



UNIVERSITAT_{DE}
BARCELONA

Afloraments fracturats digitalitzats
Avaluació de tècniques remotes en models *DFN*
i aplicació de *Machine Learning*

Laura Blanco Núñez



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- Compartitqual 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - Compartitqual 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-ShareAlike 4.0. Spain License.**



UNIVERSITAT DE
BARCELONA

AFLORAMENTS FRACTURATS DIGITALITZATS

Avaluació de les tècniques remotes en models DFN i aplicació de Machine Learning

Facultat de Ciències de la Terra

Doctorat Industrial
Laura Blanco Núñez

Directors
Dr. Josep Anton Muñoz
Dr. Òscar Gratacós Torrà

Novembre, 2022

**AFLORAMENTS FRACTURATS DIGITALITZATS.
AVALUACIÓ DE LES TÈCNIQUES REMOTES EN MODELS
DFN I APLICACIÓ DE *MACHINE LEARNING*.**

Programa de doctorat en Ciències de la Terra
Departament de Dinàmica de la Terra i l'Oceà
Grup de Geodinàmica i Anàlisi de Conques
Institut de Recerca Geomodels

Facultat de Ciències de la Terra

Doctorat Industrial presentat per

Laura Blanco Núñez

Per a optar al grau de Doctor per la Universitat de Barcelona sota la supervisió de

Dr. Josep Anton Muñoz i el Dr. Òscar Gratacós Torrà

Barcelona, novembre de 2022



UNIVERSITAT DE
BARCELONA

Anufra
Soil & Water Consulting



Institut de Recerca
Geomodels

GGAC
Grup de Geodinàmica i Anàlisi de Conques

Aquesta investigació s'ha pogut realitzar gràcies als mitjans aportats per l'Institut d'Investigació Geomodels i l'empresa Anufra, S.L.. Part del finançament s'ha realitzat amb els fons del programa de Doctorat Industrial de l'AGAUR de la Generalitat de Catalunya 2016 DI 00069 i els projectes del Ministeri d'Economia, Indústria i Competitivitat CGL2017-84720-R (AEI/FEDER, UE) i CGL2017- 85532-P (AEI/FEDER, UE).

AGRAÏMENTS

Un doctorat industrial, quina gran iniciativa! Encara recordo aquestes paraules al meu cap, on la combinació universitat i empresa, ja de per si sola, sonava a repte i a més a més doctorat, un doctorat aplicat, un doctorat industrial que tingués un interès dins de la indústria, certa aplicabilitat..., cada cop més enrevessat tot plegat, però per això existeix la creativitat. Gràcies Xavi Alemany, per ficar-me aquestes paraules al cap i per iniciar conjuntament aquesta gran iniciativa amb l'empresa Anufra i la Universitat de Barcelona, per atrevir-te, per l'empenta i les ganes de fer. Combinar la feina d'una petita consultoria i desenvolupar un doctorat, no ha estat gens fàcil, ni còmode, hem après moltíssim, però ha implicat un compromís enorme i un gran esforç de tots els companys, així que mercès Xavier Alemany, Enric Cruz, Raúl Prieto, Aida Musachs, Guillem Rodríguez, Sofia Martín i Mar Vilà, un plaer tenir-vos com a companys.

Un esforç que no l'he fet sola, he estat rodejada de molta gent al llarg d'aquest camí també a la Facultat de Ciències de la Terra. En primer lloc, les gràcies més grans són *concretament*, per a tu, David Garcia, per la paciència, per aquesta muntanya russa que tens al cap, per donar-li tantes voltes a tot, per compartir camp, oficina, converses, controvèrsies, rutines, experiències, per ajudar-me a preparar congressos i presentacions, escrits i re-escrits i re-escrits, has estat clau per aquesta tesi que comparteixo amb tu fins a la última lletra, amb detalls on si coincidim, amb detalls on no coincidirem mai, però amb l'escènica de la ciència. També vull fer menció de la Bet Beamud i la Ylenia Almar, per estar i escoltar, per donar-me consells i per dinar amb mi, a l'Eduard Roca, pels llibres enormes que m'has deixat i per la manera de pensar, pels ratets i les mini càpsules de coneixement. Al Telm Bover, per tota l'ajuda amb les làmines primes de la Fou i a l'Oriol Ferrer, per l'empenta infrenable i tanta energia, ets un crack! En general, a tots el companys i companyes de la Facultat de Ciències de la Terra i concretament al Grup de Geomodels. Gràcies també, a la Marta Guinau, a l'Oriol Falivene, la Maria Roma, la Teresa Beamud, l'Alberto Sáez, el Ramon Casas, en Joan Guimerà, la Maria Ortuño, l'Octavi, al Xabi Blanch, la Olga Magrinyar, la Glòria Furdada, i molts més.

Agrair especialment a l'Anna Puig i la Maria Salomó, de la Universitat de Matemàtiques i Informàtica, totes les facilitats, reunions i explicacions per a col·laborar amb el desenvolupament d'eines de *machine learning* aplicades a les caigudes de roques. Gràcies també a l'Ivan Sevilla, el Nicolás Pascual i en Zoupekas, per formar part de tot això.

També a Ciments Molins, SA, gràcies per haver format part d'aquest projecte. En Jose Maria Ramos, per confiar amb el nostre projecte, també a Luis Arena, per les facilitats en la organització i d'altres treballadors que a peu de pedrera ens han acompanyat i ajudat en tot lo necessari. A l'Institut Cartogràfic i Geològic de Catalunya (ICGC), per tots els precedents a Montserrat, en

Marc Janeras, Oriol Pedraza, especialment. A Jose Luís Cacho per acompanyar-nos al front de túnel de Tres Ponts, d'Organyà i mostrar-nos els coneixements i per fer moltes fotos per nosaltres, com m'agraden aquests grans sondejos del subsòl! I en Carles Mujal i tot l'equip, per totes les facilitats per entrar-hi, que no és gens fàcil.

Als meus tutors de tesi, l'Anton i l'Òscar, agrair la oportunitat d'aquesta gran experiència. Òscar, mercès per buscar sempre alguna cosa per millorar, per la paciència i l'atenció, per les llargues estones i per les correccions, modificacions, matisos, comentaris...

Durant aquesta anys he tingut l'oportunitat de retrobar-me amb gent que coneixia i de conèixer a altres persones que m'han inspirat o il·luminat i que formen en part, part d'aquesta tesi. Gràcies també a ells.

A tots els meus amics i tota la meva família, sense excepció.

A mons pares, que sempre estan allà, per escoltar-me i aguantar-me, per recollir als nens a l'escola, per ajudar-nos a casa, per cuidar-me i aguantar-me una altra vegada, gràcies! I li agraeixo especialment, en el meu marit, en Ferran, que m'hagi suportat amb els diferents estats d'ànims i haver-me donat SUPORT per aguantar tants fronts oberts. I també als meus fills, que m'han escoltat teoritzant i explicant avanços i retrocessos, detall a detall, fins a l'estrenuïtat, fins l'avorrimnt més absolut i encara així m'ha semblat que m'entenien. I també, per obligar-me a deixar de treballar quan més concentrada estava, per obligar-me a jugar amb vosaltres, per molestar-me i fer-me riure, per parlar quan no tocava, per fer-ho tot una mica més difícil però molt més divertit. Gràcies Pol i Abril, ús súper estimo tant!

En l'actual transició digital, on es generen milions de dades cada dia, l'ús de tècniques remotes digitals i els processats amb tècniques de *machine learning* és cada cop més present i necessari, sense que l'àmbit de les geociències en sigui una excepció. En aquest context, el present doctorat industrial, se centra en l'anàlisi d'afloraments fracturats utilitzant les tècniques remotes i *machine learning*, amb dos objectius específics; (1) l'avaluació de la representativitat de tècniques remotes per a identificar i caracteritzar las fractures i posteriorment representar-les en models discrets de xarxes de fractures (DFN), i (2) l'ús de models classificadors amb *machine learning* en processos associats als medis fracturats, com són les caigudes de blocs.

Metodologia de treball

En el primer cas d'estudi, s'avaluen les tècniques remotes; *Terrestrial Laser Scanner* (TLS) i Fotogrametria (SfM). L'experiment es desenvolupa en la pedrera activa de La Fou a Barcelona. Els afloraments fracturats s'han reconstruït a partir d'un alt nombre de dades capturades remotament (Santana et al., 2012; Geyer et al., 2015) i s'han agrupat en famílies per caracteritzar geomètricament la xarxa de fractures (García-Sellés et al., 2018). Amb aquests resultats, i els recollits amb tècniques directes amb els *scan-lines*, es construeixen models tridimensionals (DFN). Es construeix un model amb cada tècnica de captura (DFN_{TLS}, DFN_{SfM} i DFN_{scan-line}), on quedaran representades les principals propietats geomètriques de les fractures. Finalment, se generen els mapes de traces de fractures per calcular la porositat i permeabilitat de la xarxa que hi representen.

El segon cas té com objectiu automatitzar la classificació de caigudes de blocs. L'àrea d'estudi se situa en els escarpaments històricament inestables de Degotalls, prop de l'accés al Monestir de Montserrat. La monitorització d'aquests escarpaments mitjançant els escanejos periòdics amb TLS, es va iniciar l'any 2007, com a conseqüència dels episodis de caigudes de blocs iniciats l'any 2001, que van afectar a l'aparcament del monestir (limitant l'accés al Santuari en diverses ocasions). La metodologia consisteix en alinear dos escanejos o núvols de punts del mateix aflorament capturats en dos moments diferents, amb la intenció de detectar les agrupacions de diferències existents entre ambdós escanejos. Posteriorment, es revisen visualment les agrupacions de punts per identificar i classificar les caigudes de blocs. La metodologia desenvolupada en aquesta tesi, incorporant algorismes de *machine learning* proposa automatitzar aquesta última etapa d'identificació i classificació de blocs caiguts.

Resultats

En el primer cas d'estudi a La Fou les principals propietats geomètriques de cada fractura capturada i processada són la posició, orientació, alçada i longitud. Aquestes s'utilitzen per

caracteritzar les estadístiques de les fractures, com l'espaiat, intensitat, P_{10} , P_{20} , P_{21} i P_{32} , entre d'altres. Del conjunt de fractures s'identifica la Família I WNW- ESE, Família II NNW/SSE, Família III, N-S i Família IV NE-SW. La tècnica amb major quantitat de fractures caracteritzades ha estat el TLS amb més de 52.000, després la tècnica de fotogrametria amb 1.094 i finalment els scan-lines amb 398 fractures mesurades. L'alçada i longitud de les fractures són molt variables, amb valors mitjos que oscil·len entre els pocs centímetres als 18m. El valors obtinguts amb TLS no arriben al metre d'alçada i longitud. Els espaiats entre les fractures de la Família I presenten mides centimètriques i la resta de famílies valors mitjos més variables des de centimètrica fins a 6m. La fotogrametria obté valors que en ocasions dupliquen o tripliquen la resta de tècniques. Amb la introducció de les obertures de les fractures en el model s'han generat els valors de porositat i permeabilitat de la xarxa de fractures. En els tres models DFN, el valor mig de porositat obtingut està entre 5,4 i 5,7% i els valors de permeabilitat mitja entre 4×10^5 mD i $6,2 \times 10^5$ mD.

En el segon cas d'estudi, els resultats generats a partir de la combinació de 11 algorismes classificadors amb 15 mètodes de remostreig, permeten identificar la millor combinació predictiva para la classificació de blocs caiguts, segons el període i l'escarpament estudiat. A Degotalls-E, a partir d'aproximadament 5.800 agrupaments de diferències identificats per cada monitoratge, s'assoleix una reducció del 98% del nombre d'agrupaments a validar, entre Falsos positius (FP) i Veritables Positius (VP), entenent (VP) com les caigudes reals de blocs. A Degotalls-N la reducció ha estat del 80,16% per a una identificació completa, mentre que per identificar el 96% de les caigudes de blocs reals, la reducció ha estat del 90%, a partir d'una població d'uns 3.700 agrupaments inicials. És important també assenyalar, que la validació de la tasca mitjançant la comparació d'imatges d'alta resolució, ha permès augmentar notablement la fiabilitat dels resultats obtinguts.

Conclusions i perspectiva d'avanç

En l'avaluació de la representativitat de l'ús de tècniques remotes per a la construcció de models DFN, els resultats suggereixen que les dades capturades i tractades amb tècniques remotes es representen en els models DFN amb resultats molt acceptables. La tècnica remota que més s'aproxima a l'aflorament depèn de la propietat geomètrica que s'ha capturat i representat. La disposició de l'aflorament i les orientacions de les fractures també influencien en la captura de les dades. Al analitzar els resultats de la porositat i permeabilitat de la xarxa de fractures s'obté valors similars en els tres models DFN. En conseqüència les diferències geomètriques representades per cada tècnica en conjunt, no han condicionat en gran mesura l'estimació d'aquests valors, mostrant models de fractures equivalents en el cas de la Fou. Els models DFN que actualment s'obtenen mitjançant sensors remots, permeten mesurar i reduir incerteses en la caracterització dels models regionals tridimensionals. Això serà objecte de noves investigacions tant acadèmiques com industrials.

Finalment, els treballs duts a terme a Montserrat demostren que l'aplicació de tècniques remotes en el camp de la digitalització i la identificació de diferències, durant el seguiment d'afloraments,

és precisa i d'alta resolució. Aquesta nova metodologia que integra d'algorismes de *machine learning* per a la identificació i classificació de blocs caiguts demostra, especialment a Degotalls-E, que l'aplicació d'aquest sistema genera resultats acceptables en termes de Vertader Positiu (VP), Fals Positiu (FP) i Fals Negatiu (FN). Aquesta automatització, permet apropar-nos un pas més cap a implementar un sistema d'alerta de caiguda de blocs, amb un dispositiu TLS integrat, que permeti i faciliti la detecció de deformacions prèvies a caigudes, comportant una aplicació de la tècnica utilitzada durant el desenvolupament d'aquesta tesi al món industrial.

En la actual sociedad en transición digital, donde se generan millones de datos cada día, el uso de técnicas remotas digitales y los procesados con técnicas de *machine learning* es cada vez más presente y necesario, sin que el ámbito de las geociencias sea una excepción. En este contexto, el presente doctorado industrial se centra en el análisis de afloramientos fracturados mediante técnicas remotas y *machine learning*, con dos objetivos específicos; (1) la evaluación de la representatividad de las técnicas remotas para identificar y caracterizar las fracturas y representarlas posteriormente en modelos discretos de redes de fracturas (DFN); y (2) la incorporación de modelos clasificadores con *machine learning* en procesos asociados a entornos fracturados, como son las caídas de bloques.

Metodología de trabajo

El primer caso de estudio, evalúa las técnicas remotas; *Terrestrial Laser Scanner* (TLS) y Fotogrametría (SfM). El experimento se desarrolla en la pedrera activa de La Fou en Barcelona. Los afloramientos fracturados se han reconstruido a partir de un alto número de datos capturados remotamente (Santana et al., 2012; Geyer et al., 2015) y se han agrupado en familias para caracterizar geoméricamente la red de fracturas (García-Sellés et al., 2018). Con estos resultados, y los recogidos con técnicas directas con los *scan-lines*, se construyen modelos tridimensionales (DFN). Se construye un modelo con cada técnica de captura (DFN_{TLS}, DFN_{SfM} y DFN_{scan-line}), donde se representan las principales propiedades geométricas de las fracturas. Por último, se generan los mapas de trazas de fracturas para calcular la porosidad y permeabilidad de la red de fracturas representada.

El segundo caso tiene como objetivo automatizar la clasificación de los desprendimientos de bloques. La zona de estudio se sitúa en los escarpes históricamente inestables de Degotalls, cerca del acceso al Monasterio de Montserrat. La monitorización de dichos escarpes mediante el escaneo periódico con TLS, se inició en el año 2007, como consecuencia de los episodios de caídas de bloques iniciados en el año 2001, que afectaron al aparcamiento del Monasterio (limitando el acceso al Santuario en varias ocasiones). La metodología consiste en alinear dos escaneos o nubes de puntos del mismo afloramiento capturados en dos momentos diferentes, con la intención de detectar las agrupaciones de diferencias entre ambos escaneos. Posteriormente, se revisa visualmente cada agrupación de diferencias para identificar y clasificar las caídas de bloques. La metodología desarrollada en la presente tesis, incorpora algoritmos de *machine learning*, para automatizar esta última etapa de identificación y clasificación de bloques caídos.

Resultados

En el primer caso de estudio, a partir del análisis de los datos de las técnicas remotas (TLS y fotogrametría) y los scan-lines se obtienen las principales propiedades geométricas de cada fractura capturada y procesada; posición, orientación, altura y longitud. Éstas pueden utilizarse para caracterizar la estadísticas de las fracturas como el espaciado, intensidad, P10, P20, P21 y P32, entre otras. Del conjunto de fracturas se identifica la Familia I WNW-ESE, la Familia II NNW/SSE, la Familia III, N-S y la Familia IV NE-SW. La técnica con mayor número de fracturas caracterizadas ha sido el TLS con más de 52.000, después la técnica de fotogrametría con 1.094 y finalmente los *scan-lines* con 398 fracturas medidas. La altura y longitud de las fracturas son muy variables, los valores medios oscilan entre los pocos centímetros a los 18m. Las alturas y longitudes con el TLS son inferiores al metro. Los espaciados entre las fracturas de la Familia I presentan tamaños centimétricos y el resto de familias valores medios más variables desde centimétrica hasta 6m. La fotogrametría obtiene valores que en ocasiones duplican o triplican el resto de técnicas. Además con la introducción de las oberturas de las fracturas en el modelo, se han generado mapas de trazas y se obtienen valor de porosidad y permeabilidad de la red de fracturas. En los tres modelos DFN, el valor medio de porosidad obtenido está entre 5,4 y 5,7% y los valores de permeabilidad media entre 4×10^5 mD y $6,2 \times 10^5$ mD.

Por otro lado, en el segundo caso de estudio, los resultados generados a partir de la combinación de 11 algoritmos clasificadores y 15 métodos de remuestreo, permiten identificar la mejor combinación predictiva para la clasificación de bloques caídos, según el período y el tipo de escarpado. En Degotalls-E. A partir de aproximadamente 5.800 agrupaciones de diferencias identificadas por cada monitorización, se alcanza una reducción del 98% del número de agrupaciones a validar, entre Falsos Positivos (FP) y Verdaderos Positivos (VP), entendiendo (VP) como las caídas reales de bloques. En Degotalls-N la reducción sería del 80,16% para una identificación completa, mientras que para identificar el 96% de las caídas de bloques reales, la reducción sería del 90%, a partir de una población de unos 3.700 agrupaciones iniciales. Es importante también señalar que la validación de la tarea mediante la comparación de imágenes de alta resolución ha permitido aumentar notablemente la fiabilidad de los resultados obtenidos.

Conclusiones y perspectiva de avance

En la evaluación de la representatividad del uso de técnicas remotas para la construcción de modelos DFN, los resultados sugieren que los datos capturados y tratados con dichas técnica se representan en los modelos DFN con resultados muy aceptables. Las técnica remota que más se aproxima al aforamiento depende de la propiedad geométrica que hay que capturar y representar. La disposición del aforamiento y las orientaciones de las fracturas también influyen en la captura de datos. Al analizar los resultados de la porosidad y permeabilidad de la red de fracturas, se obtienen valores similares en los tres modelos DFN. En consecuencia las diferencias geométricas representadas por cada técnica en su conjunto, no han condicionado en gran medida la estimación de estos valores, mostrando modelos de fracturas equivalentes en el caso de la Fou. Los modelos DFN, que actualmente se obtienen mediante sensores remotos,

permiten medir y reducir las incertidumbres generadas en la caracterización de los modelos regionales tridimensionales. Esto será objeto de nuevas investigaciones tanto académicas como industriales.

Por último, señalar que los trabajos realizados en Montserrat demuestran que la aplicación de técnicas remotas en el campo de la digitalización y la identificación de diferencias durante la monitorización de afloramientos es precisa y de alta resolución. Esta nueva metodología que integra algoritmos de *machine learning* para la identificación y clasificación de bloques caídos demuestra, especialmente en Degotalls-E, que aplicando este sistema se generan resultados aceptables en términos de Verdadero Positivo (VP), Falso Positivo (FP) y Falso Negativo (FN). Esta automatización permitirá dar un paso más hacia la implementación de un sistema de alerta de caída de bloques, con dispositivo TLS integrado, que permita y facilite la detección de precursores, comportando una aplicación de la técnica utilizada durante el desarrollo de esta tesis en el mundo industrial.

ABSTRACT

The ongoing digital transition results in millions of data points gathered every day. Within the context of geosciences, the use of digital remote sensing techniques is increasing. Machine learning techniques are interesting alternative to support the processing and interpretation of large amounts of collected data. The present industrial PhD focuses on the analysis of fractured outcrops using remote sensing techniques and machine learning with two specific objectives; (1) the evaluation of how representative are remote sensing techniques to identify fractures and subsequently represent them in discrete fracture network (DFN) models; and (2) the use of classification machine learning models in processes associated with fractured environments, such as block falls.

Methodology

The first case study, evaluates remote techniques for data gathering (Terrestrial Laser Scanner (TLS) and Photogrammetry (SfM)). The experiment was carried out in active La Fou quarry in Barcelona. Fractures surfaces were reconstructed from a large number of remotely captured data (Santana et al., 2012; Geyer et al., 2015), the fractures network is characterized by multiple families (García-Sellés et al., 2018). With these results, and those collected with manual scanline techniques, three-dimensional (DFN) models are constructed. One model is built using data from each capture technique (DFN_{TLS}, DFN_{SfM} and DFN_{scan-line}), where the main geometric properties of the fractures will be represented. Finally, the fracture trace maps are generated to calculate the porosity and permeability using aperture data for each of the scenarios.

The second case study is aimed at automating the classification of block falls. The study area is located on the historically unstable Degotalls escarpments, close to the access road of the Monastery of Montserrat. Monitoring of these escarpments using periodic TLS scanning began in 2007, as a result of multiple falling block episodes starting back in 2001, which affected the Monastery car park (limiting access to the Sanctuary on several occasions). The methodology consists of aligning two scans or point clouds of the same outcrop, captured at two different times, aiming to detect clusters of differences between the scans. Subsequently, each cluster of differences is visually reviewed to identify and classify block falls. The methodology developed in this Thesis, incorporates machine learning algorithms to automate this last stage of identification and classification of fallen blocks.

Results

In the first case study, the main geometric properties for each fracture captured and processed are position, orientation, height, and length. These can be used to characterize fracture statistics

such as spacing, intensity, P10, P20, P21 and P32, among others. For individual families are identified Family I WNW-ESE, Family II NNW/SSE, Family III, N-S and Family IV NE-SW. The technique with the highest number of characterised fractures was TLS with more than 52,000, followed by photogrammetry with 1,094 and finally scan-lines with 398 fractures measured. Fracture height and length are very variable, average values range from a few centimetres to 18m. TLS average lengths are less than one metre. The spacings between fractures in Family I have centimetric sizes and the rest of the families have more variable spacing from centimetric to 6m. Photogrammetry obtains values that sometimes double or triple the rest of the techniques. Furthermore, with the input of the fracture openings into the model, trace maps have been generated and porosity and permeability value of the fracture network has been obtained. In the three DFN models, the average porosity value obtained is between 5.4-5.7% and the average permeability values between 4×10^5 mD and 6.2×10^5 mD.

In the second case study, the results generated from the combination of 11 classifier algorithms and 15 resampling methods, allow us to identify the best predictive combination for the classification of fallen blocks, according to the temporal period and steepness. In Degotalls-E, from approximately 5,800 clusters identified by each monitoring, a 98% reduction is achieved in the number of clusters to be validated, between False Positives (FP) and True Positives (VP), understanding (VP) as the real block falls. In Degotalls-N the reduction would be 80.16% for a complete identification, while for identifying 96% of the real block drops, the reduction would be 90%, from a population of about 3,700 initial clusters. It is also important to note that the validation of the task by comparing high-resolution images has significantly increased the reliability of the results obtained.

Conclusions and the way forward

In the evaluation of the representativeness of using remote sensing techniques for the construction of DFN models, the results suggest that data captured and processed with remote techniques are represented in DFN models with very acceptable results. The remote sensing technique that more accurately approximates the outcrop depends on the geometric property that need to be captured and represented. The layout of the outcrop and fracture orientations also influence data capture. When analysing the results for permeability and porosity of the fracture network, similar values are obtained for all three DFN models. Consequently, the geometrical differences of the models by each technique do not greatly influence the estimation of these values, showing equivalent fracture models in the case of the Fou. The DFN models, which are currently obtained using remote sensors allow the uncertainties generated in the three-dimensional regional models to be characterized measured and reduced. This will be the object of further research both academic and industrial.

Finally, the work carried out at Montserrat shows that the application of remote techniques in the field of digitisation and the identification of differences during the monitoring of outcrops is precise and of high resolution. This new methodology integrating machine learning algorithms for the

Afloraments fracturats digitalitzats.

Avaluació de les tècniques remotes en models *DFN* i aplicació de *Machine Learning*.

identification and classification of fallen blocks demonstrates, especially in Degotalls-E, that applying this system generates acceptable results in terms of True Positive (VP), False Positive (FP) and False Negative (FN). This automation will eventually allow to implement a block drop alert system, with an integrated TLS device, which allows and facilitates the detection of precursors, involving an application of the technique used during the development of this thesis in the industrial world.

