



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE GEOLOGÍA

**ESTIMACIÓN DE LA ORIENTACIÓN Y FRECUENCIA DE FRACTURAS EN
SONDAJES A PARTIR DE ANÁLISIS DE IMÁGENES Y MACHINE
LEARNING**

MEMORIA PARA OPTAR AL TÍTULO DE GEÓLOGO

FRANCO GIOVANNI HORMAZÁBAL GODOY

PROFESOR GUÍA:
MANUEL GARCÍA HURTADO

MIEMBROS DE LA COMISIÓN:
SOFÍA REBOLLEDO LEMUS
SERGIO SEPÚLVEDA VALENZUELA

SANTIAGO DE CHILE
2022

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE GEÓLOGO
POR: FRANCO GIOVANNI HORMAZÁBAL GODOY
FECHA: 2022
PROF. GUÍA: MANUEL GARCÍA HURTADO

ESTIMACIÓN DE LA ORIENTACIÓN Y FRECUENCIA DE FRACTURAS EN SONDAJES A PARTIR DE ANÁLISIS DE IMÁGENES Y MACHINE LEARNING

Este trabajo presenta, evalúa y analiza métodos para estimar los ángulos alfa y beta de fracturas, y la frecuencia de fracturas equivalente en tramos de roca muy fracturada en testigos de sondajes diamantinos, mediante análisis de imágenes y Machine Learning. Para la estimación de los ángulos alfa y beta en sondajes orientados, se propone un método llamado -en este estudio- “de los 3 puntos”, en el cual se marcan 3 puntos sobre la traza frontal de una fractura en el plano de la imagen, y mediante transformaciones geométricas, se obtienen los ángulos alfa y beta. Para la estimación de la frecuencia de fracturas se utilizó la información de sondas Televiwer, contrastando la ocurrencia de fracturas in-situ en el macizo con las regiones de roca fracturada presentes en el sondaje. Diversas técnicas de análisis de imágenes se aplicaron a dichas regiones para extraer características que luego se intentaron correlacionar con la frecuencia de fracturas observada, mediante el uso de Machine Learning.

Los resultados obtenidos, sobre 439 fracturas, mostraron que el método de los 3 puntos aproxima con una precisión del 88 % el cálculo de ángulos alfa, respecto a los obtenidos por medición directa, presentando una aproximación más uniforme para ángulos entre 40 y 80°, una ligera sobrestimación para los menores a 40° y una ligera subestimación para los mayores a 80°, mientras que, por otra parte, aproxima con una precisión del 94 % el cálculo de los ángulos beta, siendo uniforme en todo su rango. De igual forma, los sets de fracturas reconocidos en stereonet a partir del método de los 3 puntos son similares a los obtenidos por medición directa.

La estimación de la frecuencia de fracturas consideró parámetros geométricos de los fragmentos presentes en los tramos de roca fracturada, tales como la distribución de sus tamaños y su esfericidad. Los resultados obtenidos, sobre 448 tramos de roca fracturada, mostraron que el modelo de Machine Learning implementado predijo con un 65-68 % de precisión la frecuencia de fracturas observada, presentando una aproximación más uniforme en el rango de 10 a 20 fracturas/m, una ligera sobrestimación en frecuencias de fracturas menores a 10 fracturas/m y una subestimación importante en mayores a 20 fracturas/m, correlacionando los mayores errores con la escasez de datos para entrenar el modelo en dichos rangos.

*A mis padres,
los mejores de la Tierra.*

Gracias por todo

Agradecimientos

Agradezco en primer lugar a mi familia; a mi madre por enseñarme a luchar siempre con una sonrisa, a mi padre por mostrarme lo que es la bondad y solidaridad, y a mi hermano por su eterna compañía. Agradezco también a los miembros no humanos de mi familia, a quienes están y a quienes ya partieron, por su infinito amor incondicional.

Agradezco a mis amigas y amigos del colegio, la universidad, la Compañía Gris y la vida, con quienes he vivido incontables e imborrables momentos, y sin los cuales no hubiera podido sobrellevar este proceso.

Agradezco a los profesores, docentes y funcionarios que forjaron mi crecimiento personal y educativo, y que dedican su día a día a mejorar nuestra sociedad. Agradezco a Sofía Rebolledo y Sergio Sepúlveda por su disposición, conocimientos y apoyo en este trabajo.

Finalmente, agradezco a todos los miembros de Graiph Inteligencia Artificial por darme la oportunidad de desarrollar mis primeros pasos como profesional; a Manuel García y Marcelo Surjan por otorgarme su total confianza y generar un grato ambiente laboral, y a Diego Opazo y Emilio Carreño por sus aportes en este estudio, en el desarrollo de aplicaciones y etiquetado de imágenes, respectivamente.

Tabla de Contenido

1. Introducción	1
1.1. Objetivos e hipótesis de trabajo	6
1.2. Metodología	6
2. Cálculo de ángulos alfa y beta con el método de los 3 puntos	9
2.1. Preprocesamiento de datos	9
2.2. Método de los 3 puntos	10
2.2.1. Corrección de perspectiva	13
2.3. Procesamiento de datos	17
2.4. Resultados	22
2.4.1. Ángulos alfa	22
2.4.2. Ángulos beta	23
2.4.3. Errores cruzados	25
2.4.4. Proyecciones estereográficas (stereonet)	27
3. Estimación de frecuencia de fracturas con Machine Learning	35
3.1. Preprocesamiento de datos	35
3.2. Procesamiento de datos	38
3.2.1. Procesamiento de imágenes	38
3.2.2. Obtención de parámetros asociados a los fragmentos	39
3.3. Modelos de Machine Learning	43
3.3.1. Modelo de Regresión Básica	44
3.3.2. Validación cruzada de 10 iteraciones	45
3.4. Resultados	47
3.4.1. Mejor iteración	47
3.4.2. Peor iteración	48
3.4.3. Total de iteraciones	50
4. Discusión	53
4.1. Cálculo de ángulos alfa y beta con el método de los 3 puntos	53
4.2. Estimación de frecuencia de fracturas con Machine Learning	55
5. Conclusiones	58
Bibliografía	58

Índice de Tablas

2.1.	Resumen de ángulos Dip y Dip Direction de los métodos directo y de los 3 puntos, de los sets de fracturas reconocidos en los Sondajes 1, 2 y 3, previo a la aplicación de la ponderación de Terzaghi.	30
2.2.	Resumen de ángulos Dip y Dip Direction de los métodos directo y de los 3 puntos, de los sets de fracturas reconocidos en los Sondajes 1, 2 y 3, una vez aplicada la ponderación de Terzaghi.	34
3.1.	Coefficiente de determinación ajustado (R_a^2) de cada una de las iteraciones y su valor promedio.	47

Índice de Ilustraciones

1.1.	Representación gráfica de los ángulos internos alfa (α) y beta (β) que definen la orientación de una estructura plana intersecada por un testigo de sondaje (Extraído de Marjoribanks, 2010).	2
1.2.	Esquemas de pozo en macizo rocoso intersecando fracturas, con sus respectivas proyecciones geométricas captadas por la sonda Televierer (Extraído de Li et al., 2013).	3
1.3.	Esquema del sistema de instalación y conexión de sonda Televierer (Extraído de Li et al., 2013).	4
1.4.	Ejemplo de Transformación de perspectiva de la librería OpenCV, en Python.	7
2.1.	Aplicación desarrollada para etiquetar fracturas y tramos de roca muy fracturada a partir de imágenes de testigos de sondajes, con imagen del estudio de fracturas cargada.	9
2.2.	Sistema de referencia de fractura intersecada por un testigo de sondaje (Elaboración propia).	11
2.3.	Esquema y sistema de referencia de caja de testigos de sondajes y su imagen virtual capturada por una cámara fotográfica (Elaboración propia).	14
2.4.	Ejemplo de cómo se observaría una caja de testigos de sondajes, con corrección de perspectiva para retornar de una imagen virtual a una real.	16
2.5.	Histograma de ángulos Alfa Método directo.	22
2.6.	Comparación de ángulos Alfa: Método de los 3 puntos versus Método directo.	22
2.7.	Análisis de errores: Error Alfa versus Alfa Método directo.	23
2.8.	Histograma de ángulos Beta Método directo.	23
2.9.	Comparación de ángulos Beta: Método de los 3 puntos versus Método directo.	24
2.10.	Análisis de errores: Error Beta versus Beta Método directo.	24
2.11.	Análisis de errores: Error Beta versus Alfa Método directo.	25
2.12.	Análisis de errores: Error Alfa versus Beta Método directo.	25
2.13.	Análisis de errores: Error Beta versus Error Alfa.	26
2.14.	Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 1, con sus respectivos sets de fracturas, previo a la aplicación de la ponderación de Terzaghi.	27
2.15.	Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 2, con sus respectivos sets de fracturas, previo a la aplicación de la ponderación de Terzaghi.	28
2.16.	Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 3, con sus respectivos sets de fracturas, previo a la aplicación de la ponderación de Terzaghi.	29

2.17.	Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 1, con sus respectivos sets de fracturas, una vez aplicada la ponderación de Terzaghi.	31
2.18.	Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 2, con sus respectivos sets de fracturas, una vez aplicada la ponderación de Terzaghi.	32
2.19.	Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 3, con sus respectivos sets de fracturas, una vez aplicada la ponderación de Terzaghi.	33
3.1.	Aplicación desarrollada para etiquetar fracturas y tramos de roca muy fracturada a partir de imágenes de testigos de sondajes, con imágenes del estudio de tramos fracturados cargadas.	36
3.2.	Esquema de obtención de los pares de datos de los tramos fracturados; imagen del tramo y frecuencia de fracturas observada asociada.	37
3.3.	Histograma de Frecuencia de Fracturas Observada.	38
3.4.	Ejemplo de pasos y transformaciones aplicadas sobre una imagen de fragmentos de roca, reconociendo el área y contorno de cada fragmento, utilizando la librería OpenCV en Python.	39
3.5.	Relación cualitativa entre la Frecuencia de Fracturas Observada (eje y) y los parámetros asociados a la imagen de cada tramo de roca fracturada (eje x).	42
3.6.	Arquitectura de una red de Machine Learning (Extraído de Adcock et al., 2015).	43
3.7.	Fragmentos de código de la implementación de un Modelo de Regresión Básica en Python, utilizando las librerías TensorFlow y Keras.	44
3.8.	Esquema de Validación cruzada de 10 iteraciones, utilizada en este estudio (Elaboración propia).	45
3.9.	Comparación de Frecuencia de Fracturas: Predicha versus Observada, iteración 2.	47
3.10.	Análisis de errores: Error Frecuencia de Fracturas versus Frecuencia de Fracturas Observada, iteración 2.	48
3.11.	Comparación de Frecuencia de Fracturas: Predicha versus Observada, iteración 5.	49
3.12.	Análisis de errores: Error Frecuencia de Fracturas versus Frecuencia de Fracturas Observada, iteración 5.	49
3.13.	Comparación de Frecuencia de Fracturas: Predicha versus Observada, para la totalidad de iteraciones recopiladas.	50
3.14.	Análisis de errores: Error Frecuencia de Fracturas versus Frecuencia de Fracturas Observada, para la totalidad de iteraciones recopiladas.	51
3.15.	Relación cualitativa entre el Error Frecuencia de Fracturas (eje y) y los parámetros asociados a la imagen de cada tramo de roca fracturada (eje x).	52

Capítulo 1

Introducción

Los sondajes son perforaciones de pequeño diámetro y gran longitud que se efectúan para alcanzar zonas inaccesibles desde la superficie o laboreos mineros, los cuales son posteriormente descritos y analizados por geólogos. Estos se realizan, por ejemplo, para estimar los recursos de yacimientos, determinar las propiedades geomecánicas del subsuelo o determinar el potencial hidrogeológico de acuíferos. El proceso de descripción visual induce una gran variabilidad entre las descripciones, producto de los distintos criterios formativos y la experiencia de cada profesional a cargo, afectando la calidad de los modelos geológicos. Además, este proceso consume una gran cantidad de tiempo y tiene un elevado costo.

Actualmente existen empresas que ofrecen software de apoyo a la descripción geológica de sondajes, basado en las fotografías de estos mismos, lo que se presenta con algunas ventajas con respecto al método tradicional de descripción directa, tales como la unificación de criterios descriptivos, la automatización de cálculos y el ahorro sustancial de tiempo. Sin embargo, estos software se encuentran en constante desarrollo y presentan algunas limitaciones. Aunque describir la litología, alteración y mineralogía de un sondaje a partir de sus fotografías es complejo, realizar por ejemplo, el conteo de fracturas, la medición del largo de tramos y la obtención de otras variables resulta mucho más rápido y automático con un software adecuado. Con esto se puede estimar rápidamente distintos parámetros geotécnicos, como el RQD, el espaciamiento medio de estructuras, entre otros, claves para determinar las propiedades geomecánicas de la roca o suelo.

Uno de los componentes importantes en la descripción de sondajes geotécnicos es la caracterización de fracturas y su orientación. Los ángulos que forman las estructuras planares como fracturas con líneas y planos de referencia en un sondaje orientado se pueden utilizar para calcular la orientación de estas con respecto a los ejes geográficos vertical y horizontal; norte-sur y este-oeste, respectivamente (Goodman, 1976, 1980; Reedman, 1979). Las líneas de referencia utilizadas en un sondaje son su eje central y la línea BOH (siglas de Bottom of Hole Line, en inglés), donde esta última corresponde a una línea de rumbo constante dibujada en la superficie del sondaje. Los planos de referencia del sondaje son el plano vertical y su plano de circunferencia (Figura 1.1). Los ángulos que forman una estructura planar con las líneas de referencia se miden directamente en el sondaje y luego se convierten en mediciones de rumbo y manto mediante cálculos matemáticos o gráficos, según la orientación de los planos y líneas de referencia, conocidos a partir del levantamiento del sondaje.

La intersección del testigo de sondaje cilíndrico con una estructura planar corresponde a una elipse. El eje mayor de la elipse está marcado por puntos de máxima curvatura denominados puntos de inflexión, ubicados en lados opuestos de la traza del plano en la superficie del sondaje. Los puntos de inflexión de varios planos paralelos estrechamente espaciados definen una línea de inflexión (Marjoribanks, 2010). Los extremos del eje mayor de la elipse de intersección se denominan E-E', donde E es el extremo inferior del eje, es decir, el que forma un ángulo agudo con la dirección hacia abajo del eje central, y E' es el extremo superior del eje mayor de la elipse (Figura 1.1).

El ángulo alfa (α) es el primero de los ángulos centrales internos que se miden para determinar la orientación de una estructura planar. Corresponde al ángulo agudo entre el eje central (CA) y el eje mayor de la elipse de intersección E-E' (Figura 1.1).

El plano geométrico en ángulo recto con el eje del sondaje es el plano de circunferencia y su trazo hipotético en la superficie del sondaje es circular. El punto BOH se define como la intersección de la traza del plano de circunferencia con la línea BOH a lo largo del sondaje (Figura 1.1).

El ángulo beta (β) es el segundo de los ángulos centrales internos que se miden para determinar la orientación de una estructura planar. Corresponde al ángulo, medido en el sentido de las agujas del reloj, alrededor del plano de la circunferencia, perpendicular al eje CA, desde el punto BOH al punto E (Figura 1.1).

A partir de los ángulos alfa y beta de una estructura, y el azimut y manto del eje del sondaje a la profundidad donde se tomaron las medidas, se puede determinar el rumbo y manto de la estructura con respecto a los ejes geográficos (Stigsson & Munier, 2013).

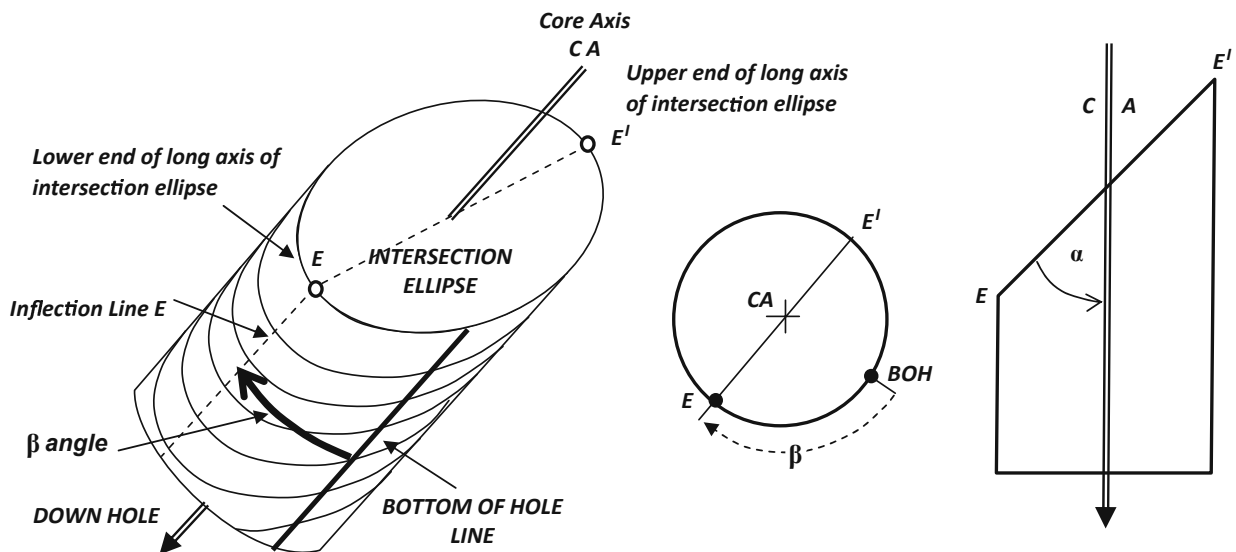


Figura 1.1: Representación gráfica de los ángulos internos alfa (α) y beta (β) que definen la orientación de una estructura plana intersecada por un testigo de sondaje (Extraído de Marjoribanks, 2010).

La medición de ángulos alfa y beta de fracturas se ha realizado tradicionalmente en sondajes de forma directa, con instrumentos como goniómetros o transportadores. Además, existen dispositivos electrónicos que estiman estos ángulos mediante láser.

Otro de los parámetros geotécnicos importantes considerados en la descripción de sondajes es la frecuencia de fracturas (FF). Esta corresponde a la cantidad de fracturas presentes por unidad de largo para un tramo de testigo seleccionado, se mide usualmente en fracturas/m y se calcula según la ecuación

$$FF = \frac{\text{Número de Fracturas}}{\text{Largo del Tramo}}. \quad (1.1)$$

Obtener la frecuencia de fracturas real de tramos de roca muy fracturada a partir de los testigos de sondaje es prácticamente imposible, ya que en estos tramos usualmente la roca ha perdido su disposición original al salir del pozo, haciendo complejo el reconocimiento y conteo de fracturas individuales. Sin embargo, es posible obtener la frecuencia de fracturas real a partir del estudio del pozo donde se ha realizado el sondaje, por ejemplo, mediante el uso de Televiewer.

El Televiewer es una sonda geofísica que se utiliza en el registro de pozos diamantinos (Figura 1.2.a) y puede complementar la caracterización geotécnica y geológica de sondajes. Este genera una imagen digital continua de alta resolución de las paredes del sondaje en 360° (Figuras 1.2.b y 1.2.c), la que permite determinar la dimensión y orientación de estructuras (Kamewada et al., 1989; Wang et al., 2002; Williams & Johnson, 2004), así como también posibles cambios litológicos (Paillet et al., 1990; Pusch, 1998).

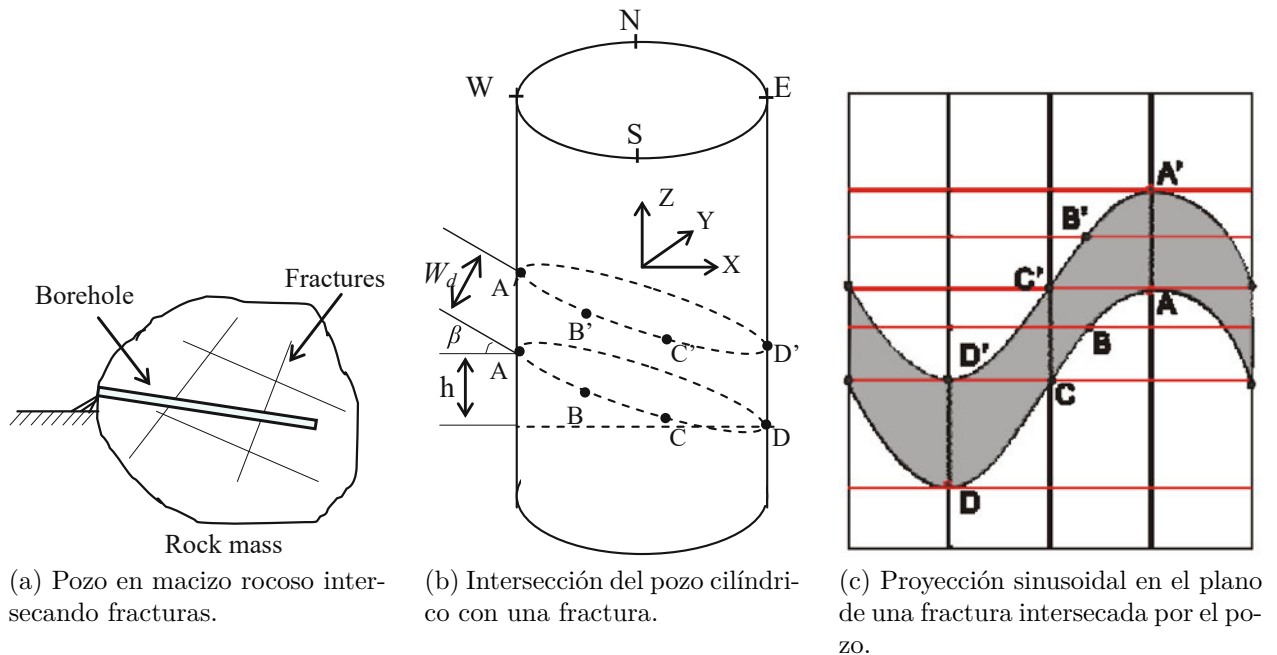


Figura 1.2: Esquemas de pozo en macizo rocoso intersecando fracturas, con sus respectivas proyecciones geométricas captadas por la sonda Televiewer (Extraído de Li et al., 2013).

La tecnología se ha aplicado ampliamente en exploración geológica (Li et al., 2013), especialmente en petróleo (Maddox, 1998; Tague, 1999; Palmer & Sparks, 1990), minería (Gochioco et al., 2002; Deltombe & Schepers, 2000), glaciares (Engelhardt et al., 1978), geotecnia e ingeniería ambiental (Lau et al., 1987; Miyakawa et al., 2000; Lahti, 2004; Cunningham, 2004; Cunningham et al., 2004; Schepers et al., 2001; Roberson & Hubbard, 2010; Uchita & Harada, 1993; Li et al., 2012a).

Aunque existen diferentes tipos de sistemas de cámaras digitales, el principio básico, los componentes y las operaciones de estos sistemas de prueba son casi los mismos (Li et al., 2013). En la Figura 1.3 se puede observar un esquema de ejemplo del sistema de instalación y conexión de sonda Televiewer.

