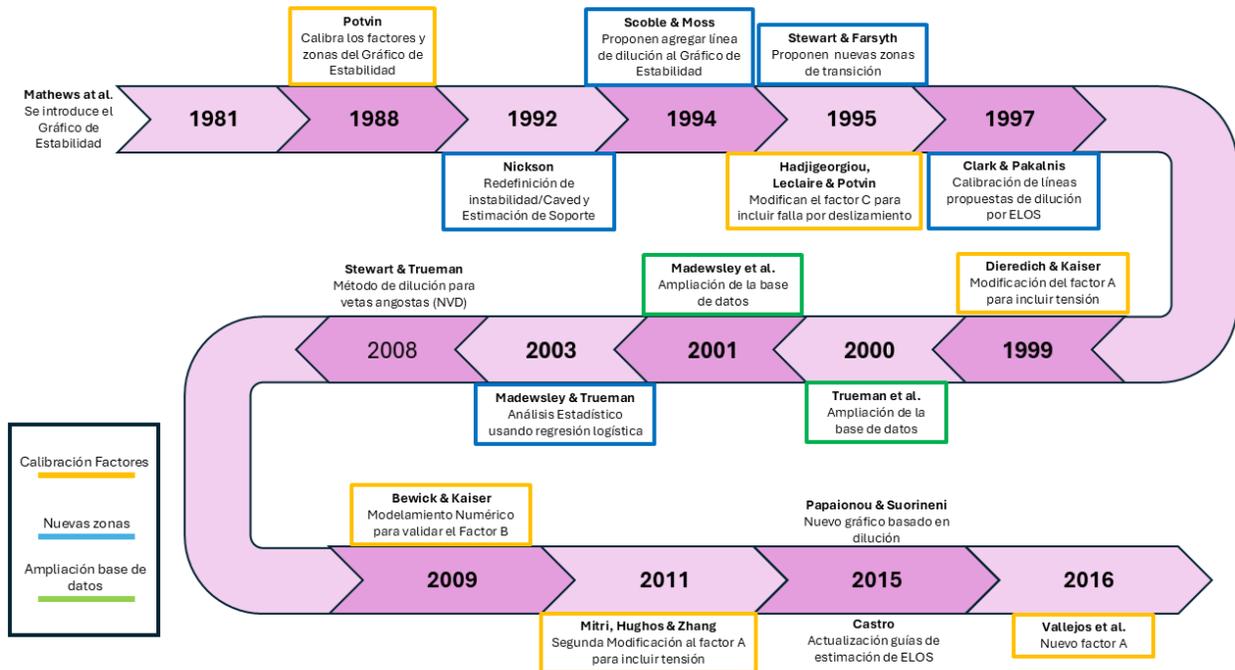


Figura 2.5: Línea de tiempo con la recopilación de los trabajos principales sobre modificaciones en los parámetros propuestos por Mathews et al. (1981), además de nuevos gráficos para estimar estabilidad en caserones (Elaboración propia).

Los métodos de estimación de estabilidad mediante gráficos empíricos, son herramientas ampliamente usadas en el diseño de caserones explotados mediante Sublevel Stopping. Sin embargo, debido a su sencillez y rápido uso, a veces son utilizados sin tener en cuenta sus importantes limitaciones entre las que se encuentran: las condiciones en las que se desarrollan los casos en los que se basan, la simplificación de la geometría del caserón, estos métodos no consideran fallas geológicas mayores, algunos poseen descripciones cualitativas de la estabilidad y la subjetividad con la que las zonas fueron definidas en los gráficos. Por esta razón, existen situaciones y condiciones que necesitan un análisis a mayor detalle.

### 2.1.2. Métodos Numéricos

La falta de información es un hecho común en la mecánica de rocas y en el diseño de ingeniería, por lo que enfoques empíricos para predecir su comportamiento, son ampliamente usados. Sin embargo, el desarrollo de computadores ha sido una contribución significativa al campo de mecánica computacional, el cual encontró un amplio uso en la mecánica de rocas (Nikolic et al., 2014).

El 'modelo' y la 'computadora' son ahora componentes integrales de la mecánica de rocas y los estudios de ingeniería de rocas. De hecho, los métodos numéricos y las técnicas de computación se han convertido en herramientas diarias para formular modelos conceptuales y teorías matemáticas que integran diversa información sobre geología, física, técnicas de construcción, economía, el medio ambiente y sus interacciones. Este logro ha mejorado en gran medida el desarrollo de la mecánica de rocas moderna, desde el arte 'empírico' tradicional de la deformabilidad y estimación de resistencia de la roca y el diseño de soporte, hasta el racionalismo de la mecánica moderna, gobernada y establecida en los tres principios básicos de la física: masa, momento y conservación de la energía (Jing, 2002).

En la práctica minera moderna, los métodos computacionales o numéricos de análisis de esfuerzos y deformación se utilizan para evaluar los esfuerzos y deformaciones inducidos alrededor de los lími-

tes de la excavación y dentro de los alrededores del macizo rocoso como resultado de la excavación. Tienen un uso generalizado como ayuda para toma de decisiones al establecer diseños generales de caserones abiertos y secuencias de extracción. Además, en el diseño detallado y dimensionamiento de los componentes de la estructura minera en su conjunto, incluyendo elementos de infraestructura, accesos, caserones y pilares (Villaescusa, 2014).

Específicamente, los modelos que analizan la estabilidad de caserones se centran en estudiar sus esfuerzos y deformaciones alrededor de la zona excavada. Para esto, se utilizan técnicas de simulación para analizar cómo responden estas zonas bajo diferentes condiciones. Al considerar propiedades del material, condiciones de carga y métodos numéricos, estos modelos permiten predecir posibles volúmenes de falla y diseñar soluciones para asegurar la estabilidad y seguridad de las operaciones mineras. Los métodos numéricos más utilizados al abordar los problemas de mecánica de rocas en minería se pueden clasificar según la Figura 2.6.

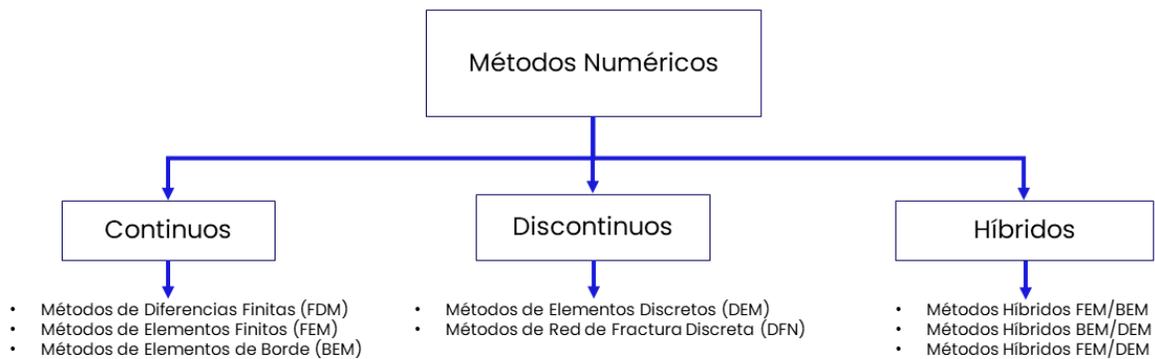


Figura 2.6: Métodos numéricos más utilizados en la mecánica de rocas (Elaboración propia).

### 2.1.2.1. Clasificación

Los modelos numéricos dividen al material en elementos, y a cada uno le asigna un modelo constitutivo y propiedades. Los elementos pueden estar conectados, formando lo que se denomina un modelo continuo, o separadas por discontinuidades, formando un modelo discontinuo. Los modelos discontinuos permiten el deslizamiento y la separación en superficies ubicadas explícitamente dentro del modelo (Wyllie et al., 2004).

#### a) Modelos discontinuos

En el modelamiento numérico de problemas de ingeniería, algunos problemas pueden representarse mediante un modelo adecuado utilizando un número finito de componentes bien definidos. El comportamiento de dichos componentes es bien conocido, o puede tratarse individualmente de forma matemática. El comportamiento global del sistema se puede determinar a través de inter-relaciones bien definidas entre componentes individuales (Jing, 2002).

#### b) Modelos continuos

En otros problemas, la definición de los componentes independientes puede requerir una subdivisión infinita del dominio del problema, y el problema solo puede ser tratado usando supuestos de

un *elemento infinitesimal*, implicando en teoría un número infinito de componentes. Esto usualmente lleva a describir el comportamiento del sistema mediante ecuaciones diferenciales en los puntos del espacio. Estos sistemas son llamados *continuos*, y tienen un número infinito de grados de libertad. Para resolver estos tipos de problemas usando computadores, el dominio del problema es usualmente subdividido en un número finito de sub-dominios (elementos), donde su comportamiento es aproximado por descripciones matemáticas más simples, con un número finito de grados de libertad. Estos sub-dominios deben satisfacer tanto las ecuaciones diferenciales que rigen el problema, como la condición de continuidad en sus interfaces con elementos adyacentes. Esto corresponde a la llamada discretización de un continuo. Es una aproximación de un sistema continuo con infinitos grados de libertad por un sistema discreto con un número finito de grados de libertad (Jing, 2002). En la Tabla 2.1, se muestran algunos programas de modelamiento numérico que funcionan con sistemas continuos.

Tabla 2.1: Comparación de algunos códigos continuos disponibles en el mercado (Modificado de Lorig, 1999).

	FLAC	FLAC3D	PLAXIS	ADINA	ABAQUS	RS2
Formulación	Diferencias Finitas (FDM)		Elementos finitos (FEM)			
Dimensiones	Dos	Tres	Dos y tres			Dos
Interfaces	Sí					
Modelos de materiales inelásticos	Mohr-Coulomb Hoek-Brown Ubiquitous joint Strain softening Creep		Mohr-Coulomb Strain softening Creep	Mohr- Coulomb Ubiquitous joint Strain softening Creep		Mohr-Coulomb Hoek-Brown
Análisis dinámicos	Sí					

FLAC: Fast Lagrangian Analysis of Continua, Itasca Consulting Group, USA

FLAC3D: Fast Lagrangian Analysis of Continua in 3 Dimensions, Itasca Consulting Group, USA

PLAXIS: Plasticity Axisymmetry, PLAXIS BV, Netherlands

ADINA: Automatic Dynamic Incremental Nonlinear Analysis, ADINA R& D, Inc. USA

ABAQUS: Superior Finite Element analysis Solution, ABAQUS, Inc. USA

RS2: Finite Element Analysis and Support Design for Excavations, Rocscience, Canada

### 2.1.2.2. Modelos constitutivos

#### a) Tipos de comportamiento

En la Figura 2.7A se presenta un ejemplo de una curva típica esfuerzo-deformación junto con diferentes idealizaciones para representar diferentes comportamientos de materiales. Según Sainsbury (2012), se describen los comportamientos respectivos a continuación:

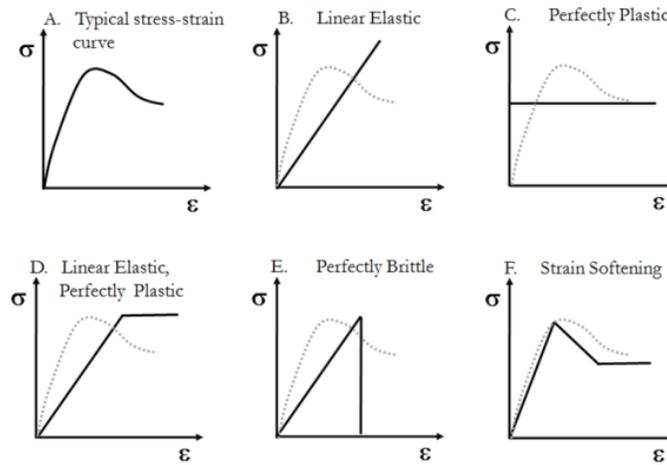


Figura 2.7: Representación de diferentes comportamientos de materiales mediante curvas idealizadas de esfuerzo-deformación (Sainsbury, 2012).

**Lineal Elástico:** Un modelo elástico e isotrópico (Figura 2.7B) entrega la representación más simple del comportamiento de un material. Es válido para materiales homogéneos, isotrópicos y continuos que exhiben comportamiento esfuerzo-deformación lineal sin histéresis de descarga. No tiene representación de la respuesta post-peak. En la realidad, la roca solo exhibe comportamiento elástico hasta que cierto esfuerzo de fluencia es superado, luego de alcanzado este valor, deja de comportarse elásticamente.

**Perfectamente Plástico:** La deformación perfectamente plástica (Figura 2.7C) continua indefinidamente bajo un esfuerzo constante. La tasa de las deformaciones plásticas está relacionada al esfuerzo de fluencia, el cual también representa el esfuerzo de falla. Un material perfectamente plástico se caracteriza por la suposición de que el esfuerzo que causa la deformación permanente no recuperable, debe alcanzar un cierto valor antes de que se pueda producirse cualquier contracción o extensión. Cuando el esfuerzo de fluencia es alcanzado, el material se deforma permanentemente y sigue fluyendo en este esfuerzo.

**Lineal Elástico Perfectamente-Plástico:** Se hace la suposición que la respuesta de la relación esfuerzo-deformación puede ser representada mediante dos líneas rectas (Figura 2.7D), para describir una rigidez lineal elástica inicial y el esfuerzo de fluencia o resistencia en la falla durante la deformación plástica. Los materiales se comportan elásticamente para esfuerzos menores que el de fluencia, para luego deformarse sin límite en el esfuerzo de fluencia. Un material que exhibe tal respuesta puede ser considerado perfectamente dúctil.

**Perfectamente Frágil:** Materiales que exhiben una respuesta similar a la Figura 2.7E, son llamados frágiles. Para estos materiales, la curva esfuerzo-deformación es cercana a la linealidad en todos los niveles de de estrés, incluso hasta el esfuerzo final de fracturamiento. Un material perfectamente frágil no exhibe para nada no-linealidad. La falla frágil es el proceso mediante el cual pérdida repentina de resistencia ocurre luego de una deformación casi imperceptible.

**Strain Softening:** Cuando el esfuerzo ha excedido el límite elástico, fluye por el fracturamiento sin perder toda la cohesión. Continuando su fluencia hasta alcanzar su resistencia peak antes de que comience a fallar. Luego, las propiedades resistivas se reducen pero el material aún tiene capacidad de resistencia residual (Figura 2.7F).

Según Hoek (2001), usualmente macizos rocoso de muy buena calidad geotécnica y roca dura como

granitos masivos o cuarcitas, bajo condiciones de altos esfuerzos se comportan de manera elástica y frágil (Figura 2.7E). Mientras que para macizos de calidad regular es razonable suponer que las características de su comportamiento post-falla pueden ser estimadas reduciendo el valor del índice GSI, desde el valor in situ a un valor más bajo, el cual caracteriza al macizo rocoso quebrado (Figura 2.7F). Y para macizos de mala calidad geotécnica se sugiere que las características post-falla queden representadas al suponer que el macizo se comporta en forma perfectamente plástica (Figura 2.7D). Esto significa que continúa deformándose bajo un nivel constante de esfuerzos, y que no se asocia ningún cambio de volumen con esta falla progresiva.

## b) Criterios de falla

El criterio de Mohr Coulomb expresa la resistencia al corte a lo largo de un plano ( $\theta$ ) en un estado triaxial de tensiones, obteniéndose la relación entre los esfuerzos normal y tangenciales actuales en el momento de la rotura mediante la expresión matemática de la Ecuación 2.5:

$$\tau = c + \sigma_n \tan \phi \quad (2.5)$$

Donde:

- $\tau$  y  $\sigma_n$  son las tensiones tangencial y normal sobre el plano de rotura.
- $c$  y  $\phi$  son la cohesión y ángulo de rozamiento de la matriz rocosa.

El criterio puede expresarse igualmente en función de los esfuerzos principales  $\sigma_1$  y  $\sigma_3$  según la Ecuación 2.6:

$$\sigma_1 = \frac{2c + \sigma_3(\sin 2\theta + \tan \phi(1 - \cos 2\theta))}{\sin 2\theta - \tan \phi(1 + \cos 2\theta)} \quad (2.6)$$

El criterio de Mohr Coulomb implica que al alcanzarse la resistencia *peak* del material, se forma una fractura al corte. La gran ventaja del criterio es su sencillez, sin embargo presenta variados inconvenientes entre los que González de Vallejo (2002) menciona:

- Las envolventes de la resistencia en roca no son lineales; se ha comprobado experimentalmente que la resistencia de las rocas aumenta menos con el incremento de la presión normal de confinamiento que lo obtenido al considerar una ley lineal.
- La dirección del plano de la fractura según este criterio no siempre coincide con los resultados experimentales.
- El criterio sobrevalora la resistencia a la tracción.

Por estas razones, otros criterios de esfuerzo *peak* se prefieren para roca intacta. Sin embargo, usando el criterio de Mohr Coulomb, se puede encontrar una buena representación de las condiciones de esfuerzo residual. Particularmente, presenta un buen ajuste del esfuerzo de corte en discontinuidades en roca, como se mostrará más adelante en el modelo constitutivo Ubiquitous Joint.

Por otro lado, el criterio de Hoek Brown fue derivado del resultado de la investigación de falla frágil de roca intacta por Hoek (1965) y en estudios del comportamiento de modelos de macizo rocoso con estructuras hechos por Brown (1970). La teoría de falla frágil fue publicada por Griffith (1924), modificada por McClintock y Walsh (1962) para dar cuenta sobre la fricción en el deslizamiento de superficies, formando la base para el criterio de falla no lineal por roca intacta publicada por Hoek y Brown (1980). La última actualización del criterio fue publicada el 2018, la cual incorpora todas las modificaciones que se han realizado en los últimos años, basadas en la experiencia ganada con la aplicación de este criterio en problemas prácticos. Por otro lado, el *geological strength index* (GSI) es un sistema de caracterización del macizo rocoso desarrollado por Hoek (1994) y Hoek et al. (1995),

para vincular el criterio de falla con observaciones ingenieriles geológicas en terreno.

El criterio generalizado de Hoek y Brown para la estimación de la resistencia de un macizo rocoso, introducida por Hoek (1994) y Hoek et al. (1995), se expresa según la Ecuación 2.7:

$$\sigma_1 = \sigma_3 + \sigma_{ci} \left( m_b \frac{\sigma_3}{\sigma_{ci}} + s \right)^a \quad (2.7)$$

Donde  $m_b$ ,  $s$  y  $a$  son constantes del macizo rocoso, dadas por:

$$m_b = m_i \exp[(GSI - 100)/(28 - 15D)] \quad (2.8)$$

$$s = \exp[(GSI - 100)/(9 - 3D)] \quad (2.9)$$

$$a = 1/2 + 1/6(e^{-GSI/15} - e^{-20/3}) \quad (2.10)$$

Para roca intacta las constantes del material son denotadas por  $m_i$ ,  $s=1$  y  $a=0.5$ .  $D$  es un factor que depende del grado de disturbancia que el macizo rocoso ha sido expuesto, puede deberse a daño por tronadura o por relajación de esfuerzos.

### c) Ubiquitous Joint

Para incluir la presencia de estructuras del macizo rocoso, se utiliza el modelo constitutivo Ubiquitous Joint, el cual da cuenta de la presencia de planos de debilidad en un modelo con Mohr Coulomb. El criterio de falla en el plano, del cual se tiene la orientación, consiste en una envolvente de Mohr Coulomb compuesta con un *cutoff* en la zona de falla por tensión. Según la documentación de Itasca (Itasca Consulting Group, Inc., 2023), la posición de un punto de esfuerzos en la envolvente de falla última, está controlada por diferentes reglas de flujo dependiendo del tipo de falla que está ocurriendo. Para fallas por corte, se utiliza una regla de flujo no asociativa, donde la dirección de la deformación plástica no coincide con el gradiente de la superficie de fluencia. Por otro lado, para fallas por tensión se utiliza una regla de flujo asociativa, donde la dirección de la deformación plástica coincide con el gradiente de la superficie de fluencia.

En este modelo numérico, primero se detecta la falla general y se aplican las correcciones plásticas relevantes, según el modelo constitutivo de Mohr-Coulomb. Luego, los nuevos esfuerzos se analizan para fallas en el plano débil y se actualizan respectivamente. Las propiedades de la matriz de la roca y de las estructuras deben especificarse dentro del modelo, como se muestra en la Figura 2.8. En la literatura existen verificaciones de la respuesta de este modelo, comparándola con la solución analítica de un solo plano de debilidad desarrollada por Jaeger (1960) y luego descrita por Jaeger & Cook (1976). La solución establece que el deslizamiento de la superficie ocurrirá siempre que  $1 - \tan\phi \tan\beta > 0$  y se cumpla la siguiente condición:

$$\sigma_1 > \sigma_3 + \frac{2(c' \sigma_3 \tan \phi)}{(1 - \tan \phi \tan \beta) \sin 2\beta}$$

Donde  $c$  es la resistencia cohesiva de la discontinuidad,  $\phi$  es el ángulo de fricción y  $\beta$  es el ángulo de la discontinuidad. La solución simple sólo permite dos modos de falla independiente: falla a lo largo de la discontinuidad y falla de la roca intacta. La Figura 2.9 compara los resultados de un modelo Ubiquitous Joint implementado en *FLAC3D* con la solución de Jaeger (1960). Se puede observar una estrecha coincidencia con las predicciones analíticas de resistencia a la compresión.

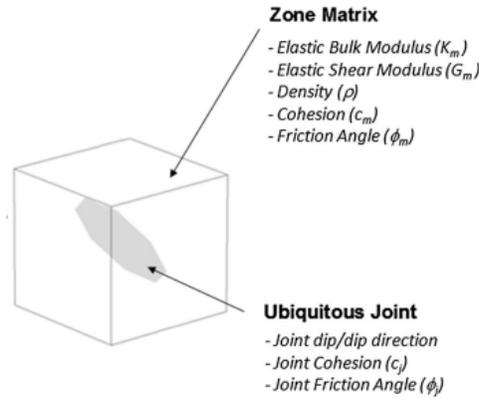


Figura 2.8: Matriz y propiedades Ubiquitous-Joint (Modificado de Sainsbury & Sainsbury, 2017).

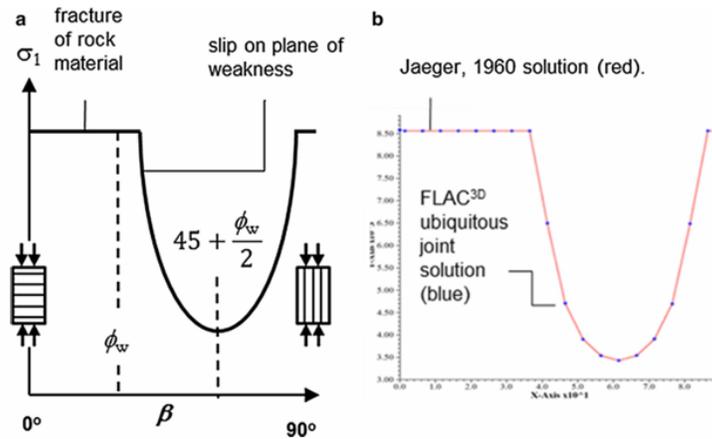


Figura 2.9: Comparación de la resistencia a la compresión uniaxial entre el modelo Ubiquitous Joint y la solución analítica de Jaeger (1960). (Sainsbury & Sainsbury, 2017).