ÍNDEX DE LA MEMÒRIA

1. INTRODUCCIÓ	11
1.1. Motivació	13
1.2. Descripció dels objectius.....	16
1.3. Antecedents.....	18
1.4. Estructura de la tesi doctoral	24
2. MARC TEÒRIC	27
2.1. Introducció	29
2.2. Medi fracturat: discontinuïtats mecàniques i medi rocós.....	31
2.2.1. Les discontinuïtats mecàniques.....	34
2.2.2. Propietats individuals de les fractures.....	35
2.2.3. Propietats col·lectives de les fractures	38
2.3. Tècniques de captura de dades en afloraments fracturats.....	45
2.3.1. Mètodes directes.....	46
2.3.2. Mètodes remots o indirectes.....	47
2.3.3. Altres mètodes de captura: mètodes geofísics	50
2.4. Aplicacions de les tècniques remotes en l'anàlisi del medi fracturat.....	53
2.4.1. Models de xarxes de fractures	54
2.4.2. Monitorització de caiguda de blocs en medis fracturats	58
2.5. Machine learning en l'anàlisi del medi rocós. Conceptes bàsics, tècniques de remostreig i models o algorismes de classificació	63
2.5.1. Conceptes bàsics	64
2.5.2. Descripció de les tècniques de remostreig o balanceig considerades en aquest estudi	66
2.5.3. Descripció dels models o algorismes de classificació supervisada considerats en aquest estudi.....	68

3. METODOLOGIES I RESULTATS.....	71
3.1. Introducció	73
3.2. AVALUACIÓ DE LES TÈCNIQUES REMOTES EN MODELS DFN APLICATS A LA SIMULACIÓ DE FLUX	75
3.2.1. Metodologia de treball	77
3.2.2. Aplicació en la Pedrera La Fou a Vallirana.....	78
3.2.3. Adquisició de dades	82
3.2.4. Resultats.....	91
3.2.5. Discussions	105
3.2.6. Conclusions.....	118
3.3. PROPOSTA METODOLÒGICA PER A LA IDENTIFICACIÓ DE BLOCS CAIGUTS AMB NÚVOLS DE PUNTS I <i>MACHINE LEARNING</i>	123
3.3.1. Proposta metodològica	125
3.3.2. Aplicació en els escarpaments de Degotalls-N i E a la Muntanya de Montserrat	136
3.3.3. Adquisició de dades	140
3.3.4. Processat de les dades i resultats	141
3.3.5. Discussions	151
3.3.6. Conclusions.....	155
4. DISCUSSIONS, CONCLUSIONS FINALS I PERSPECTIVA D'AVANÇ ...	159
4.1. Introducció	161
4.2. Resum dels principals resultats.....	161
4.3. Discussió general.....	162
4.4. Conclusions finals	163
4.5. Perspectives d'avanç.....	166
5. BIBLIOGRAFIA.....	169
6. APÈNDIX.....	187

ÍNDEX DE FIGURES

Figura 1.1: A les fotografies s'observen dos afloraments fracturats (A) A l'esquerra, parcialment l'escarpament de Degotalls, al nord-est de la Muntanya de Montserrat. Fotografia presa el maig de 2007, mostra les conseqüències d'una important caiguda de blocs que es va esdevenir el 2 de gener de 2007, catàstrofe que va afectar a les instal·lacions del pàrquing d'accés a Montserrat, sense víctimes mortals. Fotografia realitzada per l'equip de l'I.R. Geomodels. (B) La fotografia de la dreta, mostra un aflorament fracturat, que es correspon al front d'explotació de la pedrera de la Fou de Vallirana, on s'identifica els plans estratigràfics amb un suau basculament i una fracturació verticalitzada molt desenvolupada. Fotografia pròpia.....	14
Figura 1.2: Síntesi dels principals avenços relacionats amb l'aplicació de tècniques remotes i machine learning en l'anàlisi de medis rocosos, on s'hi destaquen fites rellevants des de finals del segle XIX fins l'actualitat. Veure una descripció més detallada en el text.....	22
Figura 2.1: Representació gràfica de les unitats mecàniques. Marcat en color rosa, s'identifiquen els límits de les unitats mecàniques i en color gris, les fractures contingues o no dins de la unitat mecànica. En taronja s'identifica un pla estratigràfic que no coincideix amb els límits de les unitats.	32
Figura 2.2: Esquema gràfic on es representen les principals característiques geomètriques de les fractures identificades en un aflorament de medi fracturat. S'han representats les propietats de; "orientació" representada mitjançant l'angle de màxim pendent respecte la superfície de la fractura, "espaïat" representada amb la distància entre dues fractures de la mateixa família, "alçada" és a dir, la distància de la fractura mesurada perpendicularment a la superfície de referència (pla estratigràfic), "longitud" representada amb la distància de la fractura mesurada en direcció paral·lela a aquesta superfície de referència, "obertura" representada com la distància perpendicular entre dues superfícies que son paral·leles entre elles incloent reompliment mineral o llimós i "famílies", les orientacions preferents de les fractures.....	35
Figura 2.3: Valors angulars per a la definició d'un pla de fractura. D'aquesta manera, la orientació del pla queda definida pel cabussament (angle ψ) i la direcció de cabussament (angle α). Imatge modificada de Manual de Taludes, IGME, 1987.	36
Figura 2.4: Anàlisi de les dades d'orientacions de fractures mitjançant projeccions estereogràfiques. (A) Projecció dels pols dels plans en una projecció estereogràfica considerant una falsilla en graella polar, de equal angle. La projecció es realitza a l'hemisferi inferior de l'esfera. B. Diagrames de densitat que mostren la dispersió i/o agrupació de les dades. S'observa la presència de 4 famílies de plans i les seves orientacions característiques. C. Diagrama de roses que mostren la densitat de pols respecte la direcció de capa (gràfics realitzats amb el programa FracMan)	39
Figura 2.5: Representació de la correcció entre l'espaïat aparent mesurat en un scan-line i l'espaïat real entre dues fractures. S_r = espaïat real, S_a = espaïat aparent mesurat i α = angle entre l'alineació de la cinta i el pla normal a fractura.	40
Figura 2.6: Representació gràfica de diferents distribucions de les fractures que poden quedar representades amb el coeficient de variació (C_v) descrit per Gillespie et al., (1993).....	41
Figura 2.7: Representació gràfica de la relació de les fractures respecte les unitats mecàniques. (A) Sistema de fractures "Strata-bound" i (B) Sistema de fractures "non-strata-bound" adaptat d'Odling et al., (1999). 41	41
Figura 2.8: Esquema resum de les principals tècniques de recollida de dades, utilitzades en la present tesi doctoral, per la captura de dades en afloraments fracturats. Mètodes directes (scan-lines) i remots (Terrestrial Laser Scanner (TLS), fotogrametria).....	45
Figura 2.9: Principis bàsics de l'obtenció d'un núvol de punts mitjançant TLS. A partir de pulsacions làser emeses per el dispositiu TLS, en la direcció definida "d", s'obté la senyal reflectida que computa el temps de viatge bidireccional (t_1 i t_2). Aquest punt es mesura amb l'angle de coordenades horitzontals (φ) i verticals (α). La posició del dispositiu TLS, tindrà una posició georeferenciada (x,y,z) i del reflector pot derivar-se la corresponent coordenada angular (d, φ, α).	49

Figura 2.10: Procés de digitalització de les fractures (en vermell) i els plans d'estratificació (en blanc) a partir de la visualització del núvol de punts en el programa Agisoft Metashape.	55
Figura 2.11: Esquema sintètic. (A) A l'esquerra un model dual i (B) a la dreta un model de xarxa de fractures discretes. Extret del manual de FracMan , 2018.	57
Figura 2.12: Flux de treball de la metodologia estàndard per la identificació de blocs caiguts amb tècniques remotes. El text en verd dels requadres es refereix als programaris que poden realitzar les tasques i entre parèntesis, als algorismes utilitzats. A) Filtratge previ dels núvols de punts; B) alineació dels dos núvols de punts; C) càlcul de diferències entre els núvols; D) agrupament de punts en funció de les diferències i proximitat; E) classificació dels blocs caiguts manualment, descartant els que es corresponen a vegetació o altres artefactes.	59
Figura 2.13: Flux de treball per a la classificació de dades amb machine learning. Es mostren les principals etapes. (A) L'adquisició de dades i el pre-processat d'aquestes, (B) La selecció d'algorismes (model) de classificació, (C) entrenament del model amb un grup de dades d'entrenament i (D) avaluació del model amb la resta de dades del grup. La etapa C i D, es repetirà n vegades, amb diferents configuracions, fins a obtenir la millor configuració per els objectius definits. En aquesta fase també s'aplicaran l'ajust híper-paràmetre, per reduir el processat i és molt recomanable corroborar els resultats amb la validació creuada. (E) la selecció del model de predicció a partir del que obtindrem la classificació.....	64
Figura 3.1: (veure pàgina següent) Esquema gràfic dels principals components de la present tesi, on (A) s'enumeren les característiques geomètriques dels medis fracturats (discontinuitats mecàniques i medi rocós) i (B) s'indiquen les tècniques digitals utilitzats per capturar i processar les dades dels afloraments fracturats i l'ús de machine learning. Finalment, (C) s'inclouen (marcats amb fons gris) dos fluxos de treballs resumits, on s'han utilitzat aquestes tècniques. En el cas de l'esquerra, s'han utilitzat les tècniques remotes per la caracterització de la xarxa de fractures, amb l'objectiu d'avaluar-les a l'hora de construir un model DFN, identificant les avantatges i desavantatges de cada tècnica segons la propietat representada. En el cas de la dreta, hem disposat de les dades de monitoratge d'un escarpament rocós escanejat amb TLS des de 2009, on a partir dels núvols de punts, s'identifiquen els blocs caiguts. S'ha desenvolupat una nova metodologia introduint eines de machine learning, per l'automatització de la classificació de blocs caiguts i precursors, substituint així, un procés d'identificació pràcticament manual. Aquest desenvolupament, està publicat a l'article (Blanco et al., 2022).	72
Figura 3.2: Etapes realitzades durant la investigació per avaluar les tècniques remotes aplicades en la modelització discreta de fractures (DFN) tridimensionalment. A la etapa A i B, es capturen i processen les dades per a ser analitzats. Hi ha tres grups de dades; (1) les recollides sobre els aflorament per caracteritzar la xarxa de fractures (TLS, Fotogrametria i scan-lines), (2) les obtingudes per a caracteritzar el medi rocós, un sistema sedimentari i (3) les complementàriament, geofísiques o sondejos bibliogràfics, per calibrar i corroborar les dades. A la etapa C, es mostren, en tres part, els resultats obtinguts; (C3) els que caracteritzen el sistema sedimentaris que conté les fractures, (C1) els que mostren les propietats de les fractures per a cada aflorament i tècnica i (C2) els que presenten els tres models DFN tridimensional, un per a cada tècnica aplicada. A la etapa D s'avaluen i discuteixen els resultats.....	77
Figura 3.3: Imatge A, B, C, plànols de situació de l'àrea d'estudi, situada al SE del massís del Garraf, a la Pedrera la Fou. Imatge D.; Mapa geològic 1:50.000 (Vega et al., 1973). Font: IGME. Emmarcat en vermell, els límits de la Pedrera (imatge B) i emmarcat en negre l'àrea d'estudi, en materials d'edat Cretàcic inferior.	80
Figura 3.4: Esquema sintètic de l'evolució de la fracturació a la cobertura del Massís del Garraf, extret de Guimerà, (1982).....	81
Figura 3.5: (veure pàgina següent) Situació en planta i en alçat de les captures realitzades a l'àrea d'estudi. En la part superior de la Figura, l'anomenada (A), es mostra el plànol en planta de l'àrea d'estudi, amb la situació dels fronts d'explotació, els perfils tomogràfics i el cub de GPR estudiat. La imatge es correspon a la situació final, després de totes les voladures, on s'observa el límit del frontal-N-4. Cal tenir en compte que les tomografies elèctriques i el cub GPR, es van fer prèviament a les voladures sobre el bancal superior. En la meitat inferior, la (B), es mostra la representació gràfica dels 4 frontals estudiats amb orientació N i el lateral-W, a més de situar els límits recollits en el escaneig amb TLS i la captura amb fotogrametria. També s'inclouen la situació dels scan-lines. Notar que la orientació en el apartat B és diferents que en el apartat A.	82
Figura 3.6: Imatge del Frontal-N-1, amb la posició de les fotografies captades.....	85

- Figura 3.7:** Imatge del Frontal-N-1, amb la posicions de les 3 estacions imatges i el número de imatges per estació. Els diferents colors del núvol de punts correspon a les diferents imatges. 86
- Figura 3.8:** Imatges processades del cub estudiat amb GPR. La imatge B, es correspon amb la superfície situada per sobre dels alforraments estudiats, abans del avanço d'exploració, situant-se entre el Frontal-N-2 i 3. La imatge A i C mostren la fondària en funció del temps de resposta dels polsos d'energia electromagnètica transmesos en el terreny. 89
- Figura 3.9:** Perfils tomogràfics realitzats a l'àrea d'estudi. La longitud dels perfils és de 140 metres i assoleixen una fondària màxima de 32,1m a la part central. El perfil TM-1 se situa pròxim amb el Frontal-N-1 i el perfil TM-2 està més pròxim al Forntal-N-2. 90
- Figura 3.10:** Plànol de situació dels sondejos recopilats a l'entorn de la pedrera la Fou de Vallirana i situació en planta de la columna litològica confeccionada. També es situa el tall geològic realitzat, representatiu de l'àrea d'estudi. 91
- Figura 3.11:** Microfàcies identificades i representatives de l'àrea d'estudi. (A) Fotomicrografia realitzada en llum paral·lela que mostra una textura mudstone. Mostra M-A. (B) Fotomicrografia en llum paral·lela, que mostra una textura wackestone amb fragments de mol·luscs i foraminífers. Mostra M-B. (C) Fotomicrografia en llum paral·lela d'una textura wackestone-packstone d'orbitolínids que també inclou Choffatella decipiens i altres foraminífers bentònics. Mostra M-C (D) Fotomicrografia en llum paral·lela, que mostra una textura packstone dominada per miliòlids i Choffatella decipiens. Mostra M-D. 92
- Figura 3.12:** Imatges del contingut fòssil que ha permès datar la successió estudiada a La Fou de Vallirana. (A) Microfàcies wackestone-packstone de foraminífers bentònics on s'hi observa un espècimen de Choffatella decipiens (fletxa vermella), a més de miliòlids, orbitolínids i Nautiloculina, entre altres components esquelètics. Fotomicrografia en llum paral·lela. Mostra M-B. (B) Wackestone-packstone d'orbitolínids que també inclou altres foraminífers bentònics com Choffatella decipiens o miliòlids, a més de fragments de mol·luscs, equinoderms i espícules d'esponja. Mostra M-C. (C) Exemplar de Choffatella decipiens. Fotomicrografia en llum paral·lela. Mostra M-C. (D) Exemplar de de Palorbitolina lenticularis (fletxa vermella) que mostra la cambra embrionària i permet fer un diagnòstic a nivell d'espècie. Fotomicrografia en llum paral·lela. Mostra M-C. 93
- Figura 3.13:** (A) Tall geològic representatiu de l'àrea d'estudi, de 425m de longitud, on es correlacionen les dades de l'aflorament (Frontal-N-1) amb el sondeig S-1. Sobre la secció s'identifica el sondeig S-1, el frontal-N-1 i la columna litològica projectada, Col-1. Les fractures i els plans estratigràfics són una representació reduïda d'aquestes. Es destaquen 4 plans de discontinuïtat estratigràfics, com els més penetratius, i es numeren. Les unitats estratigràfiques (B) Representació de la columna sintètica lito-estratigràfica realitzada en el Frontal-N-4, representativa de l'àrea d'estudi. S'inclou la textura i la fauna fòssil identificada. 94
- Figura 3.14:** Projeccions estereogràfiques i identificació de les distribucions de les famílies, amb les dades de tots els fronts estudiats a l'àrea d'estudi. (A) TLS (B) Fotogrametria (C) Scan-lines 96
- Figura 3.15:** Projeccions estereogràfiques i identificació de les distribucions de les famílies, amb les dades de tots els fronts estudiats a l'àrea d'estudi. (A) TLS (B) Fotogrametria (C) Scan-lines 96
- Figura 3.16:** (A) Correlació de les UM per a al Frontal-N 1, 2, 3 i el lateral-W. (B) Es projecten les unitats mecàniques identificades en els Frontals-N 1, 2, 3 i Lateral-W, sobre la columna litològica sintètica del quart Frontal-N. (C) Es mostren els gruixos de les unitats mecàniques identificades, la mitja. 97
- Figura 3.17:** Alçada i longitud entre les fractures mesurades (Frontal-N-1), separades per famílies segons cada tècnica. (A) TLS (B) Fotogrametria (C) Scan-lines. 98
- Figura 3.18:** Espaiat entre les fractures mesurades (frontal-N-1), separades per famílies segons cada tècnica. (A) Lidar (B) Fotogrametria (C) Scan-lines 98
- Figura 3.19:** Gràfics del Coeficient de variació de l'espaiat (nivell d'agrupament de fractures (Gillespie et al., 1999)) de tots els afloraments estudiats. (A) Lidar (B) Fotogrametria (C) Scan-lines. 99
- Figura 3.20:** Valors mitjos d'intensitat (P21 i P32), per a cada unitat mecànica i família de fractures al Frontal-N-1. (A) TLS, (B) Fotogrametria i (C) scan-line. 100
- Figura 3.21:** Dades mitges de les obertures obtingudes i agrupades per famílies de fractures mitjançant la tècnica d'scan-line. També es representa la distribució del reompliment descrit, segons la fractura i la família. 101
- Figura 3.22:** Volum modelitzat de 120.000m³, de 80m d'ample, per 60m de fons i 25m d'alçada. 102

Figura 3.23: Visualització de les superfícies de fractures introduïdes per a cada tècnica; TLS, Fotogrametria i scan-line.....	102
Figura 3.24: Projecció de la superfície dels plans de discontinuïtats estratigràfics més penetratius que s'han identificat com superfícies que poden influenciar en el transport de flux a través del massís fracturat. ...	103
Figura 3.25: A la part superior de la Figura, s'observa la representació gràfica del volum estudiat, separats per famílies i per tècniques. Extret del programa FracMan (GOLDER). (D) La Taula presenta els espais mitjos i el P10, per cada model DFN construït.	104
Figura 3.26: (A) Representació gràfica de les traces, (B) gràfica de la distribució de les obertures i (C) gràfica de tendència de la permeabilitat. (D) Dades mitges i desviació estàndard de l'obertura, porositat i permeabilitat a partir del mapa de traces, extretes del programa FracMan, per els models DFN construïts.	105
Figura 3.27: (A) Superposició de la tomografia elèctrica TM-1 amb el Frontal-N-1 (B) Superposició de la tomografia elèctrica TM-2 amb el Frontal-N-2. Amb fletxes vermelles, s'identifiquen les zones humides identificades després de la pluja. S'observen coincidències amb les coloracions groguenques i vermelles de la tomografia. Aquestes coloracions corresponen a les zones amb menor resistivitat.	106
Figura 3.28: Correlació de la porositat i permeabilitat amb les unitats mecàniques identificades amb la intensitat de fractures i les propietats petrogràfiques de la columna estratigràfica aixecada al Frontal-N-4.	108
Figura 3.29: Nombre de fractures identificades i mesurades amb TLS. (A) Frontal-N-1 i (B) Lateral-W..	109
Figura 3.30: Testificació del sondeig bibliogràfic S-1 on es recull el registre d'RQD.....	114
Figura 3.31: Pous virtuals generats en FracMan. De 25 m de fondària i un diàmetre de 200mm mitjançant les diferents tècniques. (A) TLS (B) Fotogrametria i (C) scan-line.....	115
Figura 3.32: Corbes de freqüència acumulativa de les obertures de les fractures per a cada tècnica de captura extrets del programa FracMan. Llegenda: color Blau: Fotogrametria, Vermell: TLS i Verd: Scan-line.	116
Figura 3.33: Corbes de freqüència de la permeabilitat per a les tècniques de captura extrets del programa FracMan. Llegenda: color Blau: Fotogrametria, Vermell: TLS i Verd: Scan-line.	117
Figura 3.34: Esquema del flux de treball del mètode actual per identificar caigudes de roques a partir de dades del núvol de punts i el flux de treball de la nova metodologia proposada en el cas exemple de la Muntanya de Montserrat desenvolupat. S'incorporen la etapa (C2) per mesurar les diferències i agrupar els punts i la etapa (D2), per classificar els blocs caiguts amb machine learning.	125
Figura 3.35: Representació de l'algorisme M3C2. A partir de cada punt del núvol de punts 3D de referència, es calcula un vector normal (VN) amb els punts veïns en un radi de cerca definit per l'usuari (Rs). El vector normal (VN) defineix la direcció en què es busca el punt més proper (PCOM) en el núvol de punts 3D comparat. La distància horitzontal màxima i mínima i les distàncies verticals màximes (MaxHd, MinHd i MaxVd) són paràmetres assignats per l'usuari per definir un doble conus.	128
Figura 3.36: a) Distribució de les diferències entre l'adquisició de dades inicial i la repetició en l'interval de temps mínim (T0-T1), intentant repetir les mateixes condicions per calcular el sistema TLS-vessant. Les mesures se centren en zones perpendiculars i distàncies mitjanes representatives per optimitzar l'error. b) Distribució de les diferències calculades durant el monitoratge en un període de temps determinat. c) Superposició de distribucions per calcular el soroll del sistema (àrea en color blau). Les zones en color vermell marquen els valors assignats als processos d'avanç i els assignats en color blau com disminució en una pèrdua de volum.	130
Figura 3.37: (A) Gràfic de distribució de referència de la presa de dades del dispositiu TLS (veure Figura 3.40). S'observen les diferències de distàncies entre punts, entre el sistema de calibratge i el monitoratge en la presa de dades. (B) Esquema gràfic dels possibles escenaris detectats en un escarpament de roca durant el monitoratge. Es relaciona amb la distribució de diferències recollides en els núvols de punts i la seva interpretació (a la dreta). (1) Soroll del sistema, les distàncies entre punts es troben dins de la corba gaussiana de calibratge. (2) Deformació. Quan els valors de les diferències es classifiquen de manera predominant com a avanços. (3) Caiguda de roques. Quan els valors de les diferències es classifiquen predominantment com a retirades. (4) Vegetació. Quan les classificacions d'avanç i retrocés tenen una distribució aleatòria i superen el LdD, matemàticament similar al predomini del soroll.	131
Figura 3.38: Diagrama de flux del model supervisat d'aprenentatge automàtic per classificar grups de blocs caiguts. La recollida de dades correspon a grups de punts creats anteriorment.....	134

- Figura 3.39:** Diagrama de roses d'azimut respecte la línia de capa de les fractures modelades amb dades TLS (García-Sellés et al., 2018)..... 137
- Figura 3.40:** Fotografia realitzada des del pàrquing de Montserrat, enfrontats amb la vessant de Degotalls. Al fons de la imatge es troba la vessant Est de Degotalls i al davant de la fotografia i a la dreta, veiem la recollida de dades mitjançant Terrestrial Laser Scanner (TLS). A partir d'aquestes dades recollides de forma freqüent i les dades històriques existents, hem pogut desenvolupar eines d'aprenentatge automàtic específiques amb la mateixa línia d'investigació que Zoumpikas, et al (2021). Fotografia pròpia realitzada el maig de 2019. 138
- Figura 3.41:** Exemple de caiguda de roques de gran volum al cingle de Degotalls-N com a conseqüència de fractures i discontinuïtats estratigràfiques. Imatge 2001: Zona degotalls abans de la caiguda de roques (en línies discontinües). L'any 2008, la superfície del penya-segat s'ha estabilitzat després de diferents episodis de desprendiments i va totalitzar un volum de caiguda de roques superior als 1000 m³. 139
- Figura 3.42:** (A) Núvol de punts de la cinglera dels Degotalls N: 2.370.000 punts. Estació 1 (B) Núvol de punts del tram sud de la cinglera dels Degotalls E: 2.860.000 punts. Estació 2, orientació 1 (C) Núvol de punts de la secció nord de la cinglera dels Degotalls E: 2.060.000 punts. Estació 2, orientació 2 (reorientació de l'escàner). Els núvols de punts i les figures corresponen a la intensitat de textura del senyal retornat TLS (1.530 nm). 141
- Figura 3.43:** Exemple a Degotalls-N de filtratge previ del núvol de punts per eliminar la informació innecessària i aconseguir una òptima alineació. (A) Núvol de punts capturat amb TLS amb la intensitat de la senyal retornada. (B) Núvol de punts posterior a l'eliminació de la vegetació més exuberant. 141
- Figura 3.44:** Mapa de les diferències calculades a mode de exemple a Degotalls-N amb el programari PCM. L'histograma mostra com la majoria dels valors estan compresos entre 0.035 i -0.035. Els valors més destacats corresponen a diferències registrades. 143
- Figura 3.45:** Mapa de les orientacions calculades amb el programari PCM com a atribut de cada punt. Les orientacions són mostrades en (A) direcció de màxim pendent o Dip Direction i; (B) Dip format. 143
- Figura 3.46:** Resultats de les "agrupacions de punts" de la comparació 2007-20017 a Degotalls N. (A) Punts que conformen les 3.712 agrupacions de punts. (B) Distribució de l'atribut "Volum" calculat pel programari PCM de les diferents agrupacions de punts en funció de les diferències calculades entre els diferents núvols de punts. (C) "Predomini" i (D) "Diferències"..... 146
- Figura 3.47:** A) La imatge correspon al núvol de punts amb textura d'Intensitat a Degotall E amb l'atribut diferències del clúster #1326 (Monitoratge 2017-2019), les dimensions d'aquest exemple són: 1,4 m (alçada) i 1 m (amplada). B) Imatge del clúster abans del monitoratge. C) Imatge post caiguda de bloc on és visible la falca i la superfície de les fractures que controlen el desprendiment. Aquest desprendiment de blocs està sub-classificat com a gran bloc. 149
- Figura 3.48:** Esdeveniments de caiguda de roques a Degotalls-N (a) i Est (b) color porpra). La línia marró representa el nombre d'esdeveniments registrats amb la metodologia actual en comparació amb la metodologia proposada (color blau i morat). 150
- Figura 3.49:** Relació entre el volum dels agrupaments de caiguda de blocs, les classes de caiguda de blocs i el percentatge de models que prediuen els mateixos agrupaments de caiguda de blocs validats als escarpaments de Degotalls. (A) Les classes de caiguda de roques de Degotalls-N són plaques majoritàries, normalment associades a processos de meteorització amb volums petits, i, blocs grans, a causa del gran desprendiment durant el període 2007-2009. Ambdues classes defineixen un escenari heterogeni que dona lloc a més dificultats en l'etapa d'entrenament dels models predictius. (B) Degotalls-E presenta una classe homogènia amb volums petits que faciliten la identificació de models predictius. 153
- Figura 3.50:** Magnitud-volum als dos escarpaments rocosos de Degotalls. Degotalls-N en color blau i Degotalls-E en color violeta. 154

ÍNDIX DE TAULES

Taula 2.1: Taula resum del Sistema de classificació, en funció de la intensitat de fractures (Sistema Pij) definits per (Dershowitz, 1998).	42
Taula 2.2: Recull de diferents tècniques de remostreig per corregir el desequilibri existent entre grups, agrupades segons si permeten un sotamostreig o sobremostreig.	66
Taula 2.3: Diferents algorismes desenvolupats per obtenir una classificació final, tenint en compte si son de base simple o de base múltiple (veure explicació al text).	68
Taula 3.1: Cronograma de les campanyes de camp realitzades a la Pedrera de La Fou de Vallirana. A la Taula, s'indiquen els trams temporals on queden exposats cadascun dels frontals mitjançant colors; verd (Frontal-N-1), blau (Frontal-N-2), groc (Frontal-N-3 i Lateral-W) i rosa (Frontal-N-4). Els límits responen a les dates de detonació: 27/03/2019, 7/1/2020 i 25/3/2020. Finalment, s'inclouen les tasques realitzades per mesos. Els acrònims utilitzats són els següents: TLS: Terrestrial Laser Scanner, Foto: Fotogrametria, PTM: Perfils de Tomografia elèctrica, GPR: Cub de Ground Penetrating Radar i Columna: Columna lito-estratigràfica.	82
Taula 3.2: Quadre resum de les dades capturades amb tècniques de Fotogrametria (remotes).....	84
Taula 3.3: Quadre resum de les dades capturades amb tècniques de TLS (remotes).....	86
Taula 3.4: Quadre resum de les dades capturades amb tècniques d'scan-line (directes)	87
Taula 3.5: Taula resum amb la relació de mostres seleccionades per a realitzar làmines primes, segons les observacions fetes amb la lupa de mà.....	87
Taula 3.6: Taula resum amb la relació de mostres assajades al laboratori. S'indiquen les mides del cilindre i el volum que ocupa. Per indicar la orientació de la mostra s'utilitzen els símbols "// SO" quan el cilindre de la mostra s'orienta paral·lel al pla estratigràfic i "LSO" quan la orientació és perpendiculars en aquest pla.	88
Taula 3.7: Quadre resum de les dimensions del cub estudiat amb tècniques de GPR (tècnica geofísica).90	
Taula 3.8: Quadre resum dels perfils realitzats amb tomografia elèctrica (tècnica geofísica)	90
Taula 3.9: Resultats dels assajos de porositat i permeabilitat en les mostres assajades. Les mostres amb codi M-Xa, corresponen a les mostres realitzades perpendicularment al pla d'estratificació i les mostres amb codi M-Xb, es corresponen a les mostres realitzades paral·lelament al pla d'estratificació. * by Archimedes' method porosimetry. ** Klinkenberg-corrected gas permeability. En la columna on s'indica el nom de la mostra, la lletra a, representa la mostra tallada perpendicular al pla de referència, en canvi la lletra b, és el contrari, representa la mostra tallada paral·lela en aquest pla.	93
Taula 3.10: Taula resum amb la relació d'afloraments estudiats i les tècniques aplicades.	95
Taula 3.11: Valors d'espaiat mig i intensitat (P_{10}) obtinguts amb el FracMan.....	113
Taula 3.12: Valors d'espaiat mig i intensitat (P_{10}) obtinguts amb el FracMan.....	113
Taula 3.13: Valors RQD extrets d'estudis geotècnics i dels models obtinguts amb el FracMan.	113
Taula 3.14: Valors mitjos obtinguts d'obertura, extretes del programa FracMan, per els models DFN construïts.	116
Taula 3.15: Característiques de la xarxa de fractures, extretes del programa FracMan, per els models DFN construïts.	117
Taula 3.16: Resum dels atributs calculats amb l'adaptació de l'algorisme M3C2 (Lague et al., 2013). Els vectors comentats són referits tant en el núvol de punts de referència com en el comparat.	128
Taula 3.17: Nous quatre atributs calculats amb l'adaptació de l'algorisme DBSCAN. Els resultats s'incorporen a l'esdeveniment i als atributs de l'agrupació de punts.....	129
Taula 3.18: Atributs associats al núvol de punts després de calcular les diferències. * No es calcula com a atributs. ¹ Només s'utilitza un ordinador amb format RGB. ² Només s'utilitza el format Intensity calculat	132