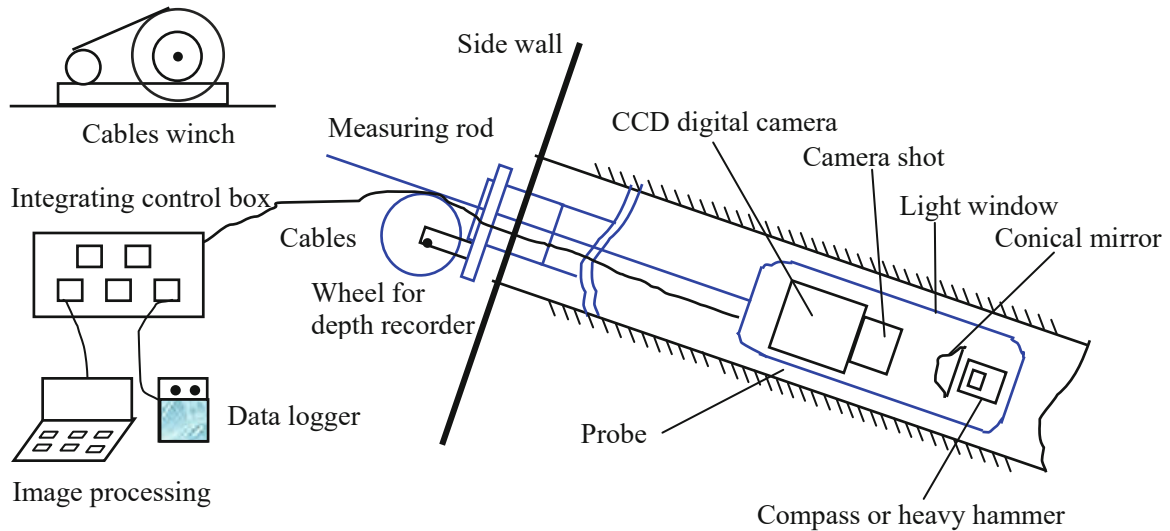


Figura 1.3: Esquema del sistema de instalación y conexión de sonda Televiewer (Extraído de Li et al., 2013).

Dependiendo de la forma en que se adquiere la imagen, el Televiewer puede ser óptico o acústico.

El Televiewer óptico captura imágenes en color real, nítidas y en alta resolución, lo que permite observar en detalle las características físicas de las paredes del pozo. En este caso, para tomar las imágenes correctamente, el pozo debe estar seco o con agua cristalina y limpia, y funciona tanto en pozos verticales como en pozos con ángulo. En términos generales, un lente “ojo de pez” permite captar una imagen continua de 360° a lo largo de todo el pozo. Estas imágenes son procesadas por un sensor de imagen digital y enviadas al operador de la sonda. Las paredes del pozo son iluminadas por LEDs para obtener imágenes más claras, y además, utiliza un espejo hiperbólico y un magnetómetro y acelerómetro, ambos de tres ejes, que permiten orientar las estructuras halladas y el pozo completo en tiempo real.

El Televiewer acústico funciona por medio de ultrasonido. Las imágenes capturadas en tiempo real son nítidas y muestran claramente las estructuras de las paredes del pozo. A diferencia del Televiewer óptico, funciona solamente en pozos llenos de agua o lodo de perforación, dado que las ondas acústicas requieren de un medio acuoso para desplazarse. El

Televiwer acústico posee un transductor fijo que emite ondas de ultrasonido, las cuales chocan con un espejo giratorio, dirigiéndolas en 360° a las paredes del pozo, luego son reflejadas al transductor, pasando por el espejo y midiendo dos parámetros importantes: el tiempo que le toma a la onda ir y regresar de la pared del pozo (tiempo de tránsito) y su amplitud. Las fracturas y sus ensanchamientos aumentan el tiempo de tránsito y disminuyen la amplitud de las ondas. Además, el Televiwer acústico, al igual que el Televiwer óptico, utiliza un magnetómetro y acelerómetro, ambos de tres ejes, que permiten, de igual forma, orientar las estructuras halladas y el pozo completo en tiempo real.

Para los casos donde en los tramos de roca muy fracturada se imposibilita el conteo de fracturas, se asume una frecuencia de fracturas equivalente de acuerdo a un protocolo establecido por cada empresa o geólogo, el cual muchas veces es de carácter cualitativo más que cuantitativo.

En el presente trabajo se definirán y estudiarán métodos alternativos para obtener los ángulos alfa y beta de fracturas y la frecuencia de fracturas equivalente para tramos de roca muy fracturada a partir de imágenes de testigos de sondajes diamantinos, con el fin de generar un aporte en el desarrollo de tecnologías que facilitan o entregan alternativas a la descripción de sondajes.

La evaluación y validación de los métodos propuestos se realizará comparando los resultados obtenidos por estos mismos, con datos tomados de métodos tradicionales para las mismas estructuras. En el caso de los ángulos alfa y beta, los resultados serán contrastados con mediciones directas con goniómetros, realizadas por un equipo de geología, y en el caso de la frecuencia de fracturas, con el conteo de fracturas in-situ realizada según datos de sonda Televiwer.

La estructura de este trabajo cuenta con 5 capítulos, que se presentan brevemente a continuación:

1. Introducción: Se presentan las generalidades y motivación del estudio propuesto, los objetivos general y específicos, y la hipótesis y metodología de trabajo.
2. Cálculo de ángulos alfa y beta con el método de los 3 puntos: Se presenta la descripción del método propuesto para calcular la orientación de una fractura a partir de una imagen, y los resultados de su evaluación al ser aplicado en un set de datos.
3. Estimación de frecuencia de fracturas con Machine Learning: Se presenta la descripción del método propuesto para estimar la frecuencia de fracturas de roca muy fracturada a partir de una imagen, y los resultados de su evaluación al ser aplicado en un set de datos.
4. Discusión: Se presenta una discusión de los resultados obtenidos en los capítulos anteriores, alcances de los mismos e implicancias para el estudio.
5. Conclusiones: Se presentan las principales conclusiones del estudio, incluyendo sus resultados, el cumplimiento de los objetivos, las contribuciones del trabajo realizado y las recomendaciones y sugerencias a futuro.

Adicionalmente, se incluye un apartado con la Bibliografía utilizada al final del documento.

1.1. Objetivos e hipótesis de trabajo

El objetivo general de este trabajo es proponer métodos estandarizados para estimar la orientación y frecuencia de fracturas en testigos de sondajes orientados a partir de imágenes.

Los objetivos específicos son los siguientes:

- Desarrollar un método de cálculo de ángulos alfa y beta para fracturas.
- Desarrollar un método de estimación de frecuencia de fracturas equivalente para tramos de roca muy fracturada.
- Comparar los resultados arrojados por los métodos propuestos con los obtenidos por otros métodos tradicionales de medición.

La hipótesis de trabajo es que los ángulos alfa y beta y la frecuencia de fracturas pueden estimarse por medio de los métodos propuestos con una precisión similar a los obtenidos por métodos tradicionales de medición.

1.2. Metodología

Revisión bibliográfica

En esta etapa se investigarán principalmente los siguientes tópicos:

- Definición de ángulos alfa y beta asociados a una estructura planar como fracturas.
- Funcionamiento y cálculo de parámetros de sonda Televier para el reconocimiento de fracturas.
- Métodos de procesamiento de imágenes acordes al estudio en cuestión, para implementar distintas técnicas y algoritmos para caracterización de texturas y elementos visuales.
- Modelos de Machine Learning acordes, para correlacionar las características visuales de tramos de roca fracturada con el número y frecuencia de fracturas reportada por la sonda Televier.
- Pruebas estadísticas para comparar dos grupos de variables, y evaluar y validar los métodos propuestos.

Recolección de datos

En esta etapa se recopilarán las imágenes a utilizar en el estudio, junto con los datos de sonda Televier y las mediciones de ángulos alfa y beta realizadas directamente por un equipo de geología. Las bases de datos serán revisadas y limpiadas para detectar inconsistencias y/o errores manuales.

Preprocesamiento de imágenes

En esta etapa se llevará a cabo la estandarización de imágenes, incluyendo la asignación de nomenclatura de nombre de archivos y el preprocesamiento visual de imágenes, que comprende correcciones de rotación, inclinación y perspectiva, de ser necesario.

La nomenclatura utilizada para asignar el nombre de cada imagen incluirá información del nombre del sondaje y de la profundidad de inicio y fin de los testigos correspondientes a las cajas en cuestión, seguidos de un guión bajo, de la siguiente forma: *SONDAJE_INICIO_FIN.JPG*. El inicio y fin se definen con un número de 6 dígitos, donde los 2 últimos corresponden a los decimales de la profundidad en metros y los 4 primeros al valor entero, rellenando con ceros a la izquierda hasta completar los 6 dígitos si corresponde y sin incluir el punto decimal. Así, por ejemplo, si se tiene una imagen de un sondaje denominado *EJEMPLO*, cuya caja comienza en la profundidad 49.55 metros y termina en los 52.70 metros, el nombre asignado a dicha imagen será *EJEMPLO_004955_005270.JPG*. Las correcciones de rotación, inclinación y perspectiva se llevarán a cabo en *Python*, con la herramienta de Transformación de perspectiva de la librería *OpenCV*. Esta herramienta permite seleccionar 4 puntos de una imagen que se transformarán en las 4 esquinas de la imagen corregida (Figura 1.4).

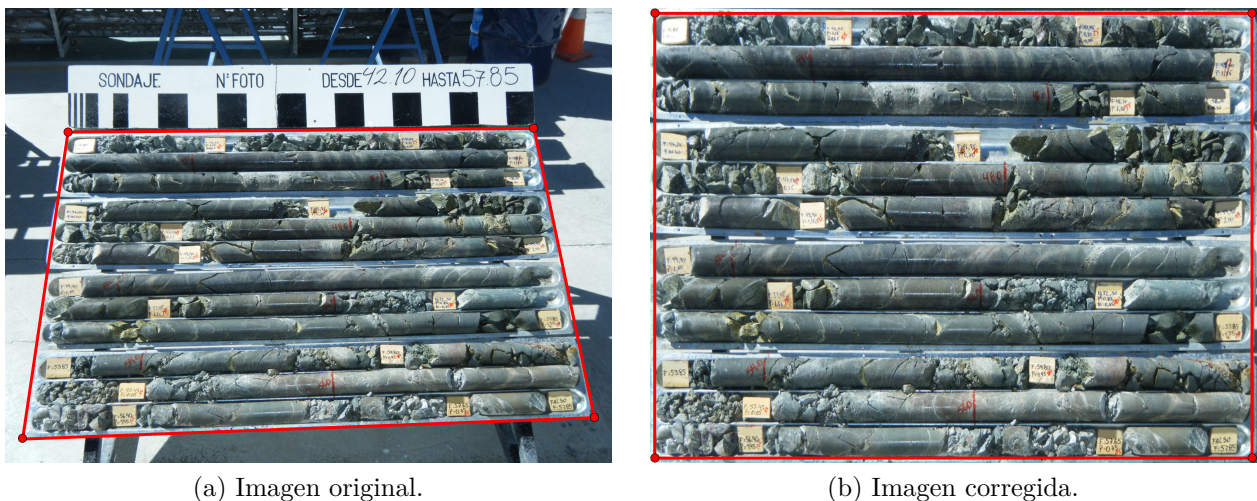


Figura 1.4: Ejemplo de Transformación de perspectiva de la librería *OpenCV*, en *Python*.

Desarrollo de métodos propuestos

En esta etapa se realizará el planteamiento e implementación de los métodos propuestos, tanto para el cálculo de ángulos alfa y beta, como para la estimación de frecuencia de fracturas, a partir de imágenes. Los algoritmos y cálculos matemáticos se implementarán mediante el uso de *Python*.

Preprocesamiento de datos

En esta etapa se desarrollará una aplicación simple para etiquetar imágenes, acorde al estudio de los métodos propuestos, la cual permitirá delimitar las estructuras en la imagen según etiquetas, incluir datos en ella y extraer información de su ubicación y disposición relativa según los píxeles de la imagen. Esta aplicación será desarrollada por medio del entorno de ejecución *NodeJS* y el framework *Angular*. Luego, se etiquetarán las estructuras correspondientes en las imágenes de los testigos de sondajes, utilizando esta aplicación.

Procesamiento de datos

Posterior al etiquetado de estructuras, se generarán paquetes de prueba y entrenamiento para evaluar el correcto funcionamiento de los métodos propuestos implementados en Python. Una vez se verifique que no hay errores, se procesará la información de las etiquetas, para cada uno de los métodos propuestos. De esta forma, para cada fractura y tramo de roca muy fracturada se obtendrán valores asociados al método propuesto y al método tradicional, para luego ser comparados.

Análisis y validación de resultados

En esta etapa se generarán gráficos y obtendrán parámetros estadísticos para comparar los resultados arrojados por los métodos propuestos, en relación con los entregados por los métodos tradicionales. Con esto se espera evaluar y validar los métodos propuestos con un cierto porcentaje de precisión y confiabilidad.

Capítulo 2

Cálculo de ángulos alfa y beta con el método de los 3 puntos

2.1. Preprocesamiento de datos

Los datos utilizados para este estudio corresponden a imágenes de testigos de 3 sondajes de 71.46, 108.52 y 135.01 metros, respectivamente, con un total de 314.99 metros.

Se desarrolló una aplicación simple para etiquetar las imágenes de los testigos de sondajes, por medio del entorno de ejecución NodeJS y el framework Angular. Esta permite cargar imágenes y navegar entre ellas, y etiquetar fracturas mediante una polilínea de 3 puntos pertenecientes a la traza de esta misma en el testigo. Los 3 puntos que se marcan se explican en la sección 2.2, así como el cálculo matemático y geométrico que se aplica sobre ellos. Además, la aplicación permite ingresar los valores alfa y beta medidos directamente en el testigo por el equipo de geología encargado, los cuales fueron registrados sobre el sondaje mismo (Figura 2.1). Las etiquetas se crean haciendo 3 clicks sobre la imagen, correspondientes a los 3 puntos de la traza de la fractura.



Figura 2.1: Aplicación desarrollada para etiquetar fracturas y tramos de roca muy fracturada a partir de imágenes de testigos de sondajes, con imagen del estudio de fracturas cargada.

2.2. Método de los 3 puntos

El método de los 3 puntos se basa en el supuesto de que las fracturas tienen geometría planar, y son intersecadas por un testigo de sondaje con geometría de cilindro. De esta manera, la intersección del plano de la fractura con el cilindro, proyectada en el plano de la imagen, contiene a todos los puntos observables que componen la traza de la fractura. Teniendo esto en cuenta y considerando que un plano se puede definir con 3 puntos distintos, es posible obtener la ecuación de este mismo ubicando 3 puntos sobre la traza de la fractura en el plano de la imagen y por lo tanto, los ángulos que lo definen y su orientación.

A partir de la traza frontal que se observa en la imagen de una fractura intersecada por un testigo de sondaje, se definen 3 puntos (Figura 2.2.a):

- *Punto superior* (P_s): Intersección entre la traza frontal de la fractura y la línea superior que delimita el testigo de sondaje.
- *Punto central* (P_c): Intersección entre la traza frontal de la fractura y la línea BOH.
- *Punto inferior* (P_i): Intersección entre la traza frontal de la fractura y la línea inferior que delimita el testigo de sondaje.

Posicionados estos 3 puntos en la fractura, se define el origen del sistema de referencia en la intersección entre el plano de fractura y el eje del sondaje. El eje x sigue la dirección del eje del sondaje, positivo hacia la derecha, el eje y es positivo hacia abajo de la imagen y el eje z es positivo entrando al plano de la imagen. El origen del sistema de referencia en x coincide con el punto medio entre la coordenada x de P_s y P_i , al igual que en y (Figura 2.2.a). El origen en z coincide con el eje del sondaje (Figura 2.2.b). El ángulo δ define la distancia angular entre la línea BOH y la proyección del eje x en la vista frontal del testigo (Figura 2.2.b).

Las coordenadas x e y de los puntos P_s , P_c y P_i se obtienen directamente de la imagen bidimensional. La coordenada z de los puntos P_s y P_i es 0, ya que están contenidos en el plano xy , es decir

$$P_s = (P_{s_x}, P_{s_y}, 0) \quad (2.1)$$

y

$$P_i = (P_{i_x}, P_{i_y}, 0). \quad (2.2)$$

Por otro lado, la coordenada z del punto P_c está dada por el radio del testigo (r) y el ángulo δ , donde

$$r = \frac{P_{i_y} - P_{s_y}}{2}, \quad (2.3)$$

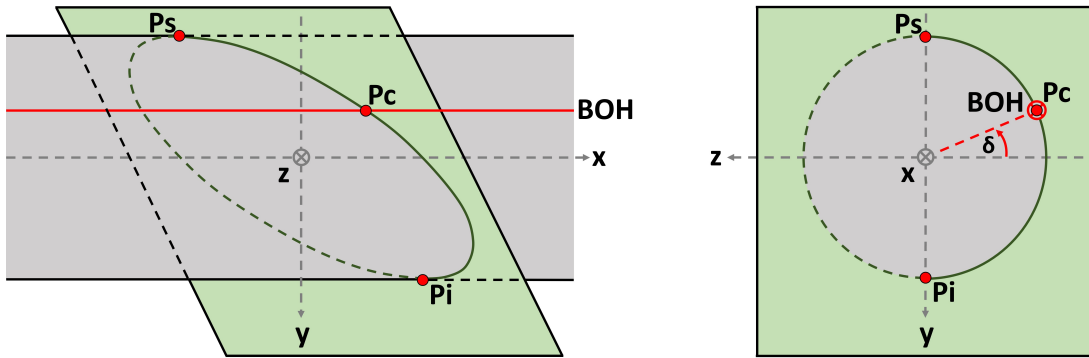
$$\delta = \sin^{-1} \left(\frac{P_{c_y}}{r} \right), \quad (2.4)$$

y

$$P_{c_z} = -r \cdot \cos(\delta), \quad (2.5)$$

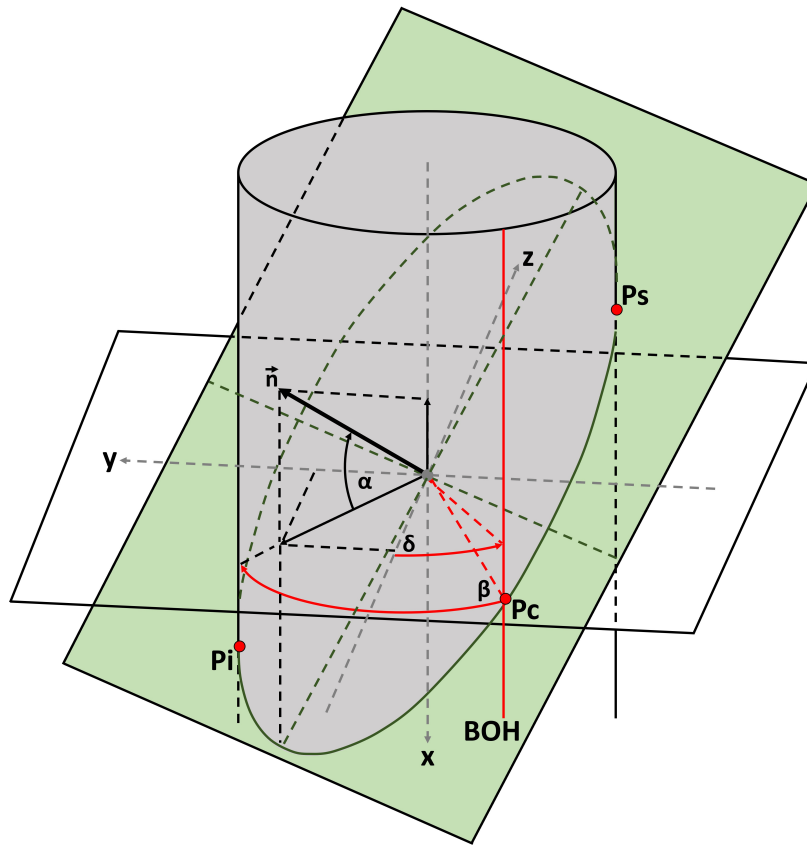
así

$$P_c = (P_{c_x}, P_{c_y}, -r \cdot \cos(\delta)). \quad (2.6)$$



(a) Vista plano xy.

(b) Vista plano yz.



(c) Vista 3D.

Figura 2.2: Sistema de referencia de fractura intersecada por un testigo de sondaje (Elaboración propia).

Por otro lado, un plano está definido por la ecuación

$$a \cdot x + b \cdot y + c \cdot z = d. \quad (2.7)$$

Los coeficientes a , b y c definen un vector \vec{n} normal al plano (Figura 2.2.c), y este a su vez se obtiene del producto cruz de dos vectores ($\vec{v1}$ y $\vec{v2}$) distintos que conectan los puntos Ps , Pc y Pi contenidos en el plano, por ejemplo

$$\vec{v1} = Ps - Pc \quad (2.8)$$

y

$$\vec{v2} = Pi - Pc, \quad (2.9)$$

luego

$$\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \vec{v1} \times \vec{v2}. \quad (2.10)$$

De esta forma, el ángulo alfa (α) corresponde al ángulo entre el vector \vec{n} y su proyección en el plano yz (Figura 2.2.c), es decir,

$$\alpha = \cos^{-1} \left(\frac{\sqrt{b^2 + c^2}}{\sqrt{a^2 + b^2 + c^2}} \right). \quad (2.11)$$

El ángulo beta (β), por otro lado, corresponde al ángulo entre la proyección de \vec{n} en el plano yz y el plano xz , menos el ángulo de desfase δ (Figura 2.2.c), que depende de en qué cuadrante se encuentra la proyección de \vec{n} en el plano yz , luego,

$$\beta = \begin{cases} \tan^{-1} \left(\frac{b}{c} \right) - \delta, & c < 0 \\ 270^\circ - \delta, & c = 0, b < 0 \\ \text{Indefinido}, & c = 0, b = 0 \\ 90^\circ - \delta, & c = 0, b > 0 \\ 180^\circ - \tan^{-1} \left(\frac{b}{c} \right) - \delta, & c > 0 \end{cases} \quad (2.12)$$

El caso $c = 0, b = 0$ corresponde a aquel en que $\alpha = 90^\circ$ y por lo tanto, no tiene un ángulo β definido.

2.2.1. Corrección de perspectiva

Dado que la luz viaja en línea recta, las imágenes tomadas por una cámara fotográfica sufren un efecto de distorsión, de acuerdo a la perspectiva desde donde se captura la imagen, considerando que las cámaras fotográficas generan una imagen virtual de 2 dimensiones, a partir de 3 dimensiones de la realidad.

Los puntos P_s , P_c y P_i definidos en la sección 2.2, así como todo el sistema de referencia esquematizado por la Figura 2.2, se consideran compatibles con la realidad de 3 dimensiones (imagen real) y corregidos de una imagen fotográfica con perspectiva de dos dimensiones (imagen virtual). Por ende, los puntos P_s , P_c y P_i se deben obtener de la corrección de los puntos de una imagen distorsionada; P_s' , P_c' y P_i' , respectivamente. La distorsión generada en una fotografía de testigos de sondajes depende de parámetros de la cámara o fotografía, y de la imagen real. Estos son los siguientes:

- *Distancia focal (f)*: Distancia entre la lente y el sensor de imagen cuando se enfoca al infinito, en una cámara fotográfica.
- *Tamaño del sensor (h)*: Ancho del aparato con forma de plano, encargado de recibir y captar la luz en una cámara fotográfica.
- *Distancia a caja (D)*: Distancia real entre la lente de la cámara fotográfica y el centro de la caja que contiene los testigos de sondajes, considerando que los bordes de la caja se encuadran perfectamente en una imagen capturada por la cámara.
- *Tamaño de caja (H)*: Ancho real de la caja que contiene los testigos de sondajes.

Los parámetros f y h se pueden obtener directamente de la información asociada a una fotografía o cámara fotográfica, mientras que H es conocido. Para obtener el parámetro D , en caso que no se cuente con esta información, se puede utilizar la siguiente relación de semejanza (Figura 2.3):

$$D = \frac{f}{h} \cdot H. \quad (2.13)$$

En la Figura 2.3 se observa un esquema de cómo una imagen real de 3 dimensiones es transformada en una imagen virtual de 2 dimensiones, por medio de una cámara fotográfica. En estas imágenes, se ubica el origen del sistema de referencia en su esquina superior izquierda, el cual coincide con la esquina superior izquierda de la caja, con el eje x positivo hacia la derecha y el eje y positivo hacia abajo. El eje z es positivo entrando a la imagen. Este sistema es con el que usualmente se trabaja en el procesamiento de imágenes, donde los puntos (x, y) definen cada pixel. Notar que este sistema de referencia no es el mismo presentado en el método de los 3 puntos. Para aplicar el cálculo del método de los 3 puntos se debe reubicar el origen en x e y .