### 2.1.2.3. Criterios para interpretar el daño

Los softwares de cálculo numérico no son herramientas automáticas que proporcionan soluciones instantáneas a los problemas planteados. Los desplazamientos y velocidades generados por los modelos numéricos, deben ser cuidadosamente analizados para interpretar el comportamiento que representan en la realidad. Entre los análisis necesarios, se incluye la evaluación de si estos parámetros están aumentando, manteniéndose estables o disminuyendo. Por ejemplo, aumento en los desplazamientos y velocidades indican una condición de inestabilidad, mientras que desplazamientos constantes y velocidades decrecientes sugieren una condición estable. También, otro análisis que se puede examinar corresponde a estado de falla (plasticidad) del modelo, donde la falla puede definirse en tensión o corte en la roca o en las discontinuidades.

#### a) Criterios de Velocidad

Según Wyllie et al. (2004), los autores encontraron que velocidades bajo  $1e-6$  indican estabilidad en *FLAC* y *FLAC3D*; en cambio, velocidades sobre  $1e-5$  indican inestabilidad. Esta velocidad no tiene

unidades, ya que no son velocidades reales, debido a la amortiguación y el escalamiento de la masa usada para alcanzar la solución estática. Mientras que los desplazamientos son reales, las velocidades no lo son, y no existe información en el 'tiempo' que los desplazamientos ocurren.

## b) Criterios deformacionales

Dado que las deformaciones plásticas están acompañadas por cambios permanentes en las posiciones de los átomos, las deformaciones plásticas no pueden ser definidas únicamente en términos del estado actual de esfuerzos. Las deformaciones plásticas dependen de la historia de los esfuerzos, por lo que la teoría de la plasticidad debe utilizar un enfoque de carga incremental en el que las deformaciones incrementales se suman para obtener la deformación plástica total. En algunos problemas de ingeniería, las deformaciones plásticas son mucho más grandes que las deformaciones elásticas, que pueden incluso despreciarse. Este no es siempre el caso para deformaciones de rocas, por lo que un análisis elastoplástico puede ser necesario (Brady et al., 2006).

El incremento de deformación total (*total strain increment*)  $\{\dot{\epsilon}\}$  es la suma de los incrementos de deformación elástica y plástica, y su cálculo es el mostrado en la Ecuación 2.11.

$$\{\dot{\epsilon}\} = \{\dot{\epsilon}^e\} + \{\dot{\epsilon}^p\} \quad (2.11)$$

El criterio *Strain Increment, Total Measure* descrito por Cavieres (2003) da cuenta de la suma de las deformaciones por corte y tensión, calibrado con mediciones reales de fracturamiento de rocas en la mina El Teniente. Se define como un criterio que mide el límite de fracturamiento, y fue calibrado con modelos Mohr Coulomb, presentando buen ajuste con datos reales medidos en terreno cuando se utiliza el límite mostrado a continuación en la Ecuación 2.12.

$$\epsilon_t = \sqrt{\sum \epsilon_{ij}^2} = 0.5\% \quad (2.12)$$

### 2.1.3. Elección de herramientas a usar

Finalmente, se decide utilizar un modelo continuo al ser más eficiente computacionalmente, específicamente *FLAC3D* debido a que utiliza soluciones optimizadas y de subprocesos múltiples para simulaciones rápidas, sensibles y precisas. FISH, un lenguaje de programación usado para herramientas numéricas desarrollado por Itasca Consulting Group, junto con bibliotecas de Python, brindan una efectiva creación de secuencias de comandos para los modelos. *FLAC3D* entrega gran flexibilización en el modelamiento al ser altamente adaptable, además se pueden utilizar secuencias de comandos FISH y/o Python para la parametrización de los modelos, lo que es sumamente útil para el trabajo a realizar. Por otro lado, se define la utilización de un criterio de falla deformacional, basado en un modelo constitutivo elasto-plástico para macizos rocosos con calidad geotécnica regular a mala.

## 2.2. Modelos de Aprendizaje Automático

El aprendizaje automático se trata de extraer conocimiento de los datos. Es un campo de investigación en la intersección de la estadística, la inteligencia artificial y la informática, y también se conoce como análisis predictivo o aprendizaje estadístico. La aplicación de métodos de aprendizaje automático se ha vuelto ubicua en la vida cotidiana en los últimos años. Desde recomendaciones automáticas de qué películas ver, qué comida pedir o qué productos comprar, hasta el reconocimiento de tus amigos en tus fotos, muchos sitios web y dispositivos modernos tienen algoritmos de aprendizaje automático en su núcleo. Cuando se miran sitios web complejos como Facebook, Amazon o Netflix, es muy probable que cada parte del sitio contenga múltiples modelos de aprendizaje automático (Müller & Guido, 2016).

En simples palabras, el aprendizaje automático es una rama de la inteligencia artificial que permite a las máquinas aprender de los datos y hacer predicciones o tomar decisiones sin ser programadas explícitamente (Samuel, 1959). Funciona mediante la creación de modelos matemáticos basados en datos de entrenamiento. Estos modelos identifican patrones y relaciones en los datos, que luego se utilizan para hacer predicciones o clasificaciones sobre nuevos datos.

### 2.2.1. Conceptos básicos

Los modelos de *machine learning* pueden ser clasificados según la cantidad y tipo de supervisión que se le da durante el entrenamiento. Dentro de las categorías clásicas se encuentra el aprendizaje supervisado y el no supervisado. El primero corresponde cuando la base de datos de entrenamiento incluye la solución o valor buscado, llamado etiqueta. Y la segunda categoría, es cuando no se tienen las etiquetas correspondientes. En la Figura 2.10 se observa las dos clasificaciones antes mencionadas, con sus típicas aplicaciones a problemas.

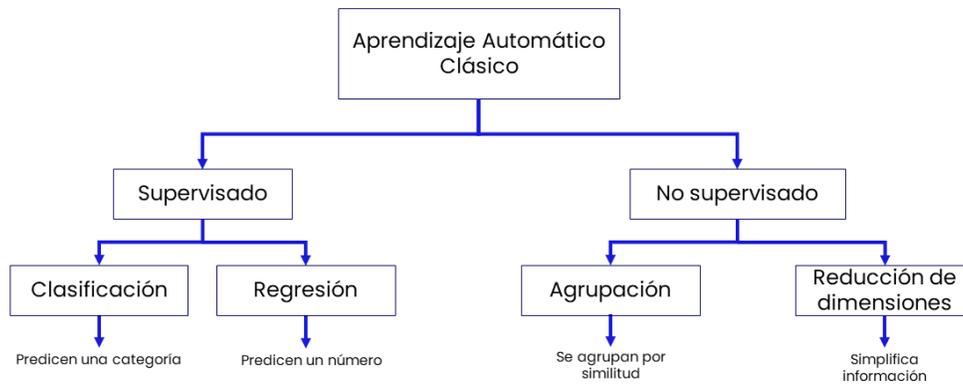


Figura 2.10: Tipos de aprendizaje automático clásico (Elaboración propia).

Otro de los conceptos fundamentales en el *machine learning* son los datos, ya que son la base sobre la cual se construyen, entrenan, validan y evalúan los modelos. Los datos pueden ser estructurados o no estructurados: Los primeros, son datos organizados en tablas, como bases de datos. Los segundos, son datos en formato de texto, imágenes, audio, entre otros.

Antes de comenzar a entrenar un modelo, se debe aplicar un conjunto de técnicas y pasos a la base de datos para mejorar la calidad y la eficacia del modelo de aprendizaje automático. La normalización y estandarización son técnicas que consisten en escalar los parámetros de entrada, para que tengan una distribución similar y sea más sencillo y rápido para el modelo procesarlos. Otra técnica es la imputación de datos, la cual trata del proceso de rellenar o estimar valores faltantes en un conjunto de datos, ya que muchos algoritmos de aprendizaje automático no pueden manejar datos incompletos. Otro paso previo del entrenamiento es la codificación de variables categóricas, la cual trata del proceso de transformar variables que representan categorías o etiquetas, en un formato numérico que los algoritmos de aprendizaje automático pueden entender. Las técnicas antes mencionadas se denominan preprocesamiento de la información, y consisten en la limpieza, normalización, selección de características y otras transformaciones a la base de datos que se utilizará posteriormente en el entrenamiento del modelo.

El entrenamiento de un modelo de *machine learning* es el proceso mediante el cual una máquina aprende a realizar tareas específicas mediante el análisis de datos. Primero, se recolecta y prepara

un conjunto de datos de entrenamiento que contiene información relevante sobre el problema que se quiere resolver. Luego este conjunto de datos se separa en dos: un conjunto de entrenamiento y otro de prueba. Con el conjunto de entrenamiento, se elige un algoritmo adecuado para el problema. Este algoritmo se usa para crear un modelo matemático que pueda aprender de los datos. Durante el entrenamiento, el modelo analiza los datos de entrada y ajusta sus parámetros internos para minimizar el error entre las predicciones del modelo y las etiquetas reales. Este proceso de ajuste se realiza iterativamente, hasta que el modelo alcanza un nivel de precisión aceptable.

Después del entrenamiento, el modelo se evalúa usando un conjunto de datos de prueba que no se utilizó durante el entrenamiento. Esta evaluación permite verificar que el modelo pueda generalizar bien a nuevos datos y no solo memorizar los ejemplos del conjunto de entrenamiento. Si el modelo funciona bien en los datos de prueba, se considera listo para hacer predicciones sobre datos nuevos y desconocidos. Sin embargo si al momento de evaluar el modelo, existen problemas que afectan la capacidad para generalizar correctamente a nuevos datos, puede deberse a dos problemas comunes. Por ejemplo, si el modelo se ajusta demasiado a los datos de entrenamiento y no generaliza bien a datos nuevos, se denomina *Overfitting* (sobre ajuste). Mientras que si el modelo no logra ni siquiera capturar los patrones en los datos de entrenamiento, el problema se llama *Underfitting* (ajuste insuficiente).

Existen diferentes métricas para evaluar el desempeño dependiendo el tipo de modelo que se utilice. En este sentido, evaluar un modelo de clasificación es usualmente más difícil que uno de regresión (Müller & Guido, 2016). Por esta razón las siguientes métricas se centrarán en las mejores para evaluar los modelos de clasificación. Una de las mejores formas de evaluar el desempeño de un modelo predictivo de categorías, es mediante una matriz de confusión (Figura 2.11), donde la idea general de este método es contar las veces que una clase A fue clasificada como una clase B.

<b>Valores Predichos</b>	Verdaderos Positivos (TP)	Falsos Positivos (FP)
	Falsos Negativos (FN)	Verdaderos Negativos (TN)
	<b>Valores Reales</b>	

Figura 2.11: Matriz de confusión (Elaboración propia).

Cada fila de la matriz representa la clase verdadera, mientras que las columnas representan la clase predicha. Es decir, la diagonal de la matriz representa las clases que fueron correctamente clasificada y la zona superior e inferior de la matriz valores mal clasificados. Esta matriz ayuda a identificar no solo cuántas predicciones fueron correctas, sino también los tipos de errores que está cometiendo el modelo.

Donde:

- Verdaderos Positivos (TP): Número de casos correctamente predichos como positivos.
- Falsos Positivos (FP): Número de casos incorrectamente predichos como positivos, también

conocidos como 'falsas alarmas'.

- Verdaderos Negativos (TN): Número de casos correctamente predichos como negativos.
- Falsos Negativos (FN): Número de casos incorrectamente predichos como negativos, también conocidos como 'omisiones'.

Si bien la matriz de confusión entrega abundante información, a veces se requiere una métrica más concisa. Para esto se puede utilizar la precisión, *recall* (exhaustividad) o *F1-score* de cierta clase (Ecuación 2.13, 2.14 y 2.15).

- La precisión es una medida de la exactitud de las predicciones positivas de un modelo. Es la proporción de verdaderos positivos (TP) entre todos los ejemplos que el modelo ha predicho como positivos (TP + FP). En otras palabras, mide cuántos de los elementos que el modelo predijo como positivos realmente lo son.
- La exhaustividad mide la capacidad del modelo para encontrar todos los ejemplos positivos reales. Es la proporción de verdaderos positivos (TP) entre todos los ejemplos que realmente son positivos (TP + FN). En otras palabras, mide cuántos de los elementos positivos reales fueron correctamente identificados por el modelo.
- El valor F1 es una medida que combina la precisión y la exhaustividad en una sola métrica, proporcionando una única puntuación que equilibra ambos aspectos.

$$\text{Precisión} = \frac{TP}{TP + FP} \quad (2.13)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.14)$$

$$\text{F1 Score} = \frac{2 \times \text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}} \quad (2.15)$$

Finalmente una de las últimas validaciones que se realiza a un modelo para entender si es robusto, corresponde a la validación cruzada. Este método divide el conjunto de datos de entrenamiento en *K-folds* (en K particiones), para luego hacer predicciones y evaluar cada pliegue usando un modelo entrenado en las particiones restantes. En palabras simples, consiste en dividir el conjunto de datos en varias partes y usar algunas para entrenar el modelo y las restantes para probarlo. Este proceso se repite varias veces, de modo que cada parte del conjunto de datos se usa para entrenar y probar el modelo en diferentes rondas. El resultado final es un promedio del rendimiento del modelo en todas las rondas, lo que proporciona una estimación más robusta y confiable de cómo el modelo se comportará con datos nuevos no vistos durante el entrenamiento. Además, los resultados que entrega realizar validaciones cruzadas, pueden dar a entender si el modelo tiene problemas al generalizar, es decir problemas de sobre-ajuste. La validación cruzada es especialmente útil cuando se dispone de un conjunto de datos limitado, ya que permite maximizar tanto el uso de los datos para el entrenamiento, como para la evaluación del modelo.

### 2.2.2. Aplicaciones de Machine Learning en la actualidad

En la actualidad el *Machine Learning* ha ganado expansión y popularización gracias la explosión de datos disponibles y la mejora en la capacidad de procesamiento en los computadores. Estos factores han permitido el uso de técnicas de *machine learning* en problemas más grandes y complejos.

Empresas tecnológicas líderes como Google, Facebook y Amazon comenzaron a utilizar *machine learning* a gran escala en productos comerciales, como motores de recomendación, asistentes virtuales y vehículos autónomos. Se ha integrado además en numerosas aplicaciones diarias, desde asistentes virtuales hasta diagnósticos médicos, finanzas y más.

En la Tabla 2.2, se pueden observar algunos trabajos del área de geomecánica en que se utilizaron técnicas de *machine learning*. De la tabla se pueden observar varios puntos, el primero es que la mayoría de los trabajos entrenan sus modelos con una base de datos acotada, menor a 300 datos, donde el trabajo que más datos utilizó fue el de Castro (2023). Otro punto interesante de la tabla, es que en la mayoría de los trabajos, el modelo de random forest fue el más eficiente para predecir los respectivos objetivos deseados.

Tabla 2.2: Resumen herramientas predictivas que usan *machine learning* en geomecánica.

Autor/Referencia	Método*, **	Objetivo/Output	#inputs	#datos	capacidad predictiva**
Castro, 2023	ANN	probabilidad de inestabilidad	19	3810	Exactitud 75.5 %
Jorquera et al., 2023	<b>RF</b> , SVM, KNN	Dilución pared colgante	11	752 (Sudamérica)	AUC 0.942 precisión 84 %
Adoko et al., 2022	ANN	Estabilidad pared caserón (clasificación)	3	225	Exactitud 91 %
Szmigiel & Apel, 2022	LR, <b>RF</b>	Estallidos de roca	7	176 (base de datos Potvin)	AUC 0.96 Exactitud 95 %
Li et al., 2021	KPCA, SVM (Híbrido)	Estabilidad caserón (clasificación)	3	104	Exactitud 92 %
Erdogan Erten et al., 2021	<b>ANN</b> , KNN, NB, SVM, DT	Estabilidad caserón (clasificación)	7	215 (Australia)	RMSE 3.25
Qi & Chen, 2018	RF	Profundidad de falla	13	125	Coef. Corr 0.93
Qi et al., 2018a	RF	Estabilidad pared colgante	13	115 (Capes, 2009)	AUC 0.884
Qi et al., 2018c	LR, MLPNN, DT, <b>GBM</b> , SVM	Dilución pared colgante (clasificación)	13	112 (Capes, 2009)	AUC 0.855
Zhou et al., 2015	LR, MLPNN, SVM, <b>RF</b> , GBM	Estabilidad pilares	6	251	Precisión 59.5-82.4 %
Dong et al., 2013	<b>RF</b> , SVM, ANN	Estallidos de roca	-	36	MR 0 %
Zhou et al., 2011	FDA, SVM	Estabilidad pilares	6	40	-
Wang et al., 2002	ANN	Luz excavaciones subterráneas	2	292 (Canadá)	-

\* en negrita se encuentra el método más eficiente

\*\*RF: Random Forest, SVM: Support Vector Machine, KNN: K Nearest Neighbor, ANN: Artificial Neural Networks

LR: Logistic Regression, KPCA: Kernel Principal Component Analysis, NB: Naive Bayes, DT: Decision Tree,

MLPNN: Multilayer Perception Neural Network, GBM: Gradient Boosting Machine, FDA: Fisher Discriminant Analysis

\*\* capacidad predictiva según el método más exitoso

Existe una extensa variedad de modelos de aprendizaje, desde modelos simples de implementar hasta modelos muy complejos de entender que trabajan como cajas negras. Los modelos de aprendizaje se pueden implementar en una amplia gama de problemas cotidianos, como también problemas específicos de diferentes áreas de ingeniería como los mostrados en este capítulo relacionados a la geomecánica. Esto muestra la versatilidad y el potencial de los modelos para abordar desafíos en diferentes campos de estudio. Sin embargo, todavía existen desafíos significativos en la aplicación de aprendizaje automático en la geomecánica, siendo uno de los más relevantes la necesidad de conjuntos más grandes de datos de alta calidad. En general la capacidad predictiva de los modelos depende de la cantidad de datos con la que el modelo aprende, mientras más cantidad de datos se tengan, mejor deberían ir respondiendo sus métricas. Al ser los datos de campo complejos de obtener, en la actualidad las bases de datos sintéticas generadas por modelos numéricos surgen como solución a estos desafíos. Por otro lado, el uso de bases de datos sintéticas podría permitir explorar el potencial de las herramientas de *machine learning* en el área de la geomecánica.

# Capítulo 3: Metodología

En la Figura 3.1, se muestra un esquema de la metodología propuesta a seguir para cumplir el objetivo general y los objetivos específicos. Además a continuación se describen las tres partes principales en las que se separa el trabajo:

- En primer lugar, de los antecedentes de estudio se definen los parámetros relevantes para la construcción del modelo numérico paramétrico. Además los potenciales usos del machine learning en el área de geomecánica.
- En la segunda parte, se aborda el modelamiento numérico, cuyo objetivo es crear una representación simplificada de la geometría de un caserón aislado. Una vez logrado este paso, se procede a verificar los resultados utilizando casos de estudio, comparándolos con las clasificaciones de gráficos empíricos.
- Finalmente, en la tercera parte se desarrolla un modelo predictivo entrenado a partir de los resultados del modelo numérico, validando y verificando sus resultados para entender si predice de forma correcta.

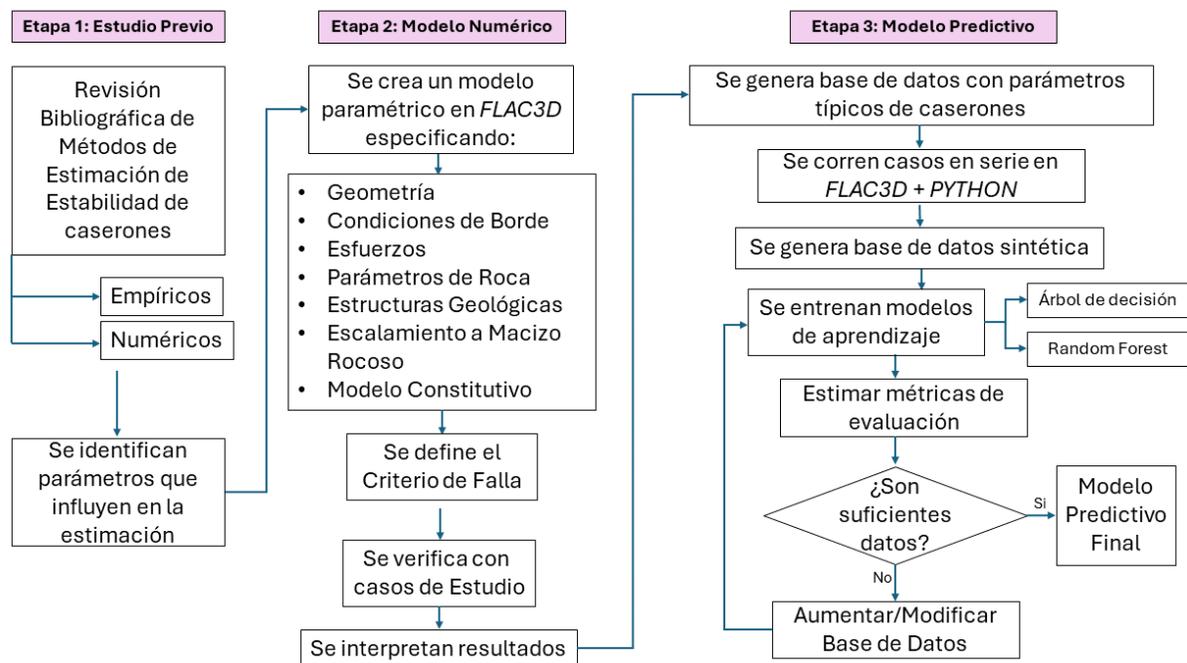


Figura 3.1: Metodología del trabajo.

## 3.1. Representación del caserón en Modelamiento Numérico

A continuación se describe la metodología a seguir para la construcción de un caserón de forma paramétrica. La idea es dejar todas las variables en función de los parámetros inputs para la configuración de un caserón. Con esto, más la implementación de códigos en Python y FISH se genera

un código que corra automáticamente todos los casos, sin necesidad de cambiar manualmente las variables.

### 3.1.1. Consideraciones del modelo

#### a) Elección modelo 2D versus 3D

Los modelos bidimensionales son una herramienta eficaz para ahorrar tiempo y recursos computacionales, sin embargo, no es posible aplicar esta metodología de forma general. Usualmente se utiliza cuando existe una geometría simétrica, y se pueden asumir condiciones de deformaciones y esfuerzos planos. Sin embargo, estas condiciones no aplican al caso a resolver, donde se tiene un caserón que no es infinito (para asumir deformaciones planas). Además, se busca conocer el volumen de material que falla en la pared en estudio, por lo que necesariamente se debe utilizar una configuración 3D.

#### b) Geometría

Para estudiar el efecto de la geometría del caserón en la respuesta de la dilución, se varía la inclinación, el alto, el ancho y el largo (en función de la razón de aspecto) del caserón. No se utiliza un rango de valores directos de el largo del caserón, para tener casos con geometrías reales y típicas. Para definir los intervalos a trabajar, se toma como referencia la base de datos Canadiense de caserones abiertos sin soporte, explotados por *Sublevel Stopping*.

Es importante mencionar que de la base de datos Canadiense, solo se consideró el rango menor y mayor, mientras que para la distribución, no se consideró ninguna en particular, con el objetivo de representar todos los posibles escenarios por igual. Los rangos de valores en que se estudia el problema, además de los intervalos para realizar el muestreo se dispone en la Tabla 3.1, mostrada a continuación.

Tabla 3.1: Parámetros geométricos de diseño.