Taula 3.19: Recull de diferents tècniques de remostreig per corregir el desequilibri existent entre grups, agrupades segons si permeten un sotamostreig o sobremostreig.	134
Taula 3.20: Diferents algorismes desenvolupats per obtenir una classificació final, tenint en compte si son de base simple o de base múltiple (veure explicació al text).	135
Taula 3.21: La taula mostra la configuració per una exhaustiva cerca dels millors paràmetres per cada model classificador amb la tècnica del híper-paràmetre. El millor paràmetre està implementat d'acord al criteri de la millor puntuació de la precisió.	135
Taula 3.22: Cronograma indicant en vermell, els mesos on es hi ha registre de dades dels afloraments de Degotalls-E i Degotalls-N, mitjançant <i>Terrestrial Laser Scanner</i> (TLS). Destaca en verd, les visites que s'han realitzat per recollir dades, únicament a l'aflorament de Degotalls-E, durant la tesi doctoral.	140
Taula 3.23: Paràmetres per establir els processos de càlcul de diferències, agrupament de punts: 1) Paràmetres utilitzats per definir la geometria del doble con truncat en la mesura de les diferències; 2) Agrupació de paràmetres amb la nomenclatura equivalent a l'algorisme DBSCAN. La distribució 3) Mitja i desviació estàndard per definir la funció de distribució dels valors de diferència per TLS ILRIS-3D a l'àrea de Degotalls per a distàncies de 175 m.	142
Taula 3.24: Sumari estadístic dels LdD a Degotalls en el període estudiat 2007–2021. El valors de Degotalls E estan calculats amb els 12 monitoratges i a Degotalls N amb els dos monitoratges (2007–2017 i 2017–2021) que s'han realitzat per cobrir el període.	144
Taula 3.25: Funcions associades als agrupaments de punts, ordre de format. * No s'utilitza en la classificació ML. ¹ Només s'utilitza un ordinador amb format RGB. ² Només s'utilitza el format Intensity calculat.	145
Taula 3.26: Resum dels resultats de l'etapa de predicció a Degotalls durant la classificació dels agrupaments de punts. A la taula s'indica el número d'agrupaments de punts utilitzats com a candidats validats en la fase d'entrenament. Els millors model de classificació, així com el de remostreig i les mètriques d'aquests. Degotalls N presenta dos solucions: a) La millor solució amb la premissa de FN =0; i b) la premissa de FN ≠ 0, acceptant un número reduït de FN. * Candidats identificats manualment, al ser inici de procés.	147
Taula 3.27: Resum dels resultats de l'etapa de predicció a Degotalls durant la classificació dels agrupaments de punts. A la taula s'indica el número d'agrupaments de punts utilitzats com a candidats validats en la fase d'entrenament. Els millors model de classificació, així com el de remostreig i les mètriques d'aquests. Degotalls N presenta dos solucions: a) La millor solució amb la premissa de FN =0; i b) la premissa de FN ≠ 0, acceptant un número reduït de FN. * Candidats identificats manualment, al ser inici de procés.	148
Taula 3.28: Resum dels atributs de l'agrupament de punts utilitzat per caracteritzar l'exemple de la Figura 3.51.	149

CAPÍTOL 1

Introducció

- 1.1. Motivació
- 1.2. Descripció dels objectius
- 1.3. Antecedents
- 1.4. Estructura de la tesis



1.1. Motivació

Les discontinuïtats presents en el medi fracturat son un element d'estudi clau en diferents disciplines de les Ciències de la Terra, com ara la hidrogeologia (Singhal i Gupta, 2010), la geofísica (Bertrand et al., 2015), l'enginyeria (Hudson, 2013), o els riscos geològics (Janeras et al., 2017), entre d'altres, així com, son elements geològics a tenir present en diferents àmbits d'aplicació, com per exemple en la construcció (tant a nivell d'edificacions privades com en les grans infraestructures civils), en l'aprofitament o emmagatzematge de recursos naturals continguts en reservoris, o com a magatzem de gasos d'efecte hivernacle, etc. En tots aquests casos i àmbits d'aplicació, la presència d'aquestes discontinuïtats en el medi fracturat, hi juga un paper cabdal, tot i que amb conseqüències que poden ser dispars. Mentre que, 1) per una banda, la seva presència pot fer que millorin certes propietats del medi fracturat (com pot ser l'augment de la porositat o de la seva inter-connectivitat en un reservori) (Naghadehi et al., 2013; Volatili et al., 2019), 2) per l'altra, introdueix inhomogeneïtats que debiliten, fragmenten i degraden el medi fracturat, fet que, per exemple, pot comportar un increment en la inestabilitat i les caigudes de blocs en vessants fracturats (Jaboyedoff et al., 2004), o generar problemes de filtracions i pèrdues en embassaments. És per aquest motiu, que el seu estudi, anàlisi i, sobretot, caracterització, ha estat i serà (com ja va introduir Dershowitz i Einstein, 1988) una tasca bàsica i imprescindible per a molts professionals i investigadors.

Entendre la seva formació i evolució, així com caracteritzar i parametritzar la seva geometria, tipologia, distribució, etc., passa a ser una necessitat i una obligació si el que es pretén és, entre d'altres, controlar i minimitzar riscos o treure el màxim profit o planificar una bona gestió d'aquests afloraments fracturats (i tot el que això comporta). S'ha de tenir present que la major part de les infraestructures existents (i futures), travessen o s'assenten sobre un substrat rocós que pot estar en major o menor grau fracturat (i per tant, condicionat per aquestes), però, a més a més, les necessitats futures de creixement de la població, així com els nous reptes de la societat en termes de, per exemple, transició energètica o de canvi climàtic, plantegen accions i infraestructures cada cop més complexes i arriscades (energia geotèrmica, emmagatzematge de residus o gasos, etc), que requereixen un coneixement cada cop més precís i acurat del substrat rocós sobre el que vivim.

L'estudi dels medis fracturats no és nou i ha anat evolucionant en el temps (ISRM, 1978; Priest i Hudson, 1981; Dershowitz i Einstein, 1988; Peacock i Mann, 2005; Assali et al., 2014; Vazaios

et al., 2017; Volatili et al., 2019; Vlachopoulos et al., 2020). Des de l'ús de tècniques clàssiques manuals realitzades in situ davant dels afloraments (o davant de testimonis continus de sondatges) i que inclouen descripcions minucioses i presa de dades sistemàtiques amb brúixola i cinta mètrica (Wines i Lilly, 2002), fins a l'ús de noves tècniques que aprofiten els avenços en el camp de la informàtica i la tecnologia (Battulwar et al., 2021) per a ampliar notablement tant: 1) la superfície dels afloraments que s'estudia (arribant a llocs inaccessibles de manera directa); 2) el nombre, qualitat i tipus de dades adquirides; i 3) les tècniques d'anàlisi i caracterització utilitzades, que, en definitiva, són més eficients i permeten obtenir resultats cada cop més fiables i precisos en menor temps (veure Figura 1.1).

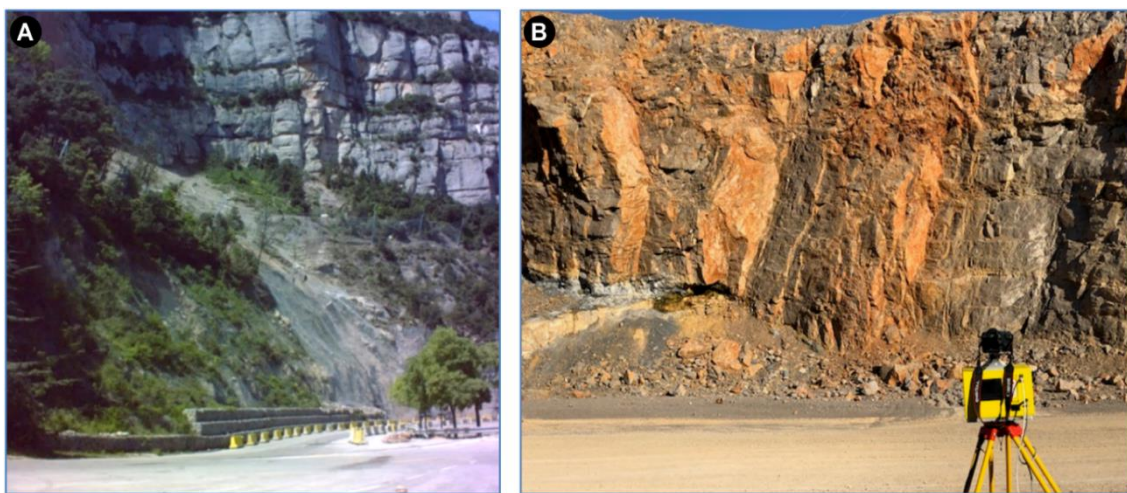


Figura 1.1: A les fotografies s'observen dos afloraments fracturats (A) A l'esquerra, parcialment l'escarpament de Degotalls, al nord-est de la Muntanya de Montserrat. Fotografia presa el maig de 2007, mostra les conseqüències d'una important caiguda de blocs que es va esdevenir el 2 de gener de 2007, catàstrofe que va afectar a les instal·lacions del pàrquing d'accés a Montserrat, sense víctimes mortals. Fotografia realitzada per l'equip de l'I.R. Geomodels. (B) La fotografia de la dreta, mostra un aflorament fracturat, que es correspon al front d'explotació de la pedrera de la Fou de Vallirana, on s'identifica els plans estratigràfics amb un suau basculament i una fracturació verticalitzada molt desenvolupada. Fotografia pròpia.

Malauradament, aquestes tècniques més avançades i innovadores, han estat molt lligades, tradicionalment, a la recerca (i aplicació) duta a terme principalment dins l'àmbit dels hidrocarburs i, per tant, a les grans companyies, deixant, la major part d'empreses "petites" (moltes centrades en el món de la geotècnia i la obra civil), desconnectades dels grans avenços tecnològics i metodològics, i fent ús principalment de les tècniques més tradicionals i assequibles. Precisament, des de la meua experiència i després d'anys treballant com a geòloga en geotècnia i la hidrogeologia, puc afirmar que aquest fet és una realitat, i que les tècniques més utilitzades per caracteritzar un massís rocós fracturat, continuen basant-se principalment en les tècniques directes com l'estudi d'afloraments amb estacions geomecàniques (Bieniawski, 1973), l'estudi dels subsòl amb sondejos geològics-geotècnics (Priest, 1993) i en alguns casos i com a suport, amb tècniques d'estudi geofísiques i assajos de laboratori (Hoek, 1981; Carcedo, 2006). A més a més, en general, l'anàlisi de les dades obtingudes i els resultats obtinguts es representen mitjançant aproximacions 2D com talls i cartografies superficials tot i que, excepcionalment, en

molts pocs casos s'utilitzen models geològics tridimensionals. El grau de fiabilitat i precisió d'aquestes representacions obtingudes i/o dels models reconstruïts, dependrà directament del nombre i qualitat de les dades que s'hagin pogut obtenir a partir de l'estudi dels afloraments existents i, per tant, estan estretament lligats a (i condicionats per) les tècniques emprades per a tal fi, que, com ja s'ha comentat, en la majoria dels casos, es veuen limitades per l'ús de metodologies clàssiques que, o no fan ús de les noves tecnologies (digitals) existents, o no integren el màxim nombre i tipus de dades que s'observen en la totalitat de la superfície dels afloraments.

Per aquest motiu, la principal motivació que m'ha portat a realitzar el present doctorat industrial, ha estat la d'apropar les noves tècniques i metodologies al món industrial, així com la de poder contribuir en la recerca, desenvolupament i millora d'eines informàtiques i/o metodologies que permetin avançar en la identificació, caracterització i comprensió de medis rocosos fracturats així com dels processos geològics associats (caigudes de blocs, inestabilitats, hidrologia, emmagatzematge, etc.), els quals juguen un paper molt important per a la societat. Partint de la hipòtesis, de que la incorporació de noves tècniques remotes i també eines de *machine learning* augmenten exponencialment el coneixement de medis fracturats, però que a la vegada, dificulten el seu tractament i control de les dades, en aquest doctorat industrial, s'avalua l'abast d'aquestes tècniques en el reconeixement de fractures i s'incorporen algorismes de *machine learning* en exemples aplicats.

Cal indicar que, el present projecte de tesi, s'emmarca dins del Pla de Doctorats Industrials, promogut per la Generalitat de Catalunya, en col·laboració amb el sistema universitari i de recerca català i l'empresa privada, amb l'objectiu de transferir la tecnologia i el coneixement més avançat i d'impacte mundial, que el nostre sistema universitari i de recerca ha estat capaç de generar en l'última dècada. Aquest Pla de Doctorats Industrials, es basa en experiències internacionals ja consolidades com, per exemple, les Conventions Industrielles de Formation per la Recherche (CIFRE) de França, l'Industrial PhD Programme de Dinamarca o els European Industrial Doctorates (EID) de la Comissió Europea.

Més concretament, per al present Doctorat Industrial, s'han unit l'empresa Anufra, Water and Soil Consulting amb una llarga experiència en el món de la hidrogeologia i obra civil en relació a, entre d'altres, recursos hídrics així com estabilitat de vessants, i la Universitat de Barcelona, representada pel Grup de Recerca consolidat de Geodinàmica i Anàlisi de Conques (GGAC), que acumula una dilatada experiència, entre d'altres, tant en l'estudi de medis fracturats i els processos geològics que hi ocorren, com en el desenvolupament d'eines digitals d'identificació, caracterització i anàlisi de les discontinuïtats geològiques presents.

1.2. Descripció dels objectius

Per tot l'exposat i tenint en compte, en primer terme, la importància de les discontinuïtats presents en els medis rocosos fracturats i el paper cabdal que hi juguen per a la societat i, en conseqüència, per a les empreses que es dediquen a la geotècnica o hidrogeologia, es defineixen els objectius. El principal objectiu de la present tesi doctoral s'ha centrat en incorporar noves tècniques remotes digitals i eines de machine learning, per estudiar i millorar tant 1) la caracterització i parametrització de les discontinuïtats presents en el medi fracturat (principalment fractures) i dels processos geològics associats, com 2) les tècniques i metodologies emprades per al seu anàlisi, per tal de poder avaluar, optimitzar i/o millorar el seu ús en l'àmbit de l'obra civil i per al control i l'explotació sostenible de recursos naturals.

Més específicament, els esforços s'han centrat en dues línies d'acció preferents (o objectius secundaris) que es deriven de l'objectiu principal, que son:

a) Identificar, caracteritzar, parametritzar i analitzar les discontinuïtats presents al medi rocós fracturat mitjançant diferents tècniques remotes avançades, per tal de valorar la seva aplicació en hidrogeologia de la forma més eficient i vàlida. En aquest sentit, s'han aplicat, sobre els mateixos afloraments fracturats, diferents tècniques remotes, com l'*Structure from Motion* o SfM (Kemény i Post, 2003) i el *Terrestrial Laser Scanner* o TLS (García-Sellés et al., 2011) i el mètode directe i tradicional de l'*scan-line* (Priest i Hudson, 1981, i Priest, 1993). Posteriorment, es podrà fer una avaluació i anàlisi de les tècniques remotes en relació a mètodes directes on es treballa directament sobre l'aflorament. Així, en aquest cas, a part de la pròpia avaluació i anàlisi de les diferents tècniques en un primer aflorament en 2D, es pretén també realitzar un model digital del sistema de fractures en 3D (un *Discrete Fracture Network* o DFN) específic de cada tècnica emprada, valorar la porositat i permeabilitat i connectivitat hidràulica obtinguda mitjançant cadascuna de les metodologies. I estudiar també, la representativitat dels afloraments en relació al medi rocós que representen.

b) Automatitzar la identificació i classificació de la caiguda de blocs en afloraments fracturats, digitalitzats amb *Laser Imaging Detection and Ranging* o LiDAR (TLS), mitjançant tècniques matemàtiques i, concretament, de *machine learning*. Per a tal fi, s'ha implementat una eina informàtica basada en una nova metodologia per processar núvols de punts provinents del monitoratge de vessants rocoses. Gràcies en aquest registre temporal, la comparació dels diferents models digitals obtinguts en temps diferents, mostra un gran nombre de diferències (que generalment és de l'ordre de milers en cada comparació) que requereixen una classificació i

identificació manual, el que es tradueix en un temps de dedicació molt elevat. Aquest tipus d'aproximacions mitjançant noves eines de *machine learning* han de permetre tant automatitzar el propi procés de comparació, com millorar la detecció i l'anàlisi (a posteriori) de les caigudes de blocs d'una manera més ràpida, àgil, objectiva i eficient.

Per a dur a terme aquests dos objectius específics, les diferents tècniques associades s'han aplicat en dos casos reals, per tal de treballar sobre dades reals, mostrar la seva aplicabilitat i donar validesa a les tècniques i metodologies emprades i/o desenvolupades. Per al primer d'ells, centrat en l'avaluació dels mètodes d'adquisició i anàlisi, l'estudi s'ha realitzat en un massís rocós fracturat situat en una pedrera activa (concretament a la pedrera de La Fou, de Vallirana, Catalunya), que ha permès adquirir dades de manera sistemàtica amb cadascuna de les tècniques emprades i en diferents fronts d'avanç de la pedrera. La distribució d'aquests fronts ha permès estudiar el medi fracturat en diferents parts del volum de roca modelitzat i mitjançant dues orientacions representatives perpendiculars entre sí. Per al segon, centrat en automatitzar la identificació i la classificació de la caiguda de blocs, l'estudi s'ha realitzat al massís de Montserrat (Catalunya central), i específicament a la zona dels Degotalls que ha presentat, històricament, inestabilitat marcada i caigudes de blocs recurrents. En aquest cas, l'adquisició de dades durant 15 anys mitjançant tècniques remotes TLS en dos vessants de la zona d'estudi que presenten orientacions perpendiculars entre elles, ha proporcionat un gran volum de dades de caire històric en format digital, ideals tant per al seu anàlisi com per a desenvolupar l'eina de *machine learning* en qüestió. Addicionalment, cal mencionar, que durant el període de la present tesi doctoral, també es van adquirir dades mitjançant tècniques fotogramètriques (Sfm) en diferents fronts d'avanç del túnel que s'estava perforant al Nord de la població d'Organyà (Catalunya), però que degut a les limitacions d'accés a la pròpia obra, així com l'elevat temps de dedicació necessari per al seu anàlisi i estudi, i per tal de centrar els esforços en els dos casos principals descrits, es va optar per no incloure'l dins la present tesi, però pot servir com a exemple per a futurs anàlisis o per desenvolupar-hi noves línies de recerca.

1.3. Antecedents

La primera descripció i classificació de les discontinuïtats presents en un massís rocós fracturat, la va realitzar Woodworth, inicialment de manera breu, el 1895, i de manera més detallada i precisa el 1896. Tot i haver treballs posteriors relacionats (Dale i Gregory, 1911; Sheldon, 1912; Parker, 1942; Raggatt, 1954), no va ser fins el 1961, en que aquesta descripció i classificació no va ser revisada, modificada i millorada per Hodgson (1961a, 1961b) i posteriorment per altres autors (Reches, 1976; Engelder, 1987; Pollard i Aydin, 1988; Hoek, 1964; Hudson i Priest, 1979; Priest i Hudson, 1981; Bahat, 1983; Bahat, 1999).

Aquestes revisions i classificacions posteriors no sempre han seguit el mateix criteri ni han perseguit el mateix objectiu, fet que ha comportat diferents interpretacions o definicions (de vegades contradictòries) en funció del punt de vista utilitzat (geomètric, cinemàtic o geològic) tal com destaca Manda i Horsman (2015) o com mostra amb gran detall Peacock et al. (2016) en el seu treball de síntesi. En aquest sentit, per exemple, no sempre el terme "fractura" s'ha utilitzat o s'utilitza com a sinònim de discontinuïtat per a referir-se a totes les discontinuïtats presents en el medi rocós (Priest i Hudson, 1976, 1981); o el terme "diàclasi" (*joint* en anglès) que presenta certa ambigüitat en la seva definició en funció de, per exemple, si es considera o no el moviment relatiu dels blocs associats (Price, 1966; Hancock, 1985), o de la magnitud d'aquest desplaçament, que en general és petit i que per aquest motiu hi ha autors que parlen de *shear joints* o "diàclasi de cisallament" (Hancock, 1985) però d'altres ho defineixen directament com a "falles petites" (Pollard i Aydin, 1988), o del sentit en que es produeix aquest desplaçament (per tant, si es produeix obertura o no) condició que es considera imprescindible per tal de ser definides com a diàclasi o com a fractures (Pollard i Aydin, 1988). En aquest sentit, son precisament aquests autors (i posteriorment també Peacock et al., (2016) els que suggereixen que "el terme fractura sigui utilitzat de forma genèrica per aquelles estructures on es desconeix el mode de fracturació, evitant així interpretacions especulatives per part del geòleg". Seguint el mateix criteri, en la present tesi doctoral, s'ha utilitzat el terme "Fractura" quan es fa referència a les discontinuïtats en genera.

Per aquest motiu, una bona identificació, definició i classificació de les discontinuïtats presents en el medi rocós, passa per una acurada descripció, caracterització i parametrització de les mateixes, recopilant, per tant, la seva posició, orientació, espaiat, alçada, longitud i obertura (Priest i Hudson, 1981; Priest, 1993). Aquestes dades tant es poden adquirir directament sobre la superfície de l'aflorament rocós (Priest i Hudson, 1981; Sturzenegger i Stead, 2009), com a

partir de la testificació de testimonis continus de sondatges (Priest, 1993) o més endavant amb les tècniques remotes (Wilkinson et al., 2016). Tradicionalment, aquesta adquisició i captura de dades i paràmetres relacionats amb les discontinuïtats, s'ha realitzat mitjançant mètodes manuals, per una banda, amb cinta mètrica i brúixola, amb *Scan-Line sampling* (Pincus, 1951; Priest i Hudson, 1981; Priest, 1993; Zeeb et al., 2013), definint un transsecte lineal (1D) sobre el que s'adquireixen les dades en aflorament i sondatge; o per altra banda, amb *Windows sampling* (Pahl, 1981) o la variant de *Circular sampling* (Watkins et al., 2015), recollint les dades sobre una superfície definida (2D) sobre fotografies aèries o imatges satel·litals (Belayneh et al., 2009; Zeeb et al., 2013). Posteriorment, amb la introducció de les tècniques remotes com en el treball d'Sturzenegger i Stead, (2009), es treballarà sobre la totalitat de la superfície de l'aflorament (*Windows sampling*), i també es realitzaran *Scan-lines sampling* virtuals. També altres tècniques geofísiques com el GPR, permetran complementar informació puntual d'alguns paràmetres de les fractures, com la longitud, estudiant un volum de fractures en (3D) (Theune et al., 2005), amb les pròpies limitacions de ser una tècnica indirecta.

Durant els anys 60 es van definir alguns dels índexs de classificació dels medis fracturats més utilitzats en el món industrial encara avui en dia, com per exemple, l'índex *Rock Quality Designation (RQD)*, definit per Deere, (1969) i, Deere i Deere, (1989) com el percentatge de trossos de testimonis recuperats majors de 10 cm en relació a la longitud total i la recuperació de la mostra d'un sondeig. O anys després, l'índex del *Rock Mass Rating (RMR)* introduït, principalment per Bieniawski, (1973) (Nicholson i Bieniawski, 1990; Hoek i Brown, 1997), l'índex Q de Barton (Barton et al., 1981; Barton, 2002) o variants com l'índex *Slope Mass Rating (SMR)* de Romana (1991) i Romana et al. (2003) o l'índex *Geological Strength Index (GSI)* descrit a (Marinos et al., 2007). Així, aquests índex de classificació, han estat utilitzats principalment per correlacionar diferents paràmetres dels medis rocosos fracturats, com els treballs de Priest i Hudson, (1976); Jiang et al., (2009) o Qureshi, (2014) que van estudiar les relacions existents entre l'índex RQD amb l'espaiat de les diàclasis. D'altres, ho han utilitzat per determinar paràmetres hidràulics, com per exemple El-Naqa, (2001) que relaciona l'índex RQD amb la conductivitat hidràulica a partir d'afloraments o El-Naqa, (2001) o Jiang et al., (2009) que ho fan a partir de testimonis de sondatges. Aquestes correlacions, tot i ser d'utilitat davant d'una clara necessitat de parametritzar i classificar els medis fracturats en l'obra civil, no sempre són suficients i presenten certes limitacions. En aquest sentit, tal com descriu Terzaghi, (1965) i més endavant Berg (2012), diferents orientacions dels sondatges en un massís rocós fracturat pot comportar variacions i errors en la disposició i identificació de les fractures. També en les correlacions entre l'índex RQD i la permeabilitat desenvolupades per Ku et al., (2009) s'identifica limitacions, ja que l'índex RQD, no inclou paràmetres com l'obertura, l'espaiat o l'orientació, podent-se obtenir grans diferències en el càlcul de la permeabilitat.