Los rayos de luz para cada uno de los puntos reales capturados por una cámara fotográfica son colineales con el centro de la lente de la cámara y su proyección en la imagen bidimensional virtual emulada en el sensor (Figura 2.3). Esto se aplica, por tanto, a los 3 puntos reales: P_s , P_c y P_i , y sus puntos virtuales proyectados el sensor: P_s' , P_c' y P_i' , respectivamente. Los puntos P_s' y P_i' no sufren distorsión con respecto a P_s y P_i ya que se encuentran contenidos

en el plano de la caja de testigos de sondajes, sin componente tridimensional, lo cual no se aplica en los puntos Pc' y Pc .

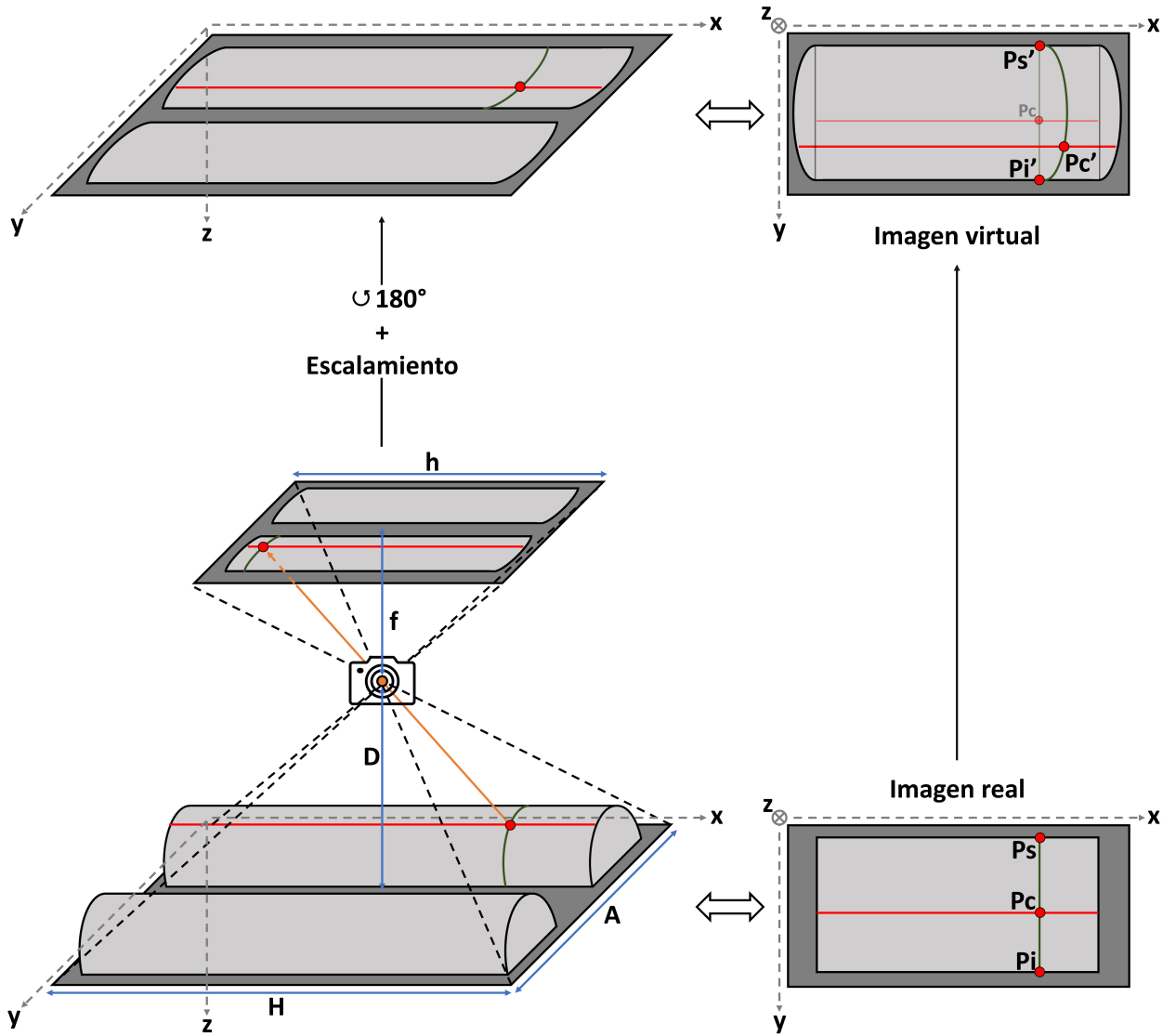


Figura 2.3: Esquema y sistema de referencia de caja de testigos de sondajes y su imagen virtual capturada por una cámara fotográfica (Elaboración propia).

Tomando en cuenta el sistema de referencia asociado a la imagen real y asumiendo que la cámara fotográfica se encuentra exactamente centrada en x e y con respecto a la caja de testigos real, entonces

$$P_{cam} = \left(\frac{H}{2}, \frac{A}{2}, -D \right), \quad (2.14)$$

donde P_{cam} es la posición de la cámara. Por otro lado, si la imagen real y virtual se encuentran igualmente escaladas, se define como \vec{d}_1 al vector director que une el punto P_{cam} y Pc , y al vector \vec{d}_2 como aquel que une el punto P_{cam} y Pc' , es decir,

$$\vec{d}_1 = Pcam - Pc \quad (2.15)$$

y

$$\vec{d}_2 = Pcam - Pc'. \quad (2.16)$$

Como los vectores \vec{d}_1 y \vec{d}_2 tienen la misma dirección, entonces

$$\vec{d}_1 = \lambda \cdot \vec{d}_2, \quad (2.17)$$

donde λ es un número real a determinar, luego

$$\begin{pmatrix} Pcam_x \\ Pcam_y \\ Pcam_z \end{pmatrix} - \begin{pmatrix} Pc_x \\ Pc_y \\ Pc_z \end{pmatrix} = \lambda \cdot \left(\begin{pmatrix} Pcam_x \\ Pcam_y \\ Pcam_z \end{pmatrix} - \begin{pmatrix} Pc'_x \\ Pc'_y \\ 0 \end{pmatrix} \right). \quad (2.18)$$

Por otra parte, el radio del testigo puede ser calculado como

$$r = \frac{|Pi'_y - Ps'_y|}{2}, \quad (2.19)$$

aproximadamente equivalente para la imagen real y virtual, ya que Ps y Pi tienen la misma posición que Ps' y Pi' , respectivamente. El valor en y por cual pasa el eje del sondaje (y_{eje}) se encuentra en el centro entre Pi'_y y Ps'_y , es decir,

$$y_{eje} = \frac{Pi'_y + Ps'_y}{2}, \quad (2.20)$$

luego, dada la geometría cilíndrica del sondaje y por trigonometría

$$Pc_z = \sqrt{r^2 - (Pc_y - y_{eje})^2}. \quad (2.21)$$

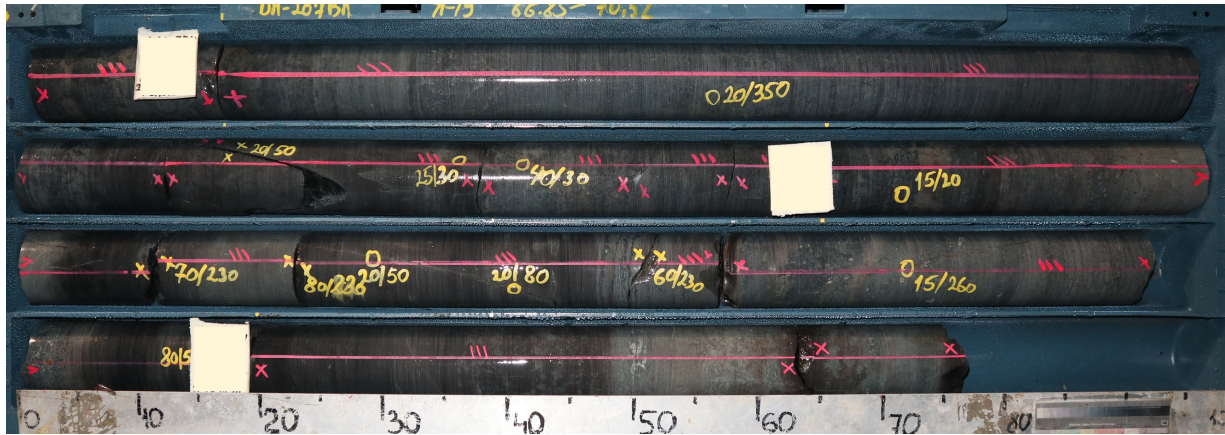
Con esto se tiene un sistema de ecuaciones de cuatro variables (Pc_x , Pc_y , Pc_z y λ) y cuatro ecuaciones, con el cual se puede obtener las coordenadas reales del punto Pc . El sistema está dado por

$$\begin{cases} Pcam_x - Pc_x = \lambda \cdot (Pcam_x - Pc'_x) \\ Pcam_y - Pc_y = \lambda \cdot (Pcam_y - Pc'_y) \\ Pcam_z - Pc_z = \lambda \cdot Pc_z \\ Pc_z = \sqrt{r^2 - (Pc_y - y_{eje})^2} \end{cases}. \quad (2.22)$$

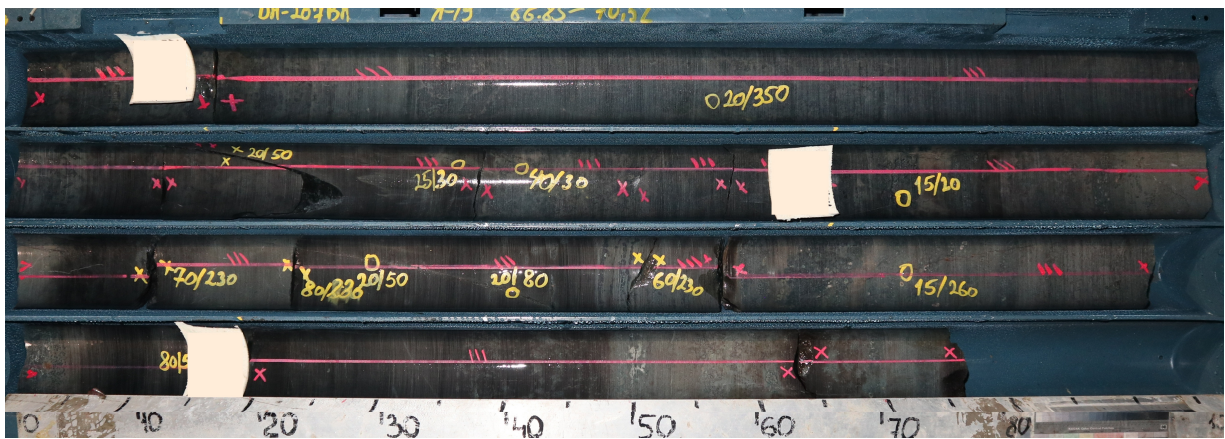
Este punto Pc luego es reubicado al origen del sistema de referencias del método de los 3 puntos.

En la Figura 2.4 se tiene un ejemplo de cómo se observaría una caja de testigos de sondajes con corrección de perspectiva para retornar de una imagen virtual a una real, transformación realizada en Python. El rasgo principal que permite evidenciar el efecto de la perspectiva,

es el borde de los testigos al principio y final de cada canaleta, los cuales no encajan en la imagen virtual, a diferencia de en la imagen real.



(a) Imagen virtual.



(b) Imagen real.

Figura 2.4: Ejemplo de cómo se observaría una caja de testigos de sondajes, con corrección de perspectiva para retornar de una imagen virtual a una real.

2.3. Procesamiento de datos

El total de fracturas etiquetadas en los 3 sondeos fue de 439. La corrección de perspectiva y el método de los 3 puntos fue implementado en Python y la información de los puntos de cada fractura fue procesada por este algoritmo, obteniendo los ángulos alfa y beta asociados al método. Por ende, cada fractura tiene 4 variables asociadas que se definieron, por efectos prácticos, de la siguiente forma:

- *Alfa Método directo* (α_{MD}): Ángulo alfa medido directamente en el testigo de sondeo.
- *Alfa Método de los 3 puntos* (α_{M3P}): Ángulo alfa calculado según los algoritmos del método de los 3 puntos.
- *Beta Método directo* (β_{MD}): Ángulo beta medido directamente en el testigo de sondeo.
- *Beta Método de los 3 puntos* (β_{M3P}): Ángulo beta calculado según los algoritmos del método de los 3 puntos.

Aquellas fracturas cuyo ángulo *Alfa Método directo* fue medido con un valor de 90° y no tienen un ángulo beta asociado, se descartaron del estudio de ángulos beta, ya que no hay forma de comparar los resultados del método propuesto con la medición original. Por ende, la cantidad de datos de ángulos beta: 414 considerados en el estudio, es menor a la cantidad total de fracturas y ángulos alfa: 439.

Por otro lado, para la evaluación de resultados de los ángulos beta, se debe considerar que su rango recorre un círculo (Figura 1.1), es decir, valores cercanos a 0° en realidad también están cercanos a 360° , y la distancia angular entre dos ángulos beta cualquiera no debería superar los 180° . Para obtener diferencias angulares acordes a lo anterior y suplir el efecto de grandes diferencias aparentes entre un valor muy cercano a 0° y otro muy cercano a 360° , se realizó la siguiente corrección, considerando equivalentes negativos:

1. Calcular el valor absoluto de la diferencia entre β_{MD} y β_{M3P} .
2. Si el valor calculado en el paso 1 es superior a 180° , continuar con los siguientes pasos, sino no considerarlos.
3. Seleccionar el mayor valor entre β_{MD} y β_{M3P} .
4. Restarle 360° al valor seleccionado en el paso 3, obteniendo su equivalente negativo.
5. Reemplazar el valor mayor seleccionado en el paso 3 por su equivalente negativo obtenido el paso 4.

Así, por ejemplo, si se tiene un par de β_{MD} y β_{M3P} de 1 y 359° , independiente de cual sea cada uno, el ángulo de 359° es transformado a su equivalente negativo de -1° , y la distancia angular entre ambos valores es luego de 2° . De esta forma, la diferencia o error entre *Beta Método directo* y *Beta Método de los 3 puntos* es representativa de su significado geométrico. El ángulo alfa, por otra parte, no sufre de este efecto, ya que corresponde a un ángulo agudo, donde mayores diferencias angulares entre los extremos 0 y 90° sí representan necesariamente mayores distancias geométricas en todo su rango.

Obtenidos los valores de alfa y beta, de los métodos directo y de los 3 puntos, se generaron gráficos y obtuvieron algunos parámetros estadísticos. El parámetro principal utilizado para evaluar la calidad de los resultados fue el coeficiente de determinación ajustado (R_a^2), el cual es un derivado del coeficiente de determinación (R^2).

El coeficiente de determinación (R^2) es un estadístico que mide qué tan cerca están los datos de un modelo o curva de regresión ajustada. Su propósito es determinar la calidad de un modelo para replicar resultados, según la proporción de variación de los resultados que puede ser explicada por el modelo (Steel, 1960). De esta forma, se pueden predecir futuros resultados o probar una hipótesis.

En el caso general de un modelo construido para explicar una variable aleatoria dependiente Y , a través de una o más variables aleatorias independientes, la variable aleatoria dependiente Y puede ser predicha mediante su media, y el máximo error cuadrático medio que puede ser aceptado por el modelo, corresponde a su varianza. Por otro lado, la diferencia entre los valores observados de la variable dependiente Y y los valores predichos por el modelo \hat{Y} corresponden a los residuos, y la varianza residual es la media cuadrática de los residuos. El coeficiente de determinación depende de la razón entre la varianza residual y la varianza de la variable aleatoria Y , y está dado por la ecuación

$$R^2 = 1 - \frac{\sigma_R^2}{\sigma^2}, \quad (2.23)$$

donde σ^2 es la varianza y σ_R^2 la varianza residual de la variable dependiente Y , dadas por

$$\sigma^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n} \quad (2.24)$$

y

$$\sigma_R^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}, \quad (2.25)$$

donde Y_i es la observación número i de la variable Y , \hat{Y}_i es la observación número i predicha por el modelo, \bar{Y} es el promedio de la variable Y , y n el número de observaciones.

El rango del coeficiente de determinación R^2 siempre se encuentra entre 0 y 1, donde 0 indica que el modelo no explica ninguna porción de la variabilidad de los datos de respuesta en torno a su media y 1 que el modelo la explica toda. En general, mientras mayor sea R^2 , mejor se ajusta el modelo a los datos. Sin embargo, hay condiciones importantes a tomar en cuenta con respecto a esta pauta y muchas veces se puede caer en errores en la interpretación del R^2 (Rodríguez, 2005), además, presenta una apreciación con respecto al número de observaciones. Cada vez que agrega un predictor a un modelo, el R^2 aumenta, incluso si se debe únicamente al azar. Nunca disminuye. En consecuencia, un modelo con más términos puede parecer que se ajusta mejor simplemente porque tiene más términos (Frost, 2013). El coeficiente de determinación ajustado (R_a^2) está diseñado para abordar este

problema.

El coeficiente de determinación ajustado (R_a^2) es una versión modificada del R^2 que se ha ajustado para el número de observaciones en el modelo y compara el poder explicativo de los modelos de regresión que contienen diferentes números de observaciones. El R_a^2 aumenta solo si el nuevo término mejora el modelo más de lo que se esperaría por azar y disminuye cuando una observación mejora el modelo menos de lo esperado por azar. El R_a^2 siempre es más bajo que el R^2 y puede ser negativo, aunque usualmente no lo es (Frost, 2013). La ecuación para obtener el R_a^2 está dada por

$$R_a^2 = 1 - \frac{(n-1)}{(n-k-1)} \cdot (1 - R^2), \quad (2.26)$$

donde k es el número de variables independientes, en este caso, una sola.

Por otro lado, el error asociado a los ángulos para cada fractura, una vez realizada la corrección para ángulos beta, fue calculada como el valor del método directo menos el valor del método de los 3 puntos, dividido por el rango posible de los ángulos (para alfa el rango es 90° y para beta 360°), en porcentaje. Estos parámetros se definen como *Error Alfa* (α_{error}) y *Error beta* (β_{error}) y están dados por las ecuaciones

$$\alpha_{error} = \frac{(\alpha_{MD} - \alpha_{M3P})}{90^\circ} \cdot 100\% \quad (2.27)$$

y

$$\beta_{error} = \frac{(\beta_{MD} - \beta_{M3P})}{360^\circ} \cdot 100\%. \quad (2.28)$$

Así, los valores de *Error Alfa* y *Error Beta* se encontrarán entre -100 y 100 % y podrán ser contrastados en la misma dimensión. Los valores absolutos de estos errores se expresan como $|\alpha_{error}|$ y $|\beta_{error}|$, respectivamente.

Por otro lado, cuando los sondajes son perfectamente verticales y su línea de referencia BOH (Figura 1.1) apunta al norte geográfico, como ocurre para los 3 sondajes de estudio (no se aplica necesariamente a otros sondajes), entonces la inclinación (Dip) y la dirección de inclinación (Dip Direction) de las fracturas con respecto a los ejes geográficos y su relación con los ángulos alfa y beta, está dado por las ecuaciones (Stigsson & Munier, 2013)

$$Dip = 90^\circ - \alpha \quad (2.29)$$

y

$$Dip\ Direction = \beta. \quad (2.30)$$

De esta forma, las variables *Dip Método directo* y *Dip Direction Método directo* (Dip_{MD} y $Dip\ Direction_{MD}$), y *Dip Método de los 3 puntos* y *Dip Direction Método de los 3 puntos* (Dip_{M3P} y $Dip\ Direction_{M3P}$) definen la orientación de las fracturas con respecto a los ejes

geográficos, para el método directo y de los 3 puntos, respectivamente, según los ángulos alfa y beta. Los errores asociados a los valores Dip (*Error Dip* o *Dip_{error}*) y Dip Direction (*Error Dip Direction* o *Dip Direction_{error}*) de las fracturas con respecto a la medición directa son, respectivamente, según el rango en que se encuentran

$$Dip_{error} = \frac{(Dip_{MD} - Dip_{M3P})}{90^\circ} \cdot 100\% \quad (2.31)$$

y

$$Dip\ Direction_{error} = \frac{(Dip\ Direction_{MD} - Dip\ Direction_{M3P})}{360^\circ} \cdot 100\%. \quad (2.32)$$

Se realizó una representación gráfica en el programa *Dips*, de los ángulos Dip y Dip Direction de las fracturas en stereonet, para cada uno de los 3 sondajes. Esto, tanto para el método directo como para el de los 3 puntos, reconociendo los sets de fracturas en cada uno y evaluando la diferencia entre ellos según las variables *Error Dip* (*Dip_{error}*) y *Error Dip Direction* (*Dip Direction_{error}*). La misma representación se realizó aplicando la ponderación de Terzaghi, utilizando el ángulo de polarización mínimo que trae por defecto el programa *Dips* (15°), para eliminar el sesgo a favor de aquellas fracturas que son perpendiculares a la dirección del sondaje.

Luego, los gráficos generados para analizar los resultados fueron los siguientes:

- Histograma de ángulos *Alfa Método directo*.
- *Alfa Método de los 3 puntos* versus *Alfa Método directo*.
- *Error Alfa* versus *Alfa Método directo*.
- Histograma de ángulos *Beta Método directo*.
- *Beta Método de los 3 puntos* versus *Beta Método directo*.
- *Error Beta* versus *Beta Método directo*.
- *Error Beta* versus *Alfa Método directo*.
- *Error Alfa* versus *Beta Método directo*.
- *Error Beta* versus *Error Alfa*.
- Proyección estereográfica (stereonet) de las fracturas de cada uno de los 3 sondajes, reconociendo los distintos sets, en el programa *Dips*.
- Proyección estereográfica (stereonet) de las fracturas de cada uno de los 3 sondajes, reconociendo los distintos sets, aplicando la ponderación de Terzaghi, en el programa *Dips*.

Los histogramas permiten conocer la distribución de datos considerados para el estudio. Los gráficos de valores método de los 3 puntos versus método directo permiten evidenciar si son similares entre sí y se acercan a una recta de pendiente 1 que pasa por el origen, y

los gráficos de error permiten reconocer si existe una correlación de aumento o disminución de este mismo con respecto otra variable, así como la posible existencia de valores atípicos. Las proyecciones estereográficas permiten identificar los sets de fracturas y sus respectivas orientaciones, comparando los resultados del método de los 3 puntos y del método directo en un significado estructural.

Por otro lado, los parámetros estadísticos más importantes a considerar fueron los siguientes:

- Promedio de *Error Alfa* (α_{error}^{prom}).
- Promedio de valor absoluto de *Error Alfa* ($|\alpha_{error}|^{prom}$).
- Máximo de valor absoluto de *Error Alfa* ($|\alpha_{error}|^{max}$).
- Coeficiente de determinación ajustado (R_a^2) entre *Alfa Método de los 3 puntos* y *Alfa Método directo*.
- Promedio de *Error Beta* (β_{error}^{prom}).
- Promedio de valor absoluto de *Error Beta* ($|\beta_{error}|^{prom}$).
- Máximo de valor absoluto de *Error Beta* ($|\beta_{error}|^{max}$).
- Coeficiente de determinación ajustado (R_a^2) entre *Beta Método de los 3 puntos* y *Beta Método directo*.
- Promedios de *Error Dip* y *Error Dip Direction* (Dip_{error}^{prom} y $Dip\ Direction_{error}^{prom}$) de los sets de fracturas entre método de los 3 puntos y método directo, previo y posterior a la aplicación de la ponderación de Terzaghi.
- Promedios de valores absolutos de *Error Dip* y *Error Dip Direction* ($|Dip_{error}|^{prom}$ y $|Dip\ Direction_{error}|^{prom}$) de los sets de fracturas entre método de los 3 puntos y método directo, previo y posterior a la aplicación de la ponderación de Terzaghi.
- Máximos de valores absolutos de *Error Dip* y *Error Dip Direction* ($|Dip_{error}|^{max}$ y $|Dip\ Direction_{error}|^{max}$) de los sets de fracturas entre método de los 3 puntos y método directo, previo y posterior a la aplicación de la ponderación de Terzaghi.