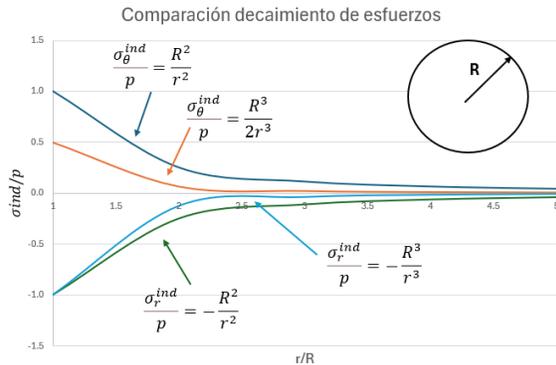
Parámetros	Min	Max
Ancho [m]	5	60
Alto [m]	10	150
Razón de aspecto	0.4	7.4
Inclinación [°]	45	90

#### c) Zona de influencia

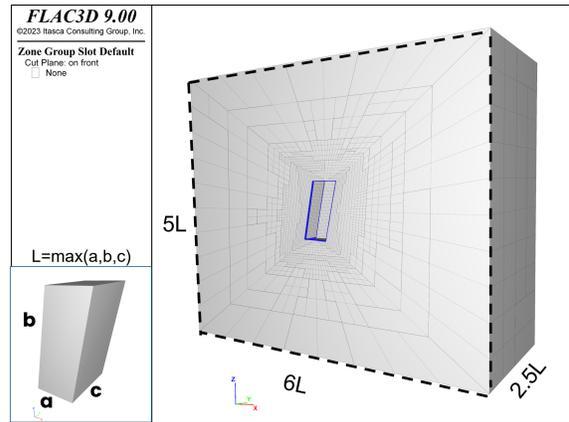
Al determinar la geometría que rodea al caserón, que corresponde a la caja exterior, se utiliza la aproximación según la solución de Kirsch. Esta solución propone que a medida uno se aleja de los bordes de la excavación, los esfuerzos se van acercando a los in situ. En la Figura 3.2.a, se compara la disminución de esfuerzos inducidos para una solución circular 3D y una 2D. Según Kirsch (1898), en la solución 2D a una distancia de tres radios del borde de la excavación, los esfuerzos inducidos corresponden a un 10 % de los in-situ. Mientras que para un modelo 3D los esfuerzos se disipan más rápidamente, y a una distancia de tres radios, los esfuerzos inducidos son prácticamente la mitad de la solución 2D.

Para usar esta primera aproximación, se modela el caserón como si fuese una esfera. Donde el radio de esta supuesta esfera corresponde a la mitad de la mayor distancia entre: el alto, largo y ancho del caserón. Para la caja exterior, finalmente se considera una distancia de tres radios desde el borde de la excavación, esto significa un tamaño de cuatro radios desde el centro del modelo, o dos veces el

diámetro del caserón ( $\max(\text{ancho}, \text{alto}, \text{largo})$ ) desde el centro del modelo. De forma conservadora, se decide utilizar 2.5 diámetros para el alto y el largo y 3 diámetros para el ancho de la caja (debido a la variación entre casos por la inclinación), como se observa en la Figura 3.2.b.



(a) Decaimiento esfuerzos inducidos.



(b) Dimensión de la caja que rodea al caserón.

Figura 3.2: Metodología para el dimensionamiento del modelo (Elaboración propia).

## d) Tamaño de zona adecuado

El tamaño de malla de un modelo numérico, se refiere a la dimensión de los elementos que componen la malla utilizada en las simulaciones. La malla divide el dominio del problema en pequeñas subregiones (elementos), permitiendo que se realicen cálculos más precisos al resolver ecuaciones diferenciales.

La densidad de la malla debe adaptarse para capturar adecuadamente las variaciones actuales de esfuerzos y desplazamientos. Usualmente los gradientes de esfuerzos son altos alrededor de los bordes de las excavaciones, especialmente en cambios abruptos de la geometría. Por esta razón es recomendable utilizar una mayor densidad de malla en áreas con altos gradientes, para capturar los detalles de manera más precisa. Si la variación de los esfuerzos es baja, no se necesitan muchos elementos para una representación aceptable. Es decir, para reducir los requerimientos de memoria y capacidad computacional, se puede emplear una menor cantidad de elementos en zonas donde los gradientes son bajos.

Existen diferentes recomendaciones de la cantidad de zonas mínimas para representar de forma adecuada diferentes problemas. Por ejemplo, según experiencia de diferentes autores (Wyllie, 2004), se ha encontrado que para taludes, al menos 20 (y preferentemente 30) zonas son requeridas en la altura de interés. Tomando como base esta recomendación, se optó por implementar una malla paramétrica más fina alrededor del caserón en las tres direcciones. Se definió un tamaño de zona igual a la altura del caserón dividida en 30 elementos. Manteniendo esta medida hasta una distancia igual a la mitad de la altura del caserón y luego duplicando su dimensión de forma exponencial hasta alcanzar los bordes de la caja.

Con una distancia máxima para la densificación más fina desde el borde de la excavación igual a la mitad de la altura del caserón, con un tamaño de zona igual a la altura dividida en 30. Además, a medida que las zonas se alejan del centro del modelo, el tamaño de la malla se incrementa de manera exponencial, duplicando el tamaño de la zona anterior.

## e) Condiciones de borde

Las condiciones de borde pueden clasificarse como reales o artificiales. Las condiciones de borde reales corresponden a áreas excavadas, como en el caso de taludes donde representan la superficie del terreno que no presenta esfuerzos in-situ. En contraste, las condiciones de borde artificiales no existen en la realidad; se introducen para simplificar la modelación numérica. En todos los problemas de geomecánica, es necesario truncar la extensión del dominio de estudio para incluir únicamente el área inmediata de interés. Esto se realiza con el fin de reducir el tiempo de cómputo y la complejidad de los modelos. Además, las condiciones de borde artificiales pueden ser de dos tipos: La primera corresponde a condiciones de desplazamiento, las cuales inhiben el desplazamiento en la dirección horizontal, vertical o ambas. En geomecánica, el desplazamiento en la base de los modelos, siempre está restringido tanto en la dirección vertical como en la horizontal, para evitar la rotación del mismo (Wyllie et al., 2004).

La condición y ubicación de esfuerzos de campo lejano, también deben especificarse en los modelos numéricos. La idea general es seleccionar la ubicación del campo lejano de manera que no influya significativamente en los resultados. Para el modelo en estudio, las condiciones de borde para la caja corresponden a caras fijas en todas las direcciones (Figura 3.3), simulando el confinamiento debido a la profundidad del caserón y la roca que lo rodea y confina.

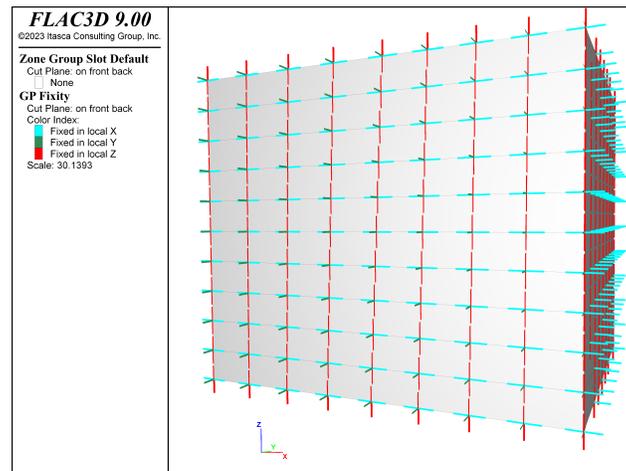


Figura 3.3: Condiciones de borde del modelo (Elaboración propia).

## f) Esfuerzos in situ

Las condiciones iniciales son aquellas que existían antes de la minería. De las condiciones iniciales más importantes se encuentran el campo de esfuerzos in-situ y las condiciones del agua subterránea. Mediciones de esfuerzos han demostrado que el esfuerzo vertical es aproximadamente el peso de la carga litostática. Por consecuencia, grandes superficies irregulares como montañas y valles, influenciarán la distribución de la carga en el macizo rocoso que se encuentra en la base. (Mathews et al. 1981). Por esta razón, en este modelo se decide dejar la carga vertical como la litostática y la horizontal en función de ella mediante el factor de esfuerzos  $k$ . Usando la convención  $k = \sigma_h/\sigma_v$ , donde  $\sigma_v$  es la carga litostática, para profundidades entre 450 a 1075 [m], se varió el  $k$  entre 0.5 a 2, con intervalos de 0.5 como se muestra en la Tabla 3.2.

Tabla 3.2: Parámetros de esfuerzos para el diseño de diseño.

Parámetros	Min	Max
Profundidad [m]	450	1075
Razón de Esfuerzos [ $k = k_x = k_y$ ]	0.5	2

### g) Parámetros de roca y escalamiento a macizo rocoso

El módulo de deformación de un macizo rocoso es un parámetro importante en cualquier análisis de comportamiento que incluya deformaciones. Los ensayos de terreno para determinar directamente este parámetro demandan tiempo, son costosos y, a veces, la fiabilidad de sus resultados es cuestionable. Por esta razón, variados autores han propuesto relaciones empíricas para estimar el valor del módulo de deformación del macizo rocoso (Hoek & Diederichs, 2006). La referencia más actual de esta relación es la planteada por Hoek & Diederichs (2006), basada en datos de terreno obtenidos de minas en China y Taiwán. Los autores definen dos fórmulas: una simplificada (Ecuación 3.1) y una detallada (Ecuación 3.2) para el escalamiento. Para el caso detallado, los autores relacionan el módulo de Young de la roca intacta con el GSI y el factor D. Sin embargo, cuando no se tiene información del módulo de Young de la roca intacta, sugieren utilizar la ecuación simplificada (Figura 3.4). Por otro lado, en trabajos anteriores, en específico por Hoek & Brown (2002), se propone la estimación según la Ecuación 3.3 cuando no se tiene información del módulo de Young de la roca intacta.

$$E_{rm}(\text{MPa}) = 100000 \left( \frac{1 - \frac{D}{2}}{1 + e^{\frac{75+25D-GSI}{11}}} \right) \quad (3.1)$$

$$E_{rm} = E_i \left( 0.02 + \frac{1 - \frac{D}{2}}{1 + e^{\frac{60+15D-GSI}{11}}} \right) \quad (3.2)$$

$$E_{rm}(\text{GPa}) = \left( 1 - \frac{D}{2} \right) \sqrt{\frac{\sigma_{ci}}{100}} \cdot 10^{\frac{GSI-10}{40}} \quad (3.3)$$

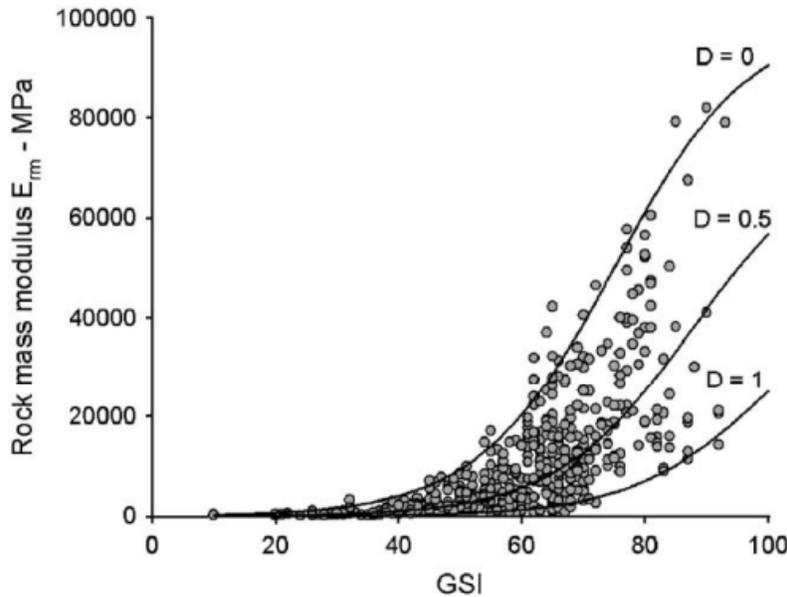


Figura 3.4: Resultados según la ecuación simplificada de Hoek & Diederichs comparada con la base de datos de China y Taiwan (Hoek & Diederichs, 2006).

En las Figuras 3.5 y 3.6 se comparan las Ecuaciones 3.1 y 3.3 respecto a la última versión propuesta de la estimación del módulo de Young (Ecuación 3.2). De estas figuras se puede observar que la Ecuación 3.1 es la más agresiva en cuanto al escalamiento. Esto tiene influencia en los resultados, ya que mientras menor es el módulo de Young, los resultados se verán afectados de forma negativa al criterio de falla sobre-estimando la dilución. Por otro lado, la Ecuación 3.3, es menos agresiva y más cercana a los valores que entrega la Ecuación 3.2, cuando se utilizan los valores promedios de la base de datos a modelar. Por esta razón, se decide utilizar la Ecuación 3.3 como escalamiento a macizo rocoso a partir del UCS, GSI y factor de daño D.

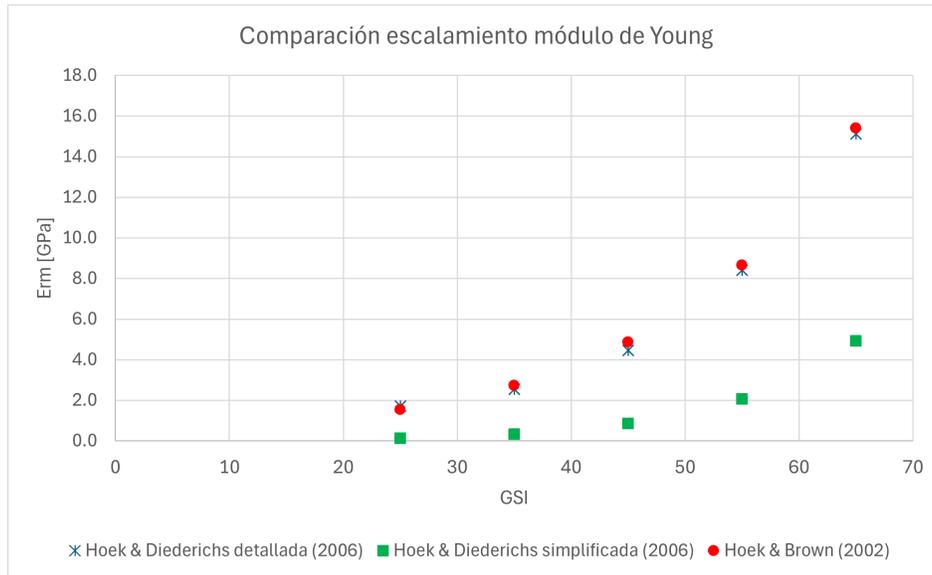


Figura 3.5: Escalamiento de Módulo de Young variando GSI, para UCS 100 [MPa].

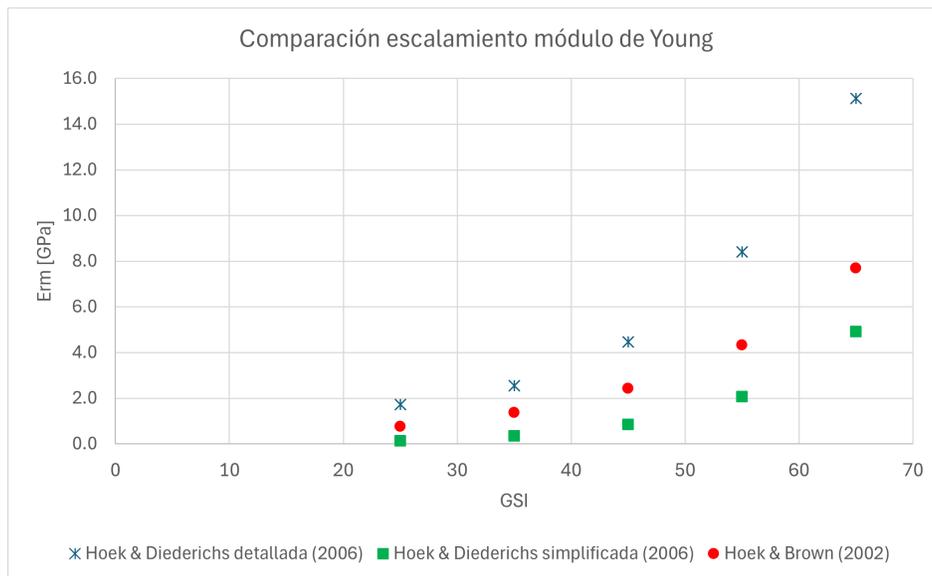


Figura 3.6: Escalamiento de Módulo de Young variando GSI, para UCS 25 [MPa].

Por otro lado, para estimar la razón de Poisson, se utilizó la conversión de Hoek et al. (1995) mostrada en la ecuación 3.4. Con los escalamientos ya definidos, las propiedades que se consideraron

para modelar la roca fueron; la densidad, UCS, GSI y  $\mu$ . Además los intervalos en los que fueron variando son los mostrados a continuación en la Tabla 3.3.

$$\nu = 0.32 - (0.0015 * GSI) \quad (3.4)$$

Tabla 3.3: Parámetros de calidad de roca para el diseño.

Parámetros	Min	Max
Densidad [ $kg/m^3$ ]	1800	3000
UCS [m]	25	225
GSI	25	65
$\mu$	5	30
D	0.7	0.7

## h) Estructuras geológicas

Para modelar las estructuras que se pueden encontrar en el macizo rocoso, se discretizaron 4 categorías que van desde una estructura con malas propiedades (simulando juntas sin fricción), a una categoría con buenas propiedades (simulando roca intacta), como se muestra en la Tabla 3.4. Además, las estructuras antes mencionadas pueden tener las inclinaciones con respecto a la cara de la pared colgante, mostradas en la Tabla 3.5. Estas orientaciones se escogieron según lo propuesto por Mathews et al. (1981), donde el factor B es más desfavorable para un valor de  $\alpha$  igual  $30^\circ$  y menos desfavorable para un  $\alpha$  igual  $90^\circ$  (Figura 2.1.b). Es importante mencionar que las estructuras se incluyen de forma implícita, debido al tipo de modelo constitutivo utilizado Ubiquitous joint, el cual se mencionará en la siguiente sección.

Tabla 3.4: Categorías y propiedades de las estructuras para el diseño.

Categoría	Cohesión [KPa]	Fricción [ $^\circ$ ]
1	0	20
2	50	25
3	100	30
4	1000	40

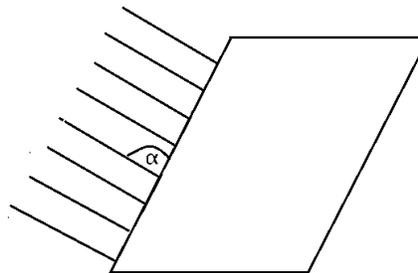


Figura 3.7: Representación ángulo entre estructuras y pared colgante.

Tabla 3.5: Orientación estructuras con respecto a inclinación pared colgante.

Ángulo $\alpha$ [°]
0
30
45
60
90

### i) Modelo Constitutivo

Con el objetivo de incluir la presencia de discontinuidades en el macizo rocoso, se utiliza el modelo constitutivo Ubiquitous Joint. Este modelo permite incluir el efecto de estructuras geológicas y también los casos con roca intacta. Para simular que no existen estructuras, se debe considerar propiedades altas de cohesión y fricción además de orientaciones de  $90^\circ$  con respecto a la pared en estudio.

Para emplear este modelo, pero considerando las propiedades de la roca según el criterio de falla Hoek & Brown, se realiza una conversión a Mohr Coulomb instantánea. Es decir, se van actualizando las propiedades cada cierta cantidad de pasos para que el modelo obtenga los parámetros de Mohr Coulomb según la tangente de la curva no lineal de Hoek & Brown, como se muestra en la Figura 3.8.

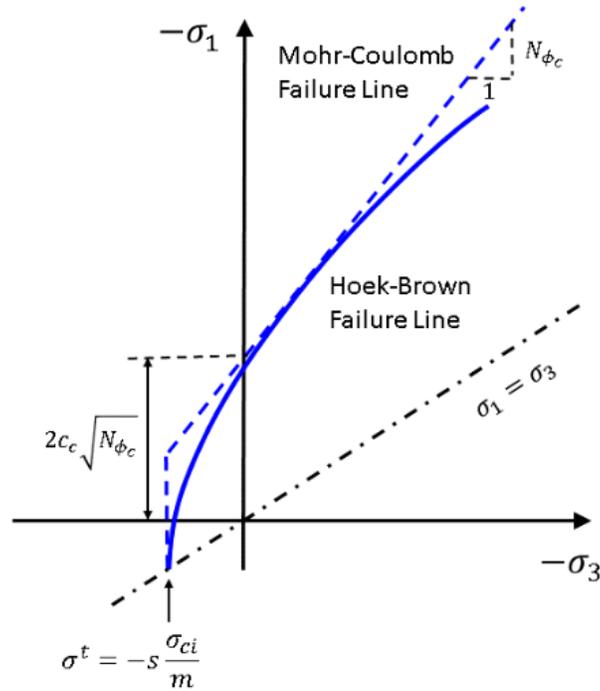


Figura 3.8: Criterio de Hoek-Brown y su aproximación a Mohr Coulomb (Itasca, 2023).

### j) Simulación de la excavación

Para simular la remoción de una gran cantidad de material es necesario primero ejecutar el modelo de manera elástica, para eliminar cualquier efecto inercial causado por la remoción repentina del material. Luego, en esta excavación gradual, lo que se realiza es una disminución progresiva en la

densidad del material, siguiendo cierto número de ciclos. Y después de estos pasos anteriores, se permite que el modelo desarrolle su comportamiento plástico.

### 3.1.2. Interpretación de Resultados

En una primera fase, se evaluaron diferentes criterios, incluyendo el criterio de velocidad y el criterio de límite de fracturamiento. Sin embargo, el criterio de velocidad no presentó resultados satisfactorios conforme a la comparación con casos de estudio. Por otro lado, el criterio de límite de fracturamiento, mostró un comportamiento consistente con lo anticipado, alineándose adecuadamente con los resultados obtenidos de las bases de datos. El criterio de límite de fracturamiento entregó resultados concordes a lo esperado mediante un análisis de sensibilidad realizado variando la profundidad y la inclinación de las estructuras. Además de tener validación de acuerdo a publicaciones y experiencia en proyectos de la de la empresa Itasca.

Para visualizar las zonas que superan el criterio de límite de fracturamiento en los modelos numéricos, se grafica una isosuperficie de las zonas donde este valor es mayor a un 0.5%, como se muestra en las zonas de color rojo en la Figura 3.9. Sin embargo, como sólo interesa el volumen que falla sobre la pared colgante, se crea un código en Python que entregue el volumen de interés en este sector (Figuras 3.10.a y 3.10.b). A continuación, para entender los modos de falla que el modelo logra representar, se estudian dos diferentes casos mostrados en las Figuras 3.11 y 3.12 y sus respectivas propiedades y parámetros geométricos en la Tabla 3.6. A la izquierda de cada imagen se muestra el estado de plasticidad y a la derecha el resultado de las zonas con *strain Increment* mayor a 0.5%.