Relacionat amb la definició d'alguns d'aquests índexs, i degut a l'increment de la necessitat de caracteritzar d'una manera més precisa i acurada el medi fracturat, a finals del segle XX,

s'incrementen el nombre d'estudis que involucren càlculs estadístics en la caracterització i parametrització geomètrica (Baecher et al., 1977; Baecher, 1983; Hudson i Priest, 1979; Dershowitz i Einstein, 1984) i comencen a definir-se les primeres propostes de models de fractures (Dershowitz i Einstein, 1988; Hancock, 1985). Aquesta parametrització cada cop més acurada i precisa, lligada amb els avenços informàtics, també va facilitar el desenvolupament de programes informàtics que permetien generar models numèrics dels medis fracturats, com el FracPAQ (Healy et al., 2006) una caixa d'eines informàtiques de codi obert que permet quantificar els patrons de fractures en 2D, o el FraNEP (Zeeb et al., 2013), software que permet caracteritzar la xarxa de fractures en 2D, a partir d'*scan-lines*, *windows* o *circular sampling* o el DigiFract (Hardebol i Bertotti, 2013), programari dissenyat per adquirir dades de fractures a partir d'afloraments de manera més eficient o altres programes. Els models matemàtics discrets de fractures o DFN (Andersson et al., 2002; Dershowitz et al., 2004; Baghbanan i Jing, 2007; Chesnaux et al., 2009; Rogers et al., 2010), van ser desenvolupats principalment en l'àmbit científic i sobretot gràcies a les inversions realitzades i els projectes engegats per grans empreses, la major part d'elles, relacionades amb la indústria de l'energia (Wilson et al., 2011; Bauer i Tóth, 2017), de la hidrogeologia (Voeckler i Allen, 2012; Jambayev, 2013), però també centrades en d'altres àmbits, com la o la geotèrmia (Doe et al., 2014), tot i que en menor grau. Així, van sorgir nombrosos programes informàtics, molts dels quals van tenir un inici dins del món de la recerca acadèmica, que, a la llarga es van convertir en programes comercials. Alguns d'aquests programes desenvolupats són, per exemple, *FracMan* (Dershowitz i Einstein, 1988), l'*Universal Distincts, Element code* (UDEC) (Board, 1989) o *Comsol* (Romano-Perez i Diaz-Viera, 2012), entre d'altres.

Paral·lelament als avenços informàtics i a l'era digital que s'iniciava cap a finals del segle XX, es van realitzar grans avenços relacionats amb els mètodes de captura de dades d'afloraments rocosos, començant-se a aplicar les tècniques remotes (ja existents, com la fotogrametria o la tècnica *Laser Imaging Detection and Ranging* (LiDAR) que permetien una captura, en format digital, de l'aflorament, fet que obria la possibilitat de desenvolupar i aplicar noves tècniques digitals sobre els models digitals del terreny obtinguts, amb el conseqüent salt qualitatiu i quantitatiu i millora (o complementació) del mètode tradicional de l'*Scan-line*.

En aquest sentit, la tècnica fotogramètrica, utilitzada ja en el segle XIX (Kenefick, 1977; Landen, 1959), durant el segle XX s'utilitza principalment en la generació de mapes topogràfics i cartografies (i posteriorment, la generació d'orto-fotografies) a partir de fotogrametria analògica, tant terrestre com àrea, i analítica. La fotogrametria consisteix en obtenir les coordenades tridimensionals a partir d'un parell de fotografies en 2D, és a dir, des de dues posicions diferents i conegudes (georeferenciades), coneixent els paràmetres interns (deformació de la lent, distància focal o punts principals) de la càmera i mitjançant les equacions de col·linealitat i coplanarietat (Wong, 1980; Wolf, 1983; Atkinson, 1996). És cap a finals del segle XX en que la irrupció de la fotografia digital, va introduir la fotogrametria digital i va expandir la seva aplicabilitat en d'altres camps on abans no s'utilitzava, com en diferents especialitats afins a les ciències de

la Terra (Hagan, 1980; Dueholm i Olsen, 1993; Chandler, 1999; Lane et al., 2000; Reid i Harrison, 2000; Baltsavias et al., 2001). L'evolució digital d'aquesta tècnica ha estat enorme i en els darrers anys han aparegut variants de la fotogrametria com la tècnica *Multi-View Photogrammetry* (MVP) o *Structure from Motion* (SfM), (Ullman, 1979) que es basa en els fonaments de la "visió artificial" (*computer vision*) (Brown, 1976; Kenefick, 1977; Longuet-Higgins, 1981). A partir de l'any 2000, amb la irrupció d'algorismes com l'*Scale-invariant feature transform* (SIFT) (Lowe, 1999) s'incrementa la automatització per facilitar la cerca automàtica de punts homòlegs en els parells fotogràfics en 2D de format digital, i per tant, la generació de núvols de punts i models digitals dels afloraments rocosos Snaveley et al., (2006) i Westoby et al., (2012).

Per altra banda, una altre tècnica remota estesa cada cop més, es basa amb la tecnologia LiDAR, utilitzada tant per dispositius d'escaneig aeris Aerial Laser Scanner o TLA, com fixes terrestres anomenats *Terrestrial Laser Scanner*, TLS. Aquesta, captura a gran velocitat un alt nombre de punts topogràfics (de l'ordre de milers de punts per segon), obtenint la digitalització de la superfície capturada amb alta resolució, a partir d'un núvol de punts.

Cap a finals dels anys 90, s'han desenvolupat metodologies amb la finalitat d'identificar les fractures en medis fracturats a partir de tècniques remotes, com les presentades per diferents autors com (Kemeny i Donovan, 2005; Slob et al., 2005; Bertotti et al., 2007; Jaboyedoff et al., 2008; Olariu et al., 2008; Sturzenegger and Stead 2009; Gigli i Casagli, 2011; Garcia-Selles et al. 2011; Hardebol i Bertotti, 2013; Seers i Hodgetts, 2014; Assali et al., 2014; Zhang et al., 2018). Durant aquests últims anys, s'han publicat diferents articles que en fan un recull i un anàlisi crític (Eitel et al., 2016; Telling et al., 2017; Battulwar et al., 2021), però cap es centra en avaluar aquestes tècniques remotes en la caracterització geomètrica de la xarxa de fractures d'un DFN, identificant les avantatges i desavantatges de cadascuna de les tècniques, incorporant també les dades que obtindriem amb els scan-lines, mètode directe i clàssic de mesures de discontinuïtats, com es presenten en aquesta tesi.

La irrupció d'aquests mètodes de captura digital, també ha obert noves vies d'anàlisi i estudi, sobretot en àmbits relacionats amb l'evolució d'afloraments rocosos, com per exemple, els relacionats amb el risc geològic o en els moviments de vessants, els quals aprofiten aquesta evolució per a observar, estudiar i classificar les diferències en el massís fracturat, observades al llarg del temps, amb l'aplicació d'eines de *machine learning*, sobretot en els últims anys.

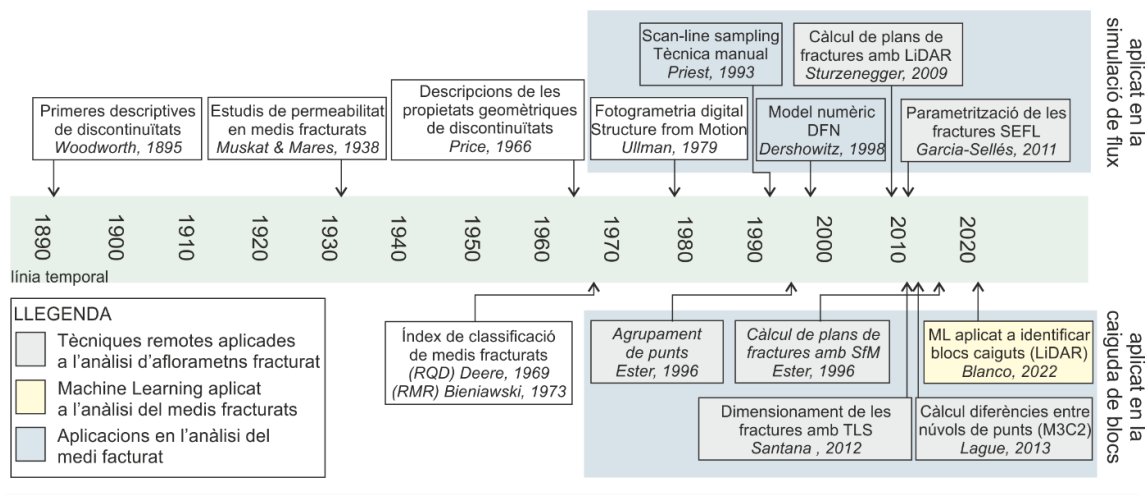


Figura 1.2: Síntesi dels principals avenços relacionats amb l'aplicació de tècniques remotes i machine learning en l'anàlisi de medis rocosos, on s'hi destaquen fites rellevants des de finals del segle XIX fins l'actualitat. Veure una descripció més detallada en el text.

L'aprenentatge automàtic o *machine learning* és un mètode per la construcció de models amb la finalitat de fer prediccions. És una divisió de la intel·ligència Artificial i utilitza algorismes i models de xarxes neuronals per progressivament anar millorant el model. Aquests algorismes construeixen un model matemàtic utilitzant dades de mostreig (dades d'entrenament) per prendre decisions sense estar específicament programat per prendre aquestes decisions. En 1959, Arthur Samuel va proposar el terme *machine learning* dins l'àmbit de la Intel·ligència Artificial i Nilsson, (1965) va demostrar les seves capacitats de classificació. Posteriorment, es van desenvolupar els primers algorismes (Nearest neighbor, Cover i Hart, 1967). La recerca d'Intel·ligència Artificial es va dirigir cap a l'enfocament lògic del coneixement i *machine learning* cap als algorismes, produint-se una divisió. Els algorismes es van desenvolupar amb models estadístics, de lògica difusa i teoria de la probabilitat. Amb el desenvolupament d'internet en la dècada dels 90, ML va evolucionar gestionant la gran quantitat de dades digitals i els algorismes de boosting (Schapire, 1990) per l'aprenentatge. No va ser fins els anys 2000, en que van sorgir els primers estudis que utilitzaven eines de *machine learning* en geociències, com el de (Vatsavai et al., 2005), aplicat en la semi-supervisió per millorar les prediccions de la cobertura de sòls o el de (Tuia et al., 2009) per a la classificació d'imatges. Centrat en estudis d'estabilitat de vessants en medis no rocosos, ja des del 2009, amb estudis com el de (Hwang et al., 2009) s'utilitzen arbres de decisions aplicats a la predicció de moviments de massa o en el 2014 en que (Micheletti et al., 2014) va fer una amb models de boscos aleatoris i Adaboost, aplicats a la identificació de la susceptibilitat d'esllavissades de vessants, augmentant així, la fiabilitat del model de prediccions de l'ocurrència de nous esllavissaments desconeguts, estudis que mostren la importància d'aquestes eines, així com d'altres. D'altres estudis més actuals, continuen utilitzant eines de *machine learning* en aquest context, obtenint bons resultats com (Xiao et al., 2018; Marjanović et al., 2018; Hemalatha et al., 2019; Prakash et al., 2020; Ma et al., 2021). Així, les eines d'aprenentatge automàtic ens proposen una bona alternativa a les feines d'anàlisi manuals sovint tedioses i són una oportunitat de millora tal com descriu (Karpatne et al.,

2019), el qual també mostra les limitacions d'aquestes eines incidint en la necessitat d'equips multidisciplinaris, per obtenir bons resultat. Concretament en l'estudi de vessants de medis rocosos, hi ha treballs que estudien la caiguda de roques mitjançant *machine learning*. Per exemple, s'apliquen diferent models d'arbres de decisió basats en la classificació de les dades de trajectòria, caiguda, velocitat i freqüència alçada de rebot, energia cinètica i ubicació de l'impacte (Fanos et al., 2018, 2020), també alguns treballs més actuals, apliquen *machine learning*, basant-se en la classificació a partir de núvol de punts (Weidner et al., 2019; Kong et al., 2020) seguint la mateixa línia d'investigació que hem fet a la present tesi doctoral, però amb importants diferències en els mètodes aplicats i els objectius de classificació. També seguint amb l'objectiu d'aconseguir metodologies d'automatisme mitjançant *machine learning*, està el treball de (Weidner et al., 2020), on es proposa una metodologia per identificar els tipus de materials de vessants rocoses, com ara vegetació, sòls, roca base i neu mitjançant núvols de punts. La proposta desenvolupada en la present tesi doctoral, presenta una metodologia fins al moment no plantejada, on a partir dels núvols de punts i la combinació de models de classificació, s'identifiquen els blocs caiguts i fins i tot, es podrien arribar a identificar aquests blocs abans de que caiguin, identificant els precursors. Aquesta línia de treball que s'ha dut a terme amb un equip pluridisciplinari presenta els primers resultats en el treball de (Zoumpikas et al., 2021) i seguint en la mateixa línia de treball (Blanco et al., 2022).

La incorporació d'eines de *machine learning* aplicat en geociències, és relativament recent i encara està en desenvolupament tal i com destaca Dramsch, (2020), obrint un camp molt important en el monitoratge i la predicció de la caiguda de blocs.

1.4. Estructura de la tesi doctoral

La present memòria en la que es recullen els principals resultats d'aquest treball, s'ha organitzat en 5 capítols. Els capítols consten d'una introducció, la fonamentació teòrica, la metodologia i desenvolupament de la investigació, i finalment els resultats obtinguts, la discussió i els treballs futurs. A continuació es resumeixen els principals capítols:

Capítol 1: INTRODUCCIÓ. S'argumenten els aspectes que han motivat la tesi doctoral (Secció 1.1 Motivació) tenint en compte que es tracta d'un doctorat industrial. Es realitza una secció de l'estat de l'art (Secció 1.3 Antecedents), introduint al lector en els principals antecedents al voltant del tema principal de la tesi doctoral. L'estructura dels antecedents ens contextualitza en els massissos rocosos i la fracturació (Peacock i Sanderson, 2018), després descriu les tècniques per la identificació de les fractures; les directes (Priest i Hudson, 1976) i les remotes (Sturzenegger i Stead, 2009) i en un tercer bloc, es mostra l'evolució de la modelització dels massissos fracturats (Dershowitz et al., 2004), amb una pinzellada final del que representa introduir eines d'aprenentatge automàtica en les geociències (Dramschi, 2020).

Capítol 2: MARC TEÒRIC. Seguint l'esquema de la secció d'antecedents (Secció 1.3 Antecedents) es descriuen els principals conceptes i l'estat actual pel que respecte a l'anàlisi dels massissos rocosos i els mètodes remots. En un primer bloc (Secció 2.2 Discontinuitats mecàniques i medi rocós), es mostra la literatura descrita *sobre la fracturació i els massissos rocosos*, mostrant les característiques individuals i col·lectives de les fractures i; les característiques del medi o matriu rocosa. En un segon bloc (Secció 2.3 Tècniques de captura de dades) es repassen les *principals tècniques per la identificació i caracterització d'un massís rocós*, incidint en els mètodes remots; LIDAR (David García-Sellés et al., 2018) i Fotogrametria (Blanch et al., 2021) i; l'adquisició de dades. Finalment, es fa un repàs per les principals *metodologies d'anàlisi dels massissos rocosos i la seva fracturació* (secció 0) incidint en modelitzacions com els *Discrete Fracture Network (DFN)* (Chesnaux et al., 2009) o les eines d'aprenentatge automàtic (Zoumpiskas et al., 2021; Blanco et al., 2022).

Capítol 3: METODOLOGIES I RESULTATS. En aquest capítol es presenten els dos casos estudiats; la Secció 3.2, amb la "Avaluació de tècniques remotes en models DFN aplicats a la simulació de flux" i la Secció 3.3, amb la "Proposta metodològica per a la identificació de blocs caiguts amb núvols de punts i *machine learning*". En cada secció s'expliquen les metodologies desenvolupades i els resultats obtingut, identificant les avantatges d'aquestes noves eines.

A la secció d'avaluació de les tècniques remotes i els scan-lines aplicades a la reconstrucció geomètrica d'una xarxa de fractures en un DFN, es presenta el flux de treball utilitzat (Figura 3.6) que permet la caracterització de les fractures a partir de tècniques remotes (García-Sellés et al., 2011) i compara els resultats amb tècniques reconegudes directes com l'*scan-line* (Chesnaux et al., 2009). Amb les dades recollides i partir de la modelització del massís rocós mitjançant un *Discrete Fracture Network* (DFN) (Dershowitz et al., 2004) s'obtenen resultats diferents, segons la tècnica de captura utilitzada.

A la secció on s'apliquen eines de *machine learning* per a la detecció de caigudes de blocs, es desenvolupa una nova metodologia amb algorismes de *machine learning*, per a al identificació de les caigudes de blocs i possibles precursors a partir de TLS. D'aquesta manera, es substitueixen processos manuals i llargs, facilitant la seva aplicabilitat dins de l'àmbit industrial. A la Secció 3.3.4 de Processat de les dades i resultats, es mostren les avantatges d'aquesta nova metodologia, sobre les anteriors utilitzades i obre noves línies d'investigació per al monitoratge més automatitzat (Zoumpekas et al., 2021; Blanco et al., 2022).

Capítol 4: DISCUSSIONS, CONCLUSIONS I PERSPECTIVA D'AVANÇ. Es presenten les conclusions generals dels avanços aconseguits en les investigacions i les propostes que se'n deriven. També es respon als objectius de la tesi doctoral i es mostren els límits de la present investigació i les noves propostes de millora o línies futures de la investigació.

Capítol 5: BIBLIOGRAFIA. Les fonts bibliogràfiques es presenten amb el format de citació estil *American Psychological Association 7th Edition* (APA). En alguns casos, no es disposava de tota la informació que identifica un escrit i s'ha inclòs la màxima disposada.

CAPÍTOL 2

Marc teòric

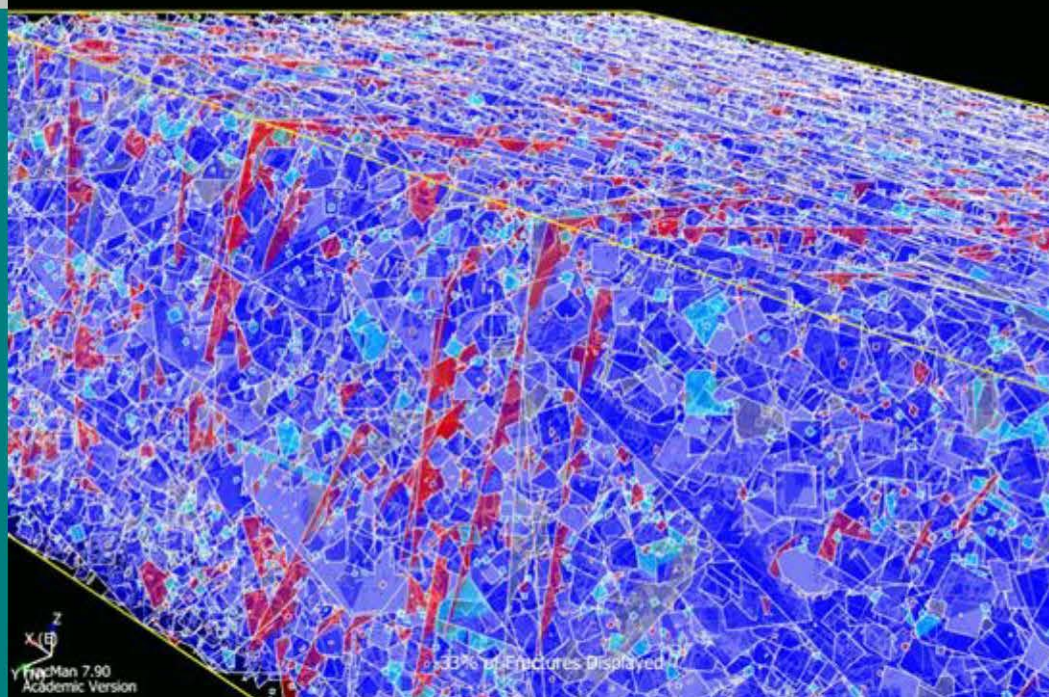
2.1. Introducció

2.2. Medi fracturat

2.3. Tècniques de captura de dades en afloraments fracturats

2.4. Aplicacions de les tècniques remotes en l'anàlisi del medi fracturat

2.5. *Machine Learning* en l'anàlisi del medi rocós



2.1. Introducció

En el present capítol es defineixen els principals conceptes teòrics relacionats tant amb les propietats i característiques que permeten parametritzar, quantificar i identificar les fractures presents en un medi rocós, com els conceptes relacionats amb les principals tècniques existents per a la captura i processat de dades, així com l'anàlisi geomètric de les fractures. Així, s'introdueixen, els anomenats "models de fractures discretes" o "*Discrete Fracture Network*" (*DFN*), com a eina per a l'anàlisi de les xarxes de fractures en medis rocosos. A més a més, es presenten algunes tècniques més noves i actuals que utilitzen aproximacions matemàtiques basades en *machine learning* aplicables a medis fracturats, i que es presenten com una alternativa valuosa als processats manuals de grans volums de dades (*big data*), com les que es generen o deriven de l'aplicació de tècniques d'adquisició remotes.

Específicament, la Secció 2.2, se centra en la descripció de les discontinuïtats mecàniques (principalment les associades a fractures), i es recullen els principals termes i definicions relacionats amb les seves propietats geomètriques i que permeten la seva caracterització, tant des d'un punt de vista individual com col·lectiva. S'introdueixen també, alguns termes relacionats amb el medi rocós que les conté i les caracteritza, com per exemple, la definició de les anomenades *unitats mecàniques*.

A continuació, en la Secció 2.3, es presenten els principals mètodes de captura de dades que permeten caracteritzar tant les fractures com la xarxa de fractures en un aflorament rocós. Per una banda, es descriuen els mètodes directes i clàssics més comuns, com l'*scan-line* tant en aflorament com en sondejos mecànics. La recollida de dades amb aquests mètodes, es realitza directament sobre el terreny i acostumen a ser mètodes puntuals o característics de sectors reduïts i específics de l'àrea estudiada. Per l'altra banda, es descriuen els mètodes remots o indirectes, com el *LiDAR* o la fotogrametria, que permeten una captura digital dels afloraments rocosos en la seva totalitat, generant milions de punts (x,y,z) en 3D representatius dels afloraments, que permetran tant la seva conservació en format digital com el seu processat i anàlisi de forma semi-automàtica. Finalment, també s'han inclòs dos mètodes geofísics indirectes com el *Ground Penetrating Radar* (*GPR*) o la Tomografia Elèctrica, que permeten caracteritzar volums de roca segons algunes de les propietats físiques que les caracteritzen. Aquests mètodes, normalment, poden complementar les observacions realitzades in situ i poden assolir diferents profunditats d'estudi en funció de la tècnica emprada i de les propietats del medi.

En la Secció 2.4, s'introdueixen els conceptes relacionats amb els models matemàtics de *Discrete Fracture Network* (DFN) utilitzats per a representar de manera digital les fractures discretes dins d'un volum de roca definit sobre el que es realitza un anàlisi de la xarxa de fractures en el seu conjunt. Aquests models s'apliquen en la simulació de fluids, l'estabilitat de talussos, l'emmagatzematge d'hidrocarburs o de residus o en d'altres escenaris on les fractures juguen un paper rellevant. Finalment, en aquesta mateixa secció, es descriuen models que utilitzen tècniques matemàtiques de *machine learning* que permeten i faciliten enormement el tractament, anàlisi i classificació de grans quantitats de dades en format digital (*big data*) mitjançant l'aprenentatge semi-automàtic. Aquestes metodologies poden ser aplicables a processos associats als medis fracturats on sigui necessari realitzar una classificació com és el cas de la caiguda de blocs i és en aquest exemple, on s'ha desenvolupat una part de la recerca duta a terme en aquesta tesi doctoral.

2.2. Medi fracturat: discontinuïtats mecàniques i medi rocós

Les propietats de les fractures dependran del medi rocós que les conté. En aquest sentit, la descriptiva geològica de detall de la roca, juntament amb la identificació de les unitats estratigràfiques, les unitats mecàniques i la porositat i permeabilitat de la roca, permetrà identificar i parametritzar el grau d'heterogeneïtat d'aquest medi, clau en els models de creixement de fractures i desenvolupament de la xarxa de fractures (Pollard i Aydin, 1988; Bai i Pollard, 2000; Olson, 2004).

Unitats estratigràfiques i unitats mecàniques

En els medis sedimentaris, s'acostumen a identificar **unitats estratigràfiques** caracteritzades per la litologia, mida de gra i/o origen de les fàcies i diferenciades unes de les altres d'una manera gradual o bé definida per superfícies estratigràfiques.

També, es poden identificar **unitats mecàniques**, que es corresponen als volums amb propietats mecàniques similars (Corbett et al., 1987; Gross, 1993; Underwood et al., 2003; Laubach et al., 2009) que no tenen perquè coincidir amb les unitats estratigràfiques identificades en aflorament. El patró de fracturació dependrà del comportament mecànic del medi rocós sota una pressió diferencial que es produeix a diferents escales (Odling, 1997; Ortega et al., 2006) i en funció de la seva heterogeneïtat. (veure Figura 2.1)

Per identificar sobre el terreny les unitats mecàniques es realitzen perfils amb el martell Schmidt a partir de les mesures successives dels rebots del martell, que es corresponen amb les dureses dels materials estudiats (Miller, 1965; Torabi et al., 2011). Així, la mesura de la duresa permet conèixer un dels principals atributs que determina el comportament mecànic en el medi. Però les propietats mecàniques de les roques, poden evolucionar en el temps degut a processos tectònics o diagenètics (Shackleton et al., 2005; Laubach et al., 2009). D'aquesta manera, les mesures realitzades amb el martell Schmidt a l'aflorament, no tenen perquè coincidir amb les propietats en el moment de la formació de les fractures (Olson et al., 2007, 2009).

Per altra banda el terme Fractura estratigràfica, introduït per Laubach et al. (2009) és utilitzat per identificar unitats mecàniques basant-se en alguna característica de les fractures (extensió o intensitat per exemple). D'aquesta manera, a partir de la identificació d'alguna característica de la xarxa de fractures és possible identificar els límits de les unitats mecàniques. L'avantatge d'identificar les unitats mecàniques amb les fractures estratigràfiques resideix en que és un mètode aplicable amb les tècniques de sensor remots (Bertotti et al., 2007).

En aquest sentit i com alternativa al concepte estrictament estratigràfic, per poder identificar les unitats mecàniques caldrà utilitzar el concepte d'**estratigrafia mecànica**, el qual inclou la intensitat de les fractures. Més concretament i segons defineix Barbier et al., (2012), l'estratigrafia mecànica és el subproducte de la composició i l'estructura deposicionals, associat a canvis químics i mecànics superposats a les roques després de la deposició, motiu per el que les unitats mecàniques resultants poden no coincidir amb les unitats estratigràfiques. El conjunt d'aquestes característiques influeixen considerablement en l'aparició de fractures, les dimensions, la persistència vertical i les relacions d'escala, tal i com ho mostren treballs com el de Gross et al., (1995); Shackleton et al., (2005) o Olson et al., (2007).

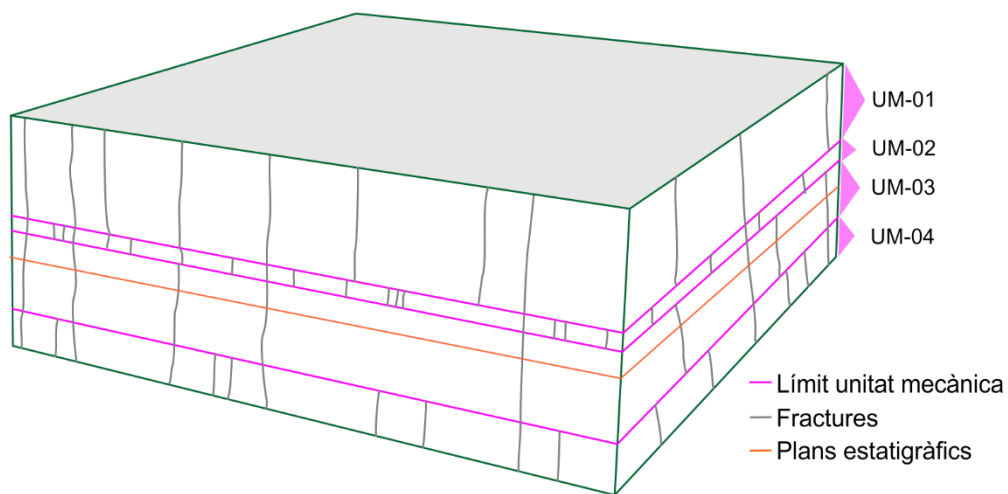


Figura 2.1: Representació gràfica de les unitats mecàniques. Marcat en color rosa, s'identifiquen els límits de les unitats mecàniques i en color gris, les fractures contingues o no dins de la unitat mecànica. En taronja s'identifica un pla estratigràfic que no coincideix amb els límits de les unitats.