Los promedios de errores permiten saber si estos se concentran positiva o negativamente, es decir si existe en general una sobrestimación o subestimación de los resultados por parte del método de los 3 puntos. Los promedios de valores absolutos de errores permiten tener una idea de qué tanto difieren los resultados entregados por el método de los 3 puntos, en relación con los datos del método directo, mientras el máximo indica la cota de este mismo. El coeficiente de determinación ajustado es el parámetro principal para medir la calidad del modelo propuesto y probar y validar la hipótesis del estudio. Los promedios de los errores, y los promedios y máximos del valor absoluto de los errores de Dip y Dip Direction permiten verificar la precisión del método de los 3 puntos desde la perspectiva estructural.

2.4. Resultados

2.4.1. Ángulos alfa

Con un total de 439 ángulos alfa considerados en el estudio, se obtuvo el histograma de *Alfa Método directo* dado por la Figura 2.5.

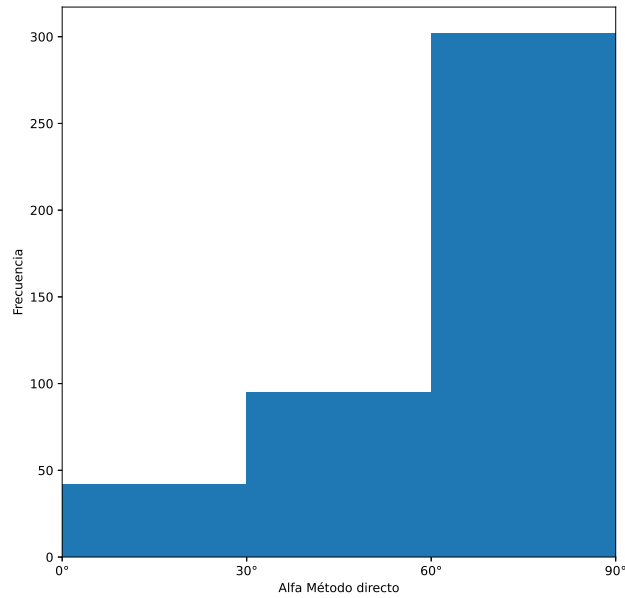


Figura 2.5: Histograma de ángulos Alfa Método directo.

La relación entre los resultados de *Alfa Método de los 3 puntos* y *Alfa Método directo* se observa en la Figura 2.6.

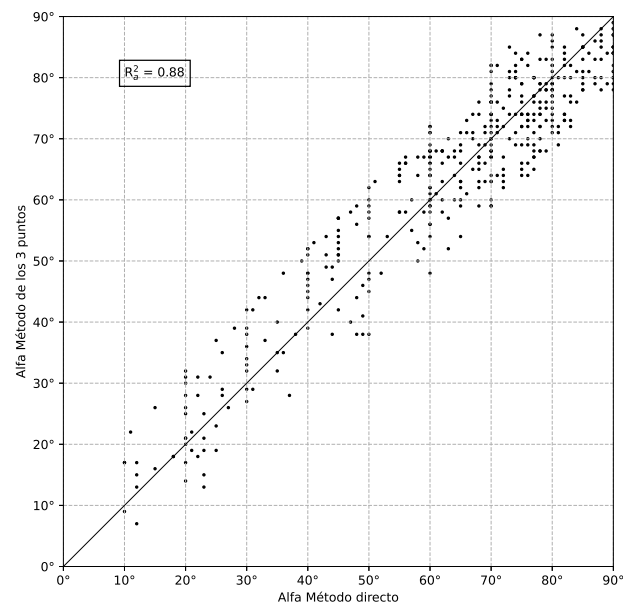


Figura 2.6: Comparación de ángulos Alfa: Método de los 3 puntos versus Método directo.

El coeficiente de determinación ajustado (R_a^2) obtenido según la Figura 2.6 fue de 0.88. Por otro lado, la relación entre *Error Alfa* y *Alfa Método directo* se observa en la Figura 2.7.

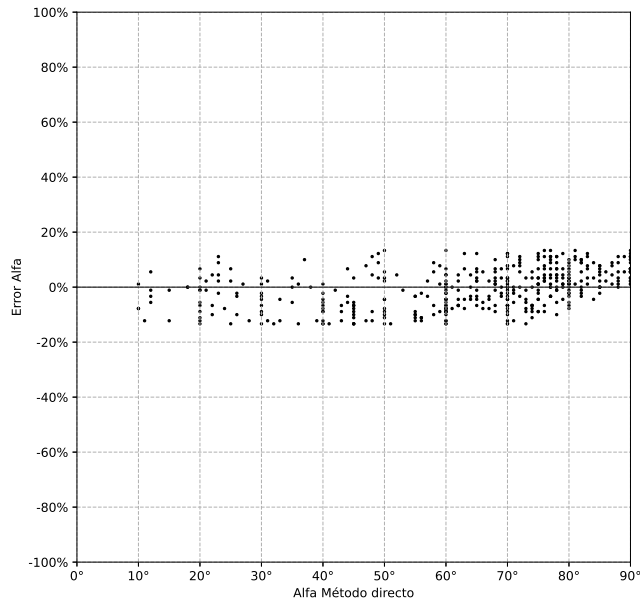


Figura 2.7: Análisis de errores: Error Alfa versus Alfa Método directo.

De los valores de *Error Alfa* (α_{error}) se obtuvo un promedio α_{error}^{prom} de 0 % (equivalente a 0°). De sus valores absolutos $|\alpha_{error}|$ se obtuvo un promedio $|\alpha_{error}|^{prom}$ de 5.56 % (equivalente a 5°) y un máximo $|\alpha_{error}|^{max}$ de 13.33 % (equivalente a 12°).

2.4.2. Ángulos beta

Con un total de 414 ángulos beta considerados en el estudio, se obtuvo el histograma de *Beta Método directo* dado por la Figura 2.8.

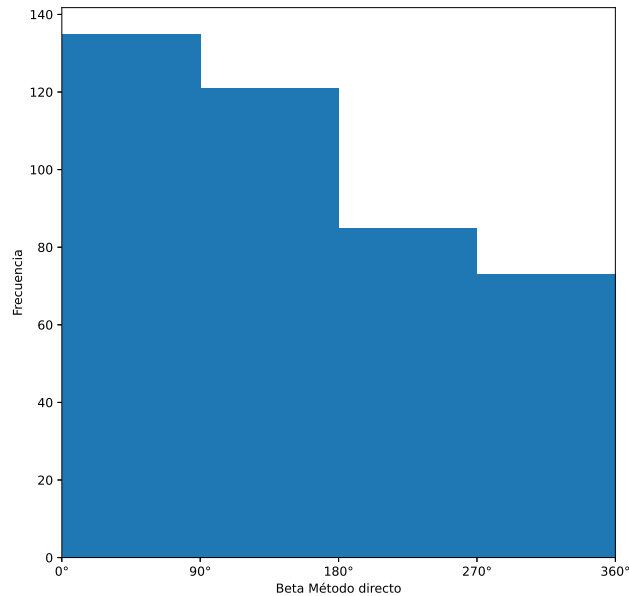


Figura 2.8: Histograma de ángulos Beta Método directo.

La relación entre los resultados de *Beta Método de los 3 puntos* y *Beta Método directo* se observa en la Figura 2.9.

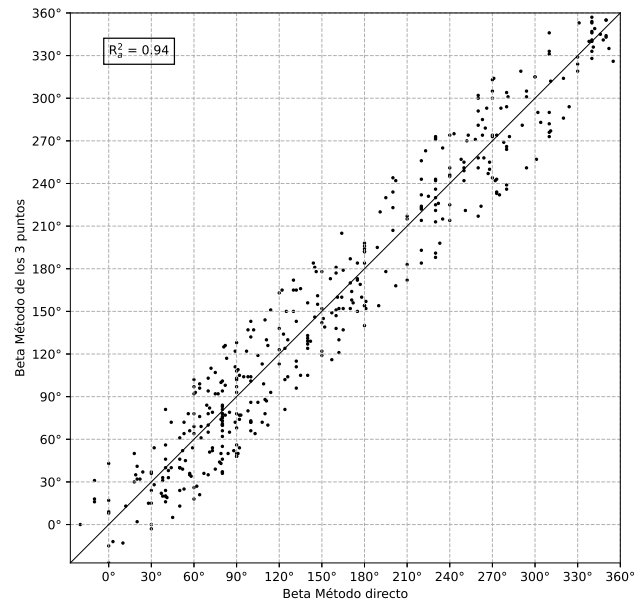


Figura 2.9: Comparación de ángulos Beta: Método de los 3 puntos versus Método directo.

El coeficiente de determinación ajustado (R_a^2) obtenido según la Figura 2.9 fue de 0.94. Por otro lado, la relación entre *Error Beta* y *Beta Método directo* se observa en la Figura 2.10.

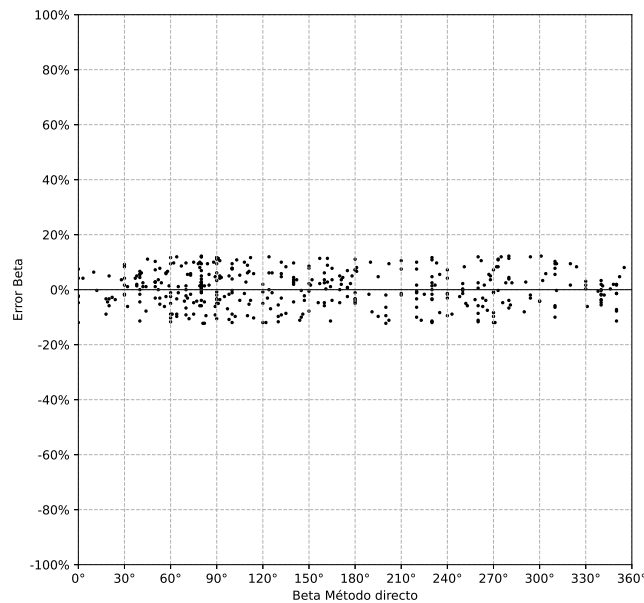


Figura 2.10: Análisis de errores: Error Beta versus Beta Método directo.

De los valores de *Error Beta* (β_{error}) se obtuvo un promedio β_{error}^{prom} de 0.56 % (equivalente a 2°). De sus valores absolutos $|\beta_{error}|$ se obtuvo un promedio $|\beta_{error}|^{prom}$ de 5.56 % (equivalente a 20°) y un máximo $|\beta_{error}|^{max}$ de 12.22 % (equivalente a 44°).

2.4.3. Errores cruzados

La relación entre *Error Beta* y *Alfa Método directo* se observa en la Figura 2.11.

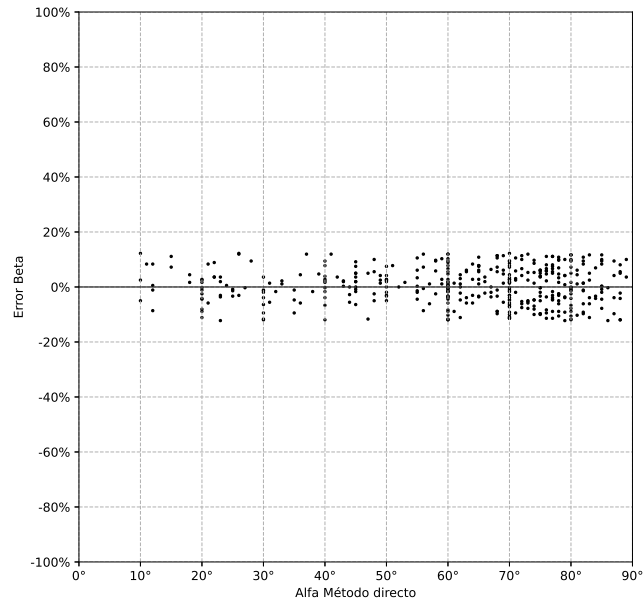


Figura 2.11: Análisis de errores: Error Beta versus Alfa Método directo.

La relación entre *Error Alfa* y *Beta Método directo* se observa en la Figura 2.12.

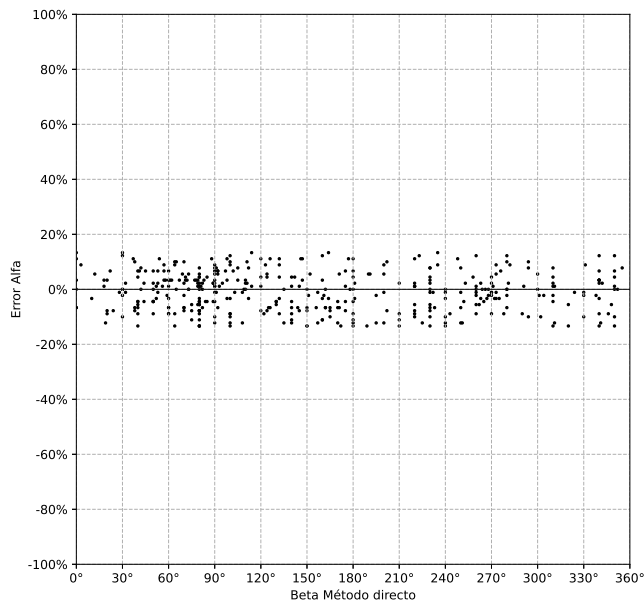


Figura 2.12: Análisis de errores: Error Alfa versus Beta Método directo.

La comparación entre ambos errores: *Error Alfa* y *Error Beta* se observa en la Figura 2.13.

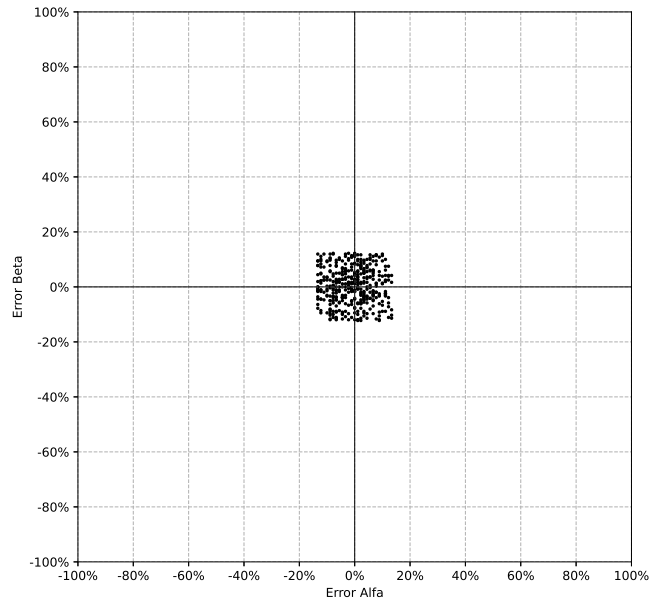
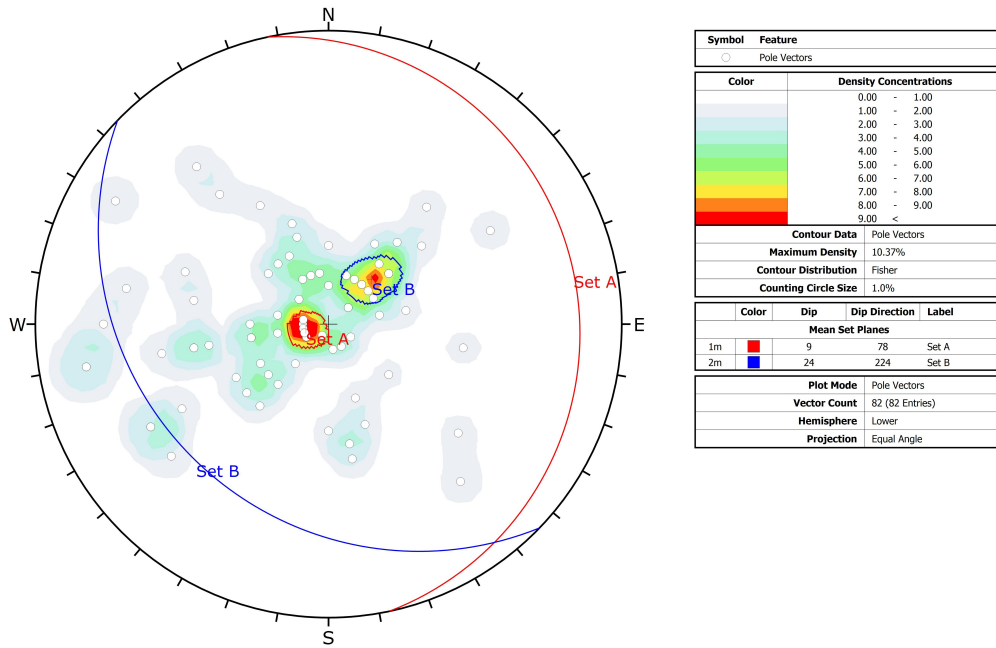


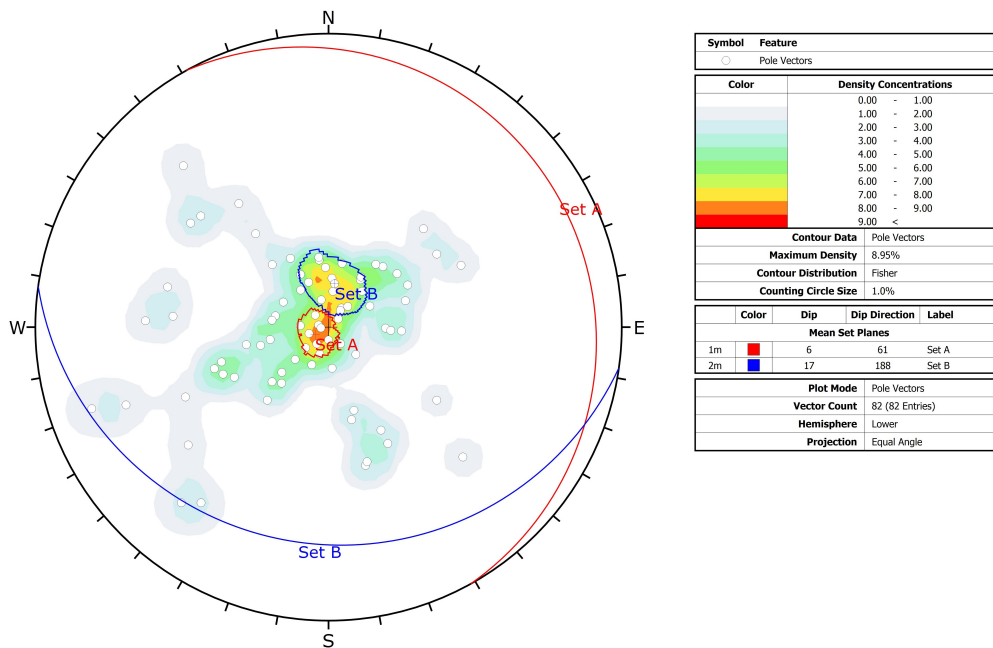
Figura 2.13: Análisis de errores: Error Beta versus Error Alfa.

2.4.4. Proyecciones estereográficas (stereonet)

Previo a la aplicación de ponderación de Terzaghi, las proyecciones estereográficas (stereonet) de los métodos directo y de los 3 puntos, de los sondajes 1, 2 y 3 se observan en las Figuras 2.14, 2.15 y 2.16, respectivamente.

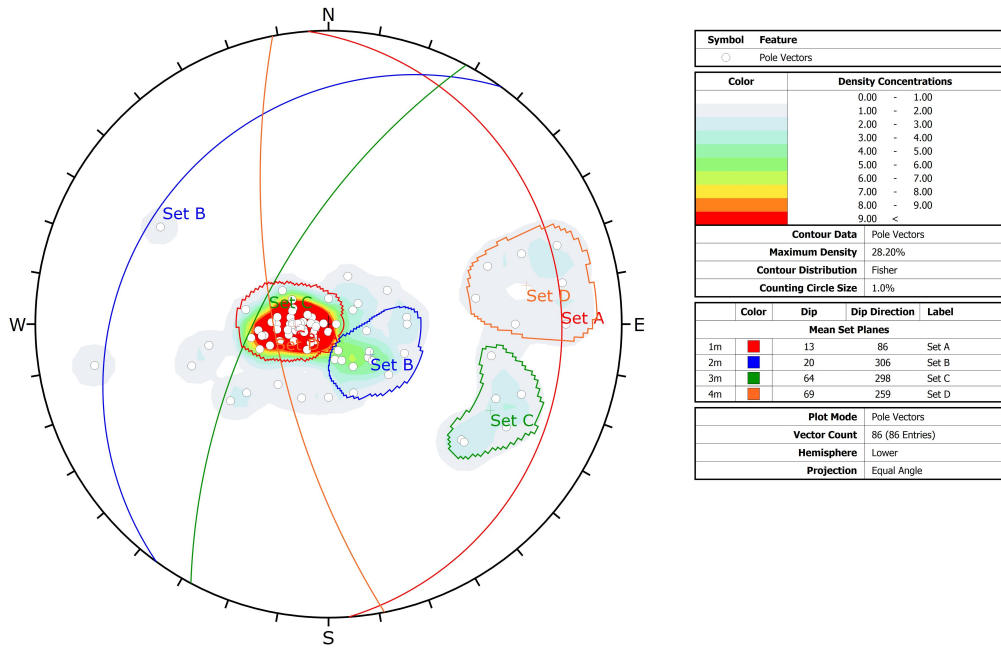


(a) Método directo.

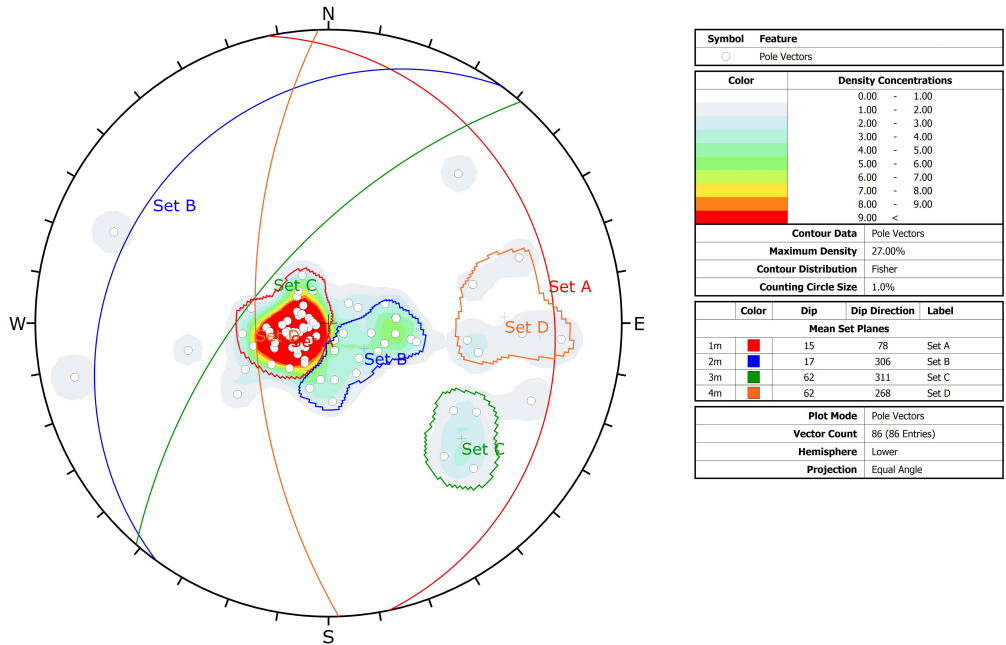


(b) Método de los 3 puntos.