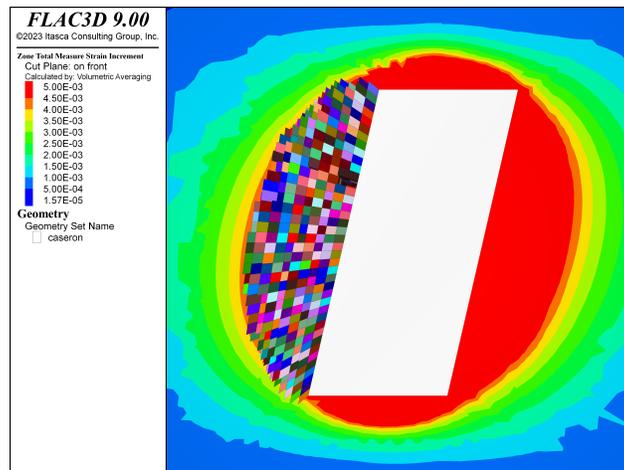
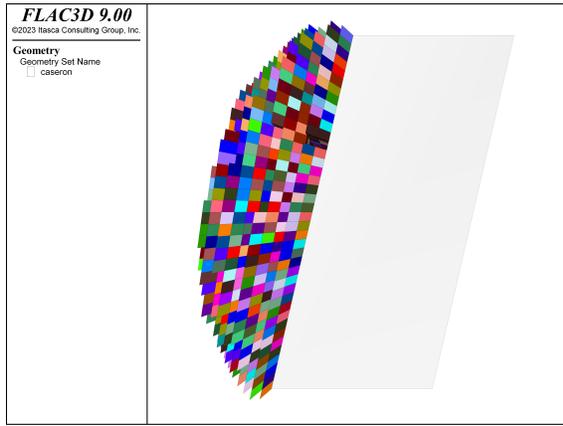
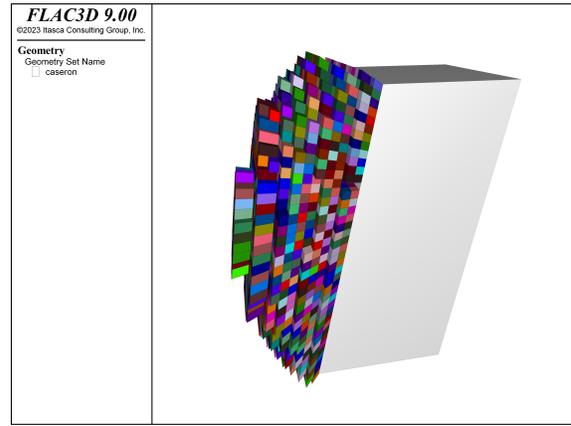


Figura 3.9: Volumen en falla sobre la pared colgante (Elaboración propia).



(a) Vista 2D.

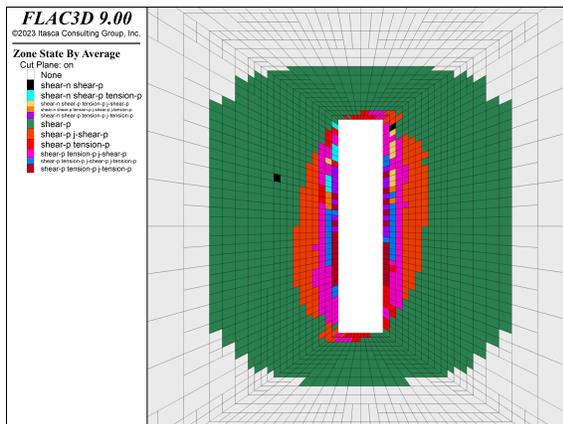


(b) Vista 3D.

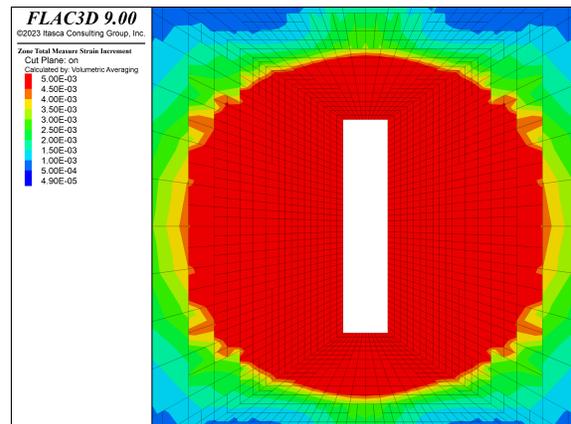
Figura 3.10: Volumen en falla sobre pared colgante (Elaboración propia).

En la Figura 3.11.a, se observa en la cercanía de los bordes de la excavación zonas en estado plástico fallando por corte, tensión en la roca y tensión en las estructuras (zonas rojas, azules y moradas). Luego, alejándose de los bordes se notan zonas fallando por corte en la roca y en las estructuras (zona naranja). Finalmente en la zona más lejana al borde de la excavación (zona verde), se encuentra fallando por corte en la roca. Comparando los resultados con la Figura 3.11.b, el criterio de falla considerado logra representar los estados de falla antes mencionados.

De igual forma, en la Figura 3.12.a se puede notar zonas en estado de corte y tensión en la roca y estructuras (zonas naranjas, azules, y verdes). Comparando estas zonas con la Figura 3.12.b, se nota una correlación entre los resultados del modelo con el criterio de límite de fracturamiento y la influencia de estructuras con bajas propiedades.

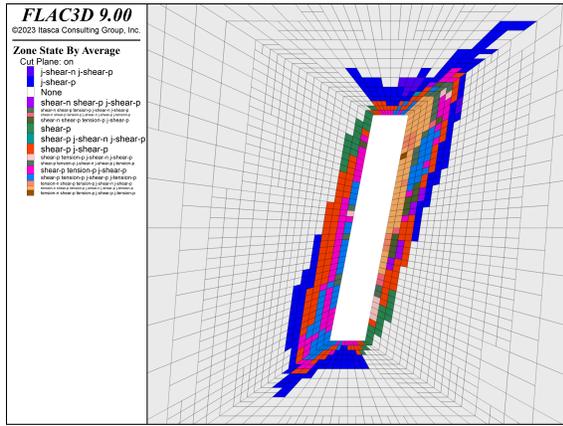


(a) Falla por corte y tensión en la roca y estructuras

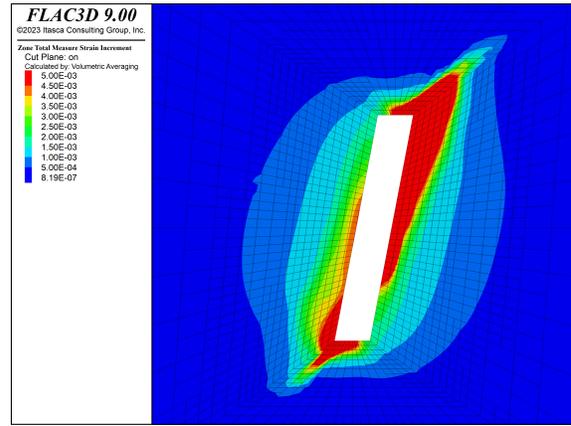


(b) Respuesta del criterio límite de fracturamiento

Figura 3.11: Comparación estado de plasticidad, con el criterio de falla para caso 1 (Elaboración propia).



(a) Falla por corte y tensión en la roca y estructuras



(b) Respuesta del criterio límite de fracturamiento

Figura 3.12: Comparación estado de plasticidad, con el criterio de falla para caso 2 (Elaboración propia).

Tabla 3.6: Parámetros utilizados para modelar casos en estudio, según modo de falla.

Caso	Ancho	Alto	RA	Dip [°]	Prof [m]	k	Dens [kg/m <sup>3</sup> ]	UCS [Mpa]	GSI	mi	Dip Estr [°]	Prop Estr	ELOS [m]
1	9	43	2.3	90	1075	2	2800	115	40	15	30	4	34.8
2	18	115	2.2	79	600	1.5	2100	175	55	11	60	1	0.3

## 3.2. Desarrollo Modelo Predictivo

### 3.2.1. Generación Base de Datos

Ahora que se tienen definidos los 12 parámetros necesarios para los modelos numéricos, se debe construir una base de datos que contenga una combinación de todos ellos para usarla como input en las simulaciones. Para realizar este paso, primero se debe definir el número de casos (combinaciones) entre estos valores. Luego, según los rangos de valores mínimos y máximos de los parámetros y su respectiva distribución, resumidos en la Tabla 3.7, se crea el set de parámetros de entrada para correr los modelos numéricos.

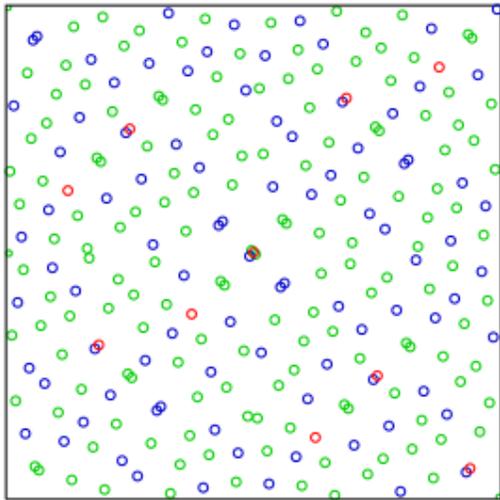
Para crear las combinaciones de los parámetros existen diferentes técnicas. El método de Monte Carlo es un método de muestreo estadístico muy útil para soluciones numéricas de problemas con alta dimensión. Este método calcula integrales de funciones utilizando un conjunto de puntos seleccionados al azar. Dado que este enfoque presenta varias deficiencias (véase, por ejemplo, Morokoff, 1995), en los años cincuenta se propuso una variación de este paradigma, denominada método cuasi Monte Carlo (véanse los detalles en Niederreiter, 1992). Es una variación no aleatoria de la simulación de Monte Carlo, en la que los puntos seleccionados al azar se sustituyen por valores deterministas cuasi-aleatorios (Vesel, 2019). Ejemplos interesantes de generadores cuasi-aleatorios son: la secuencia de Halton (Halton, 1964) y los cuasi-aleatorios más actuales basados en redes neuronales generativas (Hofert, 2019). No obstante, el ejemplo más destacado y comúnmente aplicado de un generador cuasi-aleatorio es la secuencia de Sobol (Sobol, 1967).

Para este trabajo, se empleó la secuencia de muestreo de Sobol. Esta técnica distribuye puntos de muestreo de manera uniforme en un espacio multidimensional, mejorando la eficiencia de las simulaciones y análisis numéricos. En la Figura 3.13 se muestra la diferencia entre un muestreo utilizando la

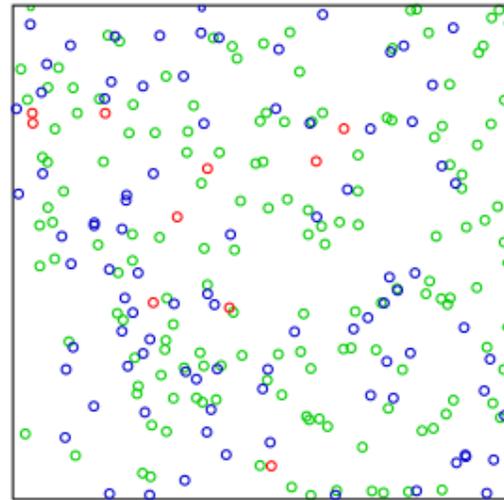
secuencia Sobol para 256 puntos y un muestreo pseudoaleatorio, donde los puntos rojos cubren desde el punto 1 al 10, los azules del 11 al 100 y los verdes del 101 al 256. Se puede notar que la secuencia Sobol logra una cobertura uniforme del espacio, en cambio el muestreo pseudoaleatorio tiende a presentar agrupamientos y huecos. Esta secuencia es especialmente útil para modelos predictivos, ya que permite la expansión progresiva del conjunto de datos, agregando puntos a los generados previamente.

Tabla 3.7: Resumen parámetros input para modelo numérico.

Parámetro	Símbolo	Min	Max	Intervalo	Distribución
Ancho	Ancho [m]	5	60	1	-
Alto	Alto [m]	10	150	1	-
Razón de aspecto	RA	0.4	7.4	-	uniforme
Inclinación	[°]	45	90	1	-
Profundidad	Profundidad [m]	450	1075	25	-
Razón de esfuerzos	k	0.5	2	0.5	-
Densidad	Densidad [kg/m]	1800	3000	100	-
UCS	UCS [MPa]	25	225	5	-
GSI	GSI	25	65	5	-
mi	mi	5	30	1	-
Propiedades Estructuras	PropEstr	1	4	1	-
Orientaciones Estructuras	DipEstr [°]	0	90	15	-



(a) Muestreo con secuencia Sobol.



(b) Muestreo pseudoaleatorio.

Figura 3.13: Comparación en la uniformidad de los datos, dependiendo el método de muestreo utilizado (Jheald, 2011).

De la Figura 3.13, es sencillo notar la uniformidad en el espacio cuando se tienen 2 dimensiones. Sin embargo, el problema que se intenta resolver es de 12 variables. Por esta razón para visualizar la distribución de los puntos utilizados como datos de entrada se disponen en la Figura 3.14. Este gráfico presenta cada variable versus las demás, de forma que es posible visualizar las diferentes dimensiones por separado. De la figura se puede notar una uniformidad para la base de datos, dependiendo la densidad de puntos que se definió simular según la Tabla 3.7.

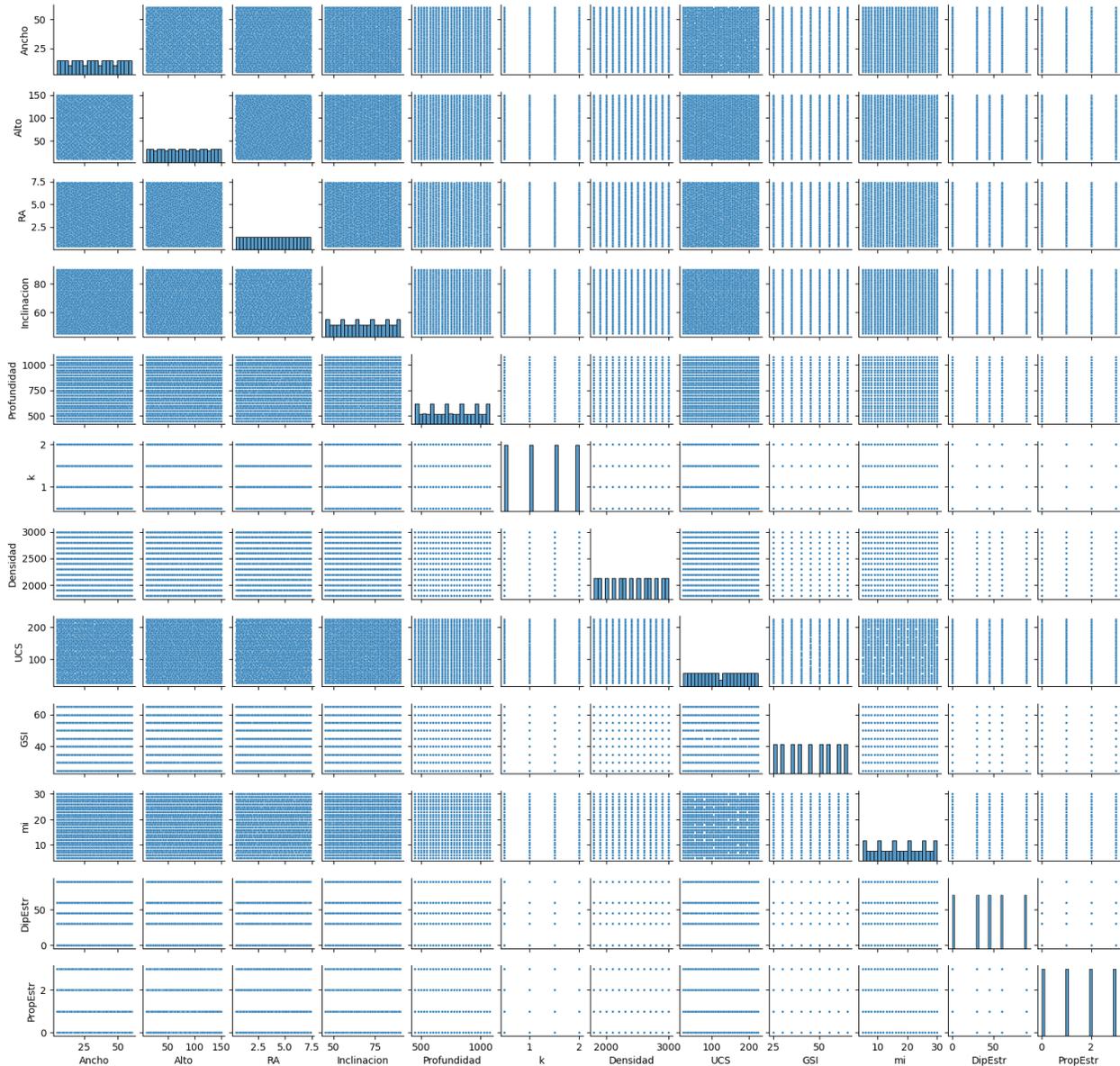


Figura 3.14: Gráfico Pairplot, para visualizar la distribución en el espacio muestral de las variables.

### 3.2.2. Preprocesamiento de datos

Antes de comenzar a entrenar el modelo, se puede hacer una limpieza a los datos eliminando valores atípicos y la realizando una imputación de valores faltantes si fuese necesario. Por ejemplo, si hay casos que den error en el modelo numérico debido a una geometría inviable, se debería eliminar este valor para que el modelo predictivo pueda funcionar correctamente. Además, algunos algoritmos requieren normalizar las variables para hacer sus predicciones, sin embargo, en el caso de árboles de decisión y random forest no es necesario. De todas formas, normalizar las variables para que tengan una escala comparable, puede mejorar el rendimiento de los algoritmos. Por último, se categorizan los resultados según las 4 clases propuestas por Clark (1998), dependiendo de la cantidad de ELOS que tiene cada caso particular.

### 3.2.3. Entrenamiento del modelo de aprendizaje

Debido a su extensa gama de librerías para cargar datos, visualización y análisis estadístico, se ha seleccionado el lenguaje de programación Python. Además, la librería *scikit-learn*, un proyecto de código abierto, ofrece una implementación gratuita y distribuible, permitiendo su acceso universal. La página web de *scikit-learn* proporciona una amplia colección de algoritmos de *machine learning*, acompañada de su documentación detallada. Gracias a su adopción tanto en la industria como en la academia, se ha decidido utilizar Python y *scikit-learn* para la construcción del modelo predictivo. En específico, este trabajo se centrará en el estudio de dos algoritmos de aprendizaje automáticos: árboles de decisión y random forest.

#### Árboles de decisión:

Los árboles de decisión son ampliamente usados para modelos de clasificación y regresión. Ellos aprenden una jerarquía de preguntas «si» y «si no» que les lleva a tomar una decisión. Para la construcción de este algoritmo, se prueban diferentes configuraciones de preguntas, tal que se llegue a la respuesta de la forma más rápida posible. Estas preguntas se llaman pruebas (no confundir con el conjunto de prueba utilizado cuando se entrena el modelo). Usualmente los datos no vienen en el formato de respuesta "si/si no", ya que es común que sean representados como datos continuos. Con datos continuos, se usan pruebas en la forma "¿es la característica  $i$  mayor que el valor  $a$ ?" (Gerón, 2017). El algoritmo puede verse representado como un árbol, donde cada nodo representa una pregunta o un nodo terminal (llamado hoja) que contiene la respuesta final. Desde cada nodo con una prueba, se separan dos posibles respuestas: a la izquierda la respuesta a la pregunta si es verdad y a la derecha si no lo es.

Para entender mejor su funcionamiento, se muestra un ejemplo del algoritmo con la base de datos de Iris, la cual es un conjunto de datos de clasificación multiclase clásico y muy sencillo. Esta base de datos fue introducida por el estadístico y biólogo británico Ronald Fisher en su artículo de 1936 "*The use of multiple measurements in taxonomic problems*". Contiene el largo y ancho de sépalos y pétalos de 150 muestras de tres especies de flores Iris (Iris setosa, Iris virginica e Iris versicolor). Para simplificar el ejemplo, solo se utilizan los parámetros de largo y ancho de los pétalos. En la Figura 3.15 se muestra un ejemplo visual de cómo trabaja el algoritmo árbol de decisión con estos datos continuos. Cada nodo en la forma "¿es la característica  $i$  mayor que el valor  $a$ ?", conlleva a otro nodo o a una hoja. Las hojas representan el final de cierta rama del árbol y es la que entrega una respuesta que puede ser una categoría o un valor continuo, dependiendo si el árbol es de clasificación o regresión. Este mismo algoritmo se puede observar en la Figura 3.16, donde se separan zonas que contienen diferentes clasificaciones de Iris. La zona roja clasifica a los puntos como Iris setosa, la zona amarilla representa a Iris versicolor y la azul a Iris virginica. De esta forma, la predicción de un nuevo punto será según región en la que se encuentre en el gráfico.

Cuando se construye un modelo con este algoritmo, existen diferentes formas de evitar el sobreajuste a los datos de entrenamiento. Una estrategia es parar de forma anticipada la creación del árbol (también llamado pre-poda). Una forma de realizar la pre-poda, consiste en la utilización de criterios, por ejemplo: limitar una máxima profundidad del árbol, definir un máximo número de hojas o el requerimiento de un número mínimo de puntos en un nodo para seguirlo dividiendo.

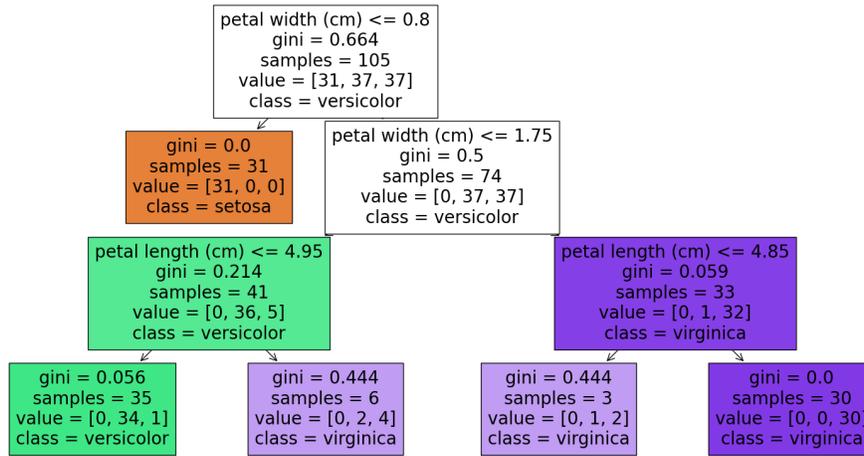


Figura 3.15: Representación visual de algoritmo árbol de decisión con base de datos de Iris (Elaboración propia).

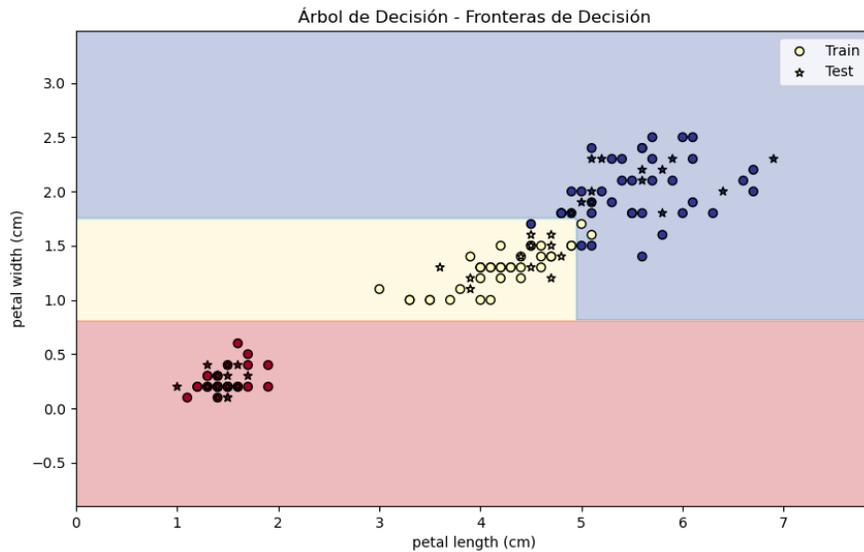


Figura 3.16: Fronteras de decisión para algoritmo árbol de decisión con base de datos de Iris (Elaboración propia).