Altres treballs mostren evidències de la relació existent entre la distribució i propietats de les unitats mecàniques i la variació de la porositat (Barbier et al., 2012), aplicables a l'estudi de l'evolució de fluids a través del medi rocós. En aquest sentit, la porositat i permeabilitat de la matriu rocosa seran paràmetres difícils de mesurar, que caldrà aproximar també a partir d'assajos de laboratori.

Porositat i permeabilitat de la roca

La porositat al laboratori es pot obtenir a partir del mètode d'*Arquímedes* (Monicard, 1982; Rasolofosaon et al., 2008). El mètode es basa en obtenir la porositat total (porositat connectada) de petits cilindres de roca mostrejada, de 5,5cm de longitud i 4cm de diàmetre, que es prenen distribuïts al llarg de l'aflorament. A laboratori els cilindres s'assequen en un forn a 60°C fins que assoleixen una massa estable (md) i després es posen en posició vertical en un recinte hermètic, durant 24 h i a temperatura ambient, al buit, a 2,6 Pa. La porositat total F (en percentatge %) ve donada aleshores per l'expressió definida a la Equació 2.1.

$$F = \left[\frac{m1-md}{m1-m2} \right] * 100 \quad \text{Equació 2.1}$$

On *F* és la porositat total, expressat en (%)
m1 és la massa de la mostra saturada d'aigua
md és la massa estable a 60°
m2 és la massa de la mostra saturada amb aigua pesada
sota l'aigua (pesatge hidrostàtic subaquàtic).

Per mesurar la permeabilitat es fa mitjançant el mètode del *flux en estat estacionari*, en el qual la permeabilitat s'obté injectant gas sobre mostres cilíndriques de roca (Cavailles et al., 2013). El mètode ens permetrà obtenir una permeabilitat absoluta representativa de la mostra estudiada, de manera que per obtenir un valor representatiu del conjunt de l'aflorament serà necessari recollir varies mostres distribuïdes al llarg de l'aflorament.

Concretament l'assaig consisteix en imposar una diferència de pressió de porus coneguda a través de la mostra i mesurar el cabal resultant que és proporcional a la permeabilitat. La proveta cilíndrica de roca es col·loca en una funda de goma en una cel·la *Hassler* (Mertz, 1991), on el gas utilitzat per a la prova és nitrogen. S'aplica una pressió de confinament (20 bars) a la funda, per unir-la fortament contra la mostra, per evitar qualsevol flux de gas entre l'endoll i la funda. La pressió de confinament i les pressions del fluid de porus a banda i banda de la mostra es controlen mitjançant manòmetres d'alta precisió. El cabal que surt de la mostra es mesura amb un mesurador de tubs de cabal. La permeabilitat absoluta (*K*) es determina mitjançant l'aplicació de la llei de Darcy expressada a l'Equació 2.2.

$$Q = \frac{K}{\eta} S \frac{\Delta P}{L} \quad \text{Equació 2.2}$$

On *Q* és el cabal de gas que surt de la mostra
K és la permeabilitat del solut
S és l'àrea de la secció transversal de l'endoll
ΔP és la diferència de pressió de porus del gas a través de la mostra
h és la viscositat dinàmica de la mostra
L és la longitud de la mostra

El caràcter laminar del flux de gas es verifica mitjançant la determinació del nombre de Reynolds. Finalment, es realitzen unes deu mesures de permeabilitat a diversos gradients de pressió de porus per a cada mostra, per corregir l'efecte *Klinkenberg* mitjançant el mètode gràfic (Rushing et al., 2004).

2.2.1. Les discontinuïtats mecàniques

Les discontinuïtats mecàniques poden definir-se (i produir-se) degut a canvis bruscs en les propietats físiques d'un medi rocós, o a partir d'aquelles estructures que marquen una ruptura en el medi rocós. De discontinuïtats mecàniques, n'hi ha de diferents tipus, però en la present tesi, ens centrarem en les anomenades "fractures" degut a la seva importància tant en les tècniques emprades i/o desenvolupades, com en els exemples estudiats.

Com ja s'ha comentat en el capítol anterior (Secció 1.3 d'Antecedents) el terme "fractura" pot presentar confusió i/o ambigüitat. Per aquest motiu, en la present tesi doctoral, el terme "fractura" s'utilitza segons la definició de Peacock et al., (2016a) on les fractures es defineixen com discontinuïtats mecàniques que es poden identificar i registrar en els afloraments, i que poden formar una xarxa en un medi fracturat, a escala local i/o regional. Així, amb aquest terme genèric, dins de les "fractures" s'hi inclouen les diàclasis, les fractures de cisalla i les falles de petit ordre.

De totes maneres, també s'han considerat importants altres discontinuïtats mecàniques com les *superfícies d'estratificació*, i que estan relacionades amb la sedimentació i els canvis litològics en els estrats associats. Aquestes tipus de discontinuïtats també poden tenir un paper significatiu en la evolució (i modelització posterior) d'un medi fracturat (Singhal i Gupta, 2010), i poden, per exemple, influir tant la circulació de fluids com en la estabilitat de vessants i caiguda de blocs.

Tal i com recomana l'*International Society for Rock Mechanics* (ISRM, 1978), la caracterització geomètrica de les fractures en un medi fracturat, consisteix en registrar un número suficientment representatiu de fractures com per poder descriure, de forma fidel, la seva morfologia i la seva interrelació dins del medi rocós. Amb aquesta finalitat, les característiques geomètriques de les fractures se simplifiquen, de manera que es puguin representar i agrupar segons les seves propietats (Hoek, 1964; Hoek i Bieniawski, 1965).

Per una banda es descriuen les **propietats individuals** de cada fractura, registrant la seva posició, orientació (Fisher, 1985), mida (Priest, 1993) i obertura (Priest i Hudson, 1981); i per l'altra, les seves **propietats col·lectives**, com són les famílies de fractures (Hancock i Kadhi, 1978), l'espaiat (Terzaghi, 1965), l'abundància (Dershowitz et al., 2004) i la connectivitat entre elles (Rossen et al., 2000). (veure Figura 2.2).

Finalment també cal definir i tenir present les principals característiques del medi rocós que conté aquestes fractures; com per exemple, la composició de la roca, heterogeneïtat, i/o les unitats mecàniques que s'hi identifiquen.

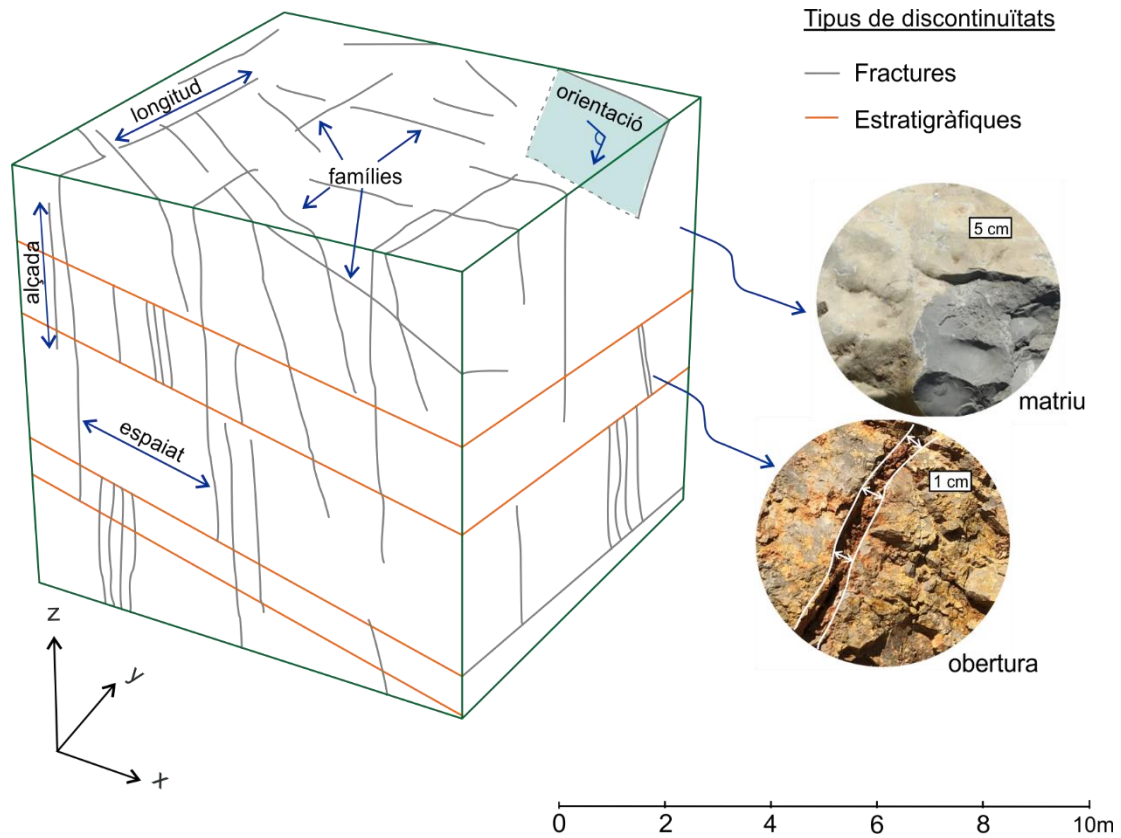


Figura 2.2: Esquema gràfic on es representen les principals característiques geomètriques de les fractures identificades en un aflorament de medi fracturat. S'han representats les propietats de; "orientació" representada mitjançant l'angle de màxim pendent respecte la superfície de la fractura, "espaïat" representada amb la distància entre dues fractures de la mateixa família, "alçada" és a dir, la distància de la fractura mesurada perpendicularment a la superfície de referència (pla estratigràfic), "longitud" representada amb la distància de la fractura mesurada en direcció paral·lela a aquesta superfície de referència, "obertura" representada com la distància perpendicular entre dues superfícies que són paral·leles entre elles incloent reompliment mineral o llimós i "famílies", les orientacions preferents de les fractures.

2.2.2. Propietats individuals de les fractures

Les propietats individuals de les fractures són tots aquells paràmetres que permeten caracteritzar i individualitzar les diferents fractures presents en el medi, a partir de les seves propietats geomètriques i de posició. Així, les principals propietats individuals a mesurar són, entre d'altres, la posició, orientació, mida, i obertura.

Posició de les fractures

Per a poder individualitzar les fractures identificades en un aflorament, el primer paràmetre a considerar és la seva posició o situació en l'espai. Aquesta, es pot definir de diferents maneres, ja sigui utilitzant un sistema de referència global (generalment a partir de dades GPS) o un sistema de referència local (com per exemple, la posició de cada fractura dins d'un *scan-line* el qual està geo-referenciat globalment en el seu punt d'inici i final).

En el cas del mètode tradicional de l'*scan-line*, la posició es mesura entre el punt d'intersecció entre la fractura i la línia d'escaneig, i la distància d'aquest amb el punt inicial de l'*scan-line* (Priest, 1993). En el cas de mètodes més avançats i moderns, com els que empren tècniques remotes (i generació de núvols de punts digitals que representen l'aflorament rocós), la posició de la fractura es pot obtenir a partir de la posició (x,y,z) del centre de massa dels punts del núvol de punts seleccionats com a representatius del pla de fractura "digital" (García-Sellés et al., 2011). Aquest núvol de punts pot estar geo-referenciat de manera global o local, proporcionant una posició exacte de les fractures identificades.

Orientacions de les fractures

Un cop definida la posició espacial d'un pla de fractura, el següent paràmetre important a mesurar, és la seva orientació. Aquesta s'obté a partir de definir un vector que sigui representatiu del pla i mesurar-ne el seus angles característics en relació, tant al Nord magnètic com al pla horitzontal. Així, tot i existir diferents notacions, la emprada en aquesta tesi és la que defineix el vector de màxim pendent (que és aquell vector perpendicular a la línia de capa en direcció descendent i que està contingut en el pla). La projecció d'aquest vector en el pla horitzontal permet mesurar, mitjançant una brúixola, l'angle en sentit horari d'aquest vector en relació tant al Nord (direcció de cabussament, α), com al pla horitzontal (cabussament, ψ) (Figura 2.3). Amb les tècniques remotes la mesura es mesura la orientació sobre les superfícies de fractures reconstruïdes seguint la metodologia descrita a la Secció 2.3.2 tècniques remotes.

Per a l'anàlisi de les dades d'orientació, s'utilitza el vector normal al pla que permet analitzar, mitjançant projeccions estereogràfiques, la dispersió i/o agrupació dels plans considerats en funció de la seva orientació (veure subsecció 2.2.2 Famílies de fractures).

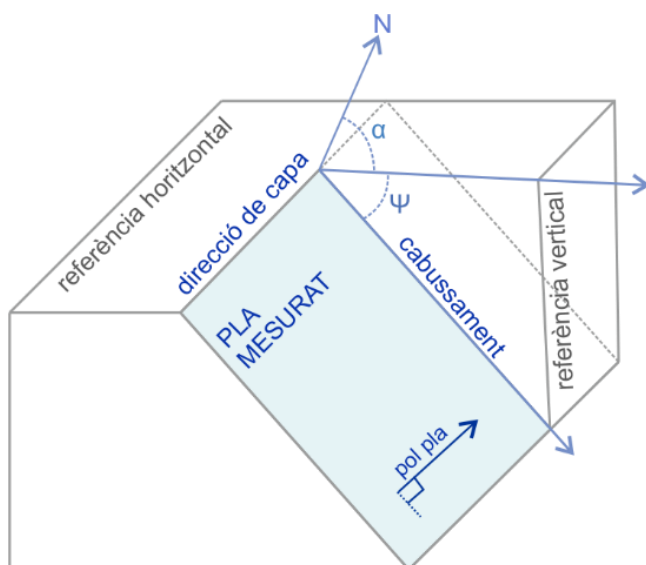


Figura 2.3: Valors angulars per a la definició d'un pla de fractura. D'aquesta manera, la orientació del pla queda definida pel cabussament (angle ψ) i la direcció de cabussament (angle α). Imatge modificada de Manual de Taludes, IGME, 1987.

Mida de les fractures

La mida d'una fractura es pot descompondre en l'alçada i la longitud, tal com descriu Priest, (1993). L'alçada és la distància de la fractura mesurada perpendicularment a la superfície de referència (scan-line) o al pla estratigràfic, i la longitud és la mida de la fractura mesurada en direcció paral·lela a aquesta superfície de referència. Aquest paràmetre és dels més difícils de mesurar amb precisió al camp, ja que en funció de l'orientació o disposició de les fractures, una de les dues direccions a mesurar acostuma a estar, el que s'anomena *censurada*, és a dir, que no es mostra o aflora, o ho fa parcialment. Generalment, la dimensió que es pot obtenir més fàcilment en afloraments verticals és l'alçada de la fractura, però no es podrà mesurar amb precisió la seva longitud, atès que quedarà censurada prolongant-se cap a l'interior del terreny.

La superfície d'una fractura té vores irregulars que, tot i que es poden representar amb polígons complexos, s'acostumen a simplificar en polígons simples rectangulars o altres simplificacions com el cercles o les el·lipses (Baecher et al., 1977). El grau d'allargament es podrà establir en funció de la relació d'aspecte de les fractures i de la seva relació alçada/longitud (*FracMan*®).

Obertura

Una fractura ideal consta de dues superfícies que son paral·leles entre elles, on la **obertura** es correspon a la distància mesurada en direcció perpendicular entre ambdues superfícies. Quan la obertura és > 0 , es genera un espai buit que es pot omplir d'aire, aigua, sediment o mineralitzacions (Priest i Hudson, 1981; Sturzenegger i Stead, 2009). En el cas d'una fractura natural, aquestes superfícies acostumen a no ser llises ni estrictament paral·leles, de manera que la distància entre els plans que defineixen la fractura no és constant (Gale et al., 2017). Aquest fet, dificulta la correcta identificació i mesura de l'obertura, fet que pot comportar un error en la estimació de la porositat secundària del massís estudiat (Snow, 1970).

El terme **obertura cinemàtica** fa referència a l'amplada total de l'obertura de la fractura, incloent reompliment mineral o llimós, de manera que la part oberta de la fractura és generalment menor que la seva obertura cinemàtica (Laubach i Gale, 2006). Per a la caracterització de les obertures es requereix la descripció completa del material que omple l'interior d'una discontinuïtat i també diferenciar entre l'amplada de la fractura plena (cinemàtica) i l'amplada buida de la fractura (Hooker et al., 2013).

Les obertures identificades varien de gruix segons la fractura i també al llarg d'una mateixa fractura (Hooker et al., 2009). En aquest aspecte, és habitual que les obertures siguin majors en superfície i que disminueixin de gruix a mesura que profunditzen en el subsòl. És a dir, en un aflorament rocós podem observar obertures majors, afectades per la descompressió o pels efectes de l'erosió, i a mesura que aprofundim en el subsol i augmenta la pressió, les obertures tendeixen a tancar-se obtenint-se en general obertures més petites. Així, a menor obertura d'una fractura, condicions més favorables d'estabilitat i menys possibilitats del pas del fluid, factors a

tenir en compte quan s'estudien problemes d'estabilitat de talussos o es generen models de flux subterrani (Larsen et al., 2010).

L'obertura d'una fractura es pot mesurar utilitzant diferents tècniques; de forma directa emprant calibradors o comparadors d'espessors (Ortega et al., 2006) o de forma indirecta a través de tests hidràulics (Barker i Black, 1983), traçadors o assajos de laboratori com Indraratna et al., (2003).

La mesura de l'obertura obtinguda a partir d'assajos o de tests in-situ com el *slug* hidràulic test (Barker i Black, 1983) o NAPL (*sunflower oil*), es realitza a l'interior de sondejos instrumentats com a piezòmetres o bé en pous. En aquests casos s'obté el que es coneix com l'**obertura equivalent**, terme que s'introdueix per explicar la variació en quant a la intensitat de la fracturació que es pot estimar mitjançant l'ús de traçadors o tests hidràulics. A més, (Tsang, 1999) introdueix altres termes més específics com "obertura de traçador" i "obertura hidràulica", depenent de la tècnica utilitzada per estimar l'obertura.

L'estudi de les obertures de les fractures tindrà un pes fonamental en la caracterització de la xarxa geomètrica de les fractures, que de forma directa o indirecta, condicionarà en gran mesura els resultats de qualsevol anàlisi del flux a través d'un massís rocós.

2.2.3. Propietats col·lectives de les fractures

A partir dels paràmetres individuals mesurats per a cada fractura, es poden derivar unes propietats col·lectives que caracteritzen el medi rocós estudiat. Aquestes propietats es deriven tant a partir de de l'anàlisi i estudi dels paràmetres individuals, com a partir d'observacions i relacions existents entre les diferents fractures observades en l'aflorament (ja sigui in situ o en afloraments digitals). Com a resultat, es parametritzen i caracteritzen una sèrie de propietats geomètriques que s'agrupen en funció de les orientacions de fractures preferents, o, el que s'anomena, les famílies de fractures identificades. Entre d'altres, les propietats col·lectives més comunes són: el nombre de fractures, el seu espaiat, la persistència vertical dels plans de les fractures, el ràtio i abundància de fractures, i la seva connectivitat.

Nombre de fractures

El nombre de fractures consisteix en sumar el nombre total de fractures que s'identifiquen per unitat de mesura. En general, la unitat de mesura, pot ser un metre lineal, o un metre quadrat, o un metre cúbic, en funció de la tècnica emprada. A part de la importància del propi valor individual, aquest paràmetre és fonamental per conèixer altres paràmetres que se'n deriven, com la densitat o la intensitat de fractures (Dershowitz i Einstein, 1988; Priest, 1993) i que s'expliquen més endavant. La quantitat de dades, serà clau, per obtenir dades representatives en els tractaments estadístics i en aquest sentit les tècniques remotes obtenen amb més facilitat major nombre de dades que les tècniques manuals o directes.

Famílies de fractures

A partir de l'anàlisi de la orientació obtinguda per a les fractures considerades, es poden agrupar aquelles fractures amb orientacions similars en diferents **famílies** o **sets** (Fisher, 1985), ja que és comú el desenvolupament de grups de fractures amb orientacions similars (agrupats en famílies o sets) en un massís rocós sotmès a un camp d'esforços (Hancock i Kadhi, 1978).

Com s'ha comentat en la subsecció anterior (2.2.1), l'anàlisi de les orientacions es realitza a partir de la representació conjunta de totes les orientacions. Aquesta representació es realitza mitjançant una projecció estereogràfica que permet la representació gràfica de diferents elements geomètrics (en general plans i línies) en un entorn comú i en relació a un mateix vector (el Nord). Així, la seva representació permet analitzar la dispersió i/o agrupació de les dades en funció de la seva orientació (Figura 2.4). Per facilitar aquesta projecció i comparació, s'acostuma a treballar amb un únic vector representatiu de la orientació pla de fractura que, en aquest cas, és el vector normal o pol del pla (Priest i Hudson, 1981).

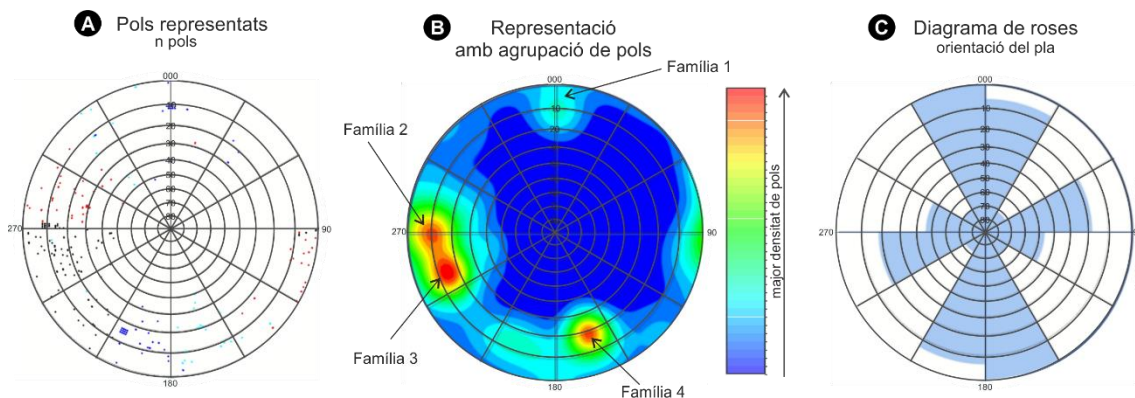


Figura 2.4: Anàlisi de les dades d'orientacions de fractures mitjançant projeccions estereogràfiques. (A) Projecció dels pols dels plans en una projecció estereogràfica considerant una falsilla en graella polar, de equal angle. La projecció es realitza a l'hemisferi inferior de l'esfera. (B) Diagrames de densitat que mostren la dispersió i/o agrupació de les dades. S'observa la presència de 4 famílies de plans i les seves orientacions característiques. (C) Diagrama de roses que mostren la densitat de pols respecte la direcció de capa (gràfics realitzats amb el programa FracMan)

Espaiat entre fractures

La distància entre dues fractures adjacents associades a la mateixa família de fractures, mesurada perpendicularment a la superfície d'una d'elles, s'anomena **espaiat** (Terzaghi, 1965; Narr i Suppe, 1991). En la pràctica, quan es mesura in situ a l'aflorament, aquesta distància sovint és difícil de mesurar en direcció perpendicular a les superfícies de fractura, de manera que el que es mesura és la distància aparent entre les fractures seguint la direcció, per exemple, de l'scan-line utilitzat (veure Figura 2.5). Així, per a poder obtenir la distància real, s'aplica, a posteriori, el que s'anomena la **correcció de Terzaghi** (1965), (Equació 2.3).

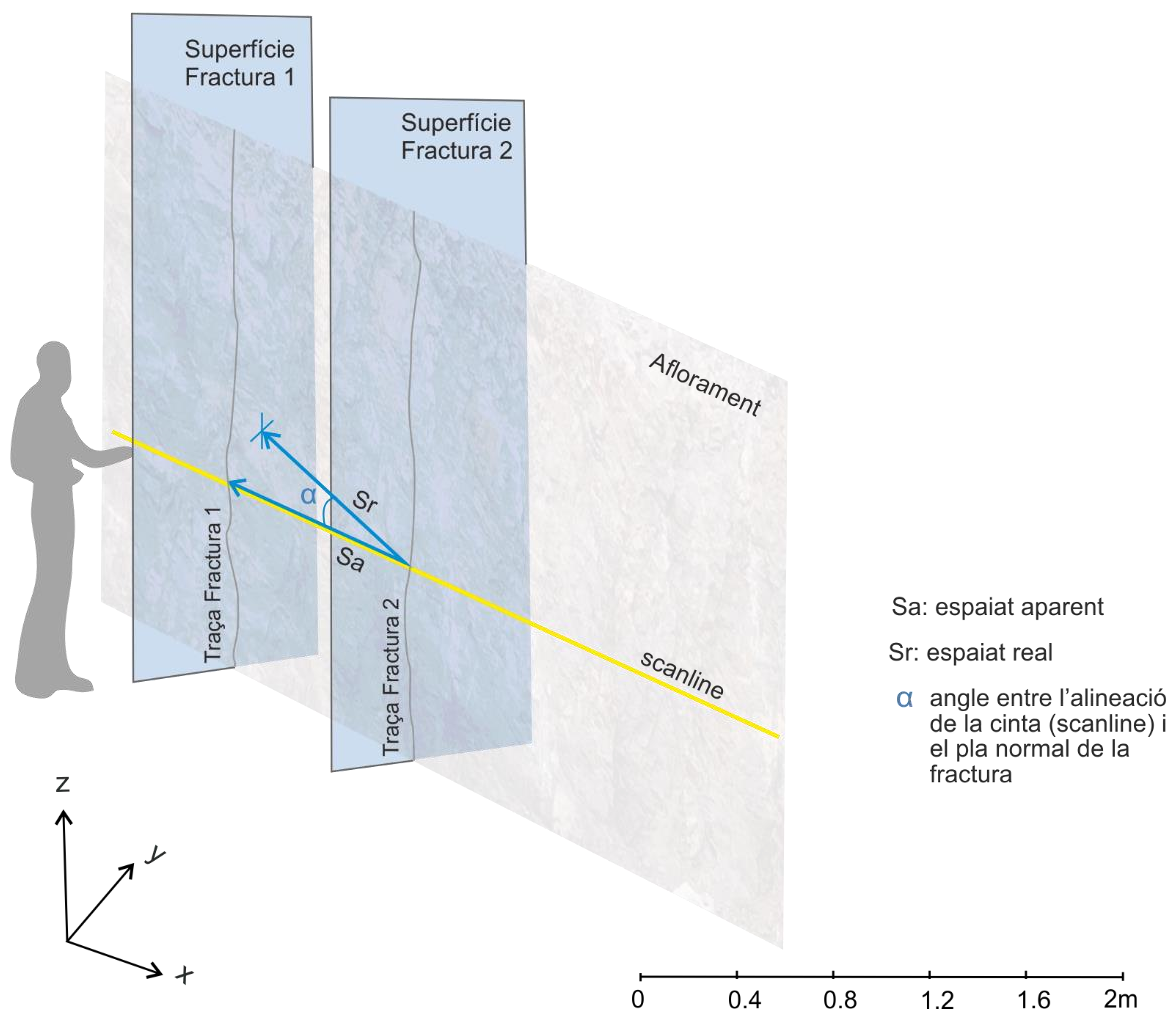


Figura 2.5: Representació de la correcció entre l'espaiat aparent mesurat en un scan-line i l'espaiat real entre dues fractures. S_r = espaiat real, S_a = espaiat aparent mesurat i α = angle entre l'alineació de la cinta i el pla normal a fractura.

$$S_r = S_a * \cos \alpha$$

Equació 2.3

On S_r és l'espaiat aparent entre fractures

S_a és l'espaiat mesurat a camp

α és l'angle que fa línia de l'scan-line i el pla normal a fractura

Els espaiats, poden presentar distribucions irregulars o seguir o no un patró, que es mesurarà a partir del coeficient de variació, descrit per Gillespie et al., (1993). Així, valors del Cv propers a 1 indiquen que la distribució de les fractures (i per tant de l'espaiat) és aleatòria; si supera l'1 s'interpreta que existeix cert grau d'agrupament entre les fractures que es concentren en certes àrees o zones; i si s'apropa a 0 significa que la distribució de les fractures és regular, el que suposa valors de la distància aproximadament constants i amb una baixa desviació estàndard (veure Figura 2.6).