Figura 2.14: Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaje 1, con sus respectivos sets de fracturas, previo a la aplicación de la ponderación de Terzaghi.

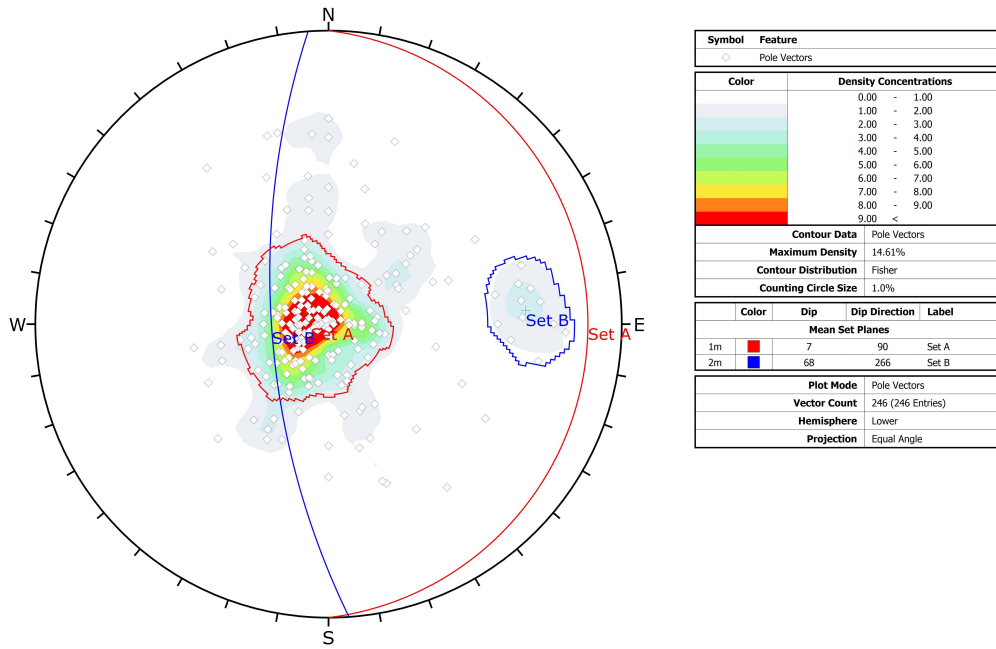


(a) Método directo.

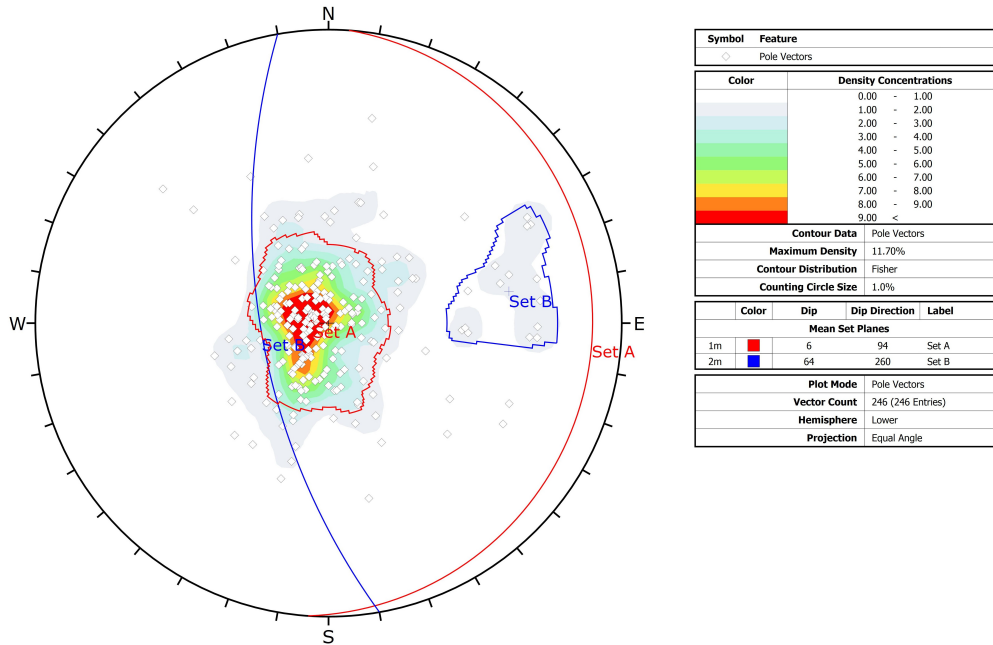


(b) Método de los 3 puntos.

Figura 2.15: Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sonda 2, con sus respectivos sets de fracturas, previo a la aplicación de la ponderación de Terzaghi.



(a) Método directo.



(b) Método de los 3 puntos.

Figura 2.16: Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaaje 3, con sus respectivos sets de fracturas, previo a la aplicación de la ponderación de Terzaghi.

La Tabla 2.1 resume los valores *Dip Método directo/Dip Direction Método directo* y *Dip Método de los 3 puntos/Dip Direction Método de los 3 puntos* obtenidos para cada set de fracturas de los 3 sondajes, previo a la aplicación de la ponderación de Terzaghi.

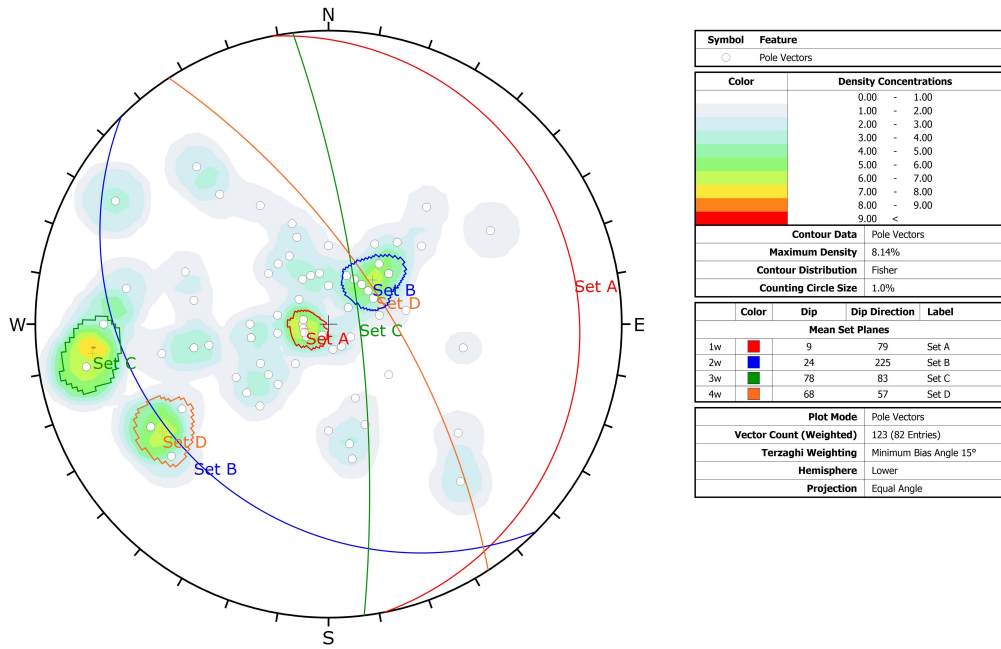
Tabla 2.1: Resumen de ángulos Dip y Dip Direction de los métodos directo y de los 3 puntos, de los sets de fracturas reconocidos en los Sondajes 1, 2 y 3, previo a la aplicación de la ponderación de Terzaghi.

Sondaje	Set	Dip/Dip Direction Método directo	Dip/Dip Direction Método de los 3 puntos
1	A	9/78	6/61
1	B	24/224	17/188
2	A	13/86	15/78
2	B	20/306	17/306
2	C	64/298	62/311
2	D	69/259	62/268
3	A	7/90	6/94
3	B	68/266	64/260

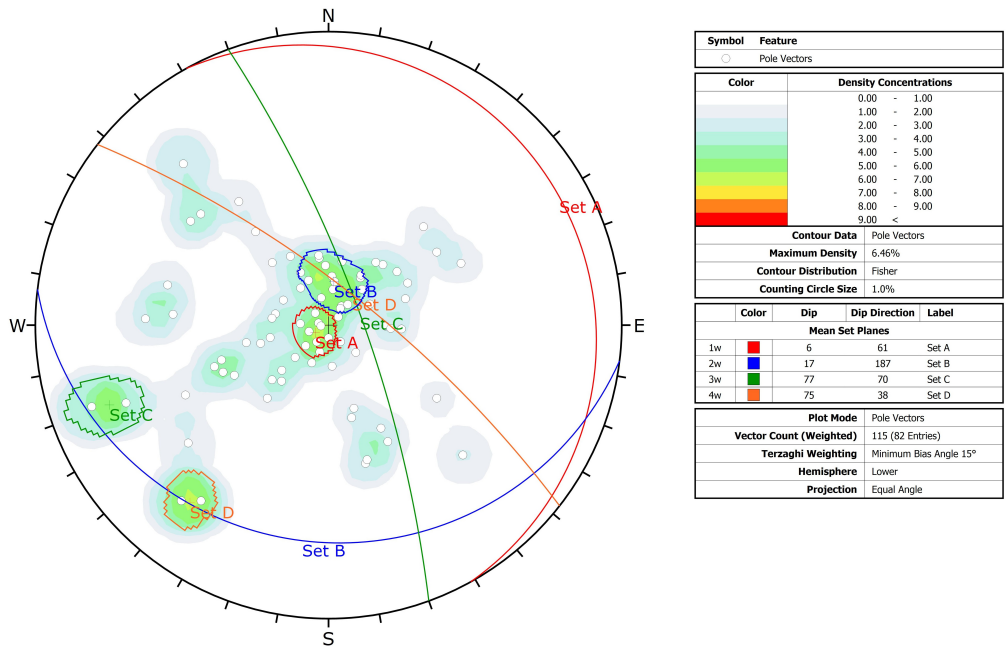
De los valores de *Error Dip* (Dip_{error}) se obtuvo un total promedio Dip_{error}^{prom} de 3.33% (equivalente a 3°). De sus valores absolutos $|Dip_{error}|$ se obtuvo un promedio $|Dip_{error}|^{prom}$ de 4.44% (equivalente a 4°) y un máximo $|Dip_{error}|^{max}$ de 7.78% (equivalente a 7°).

De los valores de *Error Dip Direction* ($Dip\ Direction_{error}$) se obtuvo un total promedio $Dip\ Direction_{error}^{prom}$ de 1.39% (equivalente a 5°). De sus valores absolutos $|Dip\ Direction_{error}|$ se obtuvo un promedio $|Dip\ Direction_{error}|^{prom}$ de 3.33% (equivalente a 12°) y un máximo $|Dip\ Direction_{error}|^{max}$ de 10.00% (equivalente a 36°).

Una vez aplicada la ponderación de Terzaghi, las proyecciones estereográficas (stereonet) de los métodos directo y de los 3 puntos, de los sondajes 1, 2 y 3 se observan en las Figuras 2.17, 2.18 y 2.19, respectivamente.

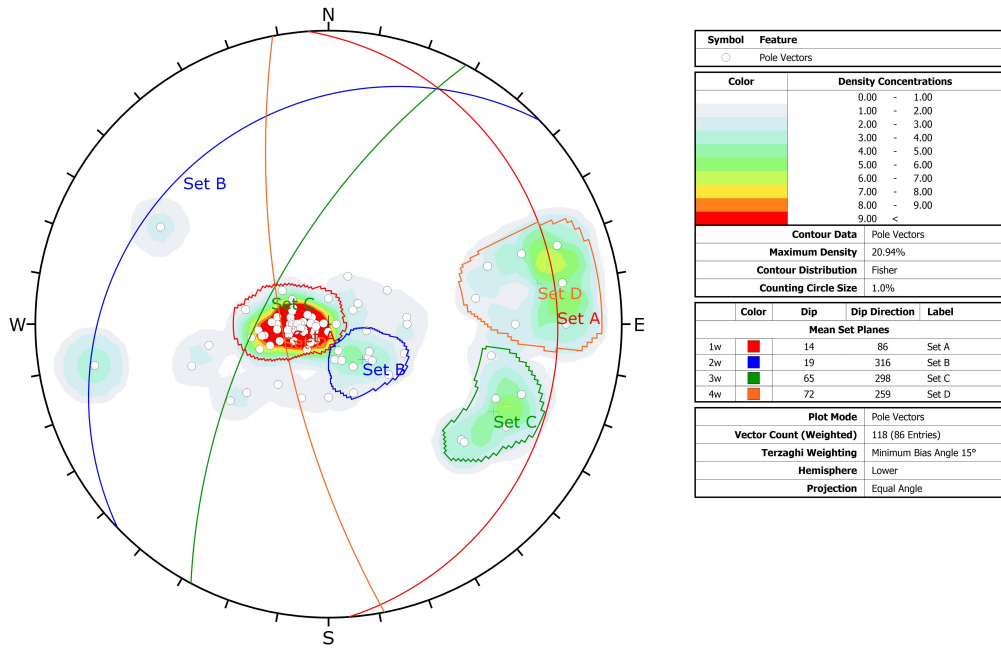


(a) Método directo.

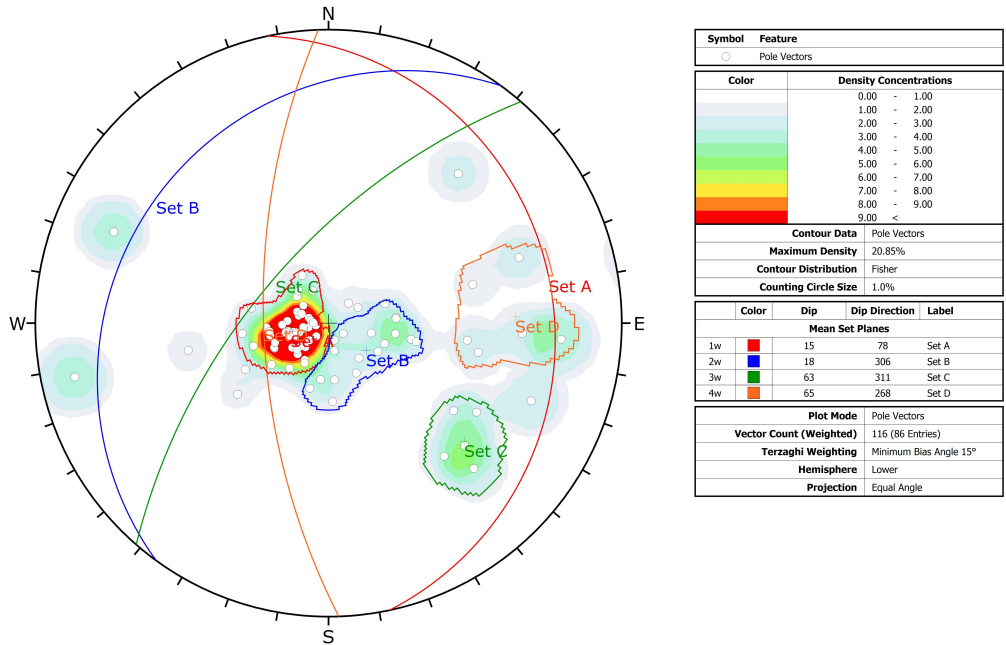


(b) Método de los 3 puntos.

Figura 2.17: Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sonda 1, con sus respectivos sets de fracturas, una vez aplicada la ponderación de Terzaghi.

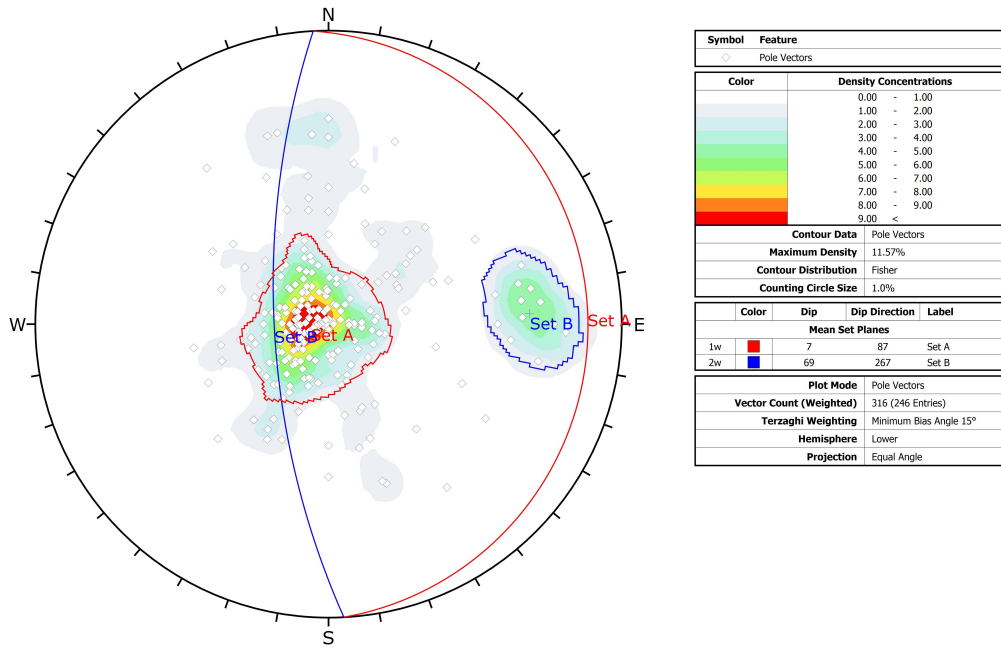


(a) Método directo.

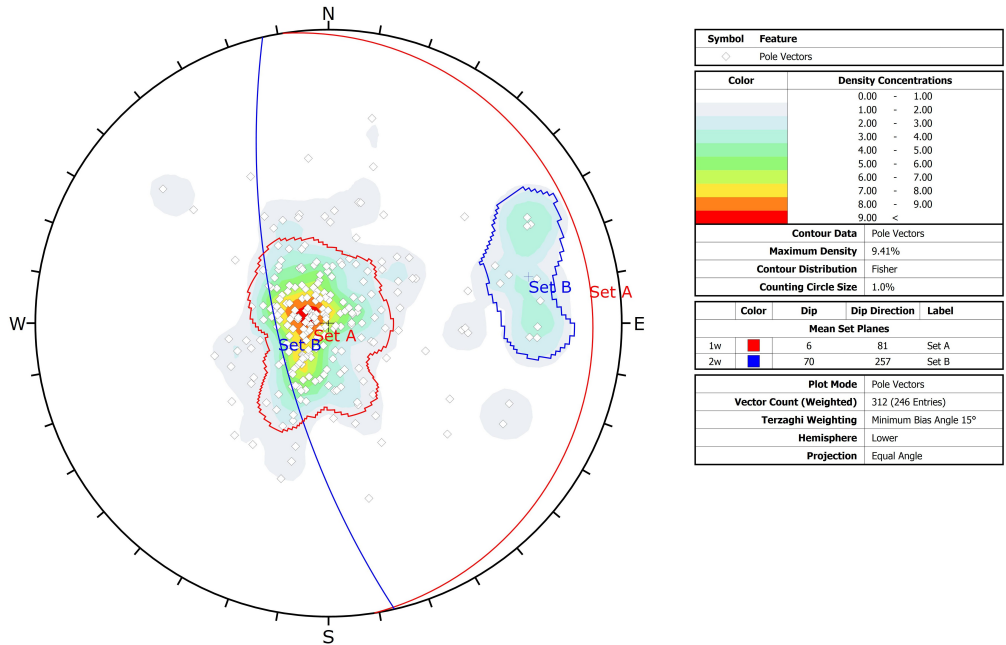


(b) Método de los 3 puntos.

Figura 2.18: Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sondaaje 2, con sus respectivos sets de fracturas, una vez aplicada la ponderación de Terzaghi.



(a) Método directo.



(b) Método de los 3 puntos.

Figura 2.19: Proyecciones estereográficas (stereonet) de orientación de fracturas del método directo y de los 3 puntos del Sonda je 3, con sus respectivos sets de fracturas, una vez aplicada la ponderación de Terzaghi.

Tabla 2.2: Resumen de ángulos Dip y Dip Direction de los métodos directo y de los 3 puntos, de los sets de fracturas reconocidos en los Sondajes 1, 2 y 3, una vez aplicada la ponderación de Terzaghi.

Sondaje	Set	Dip/Dip Direction Método directo	Dip/Dip Direction Método de los 3 puntos
1	A	9/79	6/61
1	B	24/225	17/187
1	C	78/83	77/70
1	D	68/57	75/38
2	A	14/86	15/78
2	B	19/316	18/306
2	C	65/298	63/311
2	D	72/259	65/268
3	A	7/87	6/81
3	B	69/267	70/257

De los valores de *Error Dip* (Dip_{error}) se obtuvo un total promedio Dip_{error}^{prom} de 1.11% (equivalente a 1°). De sus valores absolutos $|Dip_{error}|$ se obtuvo un promedio $|Dip_{error}|^{prom}$ de 3.33% (equivalente a 3°) y un máximo $|Dip_{error}|^{max}$ de 7.78% (equivalente a 7°).

De los valores de *Error Dip Direction* ($Dip\ Direction_{error}$) se obtuvo un total promedio $Dip\ Direction_{error}^{prom}$ de 2.78% (equivalente a 10°). De sus valores absolutos $|Dip\ Direction_{error}|$ se obtuvo un promedio $|Dip\ Direction_{error}|^{prom}$ de 3.89% (equivalente a 14°) y un máximo $|Dip\ Direction_{error}|^{max}$ de 10.56% (equivalente a 38°).

Capítulo 3

Estimación de frecuencia de fracturas con Machine Learning

3.1. Preprocesamiento de datos

Los datos utilizados para este estudio corresponden a imágenes de 2 sondajes de 100.10 y 770.05 metros, respectivamente, con un total de 870.15 metros. Además, se contaba con la información e imágenes extraídas por sondas Televiewer acústico y óptico de cada uno de los sondajes.

La misma aplicación desarrollada para etiquetar fracturas con 3 puntos fue utilizada para etiquetar tramos de roca fracturada. En el caso de estos tramos, se utilizaron etiquetas con forma rectangular, delimitando sus extremos inferior-superior e izquierdo-derecho (Figura 3.1).

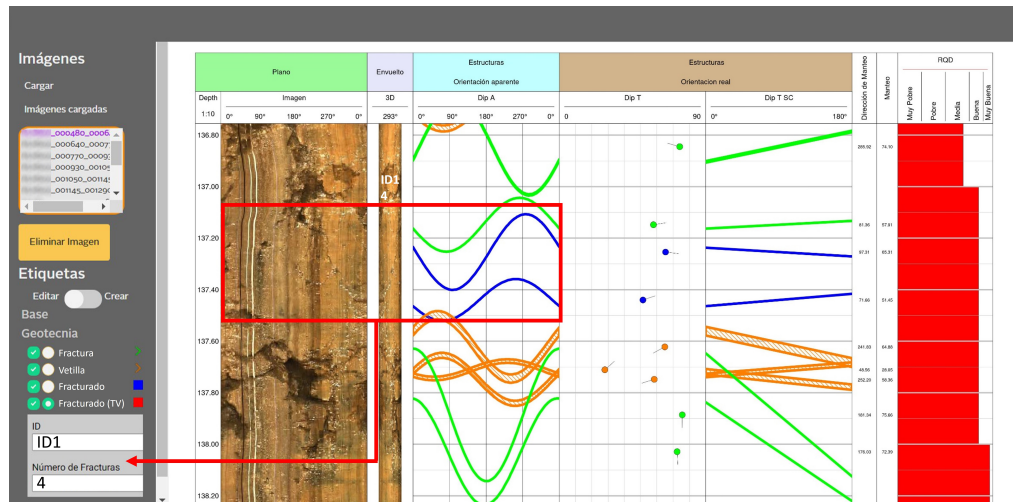
Se etiquetaron tanto imágenes de los testigos de sondajes, como las imágenes de sus respectivas secciones en Televiewer. La aplicación contaba con dos etiquetas distintas para tramos fracturados de los testigos y de los Televiewer, denominadas *Fracturado* y *Fracturado (TV)*, respectivamente (Figura 3.1).