## Random forest

Random forest es uno de los algoritmos más poderos en el *machine learning*, ya que es un método de ensamblaje de multiples modelos. El random forest es una colección de árboles de decisión, donde cada árbol difiere levemente de los otros. La idea detrás del random forest es que cada árbol hace relativamente un buen trabajo prediciendo, pero posiblemente tendrán sobre-ajuste en cierta parte de la base de datos. Si se construyen varios árboles, todos trabajarán relativamente bien, pero tendrán sobre-ajustes en diferentes partes de la base de datos. De esta forma, se pueden promediar los resultados y reducir el sobre-ajuste en los datos. Para implementar esta estrategia, es necesario construir una cantidad considerable de árboles de decisión. El algoritmo tienen su nombre por lo aleatorio en la construcción de cada árbol. Para asegurar que cada uno sea diferente, existen dos diferentes formas en que los árboles se hacen aleatorios: seleccionando diferentes partes de la base de datos y cambiando el orden de las preguntas en los nodos.

A continuación en la Figura 3.17, se muestra esquemáticamente la construcción de cada uno:

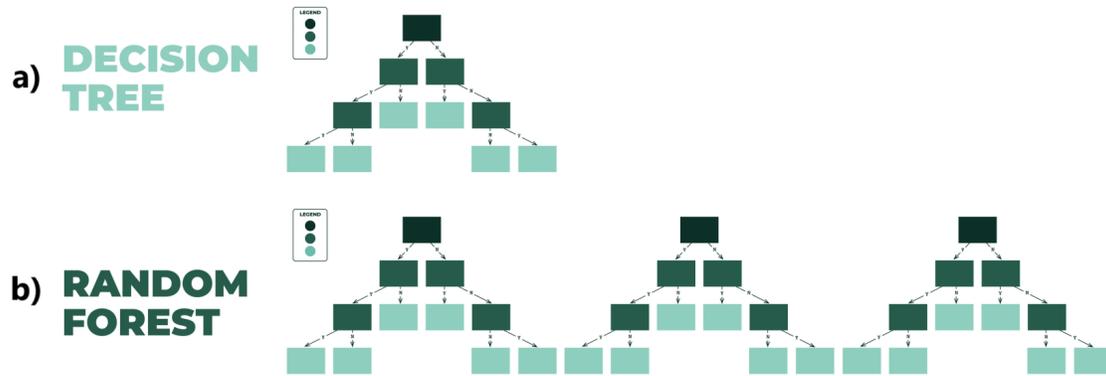


Figura 3.17: a) Esquemización algoritmo de árbol de decisión y b) random forest (Modificado de Khan, 2024).

### 3.2.3.1. Ajuste de parámetros del modelo y uso de GridSearch

Los hiperparámetros son parámetros que se utilizan para controlar el proceso de aprendizaje dentro de los modelos predictivos. En *scikit-learn* se entregan como argumentos al momento de construir los modelos y la idea es buscar la combinación de estos valores que generen el modelo con mejores métricas de rendimiento. Este proceso puede ser complicado y lento, por lo que se recomienda utilizar alguna opción automatizada. Como alternativa a la búsqueda de parámetros por fuerza bruta, existe una forma de automatizar las combinaciones llamada *GridSearch*. Esta técnica es utilizada para optimizar la búsqueda, explorando todas las combinaciones posibles dentro del espacio de búsqueda definido. En *scikit-learn* se ofrecen varios enfoques genéricos para una búsqueda automática, en particular, en este trabajo se utilizó *GridSearchCV* al ser una herramienta sencilla y de rápida aplicación. Sin embargo, existen otras opciones de búsqueda, como *RandomizedSearchCV*, la cual muestrea una cantidad dada de candidatos de un espacio de parámetros con una distribución especificada.

### 3.2.3.2. Algoritmo Árbol de Decisión

Para utilizar este modelo se usó la librería gratuita *scikit-learn*, la cual tiene la estructura siguiente, con los hiperparámetros por defecto mostrados a continuación:

```
1 class sklearn.tree.DecisionTreeClassifier (*, criterion='gini', splitter='best', max_depth=
    ↪ None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0,
    ↪ max_features=None, random_state=None, max_leaf_nodes=None,
    ↪ min_impurity_decrease=0.0, class_weight=None, ccp_alpha=0.0, monotonic_cst=
    ↪ None)
```

El hiperparámetro por defecto para '*criterion*' es '*gini*', pero otra opción es utilizar '*entropy*'. Sin embargo, según Géron (2017) la mayoría de las veces no existe gran diferencia entre usar uno o el otro. Es usual utilizar '*gini*' ya que es un poco más rápida, pero '*entropy*' produce árboles más balanceados. Por esta razón, se sugiere probar ambos hiperparámetros para analizar su efecto en las métricas del modelo.

Para evitar un sobre-ajuste en los datos de entrenamiento, existen estrategias para detener la creación de árboles muy complejos (pre-poda). Dentro de estas técnicas se encuentra la limitación de la profundidad máxima del árbol, o definir criterios de valores mínimos para no seguir agregando ramas al árbol. A continuación, se da una breve descripción de los hiperparámetros más relevantes dispuestos en el espacio de búsqueda para evitar un árbol excesivamente complejo:

- *criterion*: Es una función para medir la calidad de la división de los datos. [*'gini'*, *'entropy'*]
- *max depth*: Define la máxima profundidad del árbol. [None, 10, 20, 30, 40, 50]
- *min samples split*: El mínimo número de muestras para dividir un nodo. [2, 5, 10]
- *min samples leaf*: Es el mínimo número de muestras necesarias para estar en un nodo hoja (respuesta final). Un punto de división a cualquier profundidad sólo se considera si la hoja tiene al menos un número mínimo de hojas definido en cada una de las ramas izquierda y derecha. [2, 4, 6, 8]

### 3.2.3.3. Algoritmo de Random Forest

Para utilizar este modelo se usó la librería gratuita *scikit-learn*, la cual tiene la estructura siguiente, con los hiperparámetros por defecto mostrados a continuación:

```
1 class sklearn.ensemble.RandomForestClassifier (n_estimators=100, *, criterion='gini',
    ↪ max_depth=None, min_samples_split=2, min_samples_leaf=1,
    ↪ min_weight_fraction_leaf=0.0, max_features='sqrt', max_leaf_nodes=None,
    ↪ min_impurity_decrease=0.0, bootstrap=True, oob_score=False, n_jobs=None,
    ↪ random_state=None, verbose=0, warm_start=False, class_weight=None, ccp_alpha
    ↪ =0.0, max_samples=None, monotonic_cst=None)
```

Al igual que con el modelo de árbol de decisión, se realiza una búsqueda con los hiperparámetros que limiten la creación de árboles muy complejos. Además, se agrega un hiperparámetro para definir la cantidad de árboles que se utilizarán para la construcción del modelo. Mientras más árboles, mejor debería el modelo evitar un sobre-ajuste, pero aumenta el tiempo de resolución del algoritmo. Un parámetro crítico para el algoritmo de random forest, es el número máximo de características que cada división puede buscar entre todos los parámetros en el set de datos. Este parámetro es importante a la hora de inyectar aleatoriedad al modelo, ya que si se utiliza un número alto de características, no existirá aleatoriedad entre los árboles. Esto se explica porque todas las divisiones de los árboles buscan los mismos parámetros. En cambio, si se utiliza un valor bajo, los árboles serán muy diferentes unos con otros, por lo que será necesaria una profundidad relevante para que exista un ajuste a la base de datos. La cantidad máxima de características entonces es dependiente del número de parámetros de cada problema, por lo que su definición puede ser en función de este valor, o se puede entregar directamente si se prefiere.

- *criterion*: Es una función para medir la calidad de la división de los datos. [*'gini'*, *'entropy'*]
- *max features*: Hace referencia a la cantidad de características a considerar al buscar la mejor división. [*'auto'*, *'sqrt'*, *'log2'*]
- *max depth*: La profundidad máxima de cada árbol que lo compone. [4, 6, 8, 10]
- *n estimators*: El número de árboles que compone al modelo total. [100, 200, 300]

### 3.2.4. Evaluación del Modelo y Validación Cruzada

Entrenar un modelo y luego testarlo con los mismos datos es un error metodológico. Un modelo que solo repite las etiquetas que acaba de aprender puede tener altas métricas de rendimiento, pero cuando se le entreguen datos nuevos que nunca ha visto es probable que falle en las predicciones. Este problema se llama sobre-ajuste, y para evitarlo se realiza una práctica común que es apartar una cantidad de los datos, para utilizarlos como datos de testeo.

Luego de encontrar los mejores hiperparámetros para el modelo, se debe realizar una validación cruzada para comprobar la robustez del mismo. Este método divide el conjunto de datos de entrenamiento en  $K$ -folds (K particiones), para luego hacer predicciones y evaluar cada partición usando un modelo entrenado en las particiones restantes. En este trabajo, se dividió la base de datos original en 10, obteniendo las métricas de 10% de los datos nuevos de un modelo entrenado con un 90% de los restantes. La robustez se mide según los resultados y la varianza de las métricas de todas las particiones.

En la Figura 3.18 se dispone un diagrama de flujo para explicar el orden en que se aplica *GridSearch* y la validación cruzada en el trabajo. Y en la Figura 3.19, se explica la forma en que trabaja la validación cruzada. Además, según el manual de *Scikit-learn* (Pedregosa et al., 2011), cuando exista un gran desequilibrio en la distribución de las clases se recomienda utilizar el muestreo estratificado, por ejemplo, implementar *StratifiedKFold* o *StratifiedShuffleSplit*. Esto para garantizar que las frecuencias relativas de las clases se conserven aproximadamente en cada partición de entrenamiento y validación.

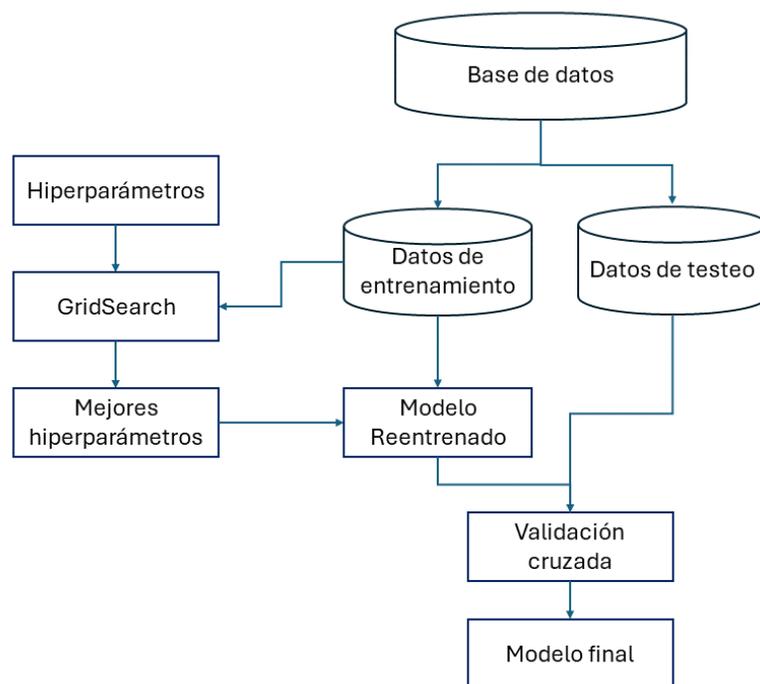


Figura 3.18: Pasos típicos a seguir para validar un modelo predictivo (Elaboración propia).

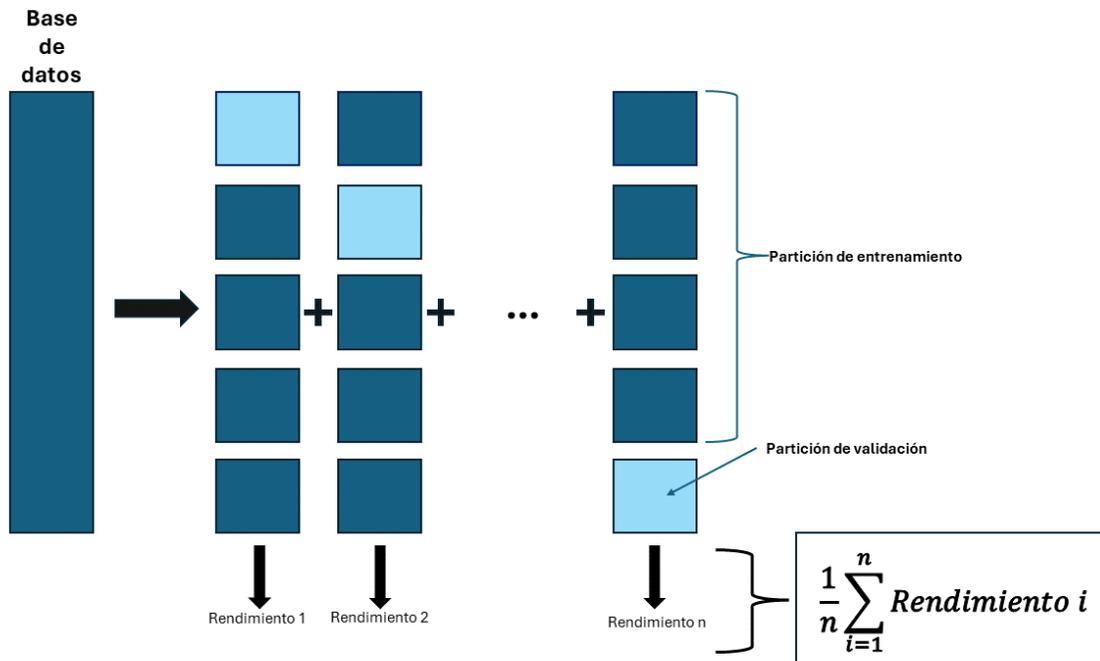


Figura 3.19: Diagrama de flujo de validación cruzada (Elaboración propia).

### 3.2.5. Ajuste de la base de datos para mejorar las métricas del modelo

Muchos dominios de problemas de aprendizaje automático, como la detección de fraudes, spam, valores atípicos y anomalías, tienden a implicar distribuciones de clases de muestras inherentemente desequilibradas. Sin embargo la mayoría de los algoritmos de clasificación asumen tamaños de muestra equivalentes para cada clase. Por tanto, los conjuntos de datos desequilibrados plantean un reto importante en el modelado de predicciones (Dong & Qian, 2022).

Los algoritmos más comunes en *machine learning* usualmente fueron pensados para bases de datos balanceadas y cuando existe un desequilibrio entre clases, no responden de la forma esperada. Los métodos para procesar bases de datos desbalanceadas se separan en dos categorías: los que operan a nivel de datos y los que operan a nivel de algoritmo. Dentro de los primeros, el sobre-muestreo y el sub-muestreo son categorías principales de los métodos a nivel de datos. El sobre-muestreo entrega un balance en el número de datos por clase al duplicar la clase minoritaria. Mientras que el sub-muestreo balancea la base de datos eliminando algunos datos de la clase mayoritaria. Sin embargo, alguna información importante se puede perder en el proceso.

La forma de implementar la solución de sobre-muestreo es directa, probando diferentes múltiplos por la que la cantidad los casos de la clase minoritaria se equilibra con las demás, y producto de esto aumenta la métrica de rendimiento general del modelo. Existen otras técnicas de sobre-muestreo, por ejemplo el algoritmo SMOTE (*Synthetic Minority Over-sampling Technique*) el cual selecciona aleatoriamente N datos de la clase minoritaria y crea nuevas muestras sintéticas, mediante interpolación lineal (Dong & Qian, 2022). También esta el algoritmo ADASYN (*Adaptive Synthetic Sampling*), que se enfoca en extender cada clase minoritaria acorde a la distribución de la densidad del set de entrenamiento con diferentes pesos para balancearlo (Dong & Qian, 2022). No obstante el algoritmo del balanceo más simple de implementar entre los mencionados, es el sobre-muestreo directo, por lo que será la forma de balancear la base de datos en este trabajo. Y para comprobar que no genere un modelo sobre ajustado, se deben obtener altas métricas de rendimiento y baja varianza en la validación cruzada para comprobar su robustez.

# Capítulo 4: Resultados Modelo Numérico

En esta sección se muestran los resultados relacionados al modelo numérico realizado en *FLAC3D*. En primer lugar, se muestra la verificación de los supuestos que se utilizaron para la construcción, comparando el resultado del modelo numérico con tres casos reales y su clasificación según el gráfico empírico de Mathews. Luego, se procede a mostrar los resultados de la distribución de clases según la clasificación de Clark del total de los 8192 casos simulados. Además, se comparan los resultados obtenidos con el gráfico empírico de Clark para entender el comportamiento general de los casos de forma visual. Finalmente, se muestran gráficos de correlación entre las variables, para entender cómo la relación entre ellas afecta en la cantidad de dilución del modelo.

## 4.1. Verificación del modelo numérico

Para verificar el comportamiento del modelo y sus consideraciones, se seleccionaron diferentes casos de estudio de literatura para evaluar la respuesta de la estabilidad. Al momento de buscar información de casos para simular, se puede notar que no existe una base de datos que contenga todos los parámetros necesarios para su representación en el modelo numérico. Por esta razón, es necesario hacer una revisión exhaustiva de literatura para encontrar los datos que se requieren. Existen criterios importantes a la hora de seleccionar los casos de estudio, por esta razón se deben buscar casos que contengan la información mostrada a continuación:

- Ser de paredes colgantes.
- Casos que reporten diferentes niveles de estabilidad, es decir, paredes estables e inestables.
- Información geométrica y de propiedades de la roca.

Por esta razón, al tener información relevante reportada por diferentes autores, se seleccionaron caserones de la mina Cannington. Este yacimiento de plata-plomo-zinc está ubicado a 135 km al sureste de Cloncurry, en Queensland, Australia. Se utilizan dos métodos principales de explotación. En cuerpos de mineral más amplios, se emplea la explotación en subniveles (*Sublevel Open Stopping*). Según lo reportado por Hustrulid y Bullock (2001) los caserones pueden tener hasta 150 m de altura, dependiendo del tamaño del cuerpo de mineral y las condiciones del terreno local. Además, los anchos típicos de cada caserón varían de 20 a 30 m en la dimensión este-oeste (a lo largo del rumbo del cuerpo) y 20 m en la dimensión norte-sur (a lo largo de la extensión). También, Hustrulid y Bullock (2001) resumen las propiedades promedio de las rocas en la parte sur de la mina, que corresponden a roca gneis para la pared colgante con un UCS de 96 [MPa] aproximadamente y un RQD de 50 %, para este tipo de roca respectivamente.

Según lo que se reporta en literatura, la extracción de los primeros 35 caserones abiertos en subniveles indicó una dilución promedio de 1.6 % de material desde fuera del contorno de los diseños de los caserones. Aproximadamente la mitad de esta dilución fue relleno de pasta aplicada durante la minería de caserones secundarios y terciarios con una a tres paredes rellenas. La dilución geológica fue mayor en caserones que contenían esquisto de moscovita en la pared colgante o en los techos, y alcanzó hasta un 15 % en dos caserones de brecha Glenholme extraídas hasta el momento. La presencia de fallas y estructuras de juntas en las paredes del caserón ha resultado en la rotura de mineral en caserones adyacentes. La orientación de la cuadrícula de diseño de los caserones se ajustó en 30° en futuros caserones para reducir el impacto negativo de esta característica.

Por otro lado, a partir de la base de datos de Mawdesley et al. (2001), se seleccionaron datos de tres caserones de la mina Cannington catalogados como estable, falla y falla mayor, según la clasificación de Mathews et al. (1981), mostrados en la Figura 4.1. Esta información incluye el número de estabilidad, el radio hidráulico, el índice  $Q'$ , los factores A, B y C, la resistencia a la compresión uniaxial (UCS) y el esfuerzo máximo en la cara de la pared en estudio. Se seleccionaron estos casos, ya que coincide con la información extraída de Hustrulid y Bullock (2001), según el RQD y el UCS. Además, por ser datos de la pared colgante (la pared relevante) y ser tres casos con diferentes niveles de estabilidad.

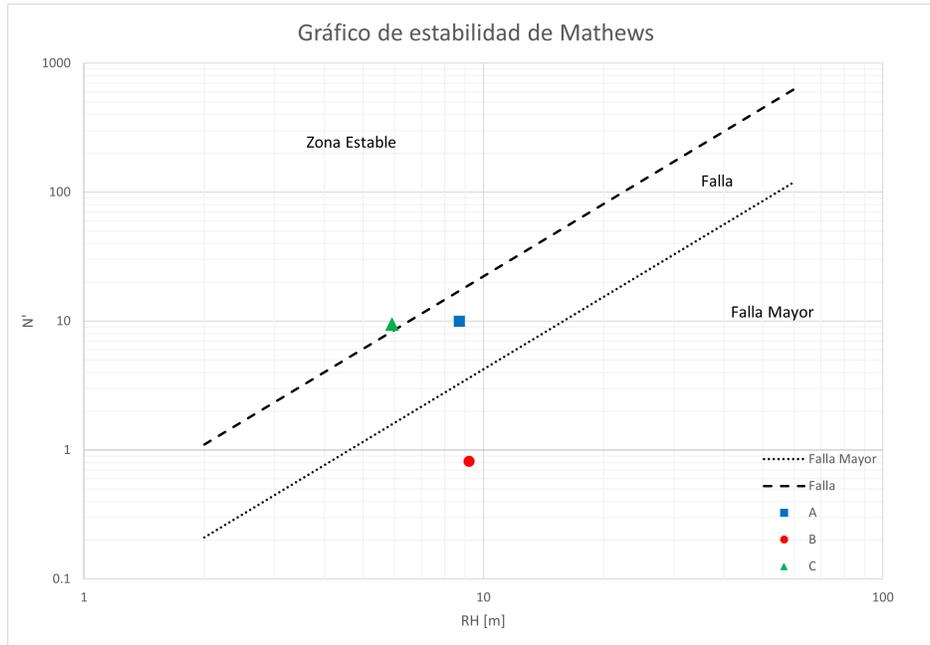


Figura 4.1: Casos de Estudio, sobre el gráfico de estabilidad de Mathews.

Ahora los parámetros que faltan para la caracterización de los caserones corresponden a los elásticos, es decir, el módulo de Young y la razón de Poisson de macizo rocoso. En Rankine (2004), se muestra una relación entre el UCS y el Módulo de Young para diferentes tipos de roca de la mina Cannington. El autor clasifica las rocas GH, GNES y SHMU según Deere et al. (1966) como rocas tipo esquisto, de resistencia a la compresión uniaxial baja a media y con una relación con el módulo de elasticidad de media a alta. Además, dice que la naturaleza brechada de estas rocas, tiende a reducir el UCS del material y devolver una relación  $E/UCS$  más alta, igual a 500:1.

Las Tabla 4.1 y 4.2 muestran un resumen de la información recolectada de cada fuente respectivamente.

Tabla 4.1: Resumen información recolectada, base de datos Mawdesley et al. (2001).

Caso	RH	N	RQD	A	B	C	UCS [MPa]	$\sigma_1$ [MPa]
A	8.7	9.9	45	0.24	0.4	8.0	100	30.6
B	9.2	0.8	30	0.54	0.3	6.8	100	17
C	6.0	9.5	57	0.24	0.3	8.0	100	30.6

Tabla 4.2: Resumen información recolectada de Hustrulid y Bullock (2001).