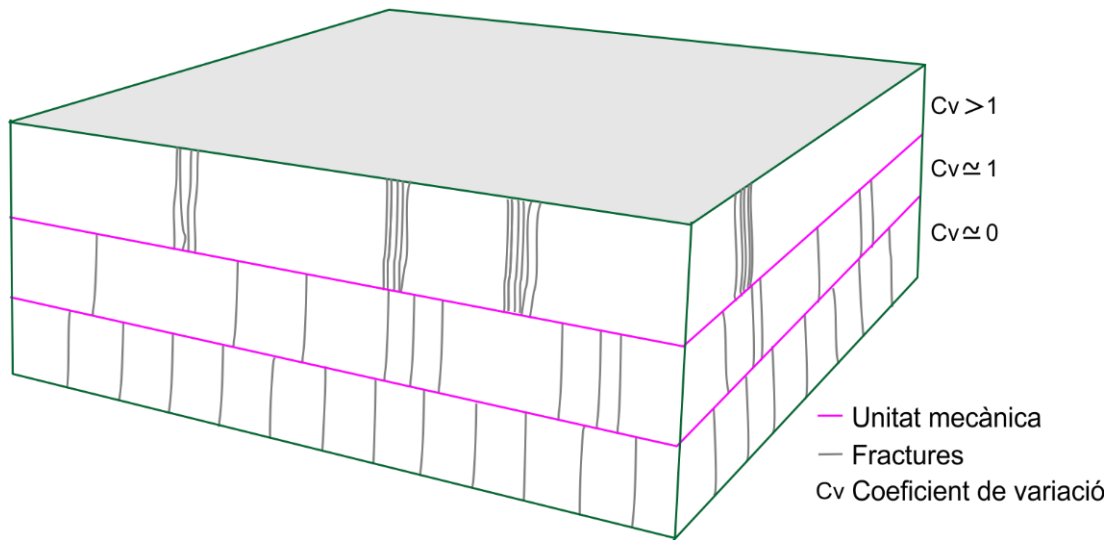


Figura 2.6: Representació gràfica de diferents distribucions de les fractures que poden quedar representades amb el coeficient de variació (C_v) descrit per Gillespie et al., (1993).

Persistència vertical de les fractures

El paràmetre de persistència vertical indica la proporció de fractures que verticalment excedeixen els límits de d'una unitat mecànica, expressat en percentatge (Petit et al., 1994; Bech et al., 2001). D'aquesta manera, les fractures que estan contingudes en els límits de la seva unitat mecànica, es poden definir com un sistema de fractures "strata-bound" i quedaran representats per un valor de persistència baix, en canvi les fractures que sobrepassin el límit vertical de la unitat, es poden anomenar com a sistema de fractures "non-strata-bound" (Odling et al., 1999) (veure Figura 2.7).

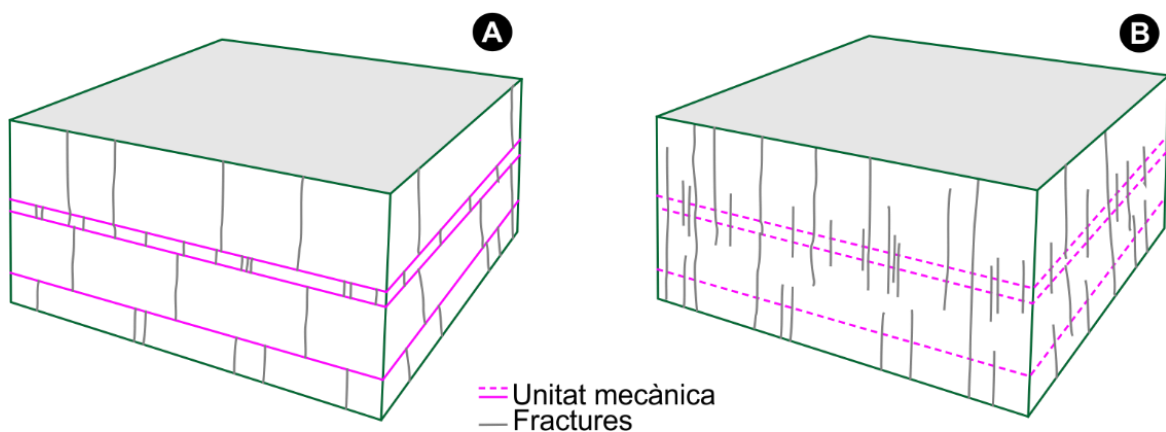


Figura 2.7: Representació gràfica de la relació de les fractures respecte les unitats mecàniques. (A) Sistema de fractures "Strata-bound" i (B) Sistema de fractures "non-strata-bound" adaptat d'Odling et al., (1999).

Relacions espaiat/gruix de la unitat mecànica: (FSR) i (FSI)

Se sap que l'espaiat pot estar condicionat pel gruix de la unitat, per tant, si tenim unitats diferents amb diferents gruixos, els espaiats, tot i tenir valors diferents, poden ser proporcionalment iguals entre unitats. Per tal motiu, es van definir dos índex: *Fracture Spacing Ratio* o FSR i *Fracture Spacing Index* o FSI (Narr i Suppe, 1991), el primer es correspon al ràtio entre el gruix de la unitat mecànica dividit per la mitjana de l'espaiat entre fractures (Gross, 1993). El segon, es calcula mitjançant el pendent de la recta que s'obté a partir de la relació entre el gruix de la unitat mecànica de les fractures i el seu espaiat mig (Narr i Suppe, 1991). Aquests índexs s'utilitzen per saber si el nombre de fractures (per tant, els espaiats) tenen alguna equivalència entre unitats que tenen altres gruixos. Ambdós paràmetres es mesuren quan les fractures es troben contingudes “**strata-bounds**” en les seves unitats mecàniques.

Intensitat de fractures (Sistema Pij)

Es tracta d'un esquema de classificació de la intensitat de fractura proposada per Dershowitz, (1989, 1998) per eliminar algunes de les ambigüitats de la terminologia. En aquest classificació es relacionen les tres dimensions de les fractures amb el medi que ocupen, proporcionant un marc senzill per moure's entre escales i dimensions diferents, conegut com el sistema P_{ij}.

Aquest sistema tal com mostra la Taula 2.1, relaciona una de les tres dimensions del mostreig (seguint una línia, una àrea o un volum) amb la dimensió de la mesura realitzada sobre la fractura (Puntual, Longitudinal, Areal i/o Volumètrica). Les dades de mesura s'expressen amb la nomenclatura P_{ij}, on el primer dígit (i) correspon a la dimensió mesurada de la fractura i el segon dígit (j) a la dimensió de l'exploració.

		Dimensió de la mesura (fractura) (j)			
		0 Puntual	1 Longitudinal	2 Areal	3 Volumètrica
Dimensió del mostreig (i)	1 <i>Mesures Lineals</i>	P10 Número de fractures per unitat de longitud	P11 Longitud de les fractures per unitat de longitud		
	2 <i>Mesures Areal</i>	P20 Número de fractures per unitat d'àrea	P21 Longitud de les fractures per unitat d'àrea	P22 Àrea de fractures per àrea	
	3 <i>Mesures volumètriques</i>	P30 Número de fractures per unitat de volum		P32 Àrea de fractures per unitat de volum	P33 Volum de fractures per unitat de volum
		Densitat	Intensitat	Porositat	

Taula 2.1: Taula resum del Sistema de classificació, en funció de la intensitat de fractures (Sistema P_{ij}) definits per (Dershowitz, 1998).

A continuació s'anomenen i es descriuen, els principals epígrafs utilitzats i que es deriven de la Taula proposada:

(P₁₀) Es tracta d'una mesura unidimensional, que normalment es registra a partir de la posició de les interseccions de les fractures al llarg d'una línia en l'espai, tal com la línia central d'un sondeig o un *scan-line* en afloraments superficials o parets de túnels. Així, aquesta mesura fa referència a la intensitat de les fractures presents al llarg d'una línia, és a dir, el número de fractures per unitat de longitud. (Ogata et al., 2017)

(P₂₁) En aquest cas, es tracta d'una mesura bidimensional al ser emprada, enlloc de al llarg d'un línia, en una àrea de referència. Per tant, indica la intensitat de les fractures presents en una àrea).

(P₃₂) Es tracta d'una mesura tridimensional, que s'utilitza per mesurar la densitat volumètrica de les fractures i s'expressa com la suma de les àrees de les fractures, per unitat de volum. P₃₂ és la mesura d'intensitat més útil per dur a terme descripcions geomètriques, ja que la major part de les tècniques de modelatge en *Discrete Fracture Network* (DFN) descomponen les fractures en polígons simples localitzats en l'espai.

(P₃₃) es tracta d'una mesura tridimensional, entesa com el percentatge de volum de les fractures en un volum mostrejat, equivalent a la porositat de les fractures. El P₃₃ és difícil de mesurar, atès que requereix, no només conèixer la distribució de les superfícies de les fractures i la intensitat de la fractura, sinó també d'informació addicional sobre les obertures de les fractures. L'obertura, és un paràmetre que encara poder-se mesurar en superfície amb mesures directes, no sempre serà representatiu en fondària (veure Subsecció 2.2.1 Obertura).

El sistema de classificació fins aquí descrit és una referència en quant a la reconstrucció de models discrets de fractures, per exemple (Rogers et al., 2014; Boro et al., 2014; Rasmussen et al., 2019).

Connectivitat entre les fractures

El terme connectivitat es refereix en aquest cas a les interseccions entre plans que formen la xarxa de fractures (Rossen et al., 2000). Aquesta connectivitat augmentarà quan les longituds i la densitat de les fractures siguin majors, ja que s'incrementarà la probabilitat de que es connectin entre elles (Robinson, 1983). Aquesta connectivitat de les fractures quan es desenvolupa en formacions geològiques impermeables o de baixa permeabilitat, és una característica crítica que controla el moviment dels fluids (Berkowitz, 2002).

Quan es parla de connectivitat, a més de les fractures (Massiot et al., 2017), també cal considerar les estructures sedimentàries primàries, com per exemple l'estratificació, ja que poden representar plans de discontinuïtat que faciliten la connexió entre tots els plans de discontinuïtat, i, per tant, condicionen i defineixen també el flux de fluid subterrani resultant. Es tracte de plans

que habitualment es disposen paral·lels a l'orientació de la unitat estratigràfica i sovint són perpendiculars a les fractures desenvolupades, de manera que és fàcil que interceptin i connectin la xarxa de fractures.

Permeabilitat del medi rocós

La permeabilitat és la capacitat que té un medi rocós en transmetre l'aigua a través dels seus intersticis (Mook, 2002). Segons Narr i Suppe, 1991, en un medi fracturat, existeixen dues fases de permeabilitat relativa; la que s'associa a les fractures i la que correspon a la matriu.

La **permeabilitat matricial** s'acostuma a mesurar en el laboratori i habitualment, ofereix valors baixos i homogenis. En canvi, la **permeabilitat de les fractures**, associada als sistemes de fractures, és molt més variable i dependrà de les propietats individuals i col·lectives de les fractures i la seva relació amb el medi rocós.

La **permeabilitat** d'un medi fracturat és molt variable i difícil d'estimar. Per determinar-la s'acostumen a realitzar assajos hidràulics dins de pous o sondejos executats; essent els més coneguts l'assaig Lugeon (Öge, 2017) i l'assaig de bombeig (Saleem, 1970), a part d'altres tests hidràulics com l'assaig Slug (Barker i Black, 1983). Els sondejos o pous de diàmetres que acostumen a variar entre 89mm i 400mm de diàmetre, poden arribar a fondàries de centenars de metres. La seva representativitat dependrà de la posició i fondària del tram assajat del sondeig i/o pou en relació al medi i la xarxa de fractures.

L'assaig Lugeon permet calcular semi-quantitativament la permeabilitat dels massissos rocosos, en qualsevol tipus de litologia i estat de fracturació. Aquest tipus d'assaig utilitza aigua injectada a pressió constant dins d'un tram obturat del sondeig mentre es mesura l'admissió d'aigua per determinar el cabal d'entrada en aquell tram obturat. A partir dels cables mesurats i amb les testificacions de les fractures, s'estima semi-quantitativament la permeabilitat. (p.e. Öge, (2017)). Els assaigs de bombeig en pous d'aigua, en canvi, són utilitzats per determinar el coeficient de permeabilitat mitjà del tram d'aquífer estudiat, quan el nivell freàtic es troba per sobre d'aquest tram (Dershowitz i Doe, 1997).

Altres propietats

Existeixen més atributs per descriure les característiques de les fractures, com per exemple la rugositat de la superfície de les fractures o les terminacions de les fractures. Aquests paràmetres no s'han estudiat en aquest tesi doctoral.

2.3. Tècniques de captura de dades en afloraments fracturats

Tots aquests paràmetres característics introduïts en els apartats anteriors i relatius a les fractures, es poden capturar mitjançant diferents mètodes (veure Figura 2.8). Aquests es poden agrupar en mètodes **directes**, a partir de la observació directa i in situ de l'aflorament o sondeig (Subsecció 2.3.1), o **indirectes**, a partir de la observació i anàlisi de dades derivades dels afloraments o dels volum de roca a estudiar mitjançant mètodes remots (Subsecció 2.3.2) o geofísics (Subsecció 2.3.3).

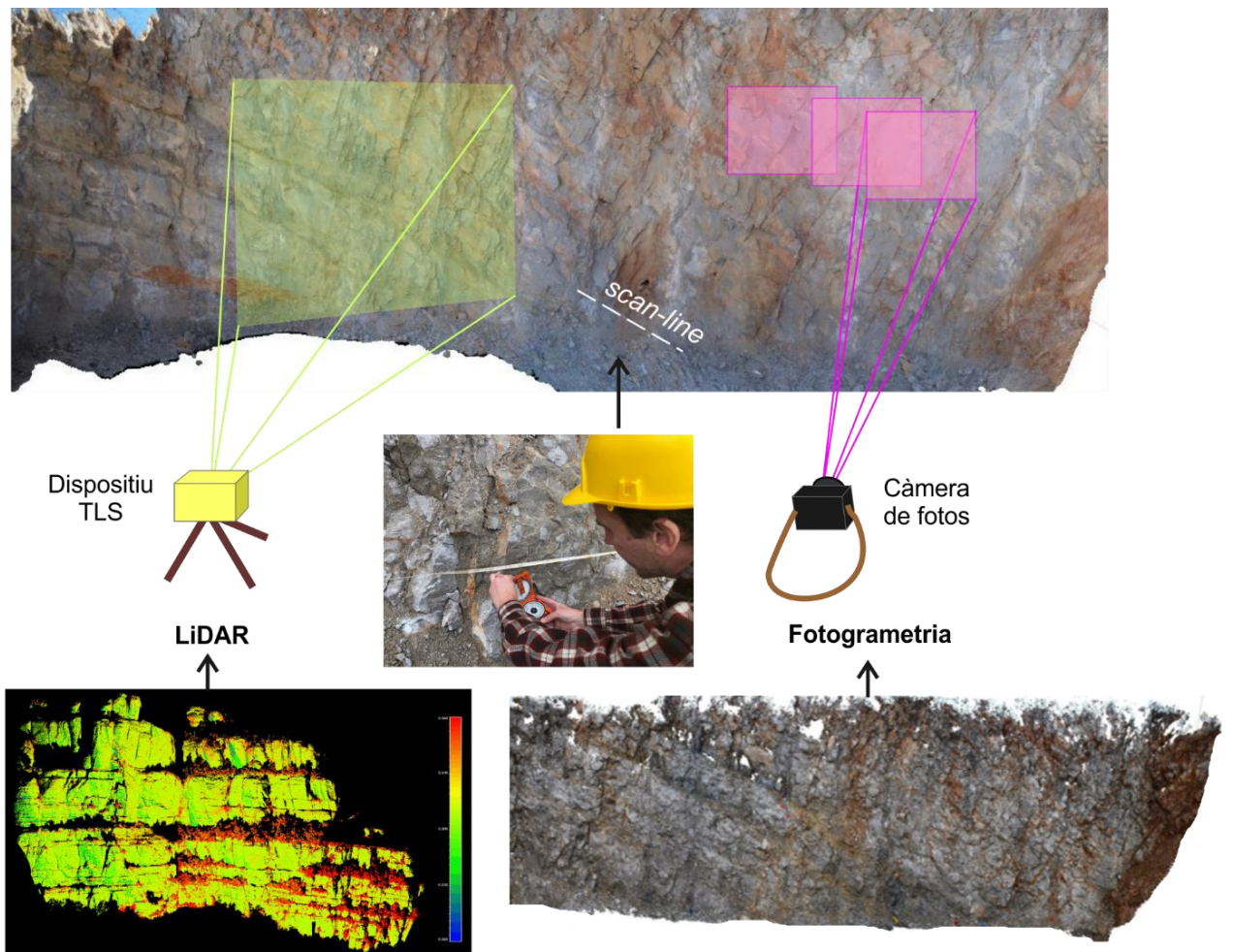


Figura 2.8: Esquema resum de les principals tècniques de recollida de dades, utilitzades en la present tesi doctoral, per la captura de dades en afloraments fracturats. Mètodes directes (scan-lines) i remots (Terrestrial Laser Scanner (TLS), fotogrametria).

2.3.1. Mètodes directes

Els mètodes directes o manuals són aquells en els quals l'observador es posa en contacte directe amb la fractura sobre el terreny. La identificació de la fractura es fa in-situ, de manera que les textures, obertures, materials de rebliment i demés paràmetres característics de la fractura es determinen a simple vista, de la mateixa manera que la presa de les mesures, que també es duu a terme directament sobre la discontinuïtat.

Quan s'empren aquests tipus de mètodes, habitualment només es pot accedir a zones pròximes a la base del talús, fet que condiona la presa de mesures atès que el seu registre no es distribueix per tota la superfície del talús. En canvi, els processats de les dades acostumen a ser ràpids, atès que, generalment, es disposa d'una quantitat de dades reduïda.

Algunes de les principals tècniques de mesures directes són l'*scan-line sampling* o *Windows sampling* (veure Secció 1.3).

A continuació, comentarem les principals característiques de l'*scan-line sampling*, tècnica que utilitzarem a la tesi doctoral per mesurar les discontinuïtats. A la Figura 2.8, es mostra esquemàticament un *scan-line*.

Tècnica de l'*scan-line*

L'*scan-line* és una metodologia àmpliament utilitzada per mesurar les propietats de les fractures, ja sigui directament sobre l'aflorament o a partir del testimoni recuperat d'un sondeig.

Descrita per Priest i Hudson, (1981) i Priest, (1993), es basa en identificar les fractures exposades sobre el terreny, al ser interceptades al llarg d'una línia: l'*scan-line* (Priest i Hudson, 1981; Priest, 1993; Zeeb et al., 2013). El nombre de paràmetres a recopilar pot variar en funció de la finalitat de de l'estudi, però el més habitual quan s'empra aquesta tècnica, és mesurar la posició, orientació, alçada, longitud, obertura i rugositat, a més de realitzar una descripció d'altres paràmetres com poden ser la relació amb altres fractures, terminacions, material de rebliment o estructures observades sobre la superfície de la fractura. El mètode destaca per la senzillesa dels requeriments tècnics necessaris per dur-lo a terme: cinta mètrica, brúixola, regla, lupa i una plantilla on recopilar les dades, en format digital i/o analògica.

Per definir la ubicació i direcció d'un *scan-line*, generalment s'escull la secció de **l'aflorament** amb la millor representació de les fractures exposades, i s'hi defineix la direcció de l'*scan-line* seguint una direcció paral·lela a la superfície de referència, com pot ser la estratificació o les unitats mecàniques. D'aquesta manera, les fractures queden orientades principalment en direcció aproximadament perpendicular a l'*scan-line* (Figura 2.3). Al ser un mètode lineal, la direcció d'escaneig estarà condicionada per la orientació preferent de l'aflorament, la qual pot no ser representativa ni la més adequada per a recopilar totes les orientacions de les fractures presents en el massís rocós (registrarà eficaçment les fractures que s'orienten perpendiculars a l'*scan-line* i en canvi, subestimarà la representativitat de les fractures que s'orienten paral·leles a

la línia d'escaneig). Aquesta limitació pot solucionar-se realitzant dos scan-lines en direccions perpendiculars entre sí. Un altre factor condicionant ve determinat per les orientacions de les superfícies de referència (estratigrafia o unitats mecàniques), les quals poden presentar un angle de cabussament elevat que impossibilita realitzar l'scan-line en direcció paral·lela a les mateixes, fet que provoca, tal i com s'ha descrit anteriorment (apartat 2.2.2), mesures de distància aparents entre fractures que, a posteriori, poden veure's corregides mitjançant la correcció de Terzaghi (1965).

Els **sondejos mecànics**, a partir de l'extracció dels testimonis (*cores*), permeten obtenir una certificació del terreny en fondària, basada en la descripció litològica i de certs paràmetres característics de les fractures (Mauldon i Mauldon, 1997; Dershowitz et al., 1998; Peacock et al., 2003). L'*scan-line* practicat sobre el testimoni d'un sondeig permetrà registrar diferents paràmetres de les fractures interceptades, com ara la posició, el cabussament respecte la vertical del sondeig, identificació del tipus de reompliment, les rugositats i en ocasions obertures. Aquestes dades s'acostumen a complementar mitjançant l'ús d'altres índex, com per exemple l'índex RQD (Deere, 1969) (veure secció 1.3 Antecedents). En canvi, la orientació de les fractures no es podrà obtenir si el sondeig no porta un orientador "dipmeter", ni tampoc obtindrem dades de la mida de les fractures. Les perforacions solen ser verticals, però també es poden executar subverticals o horitzontals. En aquest sentit, en cas que les fractures presents al massís siguin verticals o subverticals, és convenient programar un sondeig inclinat, buscant la perpendicularitat respecte els plans de fractura per tal d'obtenir la màxima informació em quant a la distribució de les fractures (Berg, 2012).

Com avantatges del mètode es poden citar la precisió dels registres que s'obtenen i el gran número de paràmetres que permet registrar (Cruden, 1977; Priest i Hudson, 1981) a escala d'aflorament o mesoscala. El principal desavantatge, en canvi, es troba relacionat amb l'accés a les dades. En aquest sentit, cal recordar que el mètode es limita a registrar dades pròximes al terreny, sense poder accedir a les parts més altes de l'aflorament i en conseqüència, la quantitat de registres duts a terme i per tant la seva representativitat, dependrà directament dels accessos a l'aflorament. Finalment, no es pot oblidar el factor seguretat en la presa de les dades, ni el paper determinant que juga en aquest àmbit l'estabilitat de les roques en l'aflorament.

2.3.2. Mètodes remots o indirectes

A diferència dels mètodes directes, els mètodes remots o indirectes, permeten extreure les dades relatives als diferents paràmetres, a partir d'imatges o de representacions digitals tridimensionals de la superfície del terreny, que es capturen mitjançant tècniques remotes o de teledetecció (principalment per Fotogrametria i/o *Light Detection and Ranging* o *LIDAR*, veure les properes seccions). A partir d'aquestes tècniques, s'obté el que s'anomena un *núvol de punts*. Aquest, es defineix per les coordenades desorganitzades en 3-D (X, Y, Z) més el valor d'un o més atributs (en funció de la tècnica utilitzada) característics de cada punt (com per exemple, el valor del color

extret a partir dels valors de les longituds d'ona del visible per cada canal RGB). Generalment està format per una quantitat dades molt elevada, és a dir, milions de punts i és considera d'alta resolució (*High Resolution*). En aquests casos la densitat pot arribar a ser de l'ordre de 1.000 a 10.0000 punt per metre quadrat.

El mètode de creació del núvol de punts, es pot realitzar de forma indiferent, per fotogrametria o TLS, encara que es pot associar a una major precisió dels punts al TLS que a la fotogrametria. Els núvols de punts adquirits, ja sigui amb tècniques TLS o de fotogrametria, acostumen a seguir un tractament d'alineació de les diferents imatges adquirides amb el software comercial *Polyworks* o altres softwares comercials. Una vegada alineades, la posició de les estacions han estat substituïts per les coordenades de *Differential Ground Positioning System* amb precisió centimètrica per geo-referenciar els núvols de punts de l'aflorament.

El gran avantatge d'aquest tipus de tècniques indirectes és que, gràcies a aquest núvol de punts digital, a part de permetre la extracció dels propietats de les fractures (veure Secció 2.2), mitjançant mètodes numèrics i eines informàtiques (veure més endavant Secció 2.4.1), permet també automatitzar diferents tasques i/o desenvolupar noves tècniques d'anàlisi que sense aquest tipus de suport digital no es podrien desenvolupar (veure Secció 2.4.2 i 2.5 *machine learning*).

Fotogrametria

La tècnica fotogramètrica permet la reconstrucció topogràfica semi-automàtica d'objectes o escenaris fotografiats que queden representats a través d'un núvol de punts. La metodologia consisteix en cobrir l'àrea d'estudi amb fotografies realitzades amb una càmera digital i que han de presentar un gran solapament (es recomanat un 60%), preferentment mantenint fixa la distància focal. Posteriorment, les imatges són processades en un programari informàtic amb capacitats *Structure from Motion* (SfM), com (*Agisoft (Metashape)*, *Pix 4Dmapper (Pix4D)*, *Context Capture (Bentley)*), obtenint com a resultat l'escena fotografiada en 3D, a partir d'un conjunt d'imatges 2D.

El procés informàtic s'inicia amb la finalitat de conèixer els paràmetres interns (distorsió de la lent, distància focal i punts principals) i externs (posició i orientació de la càmera) del sistema fotogramètric. El nombre elevat de punts homòlegs trobats amb la tècnica SfM permet conèixer per iteració aquest paràmetres amb les equacions de col·linealitat. Resolt l'ajust fotogramètric del sistema, els algorismes van reconstruint l'escenari fotografiat per crear un model tridimensional. Per escalar correctament els models creats és necessari utilitzar punts de control, és a dir, posicions o punts de la imatge de coordenades conegudes. Tot aquest procés requereix una maquinària informàtica capaç de realitzar l'esforç necessari, estant molt condicionat pel nombre d'imatges, la mida de les imatges, la qualitat de l'ajust fotogramètric i el nivell de reconstrucció de la superfície fotografiada. El format final habitual del model reconstruït és un núvol de punts, on a cada punt identificat de la imatge li correspon una coordenada 3D més la informació de la seva textura en un format Red, Green, Blue (RGB).