Los pasos a seguir fueron, en primer lugar, seleccionar y etiquetar un tramo de roca fracturada en las imágenes de los testigos. Se debe ingresar un *ID* único para cada tramo en un recuadro de la barra lateral (Figura 3.1.a), que por simplicidad fueron números correlativos. El rectángulo que enmarca el tramo se crea directamente haciendo click sobre la imagen. A continuación se identificó este mismo tramo en las imágenes del Televiewer, considerando su profundidad y extensión. Al momento de crear esta etiqueta se ingresó el mismo *ID* registrado en la etiqueta del testigo de sondaje, junto a la cantidad de fracturas encontradas en ese tramo (Figura 3.1.b). La precisión en la ubicación de estas etiquetas en la imagen no fue relevante, ya que la imagen enmarcada no fue utilizada en el estudio posterior, sin embargo, sirvió de referencia para identificar su profundidad.

Los datos del Televiewer reportan si los tramos corresponden a fracturas o secciones donde la roca es muy blanda con respecto a la roca caja. En este estudio sólo se consideraron tramos asociados a fracturas y no a roca blanda. También se detectaron tramos fracturados inducidos, ya que no fueron reconocidos en el Televiewer, los cuales también fueron descartados.



(a) Imagen de testigos de sondaje cargada.



(b) Imagen de Televiewer cargada.

Figura 3.1: Aplicación desarrollada para etiquetar fracturas y tramos de roca muy fracturada a partir de imágenes de testigos de sondajes, con imágenes del estudio de tramos fracturados cargadas.

A partir de las etiquetas rectangulares creadas en las imágenes de los testigos de sondajes con la aplicación, se extrajo una imagen recortada delimitada por la misma etiqueta. Además, se calculó la frecuencia de fracturas del tramo, contabilizando las fracturas de las etiquetas de Televiewer y dividiéndolas por el largo del tramo según la etiqueta del testigo de sondaje (Ecuación 1.1), con motivo de expresar el valor de fracturas por unidad de distancia (fracturas/m). Para determinar el largo real en metros del tramo bastó con hacer una regla de tres, conociendo el ancho de las cajas. Así se obtuvieron pares de datos que incluían una imagen del tramo de roca fracturada y su frecuencia de fracturas. Esta frecuencia de fracturas fue denominada como *Frecuencia de Fracturas Observada* (FF_{Obs}). El esquema que resume la obtención de este par de datos se observa en la Figura 3.2.

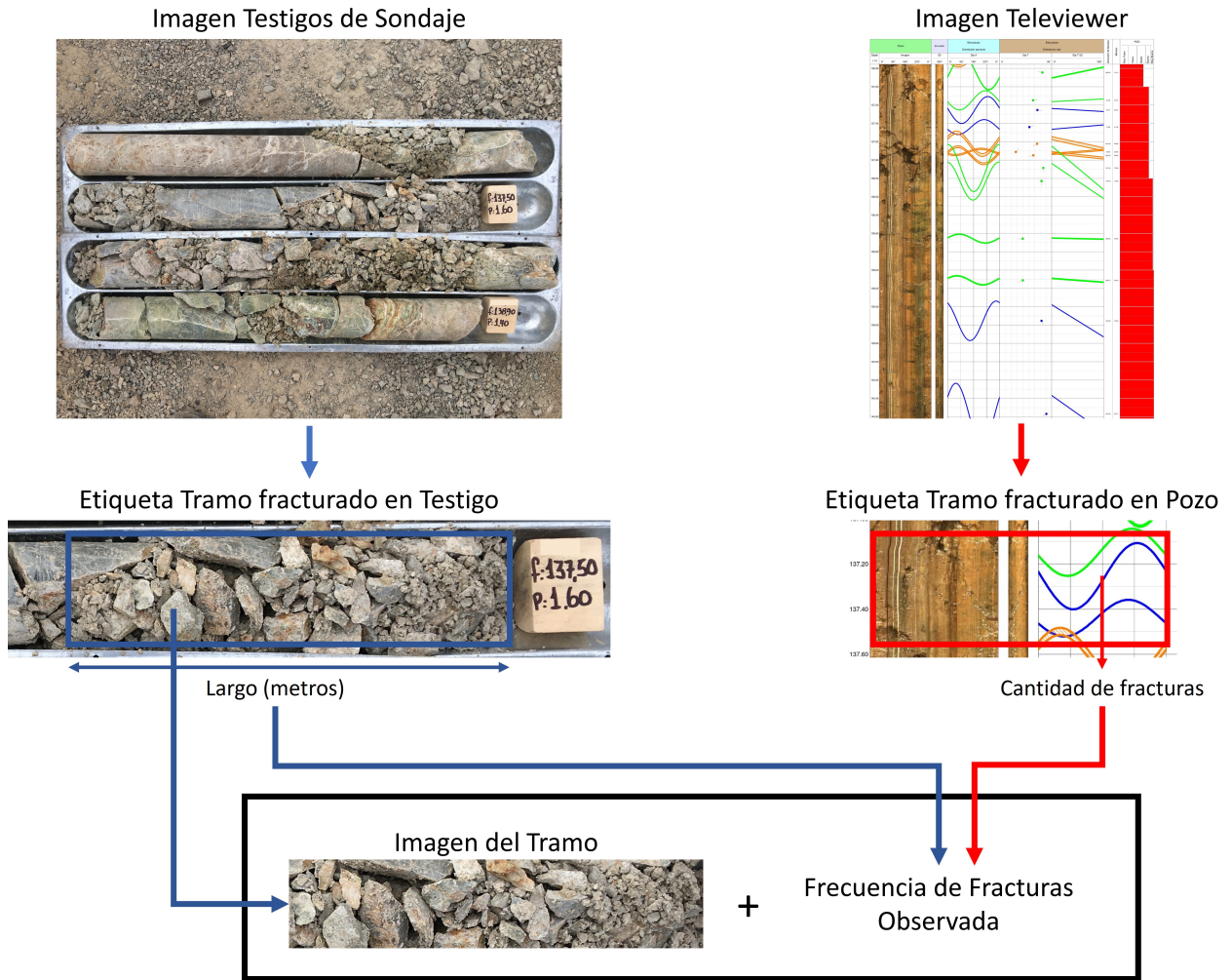


Figura 3.2: Esquema de obtención de los pares de datos de los tramos fracturados; imagen del tramo y frecuencia de fracturas observada asociada.

3.2. Procesamiento de datos

Una vez obtenida la *Frecuencia de Fracturas Observada* (FF_{Obs}) de un total de 463 tramos roca muy fracturada, se generó el histograma de la Figura 3.3 para conocer la distribución de los datos.

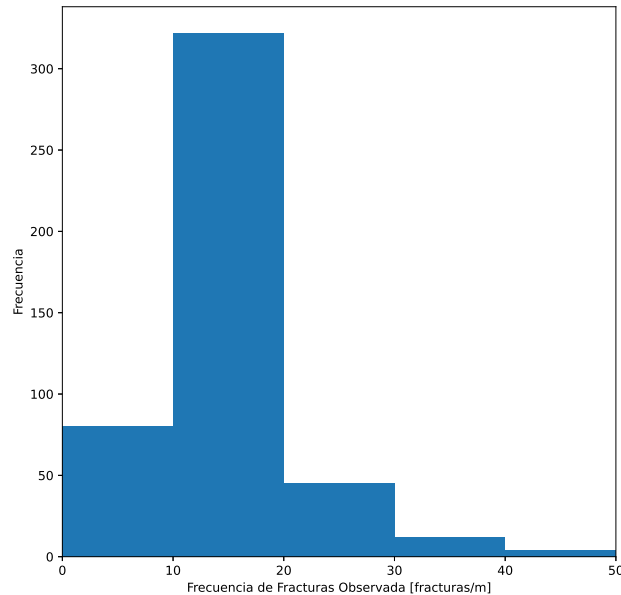


Figura 3.3: Histograma de Frecuencia de Fracturas Observada.

El valor mínimo FF_{Obs} obtenido fue de 5 fracturas/m, el máximo de 45 fracturas/m y el promedio de 14 fracturas/m.

A continuación, se procesaron las imágenes de cada imagen de los tramos de roca fracturada.

3.2.1. Procesamiento de imágenes

Para describir los tramos de roca fracturada en términos visuales, se deben reconocer y describir los fragmentos que lo componen. Para identificar cada uno de estos fragmentos a partir de la imagen del tramo, las imágenes pasaron por una serie de transformaciones utilizando la librería OpenCV en Python. A continuación, se entrega una breve descripción de cada uno de los pasos y transformaciones aplicadas a cada imagen:

1. Original: Imagen del tramo de roca fracturada sin procesar.
2. Denoising: Eliminación de ruido de imágenes utilizando el algoritmo de eliminación de ruido de medios no locales, o basado en el promedio de todos los píxeles en una imagen.
3. Escala de grises: Conversión de escala de colores RGB a escala de grises.
4. Median Blur: Desenfoque de la imagen, tomando la mediana de todos los píxeles debajo del área de un núcleo de píxeles. El elemento central se reemplaza con este valor mediano.

5. Black Hat: Transformación morfológica utilizada para realzar los objetos oscuros de interés en un fondo brillante.
6. Complemento: Convierte el valor de cada píxel por su complemento en escala de grises, es decir, colores más cercanos al blanco serán transformados en colores más cercanos al negro y viceversa.
7. Thresholding: Conversión de una imagen de múltiples canales para obtener una imagen de 2 niveles (binaria) en escala de grises.
8. Dilatación: Transformación morfológica que convoluciona la imagen con un núcleo de píxeles para expandir los colores de un tipo.
9. Encontrar contornos: Búsqueda de las curvas que unen todos los puntos continuos (a lo largo de un límite) que tienen el mismo color o intensidad, encerrando en un área respectiva cada uno.

En la Figura 3.4 se observa un ejemplo de todas las transformaciones aplicadas sobre la imagen original y el resultado visual de cada una de ellas, hasta identificar cada uno de los fragmentos de roca, detectando su área y contornos. Notar que en el caso de fragmentos de roca muy angulosos, se detecta el contorno entre las distintas caras que lo componen.

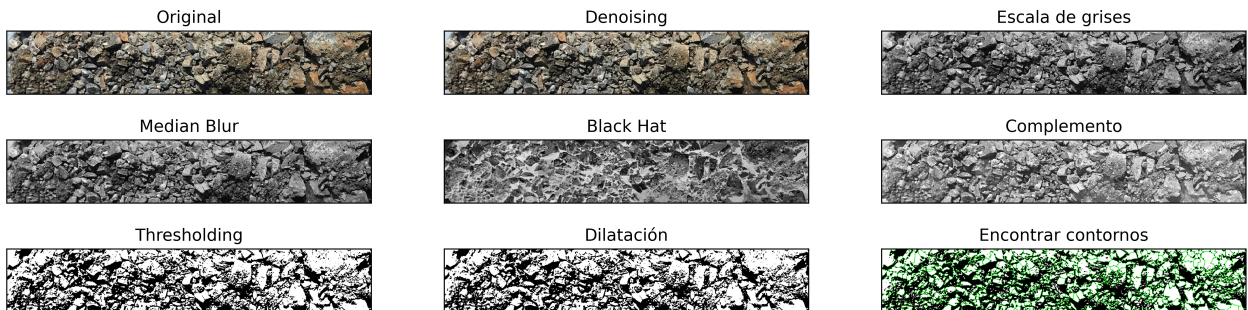


Figura 3.4: Ejemplo de pasos y transformaciones aplicadas sobre una imagen de fragmentos de roca, reconociendo el área y contorno de cada fragmento, utilizando la librería OpenCV en Python.

3.2.2. Obtención de parámetros asociados a los fragmentos

Una vez el contorno de cada fragmento es detectado, se procedió a calcular algunos parámetros para cada uno de estos fragmentos. El origen del sistema de referencia se considera en la esquina superior izquierda de la imagen, con la componente x recorriendo los píxeles de la imagen hacia la derecha y la componente y recorriendo los píxeles de la imagen hacia abajo. Sea (x_i, y_i) la tupla de coordenadas de cada uno de los píxeles que componen el contorno del fragmento, su centro de masas (x_c, y_c) entonces está dado por el promedio de estos, es decir

$$x_c = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (3.1)$$

y

$$y_c = \frac{1}{n} \cdot \sum_{i=1}^n y_i, \quad (3.2)$$

donde n es la cantidad de píxeles que conforman el contorno. Luego, la distancia o radio (r_i) entre cada píxel del contorno y el centro de masas del fragmento es

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}. \quad (3.3)$$

Una vez se cuenta con los radios r_i de los contornos del fragmento con respecto a su centro de masa, según la escala de la imagen, se obtuvieron los siguientes parámetros asociados a cada fragmento:

- *Tamaño* (D): Corresponde al doble del promedio de los radios r_i que definen el contorno del fragmento. Se mide en mm y está dado por la ecuación

$$D = \frac{2}{n} \cdot \sum_{i=1}^n r_i. \quad (3.4)$$

- *Esfericidad* (E): Corresponde a la razón entre el mínimo y máximo de los radios r_i que definen el contorno del fragmento. Es adimensional y está dada por la ecuación

$$E = \frac{\text{mínimo}\{r_i\}}{\text{máximo}\{r_i\}}. \quad (3.5)$$

Una vez se calcularon los parámetros asociados a cada fragmento, se extrajeron medidas de tendencia central y dispersión para el *Tamaño*, y el promedio de la *Esfericidad*, para la distribución total de los fragmentos. Por lo tanto, considerando que se reconocieron m fragmentos, los parámetros extraídos de la imagen de cada tramo fueron los siguientes:

- *Tamaño promedio* (D_{prom}): Corresponde al promedio de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{prom} = \frac{1}{m} \cdot \sum_{j=1}^m D_j. \quad (3.6)$$

- *Tamaño mediana* (D_{med}): Corresponde a la mediana de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{med} = \text{mediana}\{D_j\}. \quad (3.7)$$

- *Tamaño moda* (D_{mod}): Corresponde a la moda de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{mod} = \text{moda}\{D_j\}. \quad (3.8)$$

- *Tamaño máximo* (D_{max}): Corresponde al máximo de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{max} = \text{máximo}\{D_j\}. \quad (3.9)$$

- *Tamaño mínimo* (D_{min}): Corresponde al mínimo de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{min} = \text{mínimo}\{D_j\}. \quad (3.10)$$

- *Tamaño rango* (D_{rang}): Corresponde a la resta entre el máximo y mínimo de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{rang} = D_{max} - D_{min}. \quad (3.11)$$

- *Tamaño desviación estándar* (D_{desv}): Corresponde a la desviación estándar de los tamaños D_j de cada fragmento del tramo. Se mide en mm y está dado por la ecuación

$$D_{desv} = \sqrt{\frac{1}{(m-1)} \cdot \sum_{j=1}^m (D_j - D_{prom})^2}. \quad (3.12)$$

- *Tamaño asimetría* (D_{asim}): Corresponde al coeficiente de asimetría de Fisher-Pearson de los tamaños D_j de cada fragmento del tramo. Es adimensional y está dada por la ecuación

$$D_{asim} = \sqrt{m} \cdot \frac{\sum_{j=1}^m (D_j - D_{prom})^3}{\left(\sum_{j=1}^m (D_j - D_{prom})^2\right)^{\frac{3}{2}}}. \quad (3.13)$$

- *Tamaño curtosis* (D_{curt}): Corresponde al coeficiente de curtosis de Fisher de los tamaños D_j de cada fragmento del tramo. Es adimensional y está dada por la ecuación

$$D_{curt} = \frac{(m-1)^2}{m} \cdot \frac{\sum_{j=1}^m (D_j - D_{prom})^4}{\left(\sum_{j=1}^m (D_j - D_{prom})^2\right)^2}. \quad (3.14)$$

- *Esfericidad promedio* (E_{prom}): Corresponde al promedio de los esfericidades E_j de cada fragmento del tramo. Es adimensional y está dada por la ecuación

$$E_{prom} = \frac{1}{m} \cdot \sum_{j=1}^m E_j. \quad (3.15)$$

Para estudiar la relación cualitativa entre estos parámetros y la *Frecuencia de Fracturas Observada* se realizaron los gráficos visualizados en la Figura 3.5. Además, fue incluido el largo del tramo como otro parámetro a tener en cuenta para el análisis posterior.

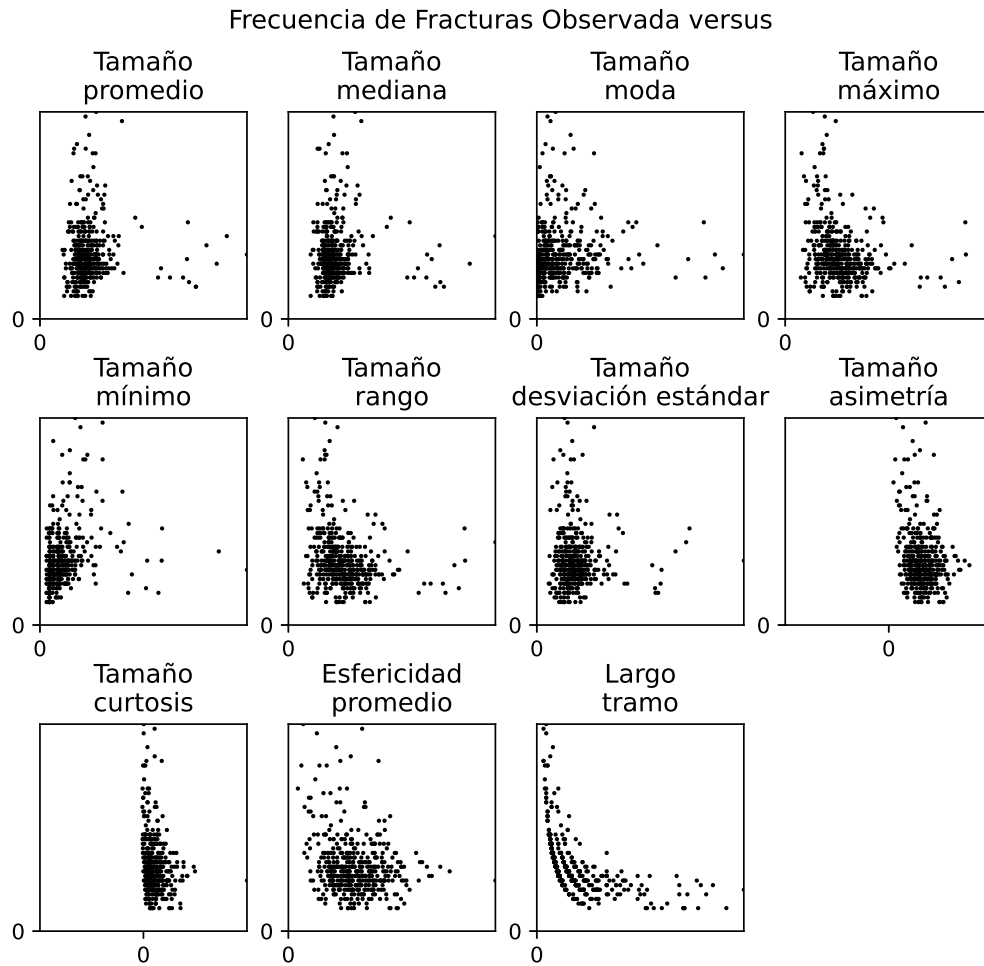


Figura 3.5: Relación cualitativa entre la Frecuencia de Fracturas Observada (eje y) y los parámetros asociados a la imagen de cada tramo de roca fracturada (eje x).

3.3. Modelos de Machine Learning

Para identificar posibles patrones y/o relaciones no evidentes a simple vista, entre los parámetros extraídos de las imágenes de tramos fracturados, con el valor de la *Frecuencia de Fracturas Observada* (FF_{Obs}) asociada, se utilizaron modelos de *Machine Learning*, implementados con las librerías *TensorFlow* y *Keras* en Python.

Machine Learning es un método analítico que permite a una máquina o sistema computacional por si mismo, de forma automática y sin intervención humana, reconocer y/o aprender relaciones, patrones o tendencias en un conjunto de datos para posteriormente clasificar o predecir los nuevos datos según los mismos patrones. Existen diversos modelos de Machine Learning y algoritmos asociados con distintos objetivos. Conforme a los alcances de este estudio, no se profundizará en los distintos tipos de algoritmos.

En términos simples, estos modelos son capaces de encontrar la relación entre un vector de datos perteneciente al espacio \mathbb{R}^n y un vector de resultados perteneciente al espacio \mathbb{R}^m , en este caso, un conjunto de datos obtenidos para cada observación (los parámetros característicos de los fragmentos de roca) y una etiqueta asignada (la frecuencia de fracturas). La aproximación de la función que relaciona estos dos espacios se realiza mediante las denominadas *redes neuronales*. Las redes neuronales son los algoritmos de aprendizaje y modelos computacionales diseñados para funcionar como neuronas en el cerebro, aproximando una función mediante una suma ponderada de coeficientes ocultos (*layers* de neuronas). Un modelo de Machine Learning tiene principalmente tres capas (Figura 3.6). La capa de entrada recibe los datos de entrenamiento. La segunda capa es una capa oculta donde se procesa un vector de datos de entrada con una función de activación, pasando el dato de salida al siguiente *layer*. Finalmente, la red usa una denominada capa de salida para generar un resultado en el espacio vectorial del conjunto de etiquetas. El modelo, internamente, hace pasar los datos desde la capa de entrada hasta la de salida (*forward-propagation*) y va evaluando qué tan bien se asemejan los resultados a la etiqueta esperada. Una función de pérdida mide esto y mediante un optimizador, modifica los parámetros que ponderan a cada neurona hacia atrás (*back-propagation*), hasta que los resultados se asemejen al esperado con un umbral de precisión o un número finito de iteraciones. En general, la precisión de un modelo define su confiabilidad y esta se mide según el porcentaje de detecciones correctas en relación con el total de predicciones.

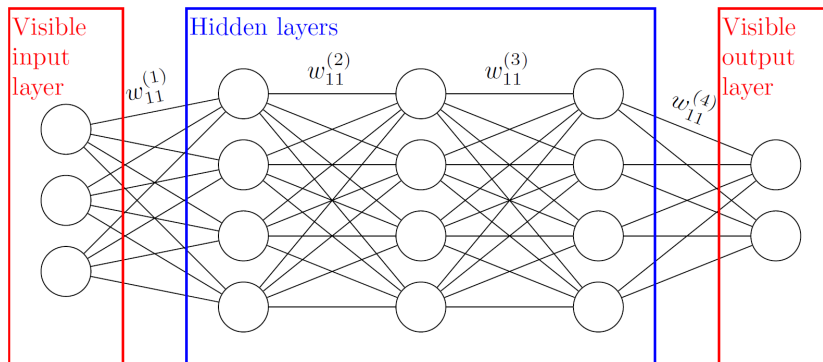


Figura 3.6: Arquitectura de una red de Machine Learning (Extraído de Adcock et al., 2015).

3.3.1. Modelo de Regresión Básica

El modelo implementado en este estudio fue basado en el ejemplo de Regresión Básica de la librería TensorFlow: *Predecir eficiencia de gasolina*. El modelo implementado en este ejemplo es uno secuencial de dos capas ocultas conectadas, cuya dimensionalidad del espacio de salida es 64, y una capa final de salida que devuelve un único valor continuo. La función de activación de las capas ocultas es de tipo *ReLU*, es decir, retorna el elemento máximo entre 0 y el tensor de entrada (Hahnloser et al., 2000). Para compilar el modelo de este ejemplo se utiliza una función de pérdida: *Error cuadrático medio* (*MSE*, siglas de *Mean Squared Error*, en inglés) y un optimizador de tipo *RMSprop*, cuya esencia consiste en mantener un promedio móvil del cuadrado de los gradientes y dividir el gradiente por la raíz de este promedio (TensorFlow). Las líneas de código asociadas a la implementación de este ejemplo en Python, con las librerías TensorFlow y Keras, se observan en la Figura 3.7.a.