Caso	Ancho [m]	Alto [m]	Profundidad [m]	UCS [MPa]
A	25	130	700	100
B	25	150	700	100
C	25	50	700	100

A partir de estos datos se pueden calcular los parámetros necesarios para construir un modelo de verificación. Del alto y el radio hidráulico se puede calcular el largo del caserón. Del UCS, la profundidad y el factor A se puede calcular la razón de esfuerzos. Del factor B se puede calcular el ángulo de las estructuras y del factor C se puede obtener la inclinación de la pared colgante. Para las propiedades de las estructuras al no tener información de ellas, se utilizan propiedades promedio, solo variando la inclinación de estas. Finalmente el módulo de elasticidad se calcula a partir del UCS.

Es importante notar que no se tiene la información de cómo se obtuvieron los datos en terreno, ni el nivel de confiabilidad. No se conoce el número de ensayos que se realizaron para respaldar la información reportada, por lo que se debe tener presente esto al momento de comparar los resultados del modelo numérico con la información de los casos de estudio.

Con las propiedades antes mencionadas, se simulan los tres casos de estudio denominados 'A', 'B' y 'C'. Esto con el objetivo de comparar la respuesta del modelo numérico con la clasificación de estabilidad según Mathews et al. (1981), versus lo reportado en la base de datos de Mawdesley (2001). Comparando el resultado del modelamiento numérico de la Figura 4.2 con la Figura 4.1, se puede notar que la dilución calculada por el modelo se acerca más a la realidad que la clasificación de estabilidad según Mathews et al. (1981). En este sentido la estabilidad se relaciona con la cantidad de ELOS en cada caso. De acuerdo a lo anterior, se verifica que las consideraciones con las que se construyó el modelo numérico, además de las clasificaciones de dilución propuestas por Clark (1998) permiten estimar la estabilidad y el comportamiento de las paredes colgantes de caserones de los casos de estudio propuestos. La comparación de estos tres casos se resume en la Tabla 4.3.

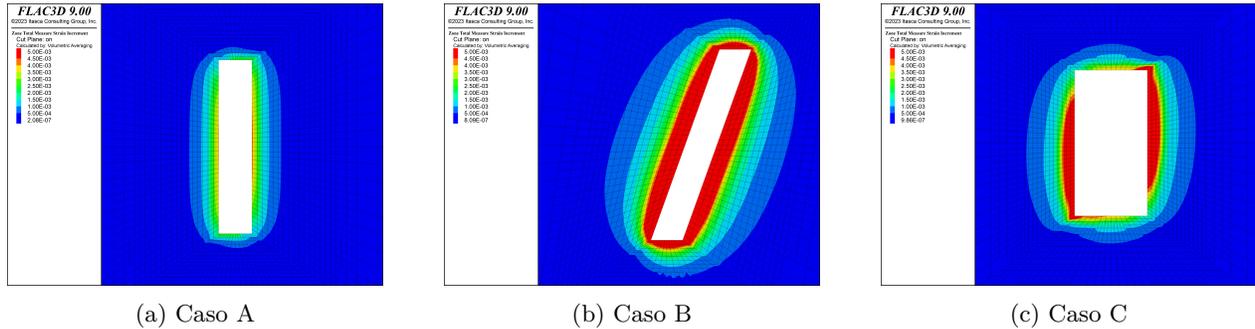


Figura 4.2: Resultados del modelamiento numérico de los tres casos de estudio de la mina Cannington.

Tabla 4.3: Comparación entre la realidad, clasificación según Mathews y modelo numérico, para los tres casos de estudio.

Caso	Realidad	Mathews	Modelo Numérico
Caso A	Estable	Falla	ELOS < 0.5 [m]
Caso B	Falla Mayor	Falla Mayor	ELOS > 2.0 [m]
Caso C	Falla	Estable	ELOS > 0.5 [m]

## 4.2. Distribución de las clasificaciones respecto al gráfico de Clark

Se puede observar de la Figura 4.3 la desproporción visual existente entre las categorías de dilución. Según la Tabla 4.4 la mayoría de los de casos están clasificados según la clase  $ELOS > 2[m]$  y  $ELOS < 0.5[m]$ , las cuales juntas son aproximadamente un 92 % del total de los casos. Mientras que las categorías  $1[m] > ELOS > 0.5[m]$  y  $2[m] > ELOS > 1[m]$  juntas son un 8 % de los casos simulados. Esto tiene sentido y es lo esperado, debido al amplio rango de valores de los parámetros que se consideraron para simular los casos, ya que se incluyeron radios hidráulicos grandes, además de malas calidades de roca y estructuras geológicas.

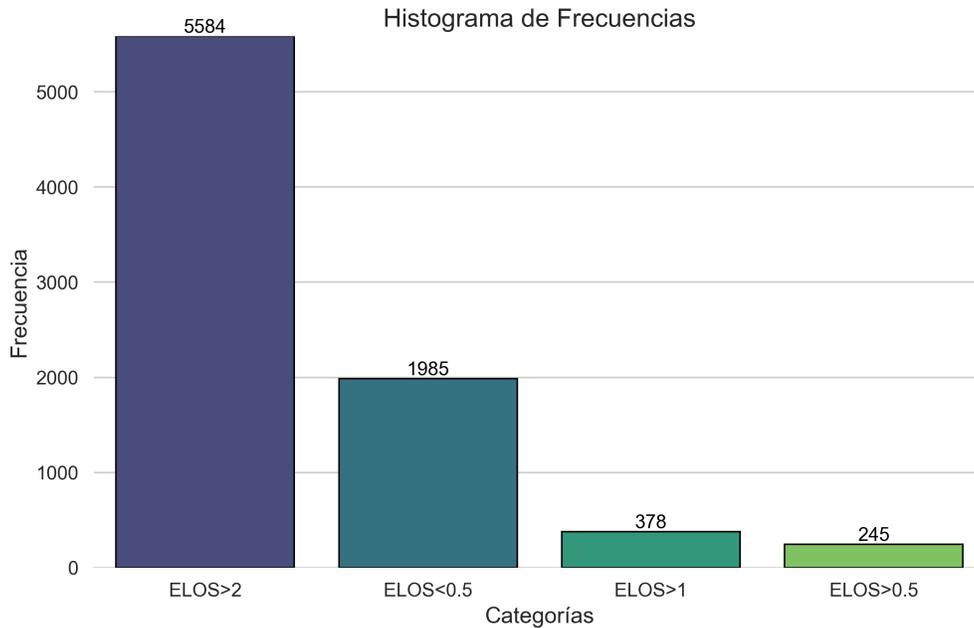


Figura 4.3: Histograma con frecuencias según categoría.

Tabla 4.4: Resumen de cantidad de casos por categoría y su respectivo porcentaje de acuerdo al total de la base de datos.

Categoría	Cantidad casos	Porcentaje [%]
$ELOS < 0.5[m]$	1985	24.2
$1[m] > ELOS > 0.5[m]$	245	3.0
$2[m] > ELOS > 1[m]$	378	4.6
$ELOS > 2[m]$	5584	68.2

En la Figura 4.4 se hace una comparación de los resultados encontrados con las curvas de estabilidad del gráfico según Clark (1998). De esta imagen se pueden evidenciar las tendencias hacia la zona estable e inestable, donde existe el comportamiento esperado según la revisión bibliográfica realizada. A mayor radio hidráulico y menor número de estabilidad, los casos tienden a tener mayor dilución. Y lo mismo en el otro sentido, a menor radio hidráulico y mejor número de estabilidad, los casos tienden a tener menor dilución. Sin embargo, en la parte intermedia de las curvas no es posible diferenciar claramente densidades de puntos que logren separar a la clasificación  $2[m] > ELOS > 1[m]$  y  $1[m] > ELOS > 0.5[m]$  (puntos verdes y amarillos, respectivamente). Si bien es posible utilizar otras

clasificaciones de dilución diferentes a las propuestas por Clark (1998), estas no tendrían necesariamente el mismo sentido y utilidad en la práctica como las estudiadas por el autor.

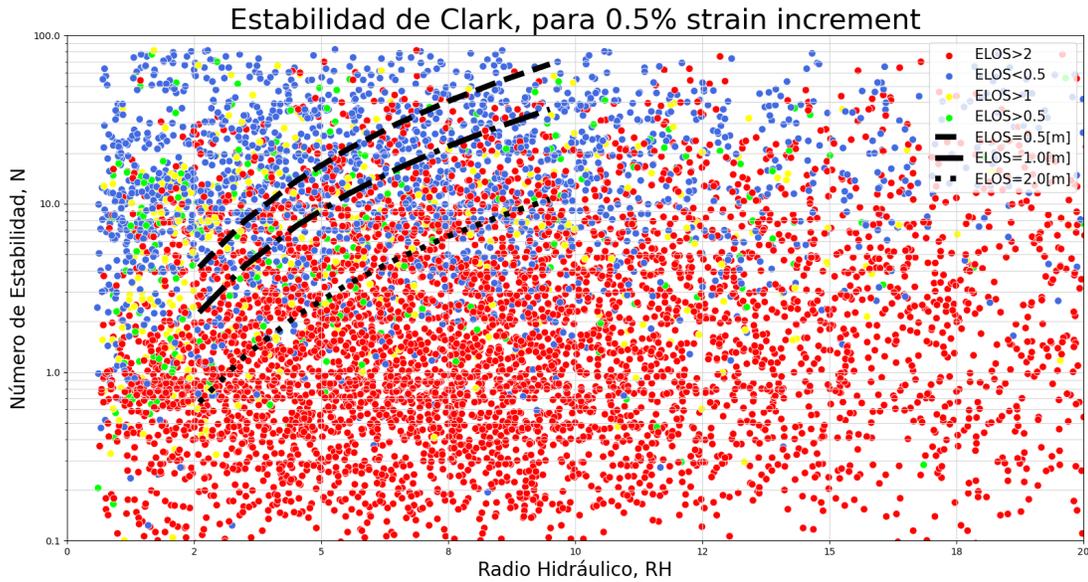


Figura 4.4: Comparación resultados versus gráfico de estabilidad de Clark.

### 4.3. Correlaciones entre variables

En la Figura 4.5, se observa el peso de cada variable en la cantidad de ELOS producida para los 8192 casos simulados. Un valor negativo cercano a -1 significa una relación inversamente proporcional, es decir, mientras menor es la variable, el ELOS es más alto. Por otro lado, si el valor es positivo cercano a 1, significa una relación proporcional, en otras palabras, mientras menor es la variable, se produce un menor ELOS. En el caso de que el valor sea cercano a cero, significará una correlación prácticamente nula. En este sentido, mientras mayor es el valor absoluto de las correlaciones, más fuerte es la relación. Siguiendo esta línea, se pueden ordenar las variables según su correlación con la cantidad de ELOS:

- Proporcionales: Razón de esfuerzos, profundidad, alto, densidad y ancho.
- Inversamente proporcionales: GSI, UCS,  $m_i$ , inclinación, razón de aspecto, propiedades y manto de las estructuras.

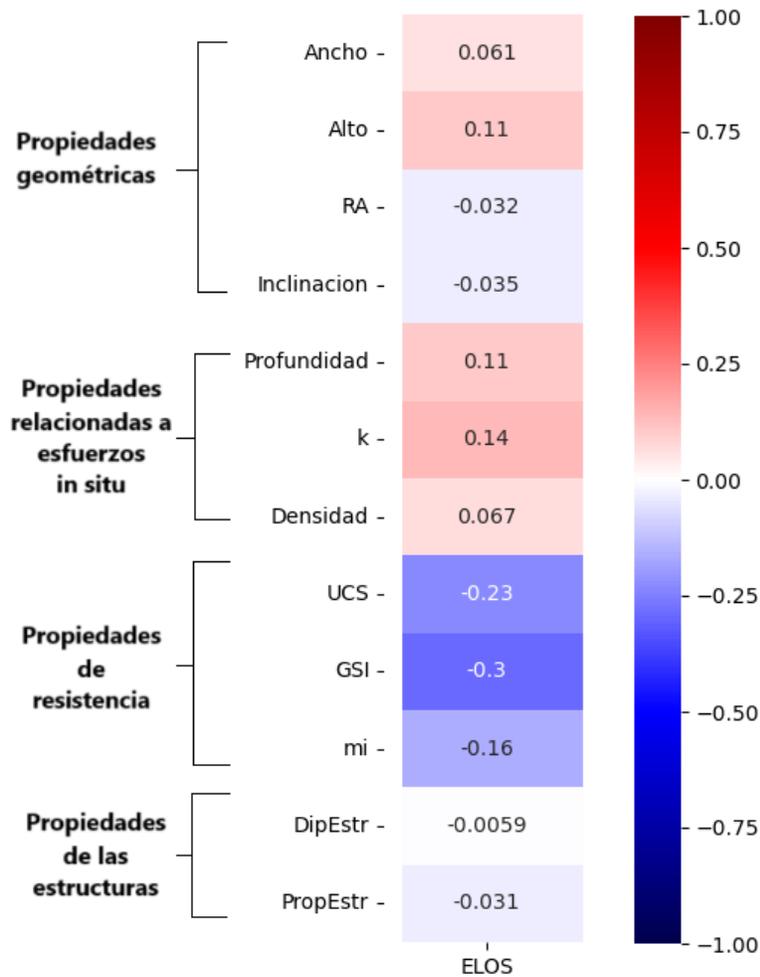


Figura 4.5: Correlación entre las variables y la cantidad de ELOS.

Además, para visualizar posibles correlaciones entre la cantidad de dilución y las variables, se grafican matrices de dispersión. Debido al número de variables, se deciden separarlas para tener una mejor visualización de las relaciones. Se agrupan según las categorías mostradas a continuación:

- Geometría del caserón: Ancho [m], alto [m], razón de aspecto, inclinación [°].
- Parámetros relacionados a los esfuerzos: Profundidad [m], razón de esfuerzos  $k$  [-], densidad [ $kg/m^3$ ].
- Propiedades de resistencia y de macizo rocoso: UCS [MPa], GSI [-], parámetro  $m_i$  de Hoek and Brown [-].
- Orientación y propiedades de las estructuras: Dip estructura [°], calidad de la estructura [-].
- Propiedades físicas y elásticas de la roca: Densidad [ $kg/m^3$ ], módulo de Young del macizo rocoso [GPa], módulo de Poisson [-].

Los resultados obtenidos para cada par de variables según su relación y su efecto en la cantidad de ELOS obtenida. Primero, para entender la información de los gráficos de correlación, se observa que se presentan como una matriz de 'n'x'n' gráficos, donde se van analizando de a 'n' variables. En la parte superior e inferior de la matriz se repiten los mismos gráficos, pero con las variables en distinta posición. Y la diagonal muestra gráficos de densidad del ELOS versus la variable respectiva. Para

lograr identificar las densidades de mejor forma, se agregan líneas que delimitan zonas con mayores densidades dentro de cada gráfico de la parte inferior de la matriz, las cuales corresponden a los siguientes colores por categoría:

- $ELOS > 2[m]$ : rojo
- $2[m] > ELOS > 1[m]$ : verde
- $1[m] > ELOS > 0.5[m]$ : morado
- $0.5[m] > ELOS$ : celeste

## a) Correlaciones de parámetros geométricos

Observando las correlaciones de los parámetros geométricos de la Figura 4.6, mientras menor es valor de la razón de aspecto (RA), mayor es la cantidad de casos inestables. Esto tiene sentido según lo reportado por diferentes autores como Hughes et al. (2010) y Vongpaisal et al. (2009), los cuales encontraron una relación entre la razón de aspecto y la cantidad de dilución en la pared colgante. Geométricamente significa que los caserones que tienen paredes colgantes más largas que altas tienen mayor dilución debido a la diferente distribución de esfuerzos en la pared colgante, según la razón de esfuerzos y la razón de aspecto.

De la Figura 4.6 se observa un mayor número de casos estables cuando los valores de ancho y alto son pequeños. También se puede notar que existen mayor cantidad de casos estables cuando la inclinación del caserón es cercana a  $90^\circ$ , lo cual tiene sentido debido a que ésta representa la geometría más favorable para la pared colgante. Además coincide con la lógica propuesta por autores como Mathews et al. (1981) y Potvin (1988), en la definición de los valores para el factor C.

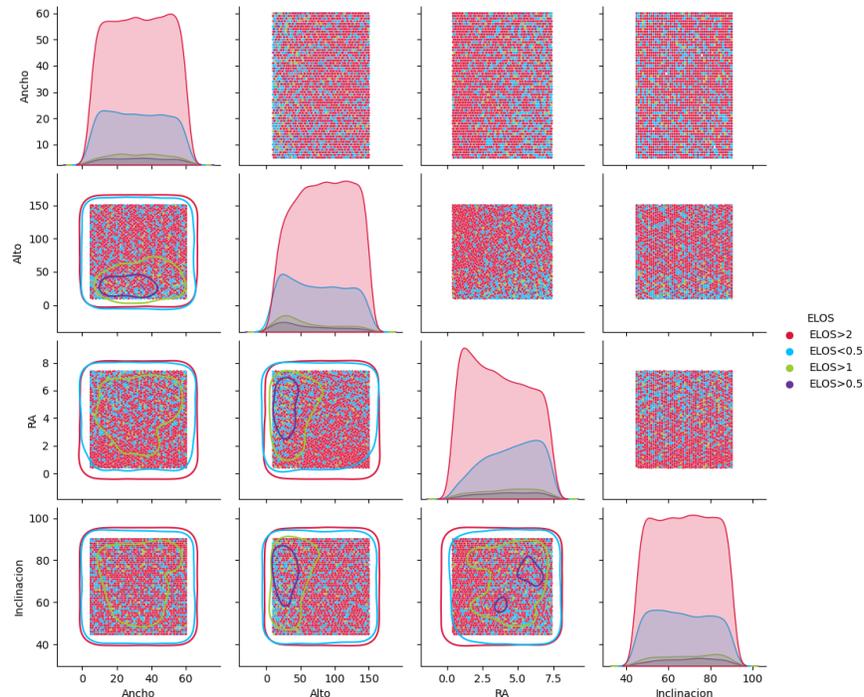


Figura 4.6: Matriz de dispersión-Geometría del caserón.

## b) Correlaciones parámetros asociados a esfuerzos in-situ

De la Figura 4.7 se puede notar que existen *peaks* y caídas para el gráfico de densidad en la diagonal en el factor k de esfuerzos. Esto se debe a la discretización de valores en los que se analizaron casos. A diferencia del gráfico de la Figura 4.6, que corresponde a las variables mejor representadas debido al amplio rango de geometrías en estudio.

Se observa de la Figura 4.7 una mayor cantidad de casos inestables cuando la profundidad, la razón de esfuerzos y la densidad tienen valores altos dentro de cada rango respectivo. Sin embargo, no se logra encontrar una correlación clara entre estas variables.

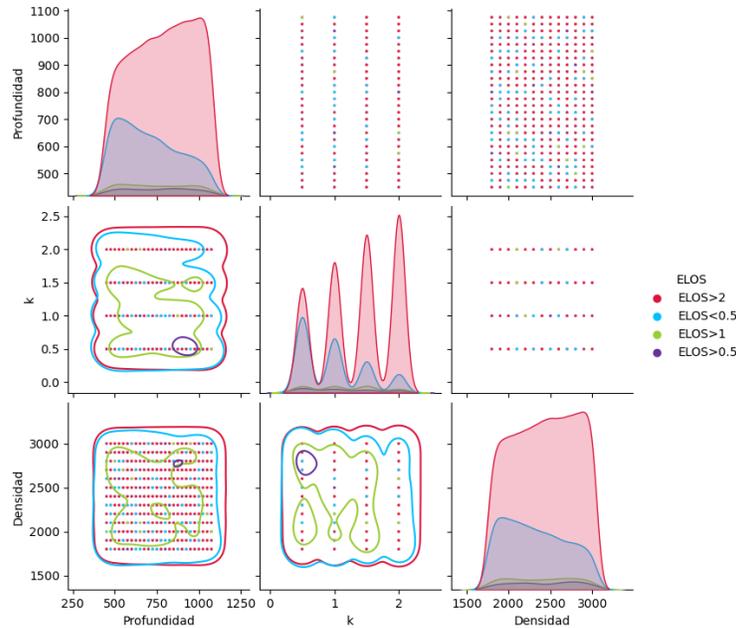


Figura 4.7: Matriz de dispersión-Parámetros relacionados al esfuerzo in situ

## c) Correlaciones de propiedades resistentes y de macizo rocoso

A continuación en la Figura 4.8 se puede notar que a mayor GSI, existe menor cantidad de casos inestables. Esto tiene sentido ya que un GSI mayor, significa un macizo rocoso menos fracturado, por lo tanto, la roca se deformará menos. Lo mismo ocurre con la resistencia a la compresión uniaxial (UCS), mientras mayor es su valor, existe una menor cantidad de casos inestables. Esto se explica ya que, al tener un valor alto de UCS es indicativo de un macizo rocoso más competente. El GSI y el UCS son buenos indicadores para separar los casos inestables y los estables, sin embargo, se puede observar en la Figura 4.8 que no se evidencia una correlación suficientemente fuerte tal que solo estos dos parámetros logren definir la estabilidad. De igual forma se puede observar que un valor de *mi* alto tiene una leve mayor densidad de casos estables.

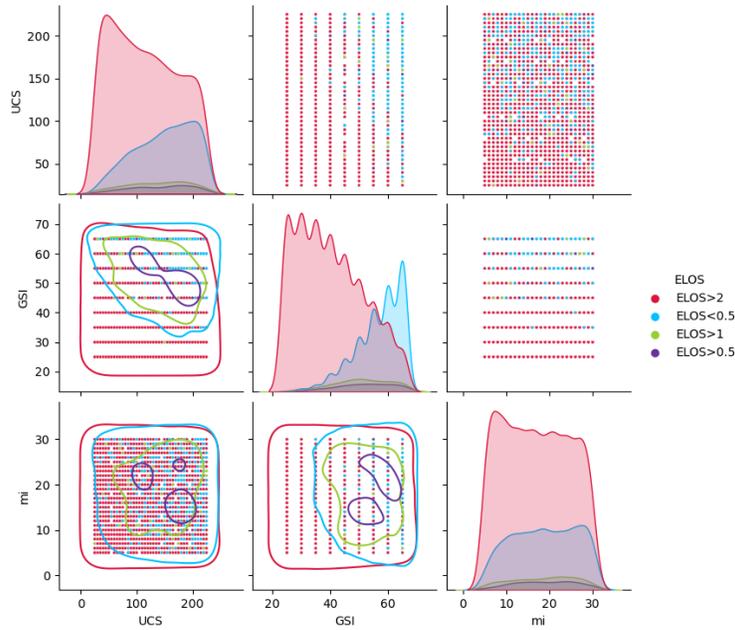


Figura 4.8: Matriz de dispersión-Propiedades de resistencia y de macizo rocoso.

#### d) Correlación propiedades de las estructuras geológicas

Siguiendo con la Figura 4.9, se observa una leve tendencia a medida que las propiedades de las estructuras tienen menor calidad hacia un aumento de casos inestables. También se puede indicar que a medida que la orientación de las estructuras con respecto a la cara del caserón es más cercana a cero, se tienen más casos inestables. En cambio, cuando mejora la calidad de las estructuras y existe una orientación perpendicular a la cara de la pared colgante, existirán más casos estables lo que coincide con lo propuesto con el factor B de los gráficos de estabilidad empíricos.