Tot aquest procés requereix d'equips informàtics amb alta capacitat de càlcul, tot i que al final, el temps de càlcul estarà molt condicionat pel nombre final d'imatges incorporades, la resolució de les mateixes (per tant, el nombre de píxels), la qualitat de l'ajust fotogramètric a realitzar pel programa informàtic amb les imatges incorporades, i el nivell de detall final (controlat per l'usuari) que es vulgui obtenir de la superfície o del núvol de punts corresponent a la zona fotografiada.

Light Detection and Ranging (LiDAR)

La tècnica LiDAR, al igual que la fotogrametria, permet la reconstrucció topogràfica semi-automàtica d'objectes, aconseguint un núvol de punts. En aquesta cas, el dispositiu *Laser Scanner*, emet un pols làser, que es reflecteix sobre la superfície del terreny (i demés objectes presents, com la vegetació) i retorna al dispositiu, obtenint així el temps que ha trigat en fer tot el recorregut i derivant-ne la distància a la que es troba. D'aquesta manera, permet la captura a gran velocitat d'un alt nombre de punts topogràfics (de l'ordre de milers de punts per segon). En el cas del dispositiu *Terrestrial Laser Scanner* o TLS, la posició del punt, es determina a partir de l'angle horitzontal i vertical del feix emès en relació al centre del sistema de referència del dispositiu (Bachman, 1979; Wehr i Lohr, 1999). (veure Figura 2.9).

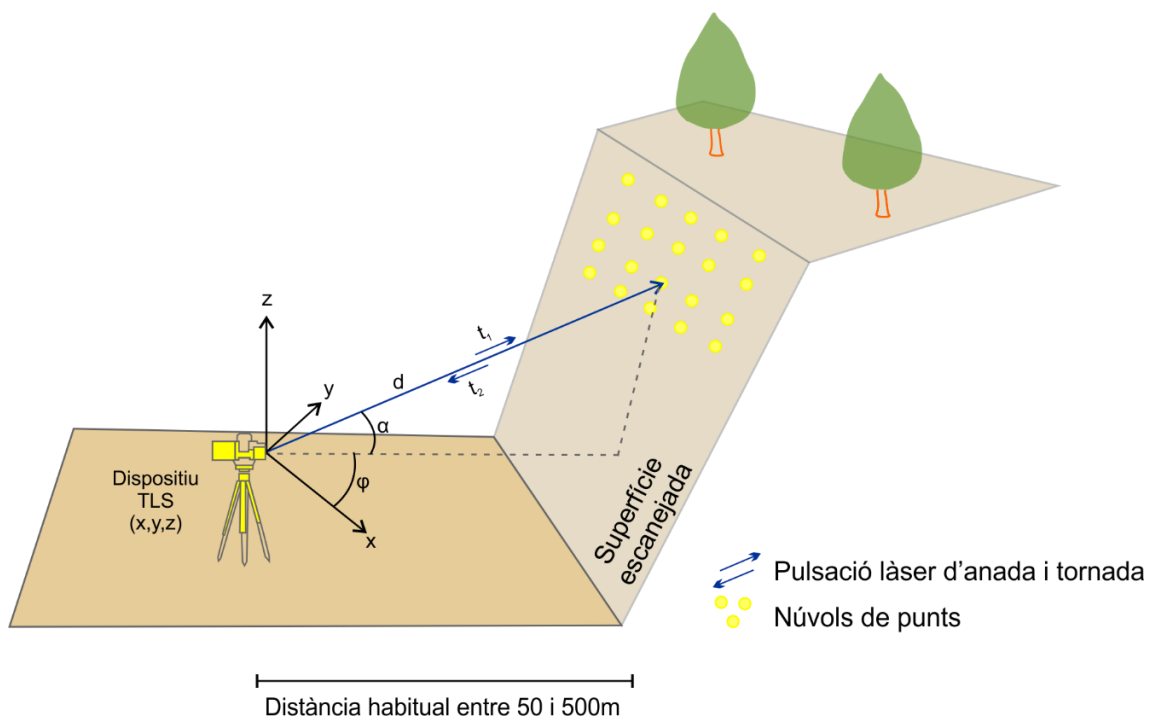


Figura 2.9: Principis bàsics de l'obtenció d'un núvol de punts mitjançant TLS. A partir de pulsacions làser emeses per el dispositiu TLS, en la direcció definida "d", s'obté la senyal reflectida que computa el temps de viatge bidireccional (t_1 i t_2). Aquest punt es mesura amb l'angle de coordenades horitzontals (φ) i verticals (α). La posició del dispositiu TLS, tindrà una posició georeferenciada (x, y, z) i del reflector pot derivar-se la corresponent coordenada angular (d, φ, α).

Aquest procés es realitza milers de vegades per segon, i acaba cobrint tota la superfície estudiada. Depenent de la mida de l'aflorament, la distància on se situa l'aparell i les característiques d'aquest, serà necessari repetir la mateixa operació en diferents estacions assolint així, recobriments del terreny amb una alta resolució i aconseguint també, cobrir les zona d'ombres (Sturzenegger i Stead, 2009). (veure figura 2.9). Com a resultat final, s'obté un núvol de punts en 3D representatiu de la superfície del terreny i els objectes capturats amb el TLS, que es defineixen per les coordenades locals angulars (d , φ i α) o cartesianes (x , y , z), segons la relació de l'equació 2.6. A partir dels diferents sensors que pot tenir el TLS, també es poden enregistrar altres paràmetres com la intensitat del feix làser de retorn o, el que equival a la textura de la superfície reflectida capturada mitjançant les longituds d'ona visibles en el canal vermell, verd i blau (RGB).

La densitat dels núvols de punts del TLS ve donada per les característiques del dispositiu utilitzat i la distància a l'aflorament. Durant l'escaneig, l'àrea del punt làser al impactar sobre el terreny condicionarà igualment la qualitat dels punt adquirits i per tant del núvol de punts.

Aquests dues tècniques remotes; fotogrametria i TLS, permeten l'adquisició digitals dels afloraments i venen condicionades per les seves característiques. El TLS permet una adquisició directa i precisa del núvol de punts, a diferència de la fotogrametria, que necessita de l'ajust fotogramètric previ a l'obtenció d'un núvol de punts menys precís que el TLS. D'altra banda, la fotogrametria requereix d'una logística i una inversió econòmica més assequible.

2.3.3. Altres mètodes de captura: mètodes geofísics

Les anteriors tècniques tenen en comú que es poden aplicar sobre la superfície de la roca on s'hi identifiquen les discontinuïtats presents en aquesta part més exterior del massís rocós. En canvi, l'aplicació de mètodes geofísics, permet penetrar a l'interior del massís rocós i derivar informació sobre les fractures mitjançant alguna propietat física mesurada en el medi rocós. Per tant, la informació que s'obté de les fractures serà indirecta a partir de la interpretació de les dades obtingudes mitjançant el mètode escollit. A continuació, es detallen les característiques i principis bàsics de dos d'aquests mètodes que s'han pogut utilitzar durant la tesi: el **Ground Penetrating Radar** (GPR) i la **tomografia elèctrica**.

Ground Penetrating Radar (GPR)

Una altra tècnica geofísica és el GPR o *Ground Penetrating Radar*, un mètode electromagnètic que a partir de generar polsos d'energia electromagnètica (EM) al subsòl, permet mesurar el temps que triga entre la transmissió i la recepció dels polsos reflectits de tornada a l'instrument (Conyers i Cameron, 1998). L'ona EM generada viatja esfèricament primer cap a l'interior i després a l'exterior a través del medi, on la penetració d'aquesta dependrà principalment de la conductivitat elèctrica, la constant dielèctrica dels materials al subsòl i de la freqüència del camp de transmissió, que es podrà visualitzar en un perfil en 2D o en un model 3D, generat a partir de perfils 2D.

La capacitat del GPR per detectar discontinuïtats, dependrà de la relació senyal-soroll, tal com descriu Annan (2001), on les principals variables que intervenen són (1) la freqüència, l'amplada de banda i l'eficiència de les antenes GPR (la sensibilitat generalment augmenta amb la freqüència més alta), (2) el soroll inherent del medi i (3) la reflectivitat de la fractura que depèn tant del contrast electromagnètic, entre l'obertura i/o el rebliment de la fractura i el medi rocós, com de la seva mida, forma, orientació i ubicació respecte a l'energia incident. La resolució d'imatge vertical obtinguda es considera generalment una quarta part (1/4) de la longitud d'ona predominant (Annan, 2001) i les freqüències centrals de l'equip acostumen a estar entre 250 i 700MHz, essent la de 500 Mhz, la més idònia per a la detecció de fractures a petita escala (Martinez i Byrnes, 2001; Jeannin et al., 2006; Kovin, 2011). Els estudis en diferents roques han demostrat que les discontinuïtats obertes i plenes d'argila o d'aigua són clarament visibles quan s'utilitzen longituds d'ona adequades (Da Silva et al., 2004, Theune et al., 2005), mentre que les discontinuïtats tancades sense farciment no solen aparèixer.

Per altra banda, alguns treballs com el Marchesini et al., (2013) i Guo et al., (2019) presenten propostes on amb la injecció d'aigua en una superfície coneguda i amb un volum establert i la captura de dades amb GPR, s'obtenen dades de la distribució i velocitat del flux.

Tomografia elèctrica

La tomografia elèctrica, és una tècnica geofísica d'alta resolució (Dahlin i Zhou, 2004), que permet la caracterització elèctrica del medi rocós a partir de la transmissivitat elèctrica del mateix. Per poder interpretar els resultats (generalment perfils 2D de resistivitat elèctrica), és necessari correlacionar la informació obtinguda amb dades de l'aflorament i/o sondejos.

Aquesta caracterització elèctrica del subsol es realitza a partir de la determinació de la resistivitat elèctrica del terreny. La resistivitat elèctrica del medi es mesura a partir de la resistivitat aparent obtinguda després d'injectar un corrent elèctric continu (I) al terreny mitjançant dos elèctrodes (elèctrodes de corrent) i de llegir la diferència de potencial elèctric (ΔV) a través de dos elèctrodes de mesura (elèctrodes de potencial). També intervé un factor geomètric (K) que depèn de la distància i posició dels elèctrodes (Equació 2.4).

$$\rho_a = K \cdot \Delta V / I$$

Equació 2.4

On ρ_a és la resistivitat aparent

I és el corrent elèctric continu injectat

ΔV és la diferència de potencial elèctric

K és el factor geomètric que depèn de la distància i posició dels elèctrodes

La distribució de corrents injectades i potencial elèctric en el subsòl permet reconstruir la secció de resistivitats aparents i després d'un càlcul d'inversió s'obté el model de distribució de resistivitats elèctriques del subsòl, representat mitjançant un o més perfils 2D (Sørensen, 1996).

En la present tesi, per realitzar el càlcul d'inversió s'ha utilitzat la metodologia descrita per (Loke i Barker, 1996) on mitjançant un càlcul d'iteracions successives proporciona un model de distribució de resistivitats elèctriques del subsòl. Aquest tipus de software treballen a partir d'algorismes basats en diferències finites i elements finits i contempla la possibilitat de treballar amb un ampli espectre de configuracions d'electrodes: Wenner, Pol-Pol, Pol-Dipol, Wenner Schlumberger i Dipol-Dipol.

A partir dels resultats obtinguts mitjançant tomografia elèctrica, es poden determinar diferents litologies, canvis laterals dels materials, profunditats i morfologies, identificar la posició del nivell freàtic, la presència de cavitats naturals o antròpiques, zones de discontinuïtats mecàniques, a partir de les diferències de resistivitat. A més a més, els resultats poden ser utilitzats per correlacionar sondatges a partir dels perfils tomogràfics obtinguts.

2.4. Aplicacions de les tècniques remotes en l'anàlisi del medi fracturat

La captura d'afloraments digitalitzats remotament bàsicament permet, dues gran aplicacions quan el que analitzem és el medi fracturat. Per una banda, l'anàlisi de la xarxa de fractures per a la construcció de models, ja siguin per la simulació de flux o d'altres aplicacions i, per altra banda l'anàlisi de processos associats, com les caigudes de blocs.

En el següent apartat es desenvolupen els principals conceptes d'aquestes anàlisis que tot i tenir en comú l'ús de tècniques remotes, en fases posteriors, tenen metodologies concretes i diferenciades. En la primera Subsecció 2.4.1, es descriuen els principals processos utilitzats en aquesta tesi doctoral, per a la caracterització i modelització de xarxes de fractures i en la següent Subsecció 2.4.2, la metodologia estàndard d'ús habitual per a la monitorització de blocs caiguts quan s'estudien amb TLS.

2.4.1. Models de xarxes de fractures

L'anàlisi de les fractures mitjançant tècniques remotes per la construcció models matemàtics té un camp d'aplicació bastant ampli, però sobretot juga un paper fonamental en l'estudi de la circulació i comportament de fluids en medis rocosos. Les dades obtingudes mitjançant els mètodes remots (veure Secció 2.2), i que inclouen els valors i paràmetres característics i identificadors de les fractures presents en un medi rocós, poden ser utilitzades per a obtenir nous paràmetres que son necessaris en posteriors càlculs, anàlisis o models numèrics, enfocats a la simulació de flux o altres aplicacions.

A partir del núvol de punts obtingut de les tècniques remots, serà necessari fer un pre-processat, per a la reconstrucció de les superfícies de les fractures i obtenir els principals paràmetres característics de les fractures per tal de ser derivats i utilitzats en modelitzacions posteriors. Aquests processos de reconstrucció acostumen a ser diferents per la fotogrametria i per el TLS, ja que es pot associar a una major precisió dels punts al TLS que a la fotogrametria al ser un mètode directe.

A continuació, s'expliquen els conceptes més rellevants de les metodologies aplicades en aquesta tesi doctoral; (1) per la reconstrucció de les superfícies de les fractures, (2) per a la caracterització de la xarxa de fractures obtinguda i finalment, (3) els models matemàtics que s'utilitzen comunament per a l'anàlisi de medis fracturats i més específicament el model discret de fractures, utilitzat durant la investigació.

Reconstrucció de les superfícies de les fractures mitjançant la fotogrametria

Els mètode de fotogrametria (veure Subsecció 2.3.2), no captura individualment les fractures, sinó que s'han d'identificar i aïllar de la totalitat del núvol de punts, pel que a diferència de les TLS, la reconstrucció de les superfícies de les fractures, no s'acostuma a fer directament amb el processat del núvol de punts, sinó que, prèviament, es realitzarà la digitalització de les traces o superfícies de les fractures. Aquesta digitalització consisteix a visualitzar el núvol de punts en un programari informàtic adient amb l'atribut de textura o en un renderitzat topogràfic i l'usuari ha d'anar traçant una línia o un polígon delimitant l'estructura geològica que reconeix (Figura 2.10), la qual, es pot realitzar a partir de la visualització del núvol de punts en un programari informàtic adient (*Agisoft Metashape, Cloud Compare, Microstation, PolyWorks*) amb l'atribut de textura o en un renderitzat topogràfic.

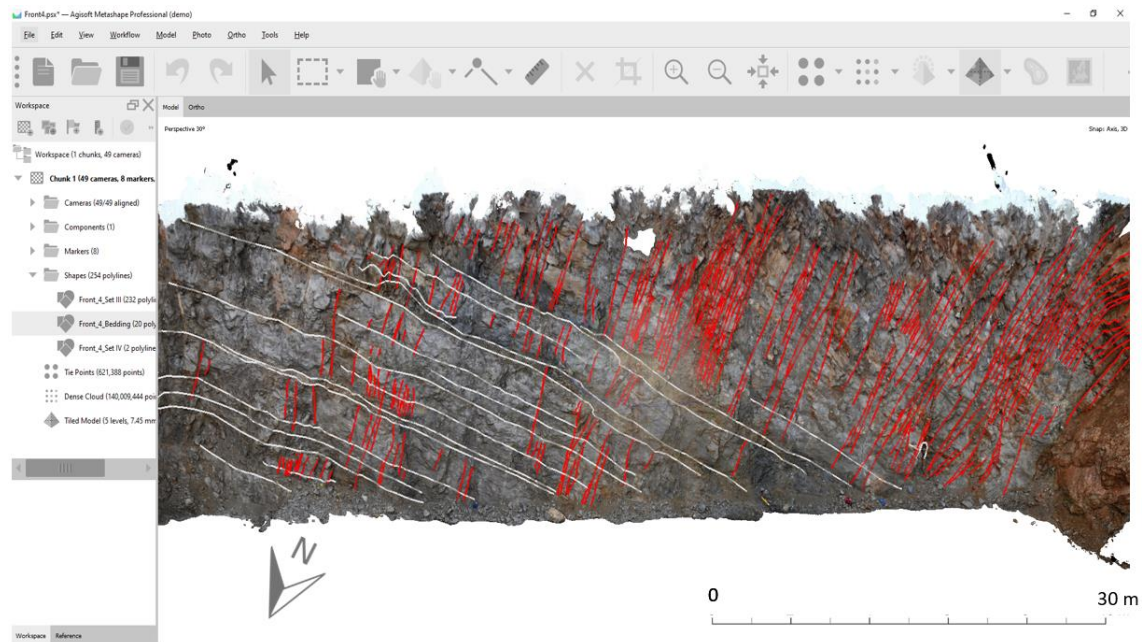


Figura 2.10: Procés de digitalització de les fractures (en vermell) i els plans d'estratificació (en blanc) a partir de la visualització del núvol de punts en el programa Agisoft Metashape.

No és un mètode ràpid i necessita molta intervenció de l'usuari, resultant dependent de l'usuari podent arribar a assolir molta qualitat en el seu resultat. En algunes solucions el programari proposa diferents alternatives buscant un patró en la digitalització. La digitalització pot anar acompanyada d'imatges o informació complementària que l'usuari pot anar contrastant (Geyer et al., 2015).

En la digitalització, l'usuari pot identificar les fractures sobre les fotografies de dues maneres diferents; per la traça (intersecció entre la fractura i la superfície de l'aflorament) o per la superfície de la fractura (que coincideix amb la superfície de l'aflorament). La primera estructura la representem per una línia i la segona per un polígon. A mesura que l'operari va reconeixent l'estructura sobre la fotografia, va inserint punts digitalment per representar la traça. Quan l'objecte és una superfície, l'operari va inserint punts fins a traçar els límits de la superfície.

- **Identificació de superfícies:** Per identificar les superfícies de la fractures i basant-se en la geometria de les fractures aflorants, s'han programat algorismes que dins d'un límits identifiquen aquesta geometria dins del núvol de punts. Aquests processos que poden arribar a ser totalment automàtics en ocasions solament poden identificar les fractures que exposen la seva superfícies i no de les traces (Roncella et al., 2004; García-Sellés et al., 2011; Gigli i Casagli, 2011; Kong et al., 2020; Lato i Vöge, 2012; Li et al., 2018).
- **Identificació de traces:** En un cas complementari amb l'anterior mètode està la identificació de les traces, que com ja em descrit abans, geomètricament és molt diferent a les superfícies. Per poder identificar les traces, el núvol de punts ha de contenir punts

registrats de la intersecció de la superfícies de l'aflorament i de la fractura, ja que si aquesta intersecció no genera cap geometria la única manera de identificar-la serà amb l'ajuda de la textura de la imatge. (Guo et al., 2019; Li et al., 2019; Umili et al., 2013).

Finalment, l'estructura digitalitzada es simplificarà en un pla per obtenir la posició (representat per un centroid) la orientació del pla i paràmetres relacionats amb la mida de la fractura digitalitzada, (Geyer et al., 2015).

Reconstrucció de les superfícies de les fractures mitjançant TLS

La identificació i reconstrucció de superfícies geològiques capturades a partir del núvol de punt geo-referenciat, necessitarà del post-processat de les dades. En aquesta tesi doctoral, s'ha seguit la metodologia descrita per García-Sellés et al., (2011) utilitzant el programari SEFL (*Surface Extraction from LiDAR*). El flux de treball descrit, consta de dues fases. La primera fase consisteix en vectoritzar cada punt del núvol de punts, relacionant-los entre els seus veïns, fins a calcular un vector que representi la superfície planar segons el mètode dels *Principal Component Analysis* (Jolliffe, 2011; Jolliffe i Cadima, 2016). Durant el càlcul s'obtenen els índexs de col·linealitat i coplanarietat (Woodcock, 1977) que permeten extreure la posició del punt tractat, ja sigui en la aresta del pla de fractura o sobre la superfície planar de la mateixa. En una segona fase, aquests atributs són utilitzats per filtrar-los i obtenir solament els punts representatius de les superfícies. D'aquesta manera les superfícies de discontinuïtat són reconstruïdes amb el mètode descrit en Santana et al., (2012), conservant els atributs de posició del centroid, alçada i longitud de cada fractura.

Models matemàtics aplicats a medis fracturats. Conceptes bàsics

Les alternatives per modelitzar un medi fracturat, són vàries, tal i com s'anomenen i es descriuen segons Carrera, (1988):

- En els models **de medi continu equivalent** (*single-porosity*) on el flux a través de les fractures és considerat similar al flux en un medi porós (Baca et al., 1984; Granet et al., 1998). Aquesta és una aproximació simple on es considera que les fractures i la matriu estan en equilibri i en un medi porós continu. Existeix una variant (*dual-porosity*), (Figura 2.11) on la permeabilitat i la porositat de la massa rocosa (matriu) són tractades com un sistema porós, i les fractures com un altre sistema porós (Barenblatt et al., 1960; Warren i Root, 1963; Pruess i Narasimhan, 1985)
- En els models **discrets de fractures**, es considera que el flux d'un fluid a través de la matriu és despreciable comparat amb el flux de les fractures (Figura 2.11). Les fractures són considerades com fractures individuals, on el flux està condicionat i controlat per la geometria i connectivitat de cada fractura dins de la xarxa de fractures que conformen, generant els models més realistes dels exposats anteriorment (Long et al., 1982; Andersson et al., 1984; Dershowitz i Einstein, 1987).

- En els **models mixtes o duals**, s'integren els dos models comentats anteriorment, representen les fractures individualment i la matriu queda representada com un medi porós (Hoteit i Firoozabadi, 2005). Alguns autors també l'anomenen models *Discrete Fracture Matrix* (DFM).

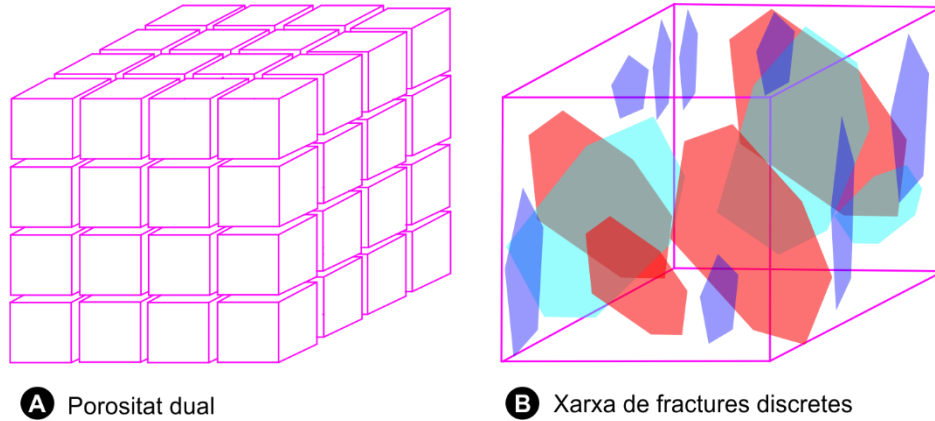


Figura 2.11: Esquema sintètic. (A) A l'esquerra un model dual i (B) a la dreta un model de xarxa de fractures discretes. Extret del manual de *FracMan*, 2018.

En aquesta tesi doctoral, s'ha treballat amb els models discrets de fractures i concretament el **model Discrete Fracture Network (DFN) tridimensional** mitjançant el programari *FracMan* de *Golder Associates*. Aquest model matemàtic ens ha permès la avaluació del sistema de fractures discretes representat a partir de les tècniques remotes, perseguint amb un dels objectius principals de la tesi doctoral.

Modelització discreta de fractures tridimensional (DFN)

La representació de les fractures implica el coneixement de les propietats de les fractures, ja sigui a partir de les dades capturades amb tècniques remotes o directes (Veure Secció 2.3) o indirectament, a partir de proves hidràuliques. El grau d'aproximació a les característiques reals de les fractures i el coneixement de la geologia del terreny són desafiaments realment importants en la creació d'aquests models informàtics. Les principals propietats geomètriques de les fractures que s'introduran per modelitzar un DFN, són les orientacions de les famílies, l'espaiat, la mida i l'obertura, propietats descrites a la Secció 2.2.1 i 2.2.2 i obtingudes a partir de la reconstrucció de les superfícies de les fractures.

Per a la modelització, s'ha utilitzat el programari *FracMan* desenvolupat per *Golder Associates Inc.* (Dershowitz, 1998), programa que construeix models discrets de fractures centrant-se en la generació de fractures (veure *FracMan® User's Manual*, 2020).

A partir d'aquests models discrets de fractures tridimensional, s'obtidran els paràmetres equivalents de les fractures per posteriors simulacions de flux amb programes com *Modflow* o *Petrel*, o directament, treballar amb el mateix programa *FracMan*.

2.4.2. Monitorització de caiguda de blocs en medis fracturats

L'avanç de les tècniques remotes (TLS i Fotogrametria) ha facilitat l'augment dels estudis sobre la detecció de canvis en la superfície d'afloraments fracturats com són els estudis de monitoratge, fet que provoca un afavoriment en el desenvolupament de mètodes per la detecció de canvis. D'altra banda, l'elaboració d'inventaris de caigudes de blocs per caracteritzar vessants comporta una important aportació en les monitoritzacions i resulten fonamental en la tasca de mitigació del risc associat (Santana et al., 2012, Schovanec et al., 2021).

Actualment, la metodologia més acceptada pel monitoratge de vessants amb núvols de punts està constituïda per diferents etapes que tenen en comú la cerca de l'automatització del processos per augmentar la productivitat. A la Figura (2.12) es representen les etapes de treball habituals o estàndard, integrades bàsicament per les següents etapes: a) la classificació prèvia dels núvols de punts per eliminar classes no desitjades (vegetació, sol, neu o sense interès), que comprometen la correcta alineació de núvols de punts; b) la alineació dels diferents núvols de punts obtinguts en cada monitoratge, situant els dos núvols de punts en el mateix sistema de referència; c) el càlcul de les diferències entre núvols de punts; d) la agrupació dels punts veïns comparats amb diferències significatives, i e) la classificació dels agrupaments de punts d'acord amb la natura dels objectes representats (p.e. bloc caigut, vegetació, efectes de vora, neu, sol o deformacions progressives del terreny com descriu Eberhardt et al., (2004), Rosser et al., (2007) i Kromer et al., (2015).

A continuació es descriuen amb més detall, els fonaments teòrics de les fases descrites que permeten la identificació i caracterització dels canvis registrats en els núvols de punts corresponents al monitoratge de l'evolució d'escarpaments rocosos, tal com s'estableix a la Figura 2.12. La part corresponent a la recerca desenvolupada en aquest àmbit en la present tesi, es recull en el la Secció 3.3.