Las modificaciones realizadas en el modelo de ejemplo de Tensorflow, con respecto al modelo implementado en este estudio fue el incluir una capa oculta extra, cuya dimensionalidad del espacio de salida también es de 64, además de modificar la función de activación de cada una de las capas ocultas, por unas de tipo *Sigmoid* (López, 2006), que está definida por la fórmula

$$P(t) = \frac{1}{1 + e^{-t}}. \quad (3.16)$$

Las líneas de código asociadas a la implementación de este modelo en Python, con las librerías TensorFlow y Keras, se observan en la Figura 3.7.b.

```
model = keras.Sequential([
    layers.Dense(64, activation='relu', input_shape=[len(train_data.keys())]),
    layers.Dense(64, activation='relu'),
    layers.Dense(1)
])

model.compile(loss='mse',
              optimizer=tf.keras.optimizers.RMSprop(0.001))
```

(a) Ejemplo de TensorFlow: Predecir eficiencia de gasolina.

```
model = keras.Sequential([
    layers.Dense(64, activation='sigmoid', input_shape=[len(train_data.keys())]),
    layers.Dense(64, activation='sigmoid'),
    layers.Dense(64, activation='sigmoid'),
    layers.Dense(1)
])

model.compile(loss='mse',
              optimizer=tf.keras.optimizers.RMSprop(0.001))
```

(b) Implementado en este estudio.

Figura 3.7: Fragmentos de código de la implementación de un Modelo de Regresión Básica en Python, utilizando las librerías TensorFlow y Keras.

Para construir este modelo, se realizaron distintas pruebas, incluyendo distintas combinaciones de variables de entrada, filtrado de datos en cierto rango, y distintos optimizadores, funciones de pérdida y funciones de activación. En particular, se filtraron aquellos tramos de

roca fracturada cuyo *Tamaño máximo* fuera inferior a 20 mm, y de un total original de 463 tramos de roca fracturada, se utilizaron 448.

Las variables de entrada utilizadas para entrenar este modelo corresponden a cada uno de los parámetros de los fragmentos, asociados a cada tramo de roca fracturada, sumado al largo del tramo, y su etiqueta a *Frecuencia de Fracturas Observada*. Cada una de las variables de entrada fue estandarizada y normalizada mediante un proceso de *encoding*, y el método utilizado para entrenar y evaluar el modelo corresponde al de *Validación cruzada de 10 iteraciones*, que se explica en la sección 3.3.2.

3.3.2. Validación cruzada de 10 iteraciones

En la implementación de modelos de Machine Learning, una porción de los datos es utilizada para entrenar el modelo (Datos de entrenamiento) y el resto para evaluar su precisión y desempeño (Datos de prueba). La Validación cruzada es una técnica muy utilizada para validar modelos de Machine Learning, la cual permite garantizar que los resultados de un análisis son independientes de la partición entre datos de entrenamiento y prueba. La Validación cruzada de k iteraciones consiste en dividir la totalidad de datos en k subconjuntos. Uno de ellos se utiliza como datos de prueba y la totalidad del resto (k-1) como datos de entrenamiento. Este proceso se repite por k iteraciones, tomando un subconjunto de prueba distinto y evaluando la precisión arrojada por el modelo en cada una de ellas.

En el caso de este estudio, se utilizaron 10 particiones e iteraciones, principalmente porque para cantidades menores el modelo reducía su precisión, a falta de datos de entrenamiento, y para cantidades mayores los datos de prueba se reducían considerablemente. Por otra parte, corresponde a un número común de iteraciones utilizadas (Joanneum, 2005). De esta forma, en cada iteración se utilizó el 90 % de los datos como datos de entrenamiento (403 tramos de roca fracturada de 448), y el 10 % como datos de prueba (45 tramos de roca fracturada de 448). Las frecuencias de fracturas arrojadas y predichas por modelo (*Frecuencia de Fracturas Predicha o FF_{Pred}*) en cada iteración fue contrastada con la *Frecuencia de Fracturas Observada*, y obteniendo el coeficiente de determinación ajustado (R_a^2) para el total de ellas. La precisión general del modelo se considera como el promedio entre los coeficientes de determinación ajustados, obtenidos en cada iteración. En la Figura 3.8 se observa el esquema de aplicación de la Validación cruzada de 10 iteraciones, utilizada en este estudio.

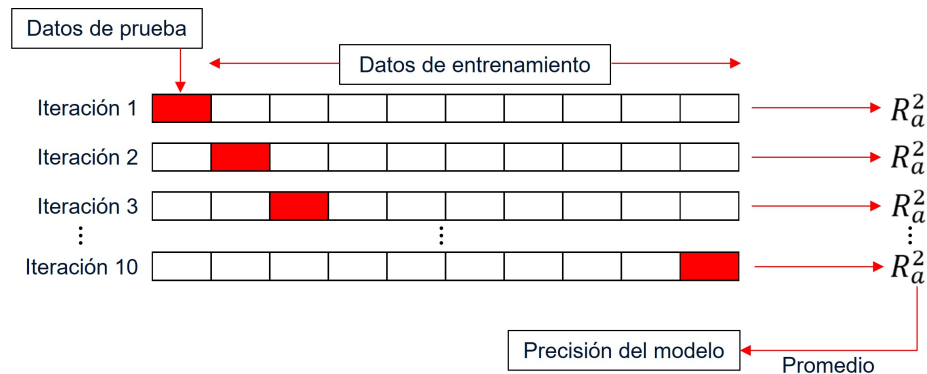


Figura 3.8: Esquema de Validación cruzada de 10 iteraciones, utilizada en este estudio (Elaboración propia).

El error asociado a la predicción (*Error Frecuencia de Fracturas* o FF_{error}) de cada dato se calcula como la resta entre el valor observado y el predicho, dividido por el rango posible de frecuencia de fracturas (máximo menos mínimo, es decir, 40 fracturas/m) y expresado en porcentaje, de la siguiente forma:

$$FF_{error} = \frac{(FF_{Obs} - FF_{Pred})}{40 \text{ fracturas/m}} \cdot 100\%, \quad (3.17)$$

mientras su valor absoluto se expresa simplemente como $|FF_{error}|$. Luego, los gráficos generados para analizar los resultados fueron los siguientes:

- *Frecuencia de Fracturas Predicha* versus *Frecuencia de Fracturas Observada* para la mejor, peor y totalidad de iteraciones.
- *Error Frecuencia de Fracturas* versus *Frecuencia de Fracturas Observada* para la mejor, peor y totalidad de iteraciones.

Los gráficos de valores predichos versus observados permiten evidenciar si son similares entre sí y se acercan a una recta de pendiente 1 que pasa por el origen, mientras los gráficos de error permiten reconocer si existe una correlación de aumento o disminución de este mismo con respecto a otra variable, así como la posible existencia de valores atípicos.

Por otro lado, los parámetros estadísticos más importantes a considerar fueron los siguientes:

- Promedio de *Error Frecuencia de Fracturas* (FF_{error}^{prom}), para la mejor, peor y totalidad de iteraciones.
- Promedio de valor absoluto de *Error Frecuencia de Fracturas* ($|FF_{error}|^{prom}$), para la mejor, peor y totalidad de iteraciones.
- Máximo de valor absoluto de *Error Frecuencia de Fracturas* ($|FF_{error}|^{max}$), para la mejor, peor y totalidad de iteraciones.
- Coeficiente de determinación ajustado (R_a^2) entre *Frecuencia de Fracturas Predicha* y *Frecuencia de Fracturas Observada*, para la mejor, peor y total de iteraciones.

Los promedios de errores permiten saber si estos se concentran positiva o negativamente, es decir si existe en general una sobrestimación o subestimación de los resultados por parte de la predicción del modelo. Los promedios de valores absolutos de errores permiten tener una idea de qué tanto difieren los resultados entregados por el modelo, en relación con los datos observados, mientras que el máximo indica la cota de este mismo. El coeficiente de determinación ajustado es el parámetro principal para medir la calidad del modelo propuesto y probar y validar la hipótesis del estudio.

3.4. Resultados

El coeficiente de determinación ajustado (R_a^2) obtenido para cada una de las iteraciones, así como el valor promedio de todas, correspondiente a la precisión del modelo implementado según el cálculo tradicional, se puede apreciar en la Tabla 3.1.

Tabla 3.1: Coeficiente de determinación ajustado (R_a^2) de cada una de las iteraciones y su valor promedio.

Iteración	Coeficiente de determinación ajustado (R_a^2)
1	0.62
2	0.78
3	0.72
4	0.61
5	0.56
6	0.63
7	0.60
8	0.62
9	0.72
10	0.64
Promedio	0.65

3.4.1. Mejor iteración

Para la iteración con coeficiente de determinación ajustado más alto obtenido (iteración 2), la relación entre los resultados de *Frecuencia de Fracturas Predicha* y *Frecuencia de Fracturas Observada* se aprecia en la Figura 3.9.

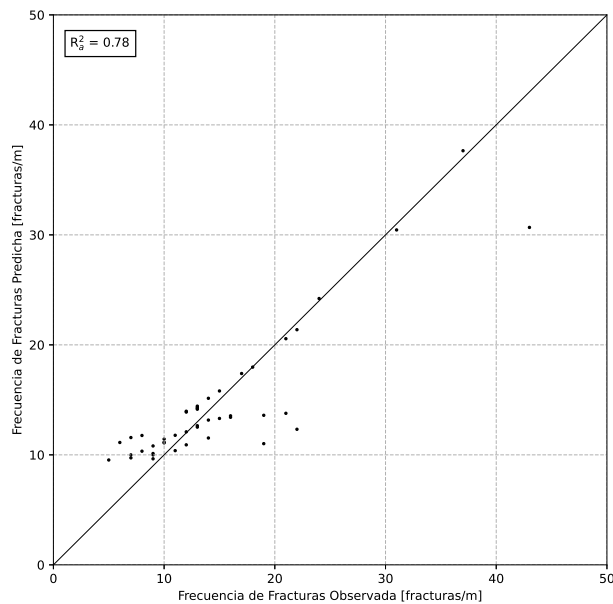


Figura 3.9: Comparación de Frecuencia de Fracturas: Predicha versus Observada, iteración 2.

El coeficiente de determinación ajustado (R_a^2) obtenido según la Figura 3.9 fue de 0.78. Por otro lado, la relación entre *Error Frecuencia de Fracturas* y *Frecuencia de Fracturas Observada* se aprecia en la Figura 3.10.

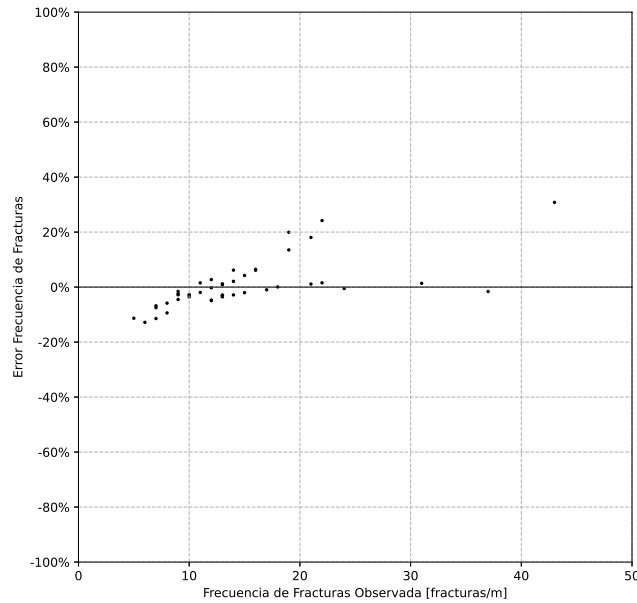


Figura 3.10: Análisis de errores: Error Frecuencia de Fracturas versus Frecuencia de Fracturas Observada, iteración 2.

De los valores de *Error Frecuencia de Fracturas* (FF_{error}) se obtuvo un promedio FF_{error}^{prom} de 0% (equivalente a 0 fracturas/m). De sus valores absolutos $|FF_{error}|$ se obtuvo un promedio $|FF_{error}|^{prom}$ de 5.00% (equivalente a 2 fracturas/m) y un máximo $|FF_{error}|^{max}$ de 30.00% (equivalente a 12 fracturas/m).

3.4.2. Peor iteración

Para la iteración con coeficiente de determinación ajustado más bajo obtenido (iteración 5), la relación entre los resultados de *Frecuencia de Fracturas Predicha* y *Frecuencia de Fracturas Observada* se aprecia en la Figura 3.11.

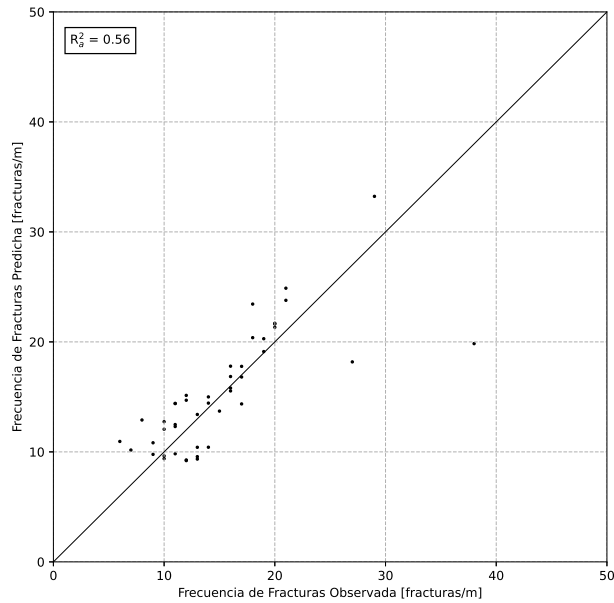


Figura 3.11: Comparación de Frecuencia de Fracturas: Predicha versus Observada, iteración 5.

El coeficiente de determinación ajustado (R_a^2) obtenido según la Figura 3.11 fue de 0.56. Por otro lado, la relación entre *Error Frecuencia de Fracturas* y *Frecuencia de Fracturas Observada* se aprecia en la Figura 3.12.

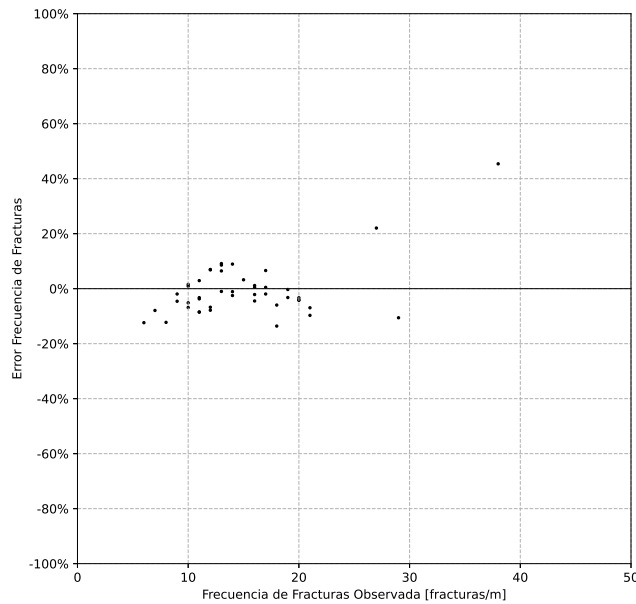


Figura 3.12: Análisis de errores: Error Frecuencia de Fracturas versus Frecuencia de Fracturas Observada, iteración 5.

De los valores de *Error Frecuencia de Fracturas* (FF_{error}) se obtuvo un promedio FF_{error}^{prom} de 0% (equivalente a 0 fracturas/m). De sus valores absolutos $|FF_{error}|$ se obtuvo un promedio $|FF_{error}|^{prom}$ de 7.50% (equivalente a 3 fracturas/m) y un máximo $|FF_{error}|^{max}$ de 45.00% (equivalente a 18 fracturas/m).

3.4.3. Total de iteraciones

Para la totalidad de datos recopilados de todas las iteraciones, la relación entre los resultados de *Frecuencia de Fracturas Predicha* y *Frecuencia de Fracturas Observada*, se aprecia en la Figura 3.13.

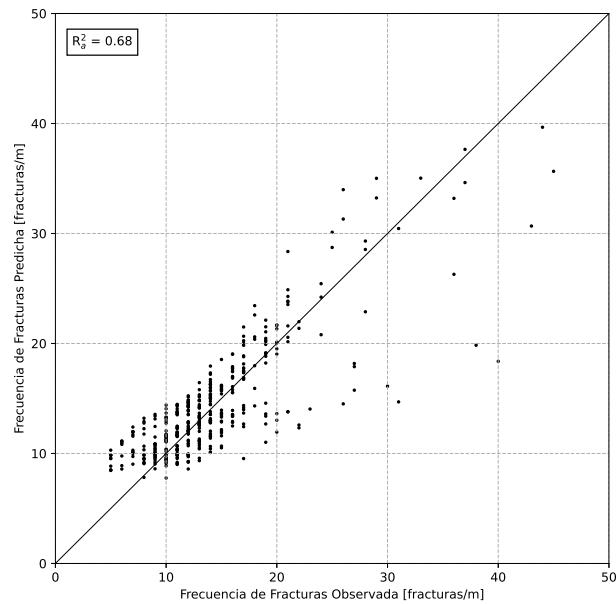


Figura 3.13: Comparación de Frecuencia de Fracturas: Predicha versus Observada, para la totalidad de iteraciones recopiladas.

El coeficiente de determinación ajustado (R_a^2) obtenido según la Figura 3.9 fue de 0.68. Por otro lado, la relación entre *Error Frecuencia de Fracturas* y *Frecuencia de Fracturas Observada* se aprecia en la Figura 3.14.

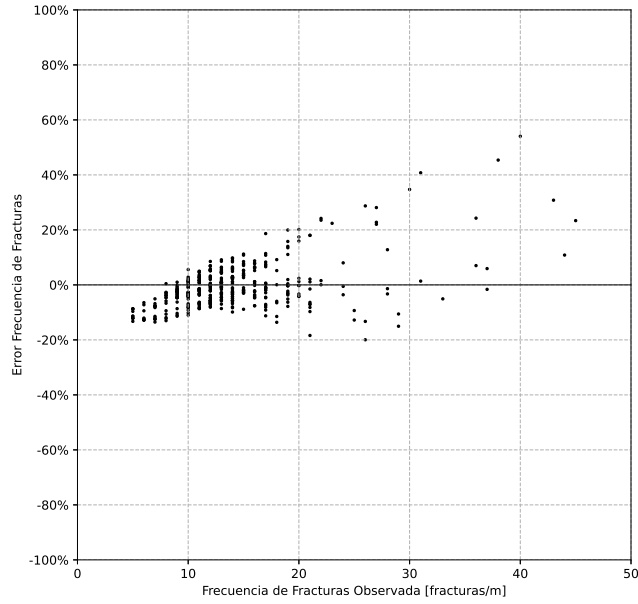


Figura 3.14: Análisis de errores: Error Frecuencia de Fracturas versus Frecuencia de Fracturas Observada, para la totalidad de iteraciones recopiladas.

De los valores de *Error Frecuencia de Fracturas* (FF_{error}) se obtuvo un promedio FF_{error}^{prom} de 0% (equivalente a 0 fracturas/m). De sus valores absolutos $|FF_{error}|$ se obtuvo un promedio $|FF_{error}|^{prom}$ de 5.00% (equivalente a 2 fracturas/m) y un máximo $|FF_{error}|^{max}$ de 55.00% (equivalente a 22 fracturas/m).

Finalmente, la relación cualitativa entre el *Error Frecuencia de Fracturas* y cada una de las variables utilizadas en el entrenamiento, se puede observar en la Figura 3.15.

Error Frecuencia de Fracturas versus

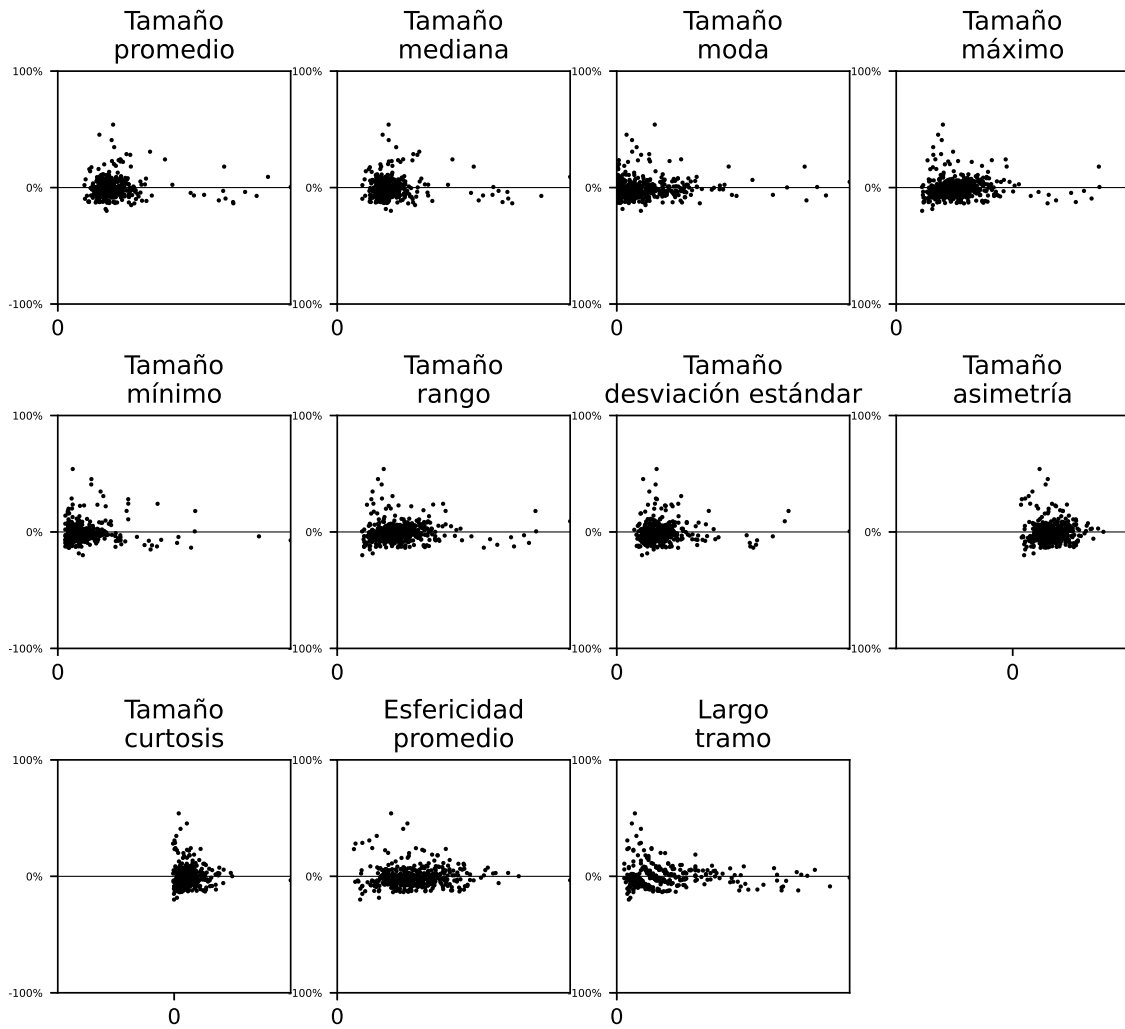


Figura 3.15: Relación cualitativa entre el Error Frecuencia de Fracturas (eje y) y los parámetros asociados a la imagen de cada tramo de roca fracturada (eje x).