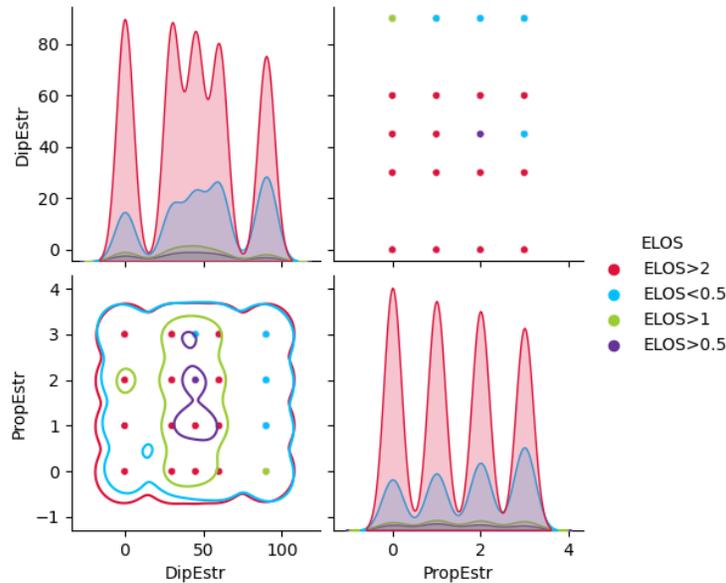


Figura 4.9: Matriz de dispersión-Orientación y propiedades de las estructuras.

### e) Correlación de las propiedades físicas y elásticas de la roca

El módulo de Young y la razón de Poisson no son variables independientes. El módulo de Young se calcula a partir del GSI, el UCS y el factor D, y la razón de Poisson se calcula a partir del GSI. Por esta razón, no presentan la distribución uniforme en los casos que se simularon, como se observa en la Figura 4.10. Por otro lado, en la diagonal de la Figura 4.10, se puede notar una mayor cantidad de casos inestables cuando el módulo de Young del macizo rocoso tiene un valor bajo. Esto es lo esperado, ya que mientras menor es el modulo de elasticidad existe una mayor probabilidad de alcanzar el umbral del criterio de deformación (límite de fracturamiento). Sin embargo, para los casos estables, la distribución del módulo de Young se mantiene relativamente uniforme a lo largo del rango de valores dado, por lo que se puede entender que la estabilidad depende de otros parámetros además del módulo de Young. Esto tiene sentido, ya que el módulo de elasticidad se calculó a partir del GSI y del UCS según la Ecuación 3.3, por lo que la variación de esos 3 parámetros esta relacionado y no se logra observa un comportamiento simple de interpretar.

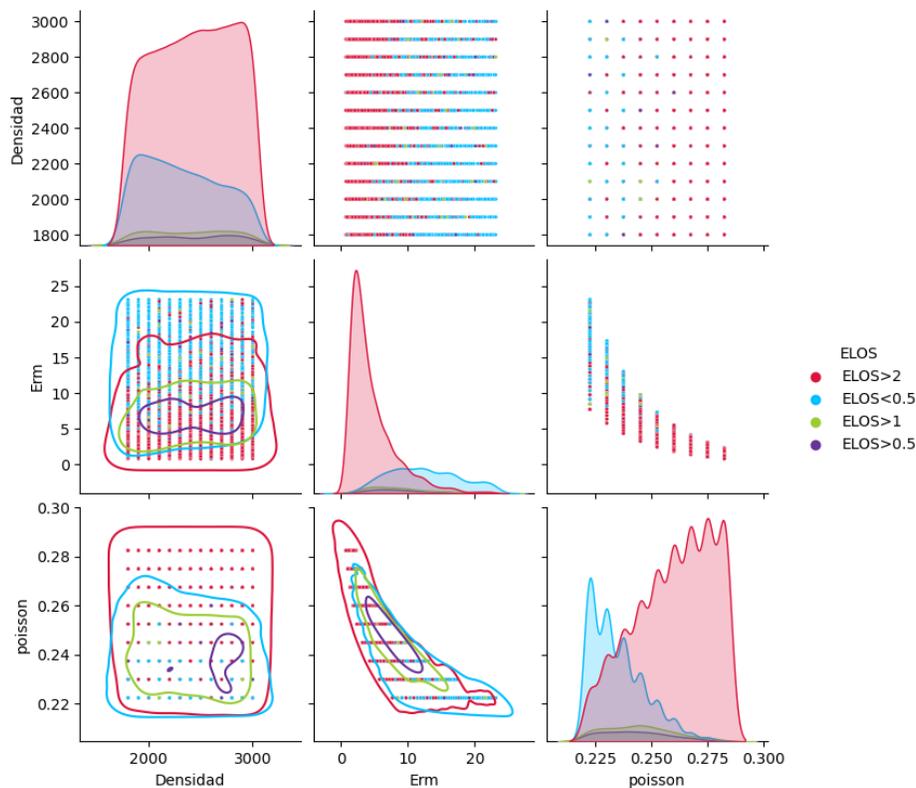


Figura 4.10: Matriz de dispersión-Propiedades físicas y elásticas de la roca.

De igual forma se puede notar que a mayor razón de Poisson, los casos inestables aumentan. Y en el otro sentido, mientras menor es la razón de Poisson existen más casos estables. También la densidad de la roca muestra tendencias que mientras mayor es su valor, existen más casos inestables y mientras menor es su valor existen más casos estables. Por último, se nota que para todos los gráficos de la diagonal, las categorías  $1[m] > ELOS > 0.5[m]$  y  $2[m] > ELOS > 1[m]$  (morada y verde respectivamente), parecen ser uniformes a lo largo del rango en estudio, sin ser afectadas significativamente por alguna variable.

## 4.4. Discusiones

La razón de comparar el comportamiento de los casos en la Figura 4.4 versus el gráfico de estabilidad de Clark, es simplemente una forma visual para entender qué tanto se acercan los resultados del modelo numérico a los casos reales. Los resultados de los modelos numéricos se deben interpretar, y también es importante comparar con fenómenos físicos esperables que ocurren en la realidad. Por otro lado se debe tener en cuenta que ajustarse a las curvas de los gráficos empíricos no significa que las clasificaciones estén bien, ya que de por sí los gráfico empíricos tiene un bajo ajuste.

En la Figura 4.4, se observa que los casos estables e inestables parecen presentar dos zonas con alta densidad, similar a las zonas propuestas en el gráfico de Clark. En cambio, para las clasificaciones intermedias no es posible visualizar claramente zonas con densidades altas de casos. Una posible causa de esto, se debe al amplio rango en el que se exploró el problema, lo cual resulta en más casos con condiciones extremas.

La construcción del modelo numérico paramétrico tuvo resultados positivos, por lo que se logró generar un modelo sencillo pero completo y automatizado. Sin embargo, para crear un código que sea capaz de generar de forma paramétrica una gran cantidad de casos con diferentes geometrías y condiciones del caserón, se debieron tomar una serie de simplificaciones y supuestos entre los que destacan:

1. En la construcción del modelo del caserón, la principal dificultad fue la creación de un código que densificara la malla alrededor de la excavación. El problema con esta densificación de malla paramétrica, fue la detección de casos que presentaban singularidades numéricas, los cuales correspondían a valores de inclinación menores a  $45^\circ$ . Una forma de solucionar esto, fue tomar la decisión de simular rangos de inclinaciones de la pared colgante desde  $45^\circ$  hasta  $90^\circ$ .
2. Otra consideración en la construcción del modelo numérico es respecto al modelo constitutivo usado Ubiquitous Joint, el cual da cuenta de las estructuras presentes en el macizo rocoso. Es importante tener claro lo que este modelo representa, ya que la palabra 'ubiquitous' (ubicuo en español) significa que está presente a un mismo tiempo en todas partes, y representa solamente al set de discontinuidades más desfavorable. No obstante, para simular macizos rocosos sin estructuras, es posible considerando las estructuras con orientación en  $90^\circ$  respecto a la pared colgante y con propiedades altas de cohesión y fricción en las estructuras.
3. Otro tema a tener presente, es el criterio de daño determinado. Este criterio fue el de límite de fracturamiento y se considera como la sumatoria de todos los componentes del tensor de deformaciones. Gracias al criterio de daño, el modelo numérico fue capaz de entregar resultados para condiciones de falla por corte o tensión en la roca y en las estructuras. Sin embargo, es importante tener en cuenta que es un criterio que debe calibrarse, por lo que los casos reales para asegurar la representividad son muy relevantes.
4. Una de las dificultades al momento de calibrar el criterio de daño, fue la insuficiencia de información de casos reales. Existen bases de datos como la de Potvin (1988) o Mawdesley et al. (2001) que entregan información de caserones en diferentes estados de falla, pero es difícil tener toda la información geométrica y geológica de sus configuraciones. Por ejemplo, la mayoría de las bases de datos no tiene la información separada de los factores que componen el número de estabilidad y el radio hidráulico, por lo que es difícil descomponer la geometría real del caserón.

# Capítulo 5: Modelo Predictivo

Con los resultados del modelo numérico se entrenan los modelos de aprendizaje árbol de decisión y random forest para escoger el que mejor resultados entregue. Se debe tener en cuenta que los resultados del modelo predictivo no se pueden visualizar, por eso su interpretación es mediante las métricas de rendimiento de precisión, recall (exhaustividad) y F1-score (valor F1). En la siguiente sección primero se comparan los resultados de dos métricas de rendimiento de modelos de aprendizaje: árbol de decisión y random forest, a medida que se va aumentando la base de datos con la que se entrenan y validan. Luego, se muestran los resultados de las nuevas métricas después de balancear la base de datos. Finalmente, se presentan los resultados de la validación cruzada del modelo, para evaluar su robustez.

## 5.1. Métricas de evaluación versus tamaño de la base de datos

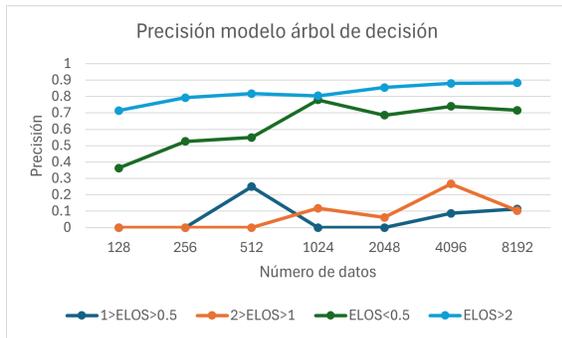
Antes de elegir el mejor modelo, es relevante identificar la métrica de interés dependiendo la necesidad del problema específico para el que se utilizará. En el caso que es más importante que se prediga correctamente por sobre la cantidad de aciertos, entonces la métrica de interés es la precisión. Un ejemplo de ello es la clasificación de vídeos para niños en Youtube, donde no interesa que sean pocos vídeos, si no que estén clasificados correctamente. En cambio, si lo que interesa es que el modelo prediga correctamente la mayor cantidad de casos de cierta categoría, entonces la métrica que se debe usar es la exhaustividad.

Para el problema en estudio, es importante analizar la precisión para los casos estables, ya que los casos que no está identificando el modelo serán clasificados como falsos positivos (considerando la inestabilidad como un caso positivo). Por el contrario, para el caso de la inestabilidad del caserón, la métrica de importancia es la exhaustividad, ya que es relevante clasificar correctamente la mayoría de los casos posibles, para no tener falsos negativos.

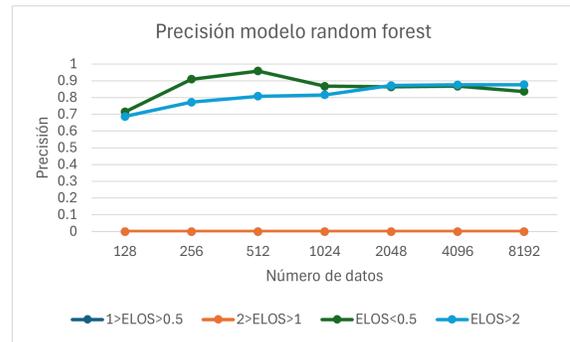
En esta sección, se analizan las métricas de rendimiento para entender el comportamiento de los modelos a medida que se van simulando casos secuencialmente, aumentando en potencia de dos la cantidad de la base de datos. Se puede observar de las Figuras 5.1, 5.2, 5.3 que las categorías  $ELOS > 2[m]$  y  $ELOS < 0.5[m]$  (inestable y estable respectivamente), muestran en general buenas métricas independiente la cantidad de datos con los que se entrenan. Es decir, presentan buena precisión, exhaustividad y valor F1, lo que significa:

- Precisión alta: el modelo predice correctamente un alto número de los casos dentro de los que clasifica como cierta clase.
- Exhaustividad alta: el modelo clasifica un alto número de casos del total de la categoría.
- Valor-F1 alto: al ser el promedio armónico de la precisión y la exhaustividad, si estas son altas, tiene sentido que este parámetro también lo sea.

Sin embargo, las categorías intermedias ( $2[m] > ELOS > 1[m]$  y  $1[m] > ELOS > 0.5[m]$ ) presentan malas métricas independiente la cantidad de datos. Esto se puede deber a la poca cantidad de casos en relación a las otras dos categorías que si presentan buena respuesta, lo que se conoce como una base de datos desbalanceada. De lo anterior se puede entender que los modelos no toman con el mismo peso las 4 categorías de estabilidad en las predicciones.

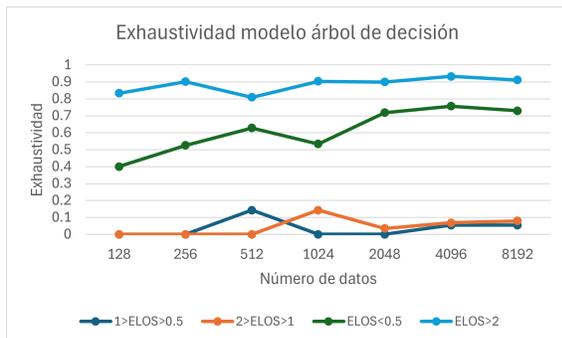


(a) Modelo árbol de decisión.

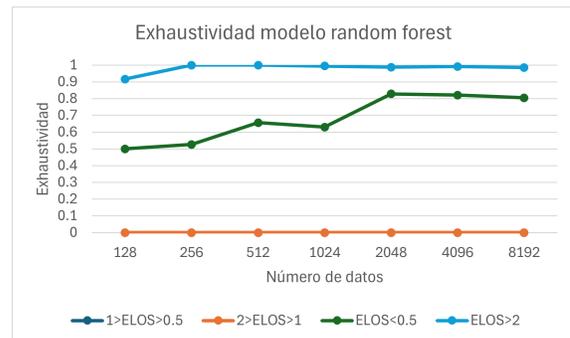


(b) Modelo random forest.

Figura 5.1: Comparación precisión según cantidad de datos.

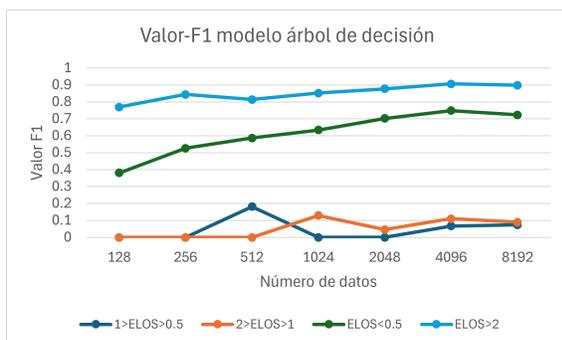


(a) Modelo árbol de decisión.

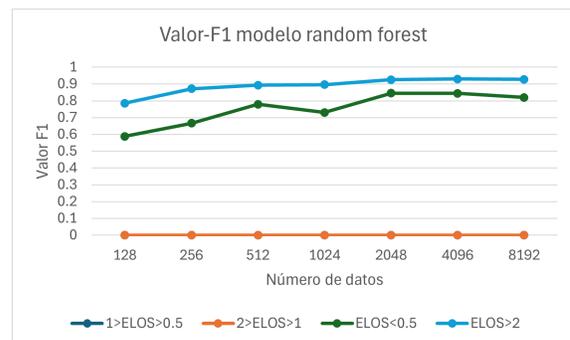


(b) Modelo random forest.

Figura 5.2: Comparación exhaustividad según cantidad de datos.



(a) Modelo árbol de decisión.



(b) Modelo random forest.

Figura 5.3: Comparación valor-F1 según cantidad de datos

En base a los resultados obtenidos, las simulaciones no se continuaron luego de los 8192 casos debido a que no se aprecia una mejora en las métricas, a pesar de aumentar la base de datos. Este problema se puede asociar al desbalance de la cantidad de casos por categoría, mostradas en el gráfico 4.3. Por esta razón, se decide hacer un balanceo mediante la técnica de sobre-muestreo, y utilizar además la técnica de GridSearch para encontrar los mejores hiperparámetros en cada caso particular.

## 5.2. Nuevas métricas de evaluación tras balance de la base de datos

Primero, al aplicar la técnica de sobre-muestreo se obtienen las nuevas cantidades de casos por categoría mostradas en la Tabla 5.1. Se puede notar que ahora los porcentajes por clases están más equilibrados que los mostrados en la Tabla 4.4. Con la base de datos ahora equilibrada más el uso de GridSearch, se buscan las combinaciones de hiperparámetros más óptimas mostradas en las Tablas 5.2 y 5.3, para cada algoritmo respectivamente con el total de casos simulados. Además en las Figuras 5.4, 5.5 y 5.6, se aprecian dos cosas: primero, que existe un aumento en las métricas para las clases intermedias de acuerdo a lo obtenido anteriormente sin el balance de la base de datos, y lo otro interesante de notar, es que el modelo de random forest tiene una mejor respuesta que el modelo de árbol de decisión para las tres métricas de rendimiento.

Tabla 5.1: Nueva cantidad de casos por categoría.

Categoría	Cantidad casos	Porcentaje [%]
ELOS<0.5[m]	1985	17.6
1[m]>ELOS>0.5[m]	1470	13.0
2[m]>ELOS>1[m]	2268	20.1
ELOS>2[m]	5584	49.4

Tabla 5.2: Mejor conjunto de hiperparámetros para el modelo árbol de decisión

Hiperparámetro	Valor
<i>criterion</i>	<i>gini</i>
<i>max depth</i>	<i>none</i>
<i>min samples leaf</i>	2
<i>min samples split</i>	2

Tabla 5.3: Mejor conjunto de hiperparámetros para el modelo random forest

Hiperparámetro	Valor
<i>criterion</i>	<i>entropy</i>
<i>max depth</i>	10
<i>max features</i>	<i>sqrt</i>
<i>n estimators</i>	300

De acuerdo a lo resultados de la Tabla 5.2 el conjunto de mejor hiperparámetros puede ser un indicador de árboles poco balanceados y un modelo sobreajustado a la base de datos. Por otro lado, de la Tabla 5.3, el modelo de random forest limita el sobreajuste, ya que éste se crea a partir del promedio de 300 árboles de decisión.

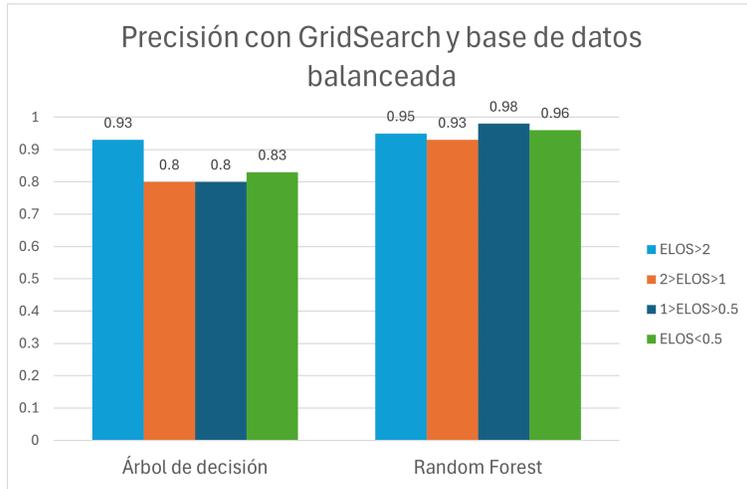


Figura 5.4: Comparación precisión para modelos árbol de decisión y random forest.

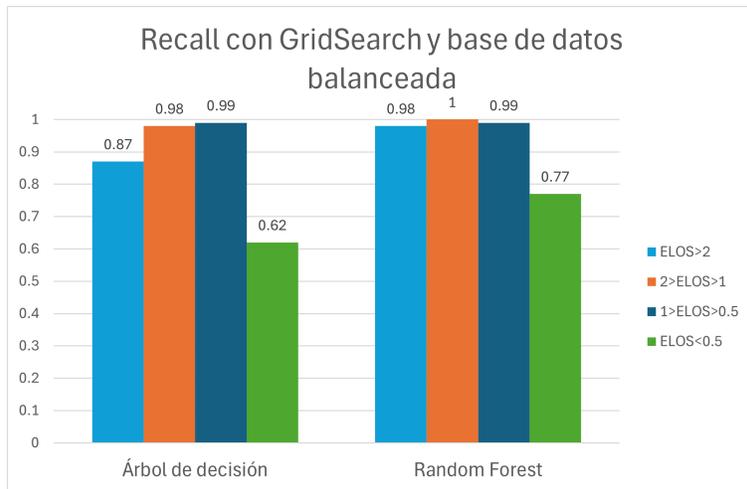


Figura 5.5: Comparación exhaustividad para modelos árbol de decisión y random forest.

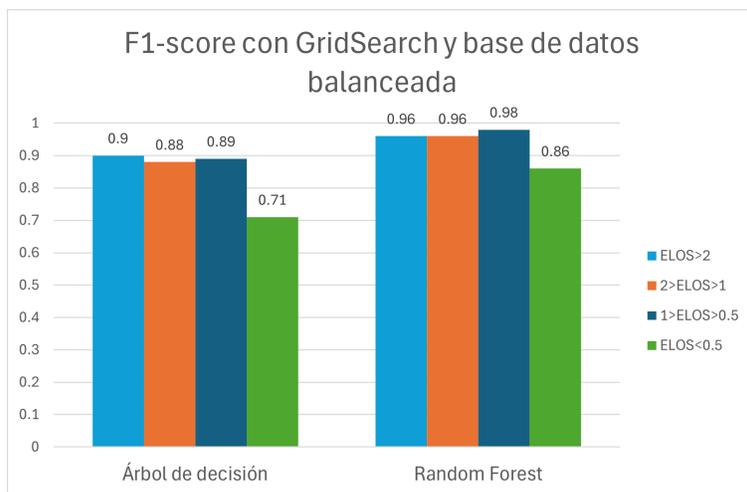


Figura 5.6: Comparación valor-F1 para modelos árbol de decisión y random forest.

Tabla 5.4: Resultados validación cruzada

Métrica	Valor
máxima exactitud	95.47 %
mínima exactitud	96.63 %
exactitud general	94.34 %
desviación estándar	0.00535

Con base a los resultados obtenidos ambas métricas son más altas para el modelo random forest en comparación al modelo de árbol de decisión, por lo que la elección del mejor modelo para el problema estudiado es directa. De igual forma, analizando los resultados del modelo random forest, se puede notar que la métrica más baja es la exhaustividad para la categoría más estable ( $ELOS < 0.5[m]$ ), lo que significa que en esta categoría existe una mayor cantidad de casos mal clasificados. Sin embargo, este es el error más conservador ya que son casos estables que se están clasificando con más dilución de la que realmente deberían presentar.

Otro fenómeno a tener en cuenta es el desbalanceo de la base de datos al momento de entrenar y validar el modelo, ya que puede afectar significativamente en los resultados. Esto es evidente, en las clases con valores menores, que corresponden a las categorías  $2[m] > ELOS > 1[m]$  con un 4.6 % y  $1[m] > ELOS > 0.5[m]$  con un 3 % del total de la base de datos, que presentaron métricas de evaluación cercanas a 0 %. Sin embargo, luego de utilizar la técnica de sobre-muestreo, sus porcentajes se incrementaron a 13 % y 20 % respectivamente, y sus métricas de precisión y exhaustividad aumentaron sobre un 93 %.