Capítulo 4

Discusión

4.1. Cálculo de ángulos alfa y beta con el método de los 3 puntos

El histograma de la Figura 2.5 muestra que la mayor cantidad de fracturas consideradas en este estudio presenta ángulos alfa superiores a 60° , una cantidad intermedia presenta entre 30° y 60° y una cantidad menor presenta menos de 30° . La mayor presencia de fracturas con ángulos alfa cercanos a 90° se explica porque existe una mayor probabilidad de intersecar fracturas más perpendiculares a la dirección del sondaje, que paralelas a su dirección.

El coeficiente de determinación ajustado obtenido para los ángulos alfa (0.88) se puede interpretar como que el método de los 3 puntos tiene una precisión del 88 % en el cálculo de este ángulo. Esto se considera un buen resultado, considerando que está sobre el 80 %, y que además el error en valor absoluto con respecto al método de medición directa obtenido, fue en promedio de 5.56 % (equivalente a 5°) y en máximo 13.33 % (equivalente a 12°), respectivamente.

En las Figuras 2.6 y 2.7 se observó que para valores de *Alfa Método directo* menores a aproximadamente 40° , los resultados de *Alfa Método de los 3 puntos* están sobrestimados en su mayoría con respecto al método directo (*Error Alfa* negativo en la Figura 2.7). Esto se puede explicar porque para ángulos alfa pequeños, la traza frontal de la fractura visualizada en la imagen tiende a ser más paralela a las líneas que delimitan el testigo de sondaje, por ende los puntos P_s y P_i tienden a ser marcados antes del extremo superior e inferior de la traza, respectivamente. En consecuencia, la separación lateral de estos 2 puntos, con respecto al punto P_c del método es menor y se sobrestima el ángulo alfa. Una solución para disminuir este error es tomar fotografías con mayor iluminación, ya sea artificial, o en horarios de mayor luz solar, y a la vez, aumentar la resolución de las imágenes. De esta forma el extremo de la traza frontal de las fracturas se podría distinguir más fácilmente de las líneas que delimitan el testigo de sondaje.

Por otro lado, para valores de *Alfa Método directo* mayores a aproximadamente 80° , los resultados de *Alfa Método de los 3 puntos* están en su mayoría subestimados con respecto al método directo (*Error Alfa* positivo en la Figura 2.7), lo que se puede explicar por la separación de ciertos tramos de testigo según la fractura, no ajustados correctamente, que aumentan la separación lateral de los puntos P_s y P_i , y en consecuencia, se subestima el

ángulo alfa. En particular, para valores de alfa muy cercanos 90° en la medición según el método directo, estos ángulos son aproximados simplemente a 90° por el equipo de geología, lo cual no se lleva cabo con el método de los 3 puntos. Una solución a este problema es ajustar mejor la separación entre los tramos de testigo, previo a la captura de la fotografía, y considerar la aproximación de ángulos alfa muy cercanos a 90° a 90° exactos, para el método de los 3 puntos.

El histograma de la Figura 2.8 correspondiente a los ángulos beta, a diferencia del histograma de ángulos alfa de la Figura 2.5, presenta una distribución más igualitaria en el rango de los ángulos posibles, sin embargo hay ligeramente una mayor presencia de fracturas con ángulo beta menor a 180° que mayores a 180° , la cual se considera azarosa, y asociada a la geología estructural de la zona en que se realizaron los sondeos.

El coeficiente de determinación ajustado obtenido para los ángulos beta (0.94) indica que el método de los 3 puntos tiene una precisión del 94% en el cálculo de este ángulo (mayor al de ángulos alfa). Esto se considera un buen resultado, al igual que con los ángulos alfa, considerando que está sobre el 90%, y que además el error en valor absoluto con respecto al método de medición directa obtenido, fue en promedio de 5.56% (equivalente a 20°) y en máximo 12.22% (equivalente a 44°), respectivamente.

En las Figuras 2.9 y 2.10 no se observaron relaciones o tendencias entre el error del método de los 3 puntos para el ángulo beta (*Error Beta*) y *Beta Método directo*. Este error se considera por tanto azaroso. De igual forma, no se observa relaciones o tendencias entre *Error Beta* y *Alfa Método directo* (Figura 2.11) y viceversa (Figura 2.12). En la Figura 2.13, donde se observa la relación entre *Error Beta* y *Error Alfa* no se distingue tampoco patrones o tendencias entre ambos y los dos se encuentran entre los -12 o -13% y los 12 o 13%.

En relación con los valores Dip y Dip Direction obtenidos de los sets de fracturas de las stereonet, previo a la aplicación de la ponderación de Terzaghi, se tiene un *Error Dip* promedio entre el método de los 3 puntos y el método directo de 3.33% (equivalente a 3°), un valor absoluto promedio de 4.44% (equivalente a 4°) y un valor absoluto máximo de 7.78° (equivalente 7°), similares en porcentaje para el *Error Dip Direction*, con un promedio de 1.39% (equivalente a 5°), un valor absoluto promedio de 3.33% (equivalente a 12°) y un valor absoluto máximo de 10.00% (equivalente a 36°). Este error se considera pequeño y permite verificar la precisión del método en el sentido estructural. Sin embargo, cabe mencionar que algunos polos de fracturas considerados en el método directo no se consideraron en el mismo set para el método de los 3 puntos (Figuras 2.14, 2.15 y 2.16).

Al aplicar la ponderación de Terzaghi se reconocieron más sets de fracturas, en particular, para el Sondaaje 1, que previo a su aplicación (Figuras 2.14 y 2.17). Posterior a la aplicación de la ponderación de Terzaghi, se tiene un *Error Dip* promedio entre el método de los 3 puntos y el método directo de 1.11% (equivalente a 1°), un valor absoluto promedio de 3.33% (equivalente a 10°) y un valor absoluto máximo de 7.78° (equivalente 7°), y para el *Error Dip Direction*, un promedio de 2.78% (equivalente a 10°), un valor absoluto promedio de 3.89% (equivalente a 124°) y un valor absoluto máximo de 10.56% (equivalente a 38°). Por lo tanto, la ponderación de Terzaghi no afectó considerablemente la distribución estadística de errores, variando en no más del 3% los valores promedio y máximo. Esto es concordante

con lo esperado, ya que la misma ponderación se aplica sobre proyecciones estereográficas cuya similitud ya fue comprobada previo a su aplicación.

Finalmente, es importante mencionar que los resultados del método de los 3 puntos se compararon con mediciones realizadas directamente en el sondaje y por lo tanto los errores considerados en este estudio representan la similitud y/o discrepancia entre ambos y no necesariamente un error con respecto a los ángulos reales de la estructura. Los errores calculados, por tanto, pueden estar asociados en algunos casos a errores o aproximaciones en la medición de los ángulos con el método directo. Posibles errores del método de los 3 puntos pueden estar asociados a que la imagen-fotografía de la caja de testigos no se encuentra perfectamente encuadrada con respecto a los bordes de la caja, o como ya se mencionó, a dificultades para visualizar correctamente la traza de la fractura en la imagen, o también, a variaciones en el plano de fractura que no se evidencia solamente con la traza frontal.

4.2. Estimación de frecuencia de fracturas con Machine Learning

El histograma de la Figura 3.3 muestra que la mayor cantidad de frecuencias de fracturas observadas y consideradas en este estudio se encuentra entre las 10 y 20 fracturas/m. Una cantidad menor entre las 0 y 10 fracturas/m, y las 20 y 30 fracturas/m, y una cantidad muy baja superiores a las 30 fracturas/m. A pesar de que sobre 30-40 fracturas/m el tramo de roca puede comenzar a considerarse molido (según la faena o proyecto en cuestión), estos valores altos de frecuencia de fracturas corresponden a casos particulares donde el largo del tramo etiquetado es pequeño (*Frecuencia de Fracturas Observada* versus *Largo tramo* en Figura 3.5), donde probablemente se cruzaron varias fracturas, lo cual no necesariamente es escalable a tramos más largos. En consecuencia, para estos tramos, la razón entre las fracturas contabilizadas y el largo del tramo arroja valores altos. Si se desea estimar la frecuencia de fracturas equivalente en estos tramos para hacer un conteo aproximado de las fracturas que pasan por él, estos valores altos no representan un problema, ya que ocurren en tramos de largo pequeño, y por ende la totalidad de fracturas contabilizadas no será sobrestimada.

Los parámetros geométricos utilizados para describir cada fragmento de roca (*Tamaño* y *Esfericidad*) corresponden a los más básicos y simples (en términos de algoritmos matemáticos) para extraer por medio de análisis de imágenes. Otros de los parámetros geométricos que pudieron haberse considerado son, por ejemplo, el redondeamiento, orientación, etc., pero un cálculo preciso de estos requiere de una muy alta resolución de las imágenes, y son de mayor complejidad. Con respecto a la distribución de los tamaños de fragmentos para cada tramo fracturado, se consideraron las medidas más comunes de tendencia central (promedio, mediana y moda) y dispersión (máximo, mínimo, rango, desviación estándar, asimetría y curtosis), pero pudieron incluirse también parámetros más complejos, por ejemplo, definiendo un tamaño para los fragmentos correspondientes a clastos y a matriz, y estudiando la relación de contacto entre los clastos: si están juntos (se tocan) o separados. La elección de la *Esfericidad promedio* como único valor representativo de la totalidad de fragmentos para un tramo, sin incluir otras medidas de tendencia central o dispersión, responde básicamente a la metodología de prueba y error al entrenar y evaluar los modelos de Machine Learning, que mostraron que al incluir más medidas para la esfericidad, la precisión de los modelos

no aumentaba, e incluso en algunos casos disminuía. Caso contrario ocurría al incluir como parámetro el largo del tramo, que sí significó un aporte en el aumento de la precisión de los modelos, ya que, como se mencionó anteriormente, las frecuencias de fracturas más altas se observaron en tramos de largo pequeño, probablemente, porque según este último, el esparcimiento de los fragmentos de roca en las canaletas varía.

Es importante considerar, además, que los parámetros geométricos de los fragmentos de roca y su distribución estadística en cada tramo fracturado, en su relación con la frecuencia de fracturas, también dependen de propiedades geológicas y geomecánicas de la roca misma, como pueden ser su litología, alteración, mineralización, meteorización, relleno, resistencia, etc. Por ejemplo, una roca más alterada y/o meteorizada, o de menor resistencia, pudiera disgregarse en más fragmentos y de menor tamaño, que una menos alterada y/o meteorizada, y de mayor resistencia. Aunque en este estudio no se incluyeron variables de ese tipo, dada la disponibilidad de datos, se recomienda incluirlas y evaluarlas en estudios posteriores, donde su adhesión pudiera complementarse con la información de los parámetros geométricos.

Por otro lado, según la Figura 3.5, los valores altos de la *Frecuencia de Fracturas Observada* se distribuyen en los menores valores de los parámetros asociados al tamaño de los fragmentos. Así mismo, para el largo del tramo, como se mencionó anteriormente. Para el caso de la *Esfericidad promedio* la distribución de los valores altos es más aleatoria para casi la totalidad del rango, salvo para el segmento mayor. Es de esperar que frecuencias de fracturas altas estén asociados a tamaños de fragmentos de roca pequeño, sin embargo, también se observan frecuencias de fracturas bajas en el mismo rango de tamaños. Esto se podría explicar por un fracturamiento posterior, ya sea por pérdida de presión litostática al extraer el testigo de sondaje, o mala manipulación de los mismos. Por otra parte, de los gráficos de esta Figura no se logra observar una tendencia o función matemática exacta que explique la relación entre los parámetros y la frecuencia de fracturas observada, por lo que se recurrió a modelos de Machine Learning.

La aplicación de la Validación cruzada de 10 iteraciones permitió evaluar la precisión del modelo en 10 casos distintos de partición de los datos de entrenamiento y prueba, donde el mejor caso arrojó un R_a^2 de 0.78, el peor de 0.56, y el promedio de los 10 casos fue de 0.65. Es decir, el valor del R_a^2 varió un rango de 0.22 según la porción de datos que se escogió para entrenar y evaluar el modelo. Este valor no es alto considerando que la distribución de datos de la *Frecuencia de Fracturas Observada* se concentra considerablemente en el rango de 10 a 20 fracturas/m, y una cantidad mucho menor en otros rangos, por lo que en algunos casos el modelo se pudo haber sesgado según la elección de la porción de entrenamiento y prueba. Al calcular el R_a^2 de la totalidad de datos recopilados de todas las iteraciones, se obtuvo un valor de 0.68, bastante cercano al promedio de todas las iteraciones (0.65). Este valor de 0.68, es otra forma de evaluar la precisión de los resultados del modelo, utilizada en este estudio, y permite visualizar la totalidad de los datos, sin necesidad de segmentar el R_a^2 en iteraciones y promediar, calculándolo una sola vez, según la Figura 3.13.

Tanto para el mejor caso, como para el peor caso y la totalidad de los 10 (Figuras 3.9, 3.11 y 3.13, respectivamente), para *Frecuencia de Fracturas Observada* menores a 10 fracturas/m hay una sobrestimación de la *Frecuencia de Fracturas Predicha* por parte del modelo de Machine Learning (*Error Frecuencia de Fracturas* negativo en las Figuras 3.10, 3.12 y 3.14,

respectivamente). Lo opuesto ocurre para *Frecuencia de Fracturas Observada* mayores a 20 fracturas/m, donde hay una subestimación de la *Frecuencia de Fracturas Predicha* por parte del modelo de Machine Learning (*Error Frecuencia de Fracturas* positivo en las Figuras 3.10, 3.12 y 3.14, respectivamente). El error en estos rangos posiblemente está asociado a la escasez de datos para estos mismos para entrenar el modelo de Machine Learning, por lo que este aún no predice con una alta precisión todos esos casos.

Por otro lado, ya que los errores más altos se encuentran asociados a *Frecuencias de Fracturas Observada* altas, estos a su vez están asociados a valores pequeños de los parámetros de tamaño y esfericidad de los fragmentos de roca, y el largo del tramo (Figura 3.15). Una posible solución para disminuir estos errores y aumentar la precisión general de los resultados predichos por el modelo es enriquecer la cantidad de datos para frecuencias de fracturas menores a 10 y mayores a 20 fracturas/m, además de incluir otras variables geológicas y geomecánicas ya mencionadas.

La precisión del modelo actualmente no es lo suficientemente alta para ser aplicado como un método para estimar la frecuencia de fracturas en tramos de roca muy fracturada, ya que aún no alcanza ni el 80% de precisión (R_a^2 de 0.68, equivalente a un 68% de precisión). Sin embargo los resultados de aplicar modelos de Machine Learning muestran que sí es posible obtener una frecuencia de fracturas predicha según distintos parámetros extraídos de la imagen para la gran mayoría de casos. Si bien el método ya está planteado y el modelo se encuentra construido, enriqueciendo la cantidad de datos de frecuencias de fracturas menores a 10 y mayores a 20 fracturas/m, y evaluando otras variables, entrenando nuevamente el modelo, se cree que es posible aumentar considerablemente la precisión para estos rangos. Así mismo, considerando datos de distintas faenas y contextos geotectónicos, se puede comprobar y generalizar el modelo para cualquier caso.

Capítulo 5

Conclusiones

Con base a los resultados obtenidos, puede concluirse que:

- El método de los 3 puntos permite obtener con un 88 % de precisión los ángulos alfa y con un 94 % de precisión los ángulos beta. Estos ángulos son similares a los obtenidos por medición directa y los sets de fracturas reconocidos en stereonet también lo son. Por consiguiente, el método propuesto para la estimación de la orientación de fracturas cumple con lo esperado de ser un buen estimador de los ángulos alfa y beta medidos directamente mediante métodos tradicionales.
- Los modelos de Machine Learning permiten estimar con un 65-68 % de precisión la frecuencia de fracturas equivalente en tramos de roca muy fracturada. Los mayores errores se correlacionan con los rangos donde hay menor cantidad de datos, y por lo tanto, el método propuesto permite estimar la frecuencia de fracturas con mayor precisión en los rangos donde el modelo ha sido entrenado con mayor enriquecimiento de estos mismos.

Además, se recomienda:

- Realizar un estudio complementario al de este trabajo, en particular, para estimar la frecuencia de fracturas, enriqueciendo la cantidad de datos de frecuencias de fracturas para entrenar el modelo de Machine Learning en los rangos donde hay menor cantidad, adhiriendo, además, distintos parámetros geológicos y geomecánicos de la roca, y considerando distintas faenas y contextos geotectónicos.
- Capturar y utilizar fotografías de alta calidad, donde los bordes de las imágenes deben ser paralelas con los bordes de las cajas de testigos, contar con buena iluminación y ser de alta resolución.

Bibliografía

- [1] Adcock, J., Allen, E., Day, M., Frick, S., Hinchliff, J., Johnson, M., Morley-Short, S., Pallister, S., Price, A., & Stanistic, S. (2015). Advances in quantum machine learning. arXiv preprint arXiv:1512.02900.
- [2] Aranda, J. T. M., Alfaro, D. A., & Flores, R. J. (2019). Gamma ray, spectral gamma ray and optical borehole imager logs applied in the exploration for ore deposits. In SEG Technical Program Expanded Abstracts 2019 (pp. 2183-2187). Society of Exploration Geophysicists.
- [3] Barton, C. A., Tesler, L. G., & Zoback, M. D. (1992). Interactive image analysis of borehole televiewer data. In Automated pattern analysis in petroleum exploration (pp. 223-248). Springer, New York, NY.
- [4] Cunningham, K. J. (2004). Application of ground-penetrating radar, digital optical borehole images, and cores for characterization of porosity hydraulic conductivity and paleokarst in the Biscayne aquifer, southeastern Florida, USA. *Journal of applied geophysics*, 55(1-2), 61-76.
- [5] Cunningham, K. J., Carlson, J. I., & Hurley, N. F. (2004). New method for quantification of vuggy porosity from digital optical borehole images as applied to the karstic Pleistocene limestone of the Biscayne aquifer, southeastern Florida. *Journal of Applied Geophysics*, 55(1-2), 77-90.
- [6] Deltombe, J. L., & Schepers, R. (2000). Combined processing of BHTV travel time and amplitude images. In Proceedings from the seventh international symposium on minerals and geotechnical logging, Golden, Colorado, October (Vol. 24, No. 26, pp. 1-13).
- [7] Engelhardt, H. F., Harrison, W. D., & Kamb, B. (1978). Basal sliding and conditions at the glacier bed as revealed by bore-hole photography. *Journal of Glaciology*, 20(84), 469-508.
- [8] Frost, J. (2013). Multiple regression analysis: Use adjusted R-squared and predicted R-squared to include the correct number of variables. *Minitab Blog*, 13(6).
- [9] Gochioco, L. M., Magill, C., & Marks, F. (2002). The borehole camera: An investigative geophysical tool applied to engineering, environmental, and mining challenges. *The Leading Edge*, 21(5), 474-477.
- [10] Goodman, R. E. (1976). *Methods of geological engineering in discontinuous rocks*. West Group.
- [11] Goodman, R. E. (1980). *Introduction to rock mechanics*. New York: Wiley.
- [12] Hahnloser, R. H., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., & Seung, H. S.

- (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*, 405(6789), 947-951.
- [13] Joanneum, F. H. (2005). *Cross-validation explained*. Graz, Austria: Institute for Genomics and Bioinformatics, Graz University of Technology.
- [14] Kamewada, S., Endo, T., Kokubu, H., & Nishigaki, Y. (1989). The device and features of BIP system [A]. In *The 21st symposium on rock mechanics proceedings, committee on rock mechanics* (pp. 196-200). sl., Japanese Society of Civil Engineers.
- [15] Lahti, M. (2004). Digital borehole imaging of the boreholes KR24 upper part and PH1 at Olkiluoto, March 2004. Posiva Oy. Working report, 28, 21.
- [16] Lau, J. S. O., Auger, L. F., & Bisson, J. G. (1987). Subsurface fracture surveys using a borehole television camera and acoustic televiewer. *Canadian Geotechnical Journal*, 24(4), 499-508.
- [17] Li, S., Feng, X. T., Li, Z., Chen, B., Zhang, C., & Zhou, H. (2012). In situ monitoring of rockburst nucleation and evolution in the deeply buried tunnels of Jinping II hydropower station. *Engineering Geology*, 137, 85-96.
- [18] Li, S. J., Feng, X. T., Wang, C. Y., & Hudson, J. A. (2013). ISRM suggested method for rock fractures observations using a borehole digital optical televiewer. *Rock mechanics and rock engineering*, 46(3), 635-644.
- [19] López, L. H. (2006). *Predicción y optimización de emisiones y consumo mediante redes neuronales*. Reverté.
- [20] Maddox, S. D. (1997). Application of downhole video technology to multilateral well completions. In *SPE Offshore Europe*. OnePetro.
- [21] Marjoribanks, R. (2010). *Geological methods in mineral exploration and mining*. Springer Science & Business Media.
- [22] Martínez Miguélez, M. (2006). Validez y confiabilidad en la metodología cualitativa. *Paradigma*, 27(2), 07-33.
- [23] Miyakawa, K., Tanaka, K., Hirata, Y., & Kanauchi, M. (2000). Detection of hydraulic pathways in fractured rock masses and estimation of conductivity by a newly developed TV equipped flowmeter. In *Developments in geotechnical engineering* (Vol. 84, pp. 19-27). Elsevier.
- [24] Paillet, F. L., Barton, C., Luthi, S., Rambow, F., & Zemanek, J. R. (1990). Borehole imaging and its application in well logging-an overview. *Borehole Imaging*, Society of Professional Well Log Analysts, Inc, 6001, 3-23.
- [25] Palmer, I. D., & Sparks, D. P. (1990). Measurement of induced fractures by downhole television camera in coalbeds of the Black Warrior basin, SPE-20660. In *SPE annual technical conference and exhibition proceedings, v. pi, Production operations and engineering: Society of Petroleum Engineers* (pp. 459-472).
- [26] Pusch, R. (1998). Practical visualization of rock structure. *Engineering geology*, 49(3-4), 231-236.
- [27] Reedman, J. H. (1979). *Techniques in Mineral Exploration*. Springer, Dordrecht.

- [28] Roberson, S., & Hubbard, B. (2010). Application of borehole optical televiewing to investigating the 3-D structure of glaciers: implications for the formation of longitudinal debris ridges, midre Lovénbreen, Svalbard. *Journal of Glaciology*, 56(195), 143-156.
- [29] Rodríguez, E. M. (2005). Errores frecuentes en la interpretación del coeficiente de determinación lineal. *Anuario jurídico y económico escurialense*, (38), 315-331.
- [30] Schepers, R., Rafat, G., Gelbke, C., & Lehmann, B. (2001). Application of borehole logging, core imaging and tomography to geotechnical exploration. *International Journal of Rock Mechanics and Mining Sciences*, 38(6), 867-876.
- [31] Steel, R. G. (1960). Principles and procedures of statistics: with special reference to the biological sciences (No. 04; QA276, S82.).
- [32] Stigsson, M., & Munier, R. (2013). Orientation uncertainty goes bananas: An algorithm to visualise the uncertainty sample space on stereonet for oriented objects measured in boreholes. *Computers & Geosciences*, 56, 56-61.
- [33] Tague, J. R., Hollman, G. F., & Chaffin, G. (1999). Downhole video optimizes scale-removal program. In SPE Western Regional Meeting. OnePetro.
- [34] Uchita, Y., Harada, T., & Urayama, M. (1993). Behavior of discontinuous rock during large underground cavern excavation. In International symposium on assessment and prevention of failure phenomena in rock engineering (pp. 807-816).
- [35] Wang, C. Y., Sheng, Q., Ge, X. R., & Law, K. T. (2002). Borehole camera technology and its application in the Three Gorges project.
- [36] Williams, J. H., & Johnson, C. D. (2004). Acoustic and optical borehole-wall imaging for fractured-rock aquifer studies. *Journal of Applied Geophysics*, 55(1-2), 151-159.
- [37] Zemanek, J., Glenn, E. E., Norton, L. J., & Caldwell, R. L. (1970). Formation evaluation by inspection with the borehole televiewer. *Geophysics*, 35(2), 254-269.