### 5.3. Discusiones

El modelo de predicción dio resultados positivos para todas las clasificación de estabilidad para el modelo de random forest, por lo que el objetivo de esta sección se considera cumplido. Sin embargo esto no hubiese sido posible sin las consideraciones siguientes:

1. Si se posee una base de datos desbalanceada, se debe tener muy presente, ya que puede afectar de forma negativa en las métricas de rendimiento del modelo predictivo.
2. En general las métricas del modelo random forest fueron mejores que las del modelo de árbol de decisión. Sin embargo, si este no hubiese sido el caso, es importante tener en cuenta las clasificaciones más desfavorables y en las que se debe tener más cuidado de clasificar mal. Por ejemplo, el peor error de clasificación es el que un caso con  $ELOS > 2[m]$  se clasifica como uno  $ELOS < 0.5[m]$ . Mientras que el error más conservador es el caso contrario, que un caso con  $ELOS < 0.5[m]$  se clasifique como uno  $ELOS > 2[m]$ .
3. Para comprobar la robustez del modelo, se sugiere realizar una validación cruzada, es decir, verificar si el modelo es capaz o no de generalizar a datos nuevos. No obstante, se debe tener en cuenta dos cosas al momento de realizar esta validación: la primera es cómo se obtienen las particiones, o en otras palabras, en qué orden se divide la base de datos. Este es un punto importante considerando que el muestreo se realizó con la secuencia sobol, por lo que la base de datos sigue cierto orden determinado al ir completando el espacio muestral. La forma correcta de extraer las particiones es primero mezclando aleatoriamente la base de datos, y luego realizar las particiones. Por otro lado, es importante mencionar que cada partición debiese tener un número proporcional de cada clase, para esto se utiliza la validación cruzada *Stratified K-Fold*, que entrega automáticamente particiones con un número proporcional de clases de cada categoría.

Para el caso de este estudio, con las consideraciones antes mencionadas, se obtiene un modelo robusto con una exactitud general de 94.34 %.

De acuerdo a lo resultados de la Tabla 5.2, se observa que en el conjunto de mejor hiperparámetros son *gini*, los modelos no tienen profundidad máxima y la cantidad de hojas es dos, lo cual es un valor bajo en relación a la cantidad de casos modelados. Este conjunto de hiperparámetros puede ser un indicador de árboles poco balanceados y un modelo sobreajustado a la base de datos, por lo que a pesar de que pueda entregar buenas métricas de rendimiento, es posible que no tenga buen ajuste a datos nuevos.

Por otro lado, de la Tabla 5.3, al utilizar el hiperparámetro de *entropy*, el modelo produce árboles más balanceados en comparación a si utilizara el criterio *gini*. Además, cada árbol tiene una profundidad máxima de 10, lo cual limita que se genere sobreajuste a los datos. El modelo random forest al ser creado a partir de 300 árboles de decisión, con una cantidad de *sqrt* características máximas que se consideran al buscar la mejor división, se generan árboles distintos y permite disminuir aún más el sobreajuste a los datos de entrenamiento.

Finalmente, aparte de los resultados positivos encontrados en el modelo predictivo, se puede destacar la metodología que se siguió para el entrenamiento y validación del modelo. Esta metodología se puede replicar para modelos en los que existan bases de datos desbalanceadas y no sea posible ampliar la base de datos, o en el caso que las métricas no parecen mejorar cuando se aumenta la cantidad de casos simulados.

# Capítulo 6: Conclusiones y Recomendaciones a Trabajos Futuros

Los parámetros que influyen en la estimación de la estabilidad se pueden clasificar en tres categorías: según la calidad de la roca (GSI, el UCS y el  $m_i$ ), la geometría del caserón (el alto, el ancho, la inclinación y la razón de aspecto) y a los esfuerzos inducidos sobre las paredes del caserón (razón de esfuerzos y la profundidad). Sin embargo, ninguna variable parece tener un nivel alto de correlación directa con la cantidad de ELOS por sí sola, lo que indica que el problema debe abordarse con modelos multivariantes. Esto tiene sentido si se consideran los parámetros de los gráficos de estabilidad (radio hidráulico y número de estabilidad), los cuales combinan las variables más relevantes de cada categoría dentro de una nueva. Sin embargo, se puede notar que estas categorías no son suficientemente claras como para clasificar el grado de estabilidad que puede presentar cierta configuración de un caserón. Además, el criterio de daño que mejor ajuste tiene en los tres casos de estudio seleccionados corresponde a un criterio de deformación de *strain increment total measure*. Este criterio logra representar fallas por tensión y corte en el macizo y en las estructuras. Sin embargo, para otros tipos de falla es necesario realizar calibraciones más detalladas y contar con más información de un caso de estudio para ello.

De los resultados del modelo predictivo, se obtuvieron valores altos en las métricas de evaluación del modelo final. Esto significa que el modelo predictivo logra representar los resultados que el modelo numérico tendría con datos nuevos, sin la necesidad de simular los casos explícitamente. El resultado exitoso fue posible gracias a la técnica de balanceo de la base de datos sintética, ya que simular más casos no parecía aumentar las métricas de las clases intermedias. De esta manera, se logró el objetivo general de generar una herramienta predictiva que logre representar más condiciones e incorpore más parámetros que los gráficos empíricos.

Los resultados obtenidos pueden aportar al campo de la geomecánica, ya que el desarrollo de este modelo predictivo puede complementar el uso de gráficos empíricos para la estimación preliminar de dimensiones de caserones. Si bien no es posible de visualizar fácilmente el resultado del modelo predictivo, este puede ser implementado en formato de una aplicación o una página web que entregue el resultado de forma visual. A diferencia de los gráficos de estimación de estabilidad utilizados en la actualidad, este modelo podría tener una ventaja al ser basado en una mayor cantidad de casos. Sin embargo, es importante tener en cuenta que se debe realizar una calibración al criterio de daño, para asegurar la representatividad del modelo.

A pesar del éxito del trabajo, existen mejoras que pueden realizarse en futuros estudios. Por ejemplo, la primera recomendación tiene que ver con que se consideraron todas las configuraciones dentro del rango de valores de las categorías, por lo que podrían existir configuraciones inviables de geometría o combinaciones de parámetros geológicos no realistas. Por esta razón, se recomienda para estudios futuros definir relaciones entre variables que entreguen casos típicos y configuraciones posibles. En relación al modelo numérico se recomienda calibrar con más casos de estudio del criterio con el que se interpreta el daño. También hacer sensibilidades variando el factor de daño y mejorar la densificación de la malla para casos extremos de geometría en los que ocurren singularidades numéricas. Por último, para el modelo predictivo se recomienda usar una secuencia de muestreo iterativa que vaya completando espacios muestrales donde se encuentren las clasificaciones de estabilidad con menos número de casos para optimizar las simulaciones. Un potencial trabajo futuro podría ser calcular la estabilidad de las otras caras del caserón, junto con su probabilidad de ocurrencia.

# Bibliografía

Adoko, A.C., Saadaari, F., Mireku-Gyimah, D., Imashev, A., (2022). A Feasibility Study on The Implementation of Neural Network Classifiers for Open Stope Design. *Geotech. Geol. Eng.* 40, 677–696. <https://doi.org/10.1007/s10706-021-01915-8>

Barton, N., Lien, R., and Lunde, J. (1974). Engineering classification of rock masses for the design of tunnel support. *Rock Mechanics*, Vol. 6, No. 4, pp. 189-236.

Banko, M., & Brill, E. (2001). Scaling to very very large corpora for natural language disambiguation. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics* (pp. 26–33). Toulouse, France: Association for Computational Linguistics. <https://doi.org/10.3115/1073012.1073017>

Bewick, R., & Kaiser, P. K. (2009). Numerical assessment of factor B in Mathews' method for open stope design. In *Proceedings of the 3rd CANUS Rock Mechanics Symposium* (Vol. 1113). Toronto, ON, Canada.

Brady, B.H.G. Brown, E.. (2006). *Rock Mechanics for underground mining: Third edition*. 10.1007/978-1-4020-2116-9.

Brown, E. T. (1970). Strength of models of rock with intermittent joints. *Journal of the Soil Mechanics and Foundations Division*, 96(SM6), 1935-1949.

Capes, G.W., (2009). Open stope hangingwall design based on general and detailed data collection in unfavourable hangingwall conditions.

Castro, C. (2015). Modelamiento numérico de la dilución por sobre excavación en minería subterránea explotada por sublevel stoping. Doctoral Thesis, Universidad de Chile.

Cavieres, P. & Gaete, S & Lorig, L. & Gómez, P. (2003). Three-dimensional analysis of fracturing limits induced by large scale underground mining at El Teniente mine. 893.

Clark, L. M., & Pakalnis, R. C. (1997). An empirical approach for estimating unplanned dilution from open stope hangingwalls and footwalls. In *Proceedings 99th Annual General Meeting* (Vol. 33).

Contreras, E. (1995). Métodos y sistemas de explotación. Guía básica-plan 1990. Departamento de Ingeniería en minas. Universidad de Santiago de Chile.

Diederichs, M. S., & Kaiser, P. K. (1999). Tensile strength and abutment relaxation as failure control mechanisms in underground excavations. *International Journal of Rock Mechanics and Mining Sciences*, 36(1), 69–96.

Dong, L., Li, X., Peng, K. (2013). Prediction of rockburst classification using Random Forest. *Trans. Nonferrous Met. Soc. China* 23, 472–477. [https://doi.org/10.1016/S1003-6326\(13\)62487-5](https://doi.org/10.1016/S1003-6326(13)62487-5)

Dong, J., & Qian, Q. (2022). A Density-Based Random Forest for Imbalanced Data Classification. *En Future Internet* (Vol. 14, Issue 3, p. 90). MDPI AG. <https://doi.org/10.3390/fi14030090>

Erdogan Erten, G., Bozkurt Keser, S., Yavuz, M. (2021). Grid Search Optimised Artificial Neural Network for Open Stope Stability Prediction. *Int. J. Min. Reclam. Environ.* 35, 600–617. <https://doi.org/10.1080/17480930.2021.1899404>

FISHER, R. A. (1936). THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS. En *Annals of Eugenics* (Vol. 7, Issue 2, pp. 179–188). Wiley. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>

Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.

González de Vallejo, L. I., Ferrer, M., Ortuño, L., Oteo, C. (2002). *Ingeniería geológica*. Pearson Educación.

Gramacy, R.B., (2021). *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences*, 1st ed. Chapman & Hall.

Griffith, A. A. (1924). Theory of rupture. *Proceedings of the 1st International Congress on Applied Mechanics* (pp. 55-63). Delft, The Netherlands.

Hadjigeorgiou, J., Leclaire, J., & Potvin, Y. (1995). An update of the stability graph method of open stope design. In *97th Annual General Meeting, CIM* (pp. 154-161). Halifax, Nova Scotia: Canadian Institute of Mining, Metallurgy and Petroleum.

Halton, J. (1964). Algorithm 247: Radical-inverse quasi-random point sequence. *Communications of the ACM*, 7, 701-702. <https://doi.org/10.1145/355588.365104>

Hoek, E. (1965). Rock fracture under static stress conditions. (Report MEG, p. 383). Council for Scientific and Industrial Research.

Hoek, E. (1994). STRENGTH OF ROCK AND ROCK MASSES. *NEWS JOURNAL OF INTERNATIONAL SOCIETY FOR ROCK MECHANICS*, 2(2), 4-16. SID. <https://sid.ir/paper/546357/en>

Hoek, E. (2001). Rock mass properties for underground mines. En W. A. Hustrulid & R. L. Bullock (Eds.), *Underground Mining Methods: Engineering Fundamentals and International Case Studies*. Society for Mining, Metallurgy, and Exploration (SME).

Hoek, E., Carranza-Torres, C., & Corkum, B. (2002). Hoek-Brown failure criterion – 2002 Edition. En *Proceedings of the NARMS-TAC Conference* (pp. 267-273). Toronto.

Hoek, E., & Brown, E. T. (1980). Empirical Strength Criterion for Rock Masses. En *Journal of the Geotechnical Engineering Division* (Vol. 106, Issue 9, pp. 1013–1035). American Society of Civil Engineers (ASCE). <https://doi.org/10.1061/ajgeb6.0001029>

Hoek, E., & Brown, E. T. (2019). The Hoek–Brown failure criterion and GSI – 2018 edition. En *Journal of Rock Mechanics and Geotechnical Engineering* (Vol. 11, Issue 3, pp. 445–463). Elsevier BV. <https://doi.org/10.1016/j.jrmge.2018.08.001>

Hoek, E., & Diederichs, M. S. (2006). Empirical estimation of rock mass modulus. En *International Journal of Rock Mechanics and Mining Sciences* (Vol. 43, Issue 2, pp. 203–215). Elsevier BV

<https://doi.org/10.1016/j.ijrmms.2005.06.005->

Hoek, E., Kaiser, P.K., & Bawden, W.F. (1995). Support of underground excavations in hard rock. CRC Press.

Hofert, M., Prasad, A., Zhu, M. (2018). Quasi-random sampling for multivariate distributions via generative neural networks (Versión 3). arXiv. <https://doi.org/10.48550/ARXIV.1811.00683>

Hughes, R., Mitri, H. S. & Lecomte, E. (2010) Examining the influence of stope strike length on unplanned ore dilution in narrow vein longitudinal mining, paper presented at the 44th U.S. Rock Mechanics Symposium and 5th U.S.-Canada Rock Mechanics Symposium, City.

Hustrulid, W. A., & Bullock, R. L. (Eds.). (2001). Underground mining methods: Engineering fundamentals and international case studies. SME.

Itasca Consulting Group, Inc. (2023) FLAC3D — Fast Lagrangian Analysis of Continua in Three-Dimensions, Ver. 9.0. Minneapolis: Itasca.

Jaeger, J.C. (1960) Shear failure of anisotropic rocks. *Geol Mag* 97:65–72

Jheald, (2011). A 2D sequence of pseudorandom numbers generated in R. Obtenido de: [https://en.wikipedia.org/wiki/Sobol\\_sequence](https://en.wikipedia.org/wiki/Sobol_sequence)

Jing, L. (2002). A review of techniques, advances and outstanding issues in numerical modelling for rock mechanics and rock engineering. *International Journal of Rock Mechanics and Mining Sciences*.

Jorquera, M., Korzeniowski, W., Skrzypkowski, K. (2023). Prediction of dilution in sublevel stoping through machine learning algorithms. *IOP Conf. Ser. Earth Environ. Sci.* 1189, 012008. <https://doi.org/10.1088/1755-1315/1189/1/012008>

Kirsch (1898), Die Theorie der Elastizität und die Bedürfnisse der Festigkeitslehre. *Zeitschrift des Vereines deutscher Ingenieure*, 42, 797–807.

Khan, S. Data Science Explained: Random Forests (2024) . Extraído de: <https://godatadrive.com/blog/random-forests>

Laubscher, D. H. (1994). Cave mining - the state of the art. *Journal of the Southern African Institute of Mining and Metallurgy*, 94, 279-293.

Li, Y., Wang, C., Xu, Jiankun, Zhou, Z., Xu, Jianhui, Cheng, J. (2021). Rockburst Prediction Based on the KPCA-APSO-SVM Model and Its Engineering Application. *Shock Vib.* 2021, 1–12. <https://doi.org/10.1155/2021/7968730>

Mathews, K. E., Hoek, E., Wyllie, D. C., & Stewart, S. B. V. (1981). *Prediction of stable excavation spans at depths below 1000m in hard rock mines* (Tech Report DSS Serial No. OSQ80-00081). Ottawa, Canada: Canada Centre for Mineral and Energy Technology.

Mawdesley, C., Trueman, R., & Whiten, W. (2001). Extending the Mathews stability graph for open stope design. *Transactions of the Institution of Mining and Metallurgy, Section A: Mining Industry*, 110, A27–A39.

- Minería Chilena. (2024, Julio). Minería subterránea: Con Codelco como punta de lanza. *Minería Chilena*, 517(44).
- McClintock, F. A. (1962). Friction on Griffith cracks in rocks under pressure. In Proc. 4th US Nat. Congr. Appl. Mech. (Vol. 2, pp. 1015-1022).
- Mitri, H. S., Hughes, R., & Zhang, Y. (2011). New rock stress factor for the stability graph method. *International Journal of Rock Mechanics and Mining Sciences*, 48, 141–145.
- Morokoff, W. J. & Caffisch R. E. (1995). Quasi-Monte Carlo Integration. *Journal of Computational Physics*. 122 (2), 218-230. <https://doi.org/10.1006/jcph.1995.1209>
- Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python: A guide for data scientists. O'Reilly Media.
- Nickson, S. D. (1992). *Cable support guidelines for underground hard rock mine operations* (Master's thesis, University of British Columbia, Vancouver, Canada).
- Niederreiter, H. (1992). Random Number Generation and Quasi-Monte Carlo Methods, Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Nikolić, M., Roje-Bonacci, T., & Ibrahimbegović, A. (2014). Overview of the numerical methods for the modelling of rock mechanics problems. ISSN 1848-6339. <https://doi.org/10.17559/TV-20140521084228>
- Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python. *JMLR* 12, pp. 2825-2830, 2011.
- Pereira, N.(2017). Metodología de diseño geomecánico para sublevel stoping con caserones auto-soportados. Trabajo para titulación presentando en conformidad a los requisitos para obtener títulos de ingeniería civil en minas.
- Potvin, Y. (1988). *Empirical open stope design in Canada* (Doctoral dissertation, University of British Columbia, Vancouver, Canada).
- Qi, C., Chen, Q. (2018). Evolutionary Random Forest Algorithm for Predicting the Maximum Failure Depth of Open Stope Hangingwalls. *IEEE Access* 6, 72808–72813. <https://doi.org/10.1109/ACCESS.2018.2880009>
- Qi, C., Fourie, A., Du, X., Tang, X. (2018a). Prediction of open stope hangingwall stability using random forests. *Nat. Hazards* 92, 1179–1197. <https://doi.org/10.1007/s11069-018-3246-7>
- Qi, C., Fourie, A., Ma, G., Tang, X. (2018b). A hybrid method for improved stability prediction in construction projects: A case study of stope hangingwall stability. *Appl. Soft Comput.* 71, 649–658. <https://doi.org/10.1016/j.asoc.2018.07.035>
- Qi, C., Fourie, A., Ma, G., Tang, X., Du, X. (2018c). Comparative Study of Hybrid Artificial Intelligence Approaches for Predicting Hangingwall Stability. *J. Comput. Civ. Eng.* 32, 04017086. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000737](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000737)
- Qi, Z., Liu, H., An, Z., Yin, Y., Liu, Z., Zhu, M., Wang, Y., & Zhou, H. (2023). Comprehensive evaluation method of stope stability and its application in deep metal mine. 18(3), e0283205.

<https://doi.org/10.1371/journal.pone.0283205>

Rankine, R. M. (2004). The geotechnical characterisation and stability analysis of BHP Billiton's Cannington Mine paste fill (PhD thesis). James Cook University. Recuperado de <http://researchonline.jcu.edu.au/1260/>

Sainsbury, Bre-Anne. (2012). A Model for Cave Propagation and Subsidence Assessment in Jointed Rock Masses. 10.13140/RG.2.1.4280.9446.

Sainsbury, B. L., & Sainsbury, D. P. (2017). Practical use of the ubiquitous-joint constitutive model for the simulation of anisotropic rock masses. *Rock Mechanics and Rock Engineering*, 50\*(6), 1507-1528. <https://doi.org/10.1007/s00603-017-1177-3>

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210-229.

Scoble, M. J., & Moss, A. (1994). Dilution in underground bulk mining: Implications for production management, mineral resource evaluation, II. *Geological Society Special Publications*, 79, 95–108.

Sobol, I.M. (1967). Distribution of points in a cube and approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7, 86–112. [https://doi.org/10.1016/0041-5553\(67\)90144-9](https://doi.org/10.1016/0041-5553(67)90144-9)

Stewart, P. C., & Trueman, R. (2008). Strategies for minimising and predicting dilution in narrow-vein mines – NVD Method. In *Narrow Vein Mining Conference 2008*. Victoria, Australia: The Australasian Institute of Mining and Metallurgy, pp. 153-164.

Stewart, S. B. V., & Forsyth, W. W. (1995). The Mathews method for open stope design. *CIM Bulletin*, 88, 45–53.

Suorineni, F.T. (2014). Reflections on Empirical Methods in Geomechanics – The Unmentionable and Hidden Risks. Sydney : AusIMM. 3rd Australian Ground Control in Mining Conference. págs. 143 – 15

Suorineni, F., Papaioanou, A., Baird, L., & Hines, D. (2016). A dilution-based stability graph for open stope design. In *Proceedings of the 7th International Conference and Exhibition on Mass Mining*. Sydney, Australia: The Australasian Institute of Mining and Metallurgy, pp. 511-522.

Szmigiel, A., Apel, D.B. (2022). Predicting the stability of open stopes using Machine Learning. *J. Sustain. Min.* 21, 241–248. <https://doi.org/10.46873/2300-3960.1369>

Trueman, R., Mikula, P., Mawdesley, C., & Haries, N. (2000). Experience in Australia with the application of the Mathews method of open stope design. *CIM Bulletin*, 93, 162–167.

Trueman, R., & Mawdesley, C. (2003). Predicting cave initiation and propagation. *CIM Bulletin*, 96, 54–59.

Vallejos, J., Delonca, A., Fuenzalida, J., & Burgos, L. (2016). Statistical analysis of the stability number adjustment factors and implications for underground mine design. *International Journal of Rock Mechanics and Mining Sciences*, 100, 104-112. <https://doi.org/10.1016/j.ijrmms.2016.06.001>

Vesel, T. (2019). EFFICIENT COMPUTATION OF SOBOL' QUASI-RANDOM GENERATOR. En International Journal of Students' Research in Technology amp; Management (Vol. 7, Issue 2, pp. 01–04). GIAP Journals <https://doi.org/10.18510/ijstrtm.2019.721>.

Villaescusa, E.(2014). Geotechnical Design for Sublevel Open Stopping. CRC Press.

Vongpaisal, S., Li, G., Pakalnis, R. & Brady, T. (2009) New 3d engineering curves for predicting stope stability and mining dilution in longitudinal blasthole mining operations. International Journal of Mining, Reclamation and Environment, 23(2), 92-102.

Wang, J., Milne, D., Pakalnis, R. (2002). Application of a neural network in the empirical design of underground excavation spans. Min. Technol. 111, 73–81.<https://doi.org/10.1179/mnt.2002.111.1.73>

Wyllie, D. C., & Mah, C. W. (2004). Rock Slope Engineering: Civil and Mining (4th ed.). CRC Press.

Zhalel, M., Adoko, A.C., Korigov, S. (2020). An approach to stope stability assessment in open stope mining environment: 54th U.S. Rock Mechanics/Geomechanics Symposium.

Zhou, J., Li, X., Mitri, H., (2015). Comparative performance of six supervised learning methods for the development of models of hard rock pillar stability prediction. Nat. Hazards J. Int. Soc. Prev. Mitig. Nat. Hazards 79, 291–316.

ZHOU, J., LI, X., SHI, X., WEI, W., & WU, B. (2011). Predicting pillar stability for underground mine using Fisher discriminant analysis and SVM methods. En Transactions of Nonferrous Metals Society of China (Vol. 21, Issue 12, pp. 2734–2743). Elsevier BV. [https://doi.org/10.1016/S1003-6326\(11\)61117-5](https://doi.org/10.1016/S1003-6326(11)61117-5)