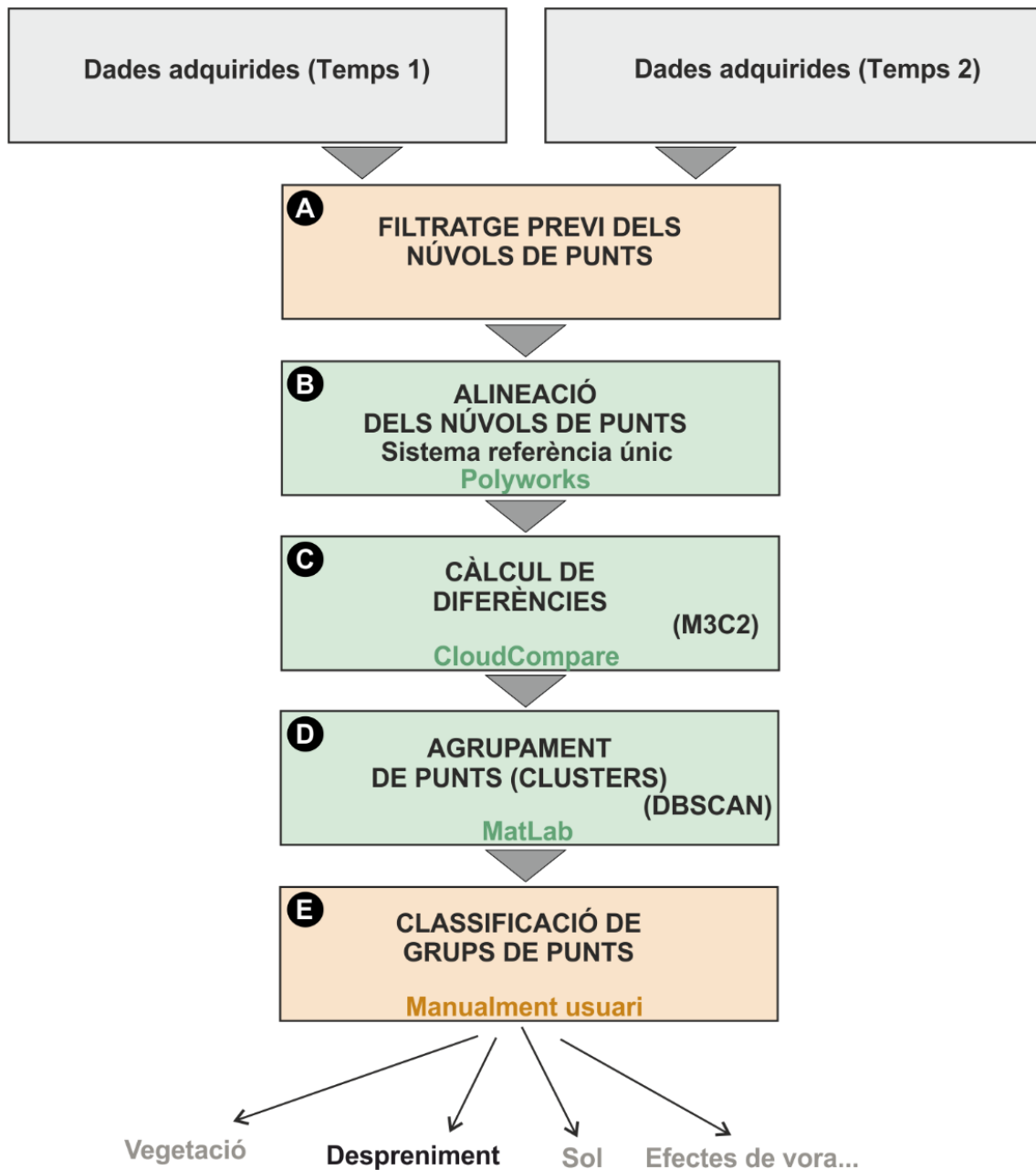


Figura 2.12: Flux de treball de la metodologia estàndard per la identificació de blocs caiguts amb tècniques remotes. El text en verd dels requadres es refereix als programaris que poden realitzar les tasques i entre parèntesis, als algorismes utilitzats. A) Filtratge previ dels núvols de punts; B) alineació dels dos núvols de punts; C) càlcul de diferències entre els núvols; D) agrupament de punts en funció de les diferències i proximitat; E) classificació dels blocs caiguts manualment, descartant els que es corresponen a vegetació o altres artefactes.

(A) (B) Filtratge previ dels núvols de punts i alineació

Per una correcta alineació dels núvols de punts es necessari disposar d'una manera majoritària de punts que siguin invariables entre el temps de monitoratge. Diferents cobertores com la vegetació o la neu, amb variacions temporals, poden arribar a ser molt abundants i dificultar molt l'alineació afegint moltes diferències sense interès. Per tant, l'eliminació d'aquestes cobertores és una condició necessària per un bon alineament, encara que sigui difícil a vegades i consumeixi molts esforços de temps i de supervisió.

Existeixen diferents mètodes per aconseguir aquest filtratge. Una primera aproximació és la generació manual de màscares amb les àrees d'interès de l'escarpament o escena per eliminar el contingut (Schovanec et al., 2021; Weidner et al., 2019). També el mètode utilitzat per Williams et al., (2018), que es basa en la classificació de punts en funció del atribut d'intensitat dels núvols de punts s'utilitza per elimina les classes que no es volen analitzar, especialment quan el núvol de punts ha estat creat mitjançant tecnologia LIDAR. O les tècniques de *machine learning* que han estat implementades per classificar regions per a ser descartades com per exemple l'aplicació CANUPO (Brodu i Lague, 2012), o la classificació de curvatura multiescala (MCC, *Multiscale Curvature Classification*, Zhang et al., 2016, Weidner et al., 2019). La condició necessària per poder comparar núvols de punts és que han d'estar en el mateix sistema de referència i compartir àrees inalterables en el temps en la mateixa posició (Figura 2.12B). El procés consisteix en considerar un dels núvols de punts fixe i a l'altre, aplicar-li una rotació i translació fins aconseguir el millor ajust entre els dos. Els mètodes matemàtics que permeten aconseguir el millor ajust són varis i en general, utilitzen aquestes àrees comuns entre núvols de punts per anar reduint les distàncies entre les seves superfícies a través de processos iteratius. A d'haver-hi una majoria de punts inalterables per poder ajustar-se a la millor solució, evitant així inexactituds.

Els mètodes més comuns alineen punt a punt, punt a model (entenent model com a núvol de punts) o model a model amb l'algorisme *Iterative Closest Point* (ICP, Besl i McKay, 1992) i les seves variants: ICP *Invariant Feature* (Sharp et al., 2002); l'*Sparse ICP* (Bouaziz et al., 2013) o el mínims quadrats 3D (Gruen i Akca, 2005). Aquests algorismes estan disponibles en llibreries lliures o incorporats en programaris comercials de tractament de dades digitals de núvols de punts.

(C) Càlcul de diferències entre núvols de punts

Mesurar els canvis que s'han produït en les superfícies dels vessants durant els període de monitoratge, és una tasca per la que s'han desenvolupat diferents solucions (Figura 2.12C). Principalment, aquests enfoc s'han dirigit a comparar les dades a través dels punts aïllats que conformen els núvols de punts o a través d'estructures de models de superfície (p.e., *Triangulated irregular network TIN*), com la proposada per Abellán et al. (2009).

La comparació es pot fer punt a punt (Girardeau-Montaut et al., 2005; Barnhart i Crosby, 2013); punt a model (Cignoni et al., 1998; Kazhdan et al., 2006; Tonini i Abellan, 2014) o model a model

(Lague et al., 2013; Kromer et al., 2015; DiFrancesco et al., 2020). La comparació es pot fer punt a punt (Girardeau-Montaut et al., 2005; Barnhart i Crosby, 2013); punt a model (Cignoni et al., 1998; Kazhdan et al., 2006; Tonini et al., 2013) o model a model (Lague et al., 2013; Kromer et al., 2015; DiFrancesco et al., 2020). La distància de la diferència es pot mesurar amb el mètode del punt més proper o trobant el punt més proper al llarg del vector normal de la superfície. Les diferents característiques de les vessants estudiades, en referència al seu grau d'escarpament, rugositat o continuïtat són les que condicionaran el mètode a escollir al adaptar-se millor al terreny.

Un dels mètodes més acceptats en els estudis de caigudes de roques (Van Veen et al., 2017; Bonneau i Hutchinson, 2019; Schovanec et al., 2021 i Hendrickx et al., 2022) és el mètode *Multiscale Model to Model Cloud Comparison* (M3C2) proposat per Lague et al., (2013) i implementat en el programari *CloudCompare* (Girardeau-Montaut, 2022). Amb aquest mètode model a model, la diferència entre el núvol de punts de referència i el comparat és calculada en la direcció del vector normal a una superfície definida per l'usuari amb un número de punts veïns. La direcció del vector és projectada fins a interceptar el núvol de punts comparat. Dos paràmetres són necessaris per aquesta configuració, el diàmetre de cerca per definir les dues superfícies (referència i comparada) i una distància màxima entre superfícies. La distància queda registrada com un nou atribut en el núvol de punts.

(D) Agrupació de diferències

Els processos anteriors tenen com a objectiu identificar punts que han estat desplaçats durant un període de monitoratge i quantificar el desplaçament. El procediment consisteix en que cada punt d'un núvol de referència es pugui relacionar amb un punt del núvol de comparació per mesurar la distància desplaçada, la qual, es convertirà en un nou atribut del punt i quedarà associada al núvol de punts comparat (Figura 2.12D). L'agrupació dels punts amb distàncies significatives i propers en l'espai són susceptibles de representar objectes amb interès geològic, i per tant es requereix d'un procés addicional amb el que agrupar-les i dotar-les d'una nova entitat, per poder ser analitzats en el conjunt.

El paràmetre, valor límit de detecció (LdD), és utilitzat per considerar si el desplaçament calculat entre cada punt del núvol de punts és prou significatiu per representar un canvi en els aflorament o es considerat com una limitació del mètode de mesura. En funció d'aquest límit es poden considerar tres situacions possibles: a) una diferència positiva; b) una diferència negativa; i c) una indeterminació entre els valors positius i negatius del LdD que es assumit com depenent de la precisió del sistema (e.i., precisió del làser escàner terrestre, distància aparell-superfície o mètode de sistema).

Geomorfològicament, el moviment de retrocés correspon a la sostracció de volum provocada per la caiguda de material, mentre l'avanç és interpretat com un desplaçament de la roca abans del

despreniment (Eberhart et al., 2004). El valor del LdD s'estableix amb el coneixement de l'usuari en l'àrea de treball o per estudis experimentals (Tonini et al., 2013).

L'algorisme *Density Based Spatial Clustering of Applications with Noise* (DBSCAN; Ester et al., 1996) compleix aquesta finalitat i ha estat aplicat en diferents estudis (Royán et al., 2014; Benjamin et al., 2020; Carrea et al., 2021; Schovanec et al., 2021; Y. Wang et al., 2021). Els paràmetres que necessita són: a) la distància màxima entre punts veïns; b) una distància mínima com per considerar-la desplaçament i; c) un nombre mínim de punts per constituir una agrupació, també anomenada clústers. L'algorisme està disponible en biblioteques d'accés lliure com Open3d (Zhou et al., 2018), Scikit-learn (Pedregosa, 2011) o implementat dins de programaris comercials com MATLAB.

(E) Classificació d'agrupacions de diferències o clústers

El darrer pas de la metodologia (Figura 2.12E), consisteix en classificar les agrupacions de punts en funció de la seva naturalesa. El darrer pas de la metodologia consisteix en classificar les agrupacions de punts en funció de la seva naturalesa. El reconeixement i classificació de les agrupacions de punts no és una tasca trivial. Les diferents classes que poden ser identificades en cada anàlisi poden ser assignades a caigudes de blocs o assignades canvis en la vegetació, erosió de sol, errors de vora, presència o canvis en la neu o humitat, canvis antròpics o qualsevol alteració de la superfície. Normalment la classificació es realitza visualment per un operari i dependrà de la seva habilitat o experiència, dimensions de l'àrea d'estudi, volum dels blocs caiguts, nombre de d'agrupacions i classes existents (Mazzanti et al., 2021; Royán et al., 2014; van Veen et al., 2017).

També s'han desenvolupat diferents mètodes per guanyar objectivitat, eficiència i productivitat amb mètodes supervisats com el proposat per Schovanec et al. (2021) basats en classificacions de models estadístics de *machine learning*. Zoumpekas et al. (2021) focalitza el seu treball en trobar el millor model de *machine learning* per casos de caigudes de blocs comparant sistemàtiques monitoritzacions de núvols de punts amb un rati desequilibrat amb el nombre d'esdeveniments de caigudes de blocs al massís de Montserrat (Espanya). En la present tesi (veure Secció 3.3), es proposa una nova metodologia que utilitza tècniques de ML per automatitzar i simplificar el procés de classificació dels agrupaments de punts, per identificar caigudes de blocs.

2.5. Machine learning en l'anàlisi del medi rocós. Conceptes bàsics, tècniques de remostreig i models o algorismes de classificació

L'aprenentatge automàtic o *machine learning* és un camp de la intel·ligència artificial que està dedicat al disseny, l'anàlisi i el desenvolupament d'algorismes i tècniques que permeten que les màquines evolucionin. El mètode és utilitzat per la construcció de models matemàtics amb dades de mostreig (anomenades dades d'entrenament) i amb la finalitat de fer prediccions sense ser explícitament programat per fer-les. La seva característica resideix en la millora de les prediccions amb l'experiència. És una divisió de la intel·ligència Artificial i utilitza algorismes i models de xarxes neuronals per progressivament anar millorant el model. Lligat amb els medis fracturats i, concretament, amb l'estudi i anàlisi de processos de caigudes de roques, ja s'estan desenvolupant algorismes de classificació, com el relacionat amb la present tesi doctoral (Blanco et al., 2022). (veure secció 1.3 Antecedents).

2.5.1. Conceptes bàsics

El camp del *machine learning* aplicat a la classificació y regressió, necessita d'un seguit d'etapes bàsiques per obtenir resultats vàlids, la recoll·lecció de dades i el pre-processat d'aquestes, la selecció dels models o algorismes de classificació, entrenar el model i avaluar-los per obtenir els resultats de la predicció (veure Figura 2.13).

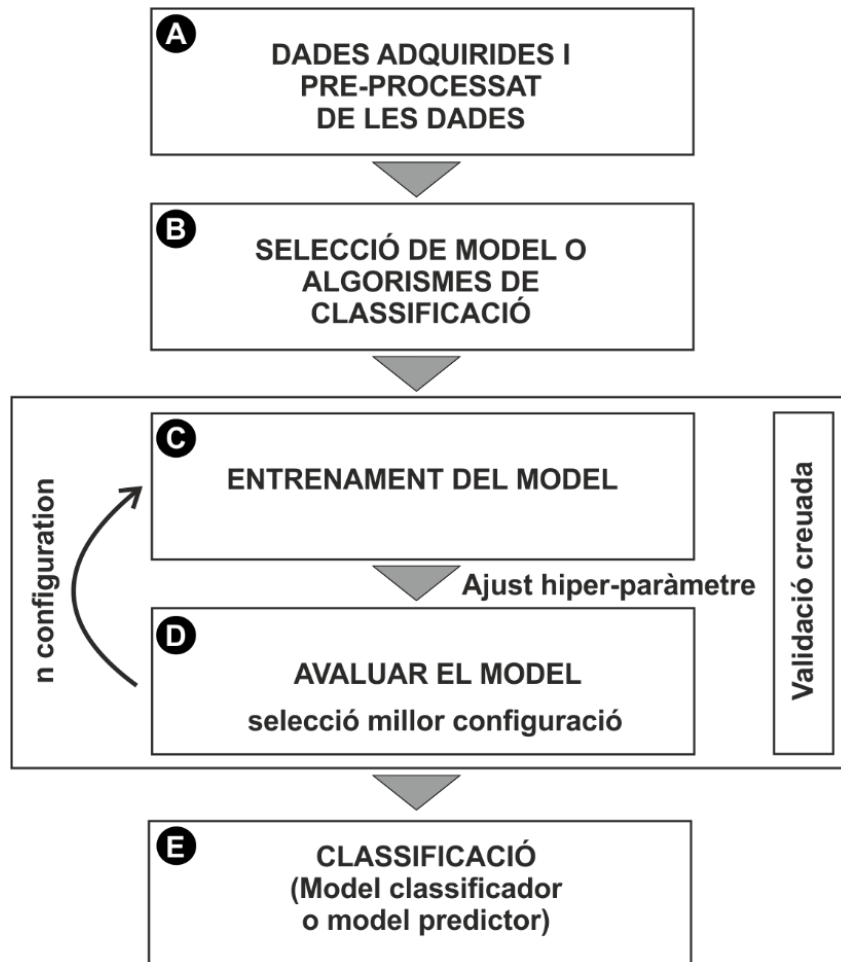


Figura 2.13: Flux de treball per a la classificació de dades amb machine learning. Es mostren les principals etapes. (A) L'adquisició de dades i el pre-processat d'aquestes, (B) La selecció d'algorismes (model) de classificació, (C) entrenament del model amb un grup de dades d'entrenament i (D) avaluació del model amb la resta de dades del grup. La etapa C i D, es repetirà n vegades, amb diferents configuracions, fins a obtenir la millor configuració per als objectius definits. En aquesta fase també s'aplicaran l'ajust hiper-paràmetre, per reduir el processat i és molt recomanable corroborar els resultats amb la validació creuada. (E) la selecció del model de predicció a partir del que obtindrem la classificació.

Dades adquirides i pre-processat de les dades.

En aquestes etapes, es recol·lecten les dades i es processen per a que estiguin acord amb els algorismes o models escollits i on s'identificaran les possibles mancances.

Una problemàtica habitual que també ens ha succeït en la present tesi, és que posteriorment a la Etapa 1 i 2, el nombre dades (elements) que integren cada classe proposada no és similar. En l'etapa d'aprenentatge l'algorisme d'aprenentatge necessita d'una quantitat d'informació equitativa per cada grup. Quan aquesta condició no es compleix, el desequilibri en la informació aportada de cada grup pot esbiaixar els resultats sobreestimant o subestimant l'existència i la informació derivada de cada grup. Per evitar aquest tipus de comportament o situacions, en el *machine learning* existeixen les anomenades **tècniques de balanceig o de remostreig** que permeten equilibrar el nombre d'elements de cada grup de forma semi-automàtica. Aquestes tècniques permeten, o sobredimensionar (*oversampling*) la classe insuficientment representada incorporant-ne nous elements, o sub-dimensionar (*undersampling*) la classe excessivament representada, eliminant-ne elements, o utilitzar eines mixtes que fusionen o combinen les dues solucions anteriors. Les tècniques de balanceig o remostreig utilitzats en aquesta tesi, es descriuen a la Subsecció 2.5.2.

Seleccionar els models supervisats de classificació

Existeixen diferents algorismes matemàtics que resolen el procés de classificació. Aquests algorismes es poden agrupar en dos conjunts (Ma et al., 2021) segons si contempen una única via o estratègia (anomenats de **Base Simple**) per a obtenir la classificació final, o els que, per obtenir-la, en combinen més d'una (de **Base Múltiple**).

Entrenar el model

Es necessita una primera etapa, denominada d'**aprenentatge o d'entrenament**, on es processa una part de les dades ja classificades amb els models o algorismes seleccionats, i s'entrena el model amb aquestes dades. Els models o algorismes seleccionats per aquesta teia, es descriuen a la Subsecció 2.5.3.

Avaluar el model i realitzar la predicció

Un cop entrenat el model, procedirem a l'avaluació de les dades que havíem reservat per aquesta etapa. (veure Figura 2.13). En aquest punt els resultats obtinguts probablement no seran els desitjats i s'haurà de tornar a repetir la etapa C, canviant la selecció de les dades d'entrenament, per a continuació, tornar a avaluar el model. La etapa C i D, es repetiran tantes vegades com sigui necessari, fins obtenir la millor configuració. També en aquesta etapa és convenient realitzar l'ajust d'hiper-paràmetres i la validació creuada, per refinar els resultats del model i que queden descrits en els apartats següents.

Ajust d'híper paràmetres. Les possibilitats de combinació dels paràmetres de cada model classificador són molt elevades, raó per la que s'acostuma a utilitzar la tècnica d'optimització amb híper parametrització dels models (*Hyper parameter tuning*). Aquesta tècnica, consisteix en definir les alternatives que es considerin importants per l'optimització de l'algorisme classificador, ja sigui identificant rangs de valors que definirem com a òptims per a cada paràmetre o indicant diferents opcions de cada paràmetre. Durant l'etapa d'aprenentatge de l'algorisme, els híper paràmetres es combinen per a ser avaluats mitjançant la mètrica resultant (definida per la exactitud, precisió, etc) per a cadascuna de les combinacions realitzades. Finalment, la millor mètrica serà la utilitzada per obtenir la predicció final. S'ha de tenir present que, en funció del nombre de paràmetres definits, dels rangs de valors, i de les possibles combinacions que en resultin, els temps de càlcul poden veure's incrementats enormement.

Validació creuada. La tècnica d'optimització amb híper paràmetres es pot combinar amb la denominada tècnica de validació creuada (*Cross Validation*) per avaluar els algorismes d'aprenentatge. La tècnica consisteix en seccionar supervisadament les dades d'entrenament en diferents subconjunts per avaluar la classificació obtinguda en funció d'aquests subconjunts. D'aquesta manera es pretén garantir la independència dels resultats en la partició de les dades per realitzar els processos d'entrenament i prova de l'aprenentatge automàtic. El nombre de particions de les dades es defineix per l'usuari, i condiciona també el temps de càlcul resultant.

2.5.2. Descripció de les tècniques de remostreig o balanceig considerades en aquest estudi

Es comenten a continuació les tècniques de remostreig contemplades en la investigació. (veure Taula 2.2).

Sotamostreig (<i>Undersampling</i>)	Sobremostreig (<i>Oversampling</i>)
<i>Cluster Centroid</i>	<i>SMOTE</i> (Chawla et al., 2002)
	<i>ADASYN</i> (He et al., 2008)
<i>Cluster Representative</i>	<i>SPIDER</i> (Stefanowski, 2013)
	<i>SWIM</i> (Sharma et al., 2018)
	<i>Polynom-Fit SMOTE</i> (Gazzah i Amara, 2008)
	<i>ProWsyn</i> (Barua et al., 2013)
	<i>SMOTE-IPF</i> (Sáez et al., 2015)
	<i>LEE</i> (Lee et al., 2015)
	<i>SMOBD</i> (Cao i Wang, 2011)
	<i>G-SMOTE</i> (Douzas i Bacao, 2017)
	<i>LVQ-SMOTE</i> (Nakamura et al., 2013)
	<i>Assembled-SMOTE</i> (Zhou et al., 2013)
	<i>TomekLinks- SMOTE</i> (Batista et al., 2004)

Taula 2.2: Recull de diferents tècniques de remostreig per corregir el desequilibri existent entre grups, agrupades segons si permeten un sotamostreig o sobremostreig.

Subdimensionament (*Undersampling*): en el cas de voler eliminar la classe excessivament representada, es podrà realitzar de dues maneres diferents: eliminant elements aleatòriament, o definint criteris que condicionin la eliminació dels elements (Yen i Lee, 2009).

Les estratègies de subdimensionament utilitzades en la tesi estan basades enfocades en agrupar elements de la classe majoritària en un nombre similar al de la classe minoritària. Un primer mètode (*Cluster Centroid*) utilitza el centroid per agrupar la classe majoritària i un segon mètode utilitza el centre del clúster, denominant-se (*Cluster Representative*)

Sobredimensionament (*Oversampling*): En el cas de voler augmentar els elements que s'inclouen a la classe menys representada, es poden utilitzar diferents mètodes (veure Taula 2.2).

SMOTE: La major part dels mètodes utilitzats, són adaptacions del mètode *Synthetic Minority Oversampling Technique* que genera elements del grup minoritari a partir dels clústers exemple per crear elements sintètic o artificial i equilibrar els elements de totes les classe (Chawla et al., 2002).

ADASYN: El mètode *Adaptive Synthetic Sampling Approach for Imbalanced Learning* fa la generació d'elements sintètics amb una adaptació progressiva del límit de decisió dels elements de la classe minoritària. La adaptació progressiva l'aconsegueix utilitzant la densitat de la distribució per generar les mostres sintètiques (He et al., 2008).

SPIDER: el mètode és el *Selective Pre-processing of Imbalanced Data*, el qual fusiona els elements que queden després de filtrar els elements més problemàtics de la classe majoritària amb el sobredimensionament de la classe minoritària (Stefanowski, 2013).

SWIM: La fonamentació del mètode *Sampling With the Majority* és que els elements sintètics de la classe minoritària s'han de generar en llocs de densitats similars respecte a la classe majoritària (Sharma et al., 2018).

Polynom-fit-SMOTE: El mètode que crea els elements sintètics de la classe minoritària ajustant una funció polinòmica (Gazzah i Amara, 2008).

ProWsyn: El mètode *Proximity Weighted Synthetic* genera valors de pes per als elements de la classe minoritària basant-se en la informació de la distància als límits de cada classe (Barua et al., 2013).

SMOTE-Iterative Partitioning Filter: El mètode genera elements sintètics amb escissions de la classe minoritària i els classifica supervisadament aplicant iterativament filtres de soroll i propers als límits de la classe. Després comprova les probabilitats de pertànyer a la classe (Sáez et al., 2015).

LEE: En aquets cas, implementa una tècnica similar a l'anterior però filtrant els elements sintètics generats amb SMOTE en funció de la seva localització amb el filtre K-Nearest Neighbor en una etapa de post-procés (Lee et al., 2015).

SMOBD: *The Synthetic Minority Oversampling Based on Samples Density* considera la informació de la densitat i de la distribució de la classe minoritària abans de generar els elements sintètics (Cao i Wang, 2011).

G-SMOTE: Dins d'aquest mètode derivat del mètode SMOTE l'algorisme, Geometric-SMOTE genera elements sintètics dins d'una distància definida amb el mètode K-Nearest Neighbor amb cada element de la classe minoritària (Douzas i Bacao, 2017).

LVQ-SMOTE: El mètode Learning Vector Quantization crea els seus elements sintètics utilitzant codi obtingut amb l'algorisme de classificació supervisada Learning Vector Quantization (Nakamura et al., 2013).

Assembled-SMOTE: Aquets, crea particions lineals locals dins i al voltant dels elements de la classe minoritària per aplicar l'algorisme SMOTE en grups de dades separades (Zhou et al., 2013).

Tomelinks-SMOTE: aplica primerament l'algorisme SMOTE per eliminar parells d'elements majoritaris i minoritaris si són identificats com soroll o propers als límits de la classe (Batista et al., 2004). La comparació d'aquest mètodes és pot consultar en la publicació de Kovács, (2019).

2.5.3. Descripció dels models o algorismes de classificació supervisada considerats en aquest estudi

Es comenten els trets principals dels models o algorismes de classificació supervisada contemplats en la investigació. (veure Taula 2.3).

Base simple	Base múltiple
<i>Decision tree</i> (Hastie, 2001)	<i>Adaptative Boost</i> (Hastie et al., 2009)
<i>Gaussian Naive Bayes</i> (Hastie, 2001)	<i>Random Forest</i> (Hastie et al., 2001)
<i>K-Nearest Neighbor</i> (Hastie, 2001)	<i>Extra Tree</i> (Geurts et al., 2006)
<i>Linear Discriminant Analysis</i> (Hastie, 2001)	<i>Extreme Gradient Boosting (XGBoost)</i> (Chen i Guestrin, 2016)
<i>Quadratic Discriminant Analysis</i> (Hastie,	
<i>Multi-Layer Perceptron</i> (Murtagh, 1991)	
<i>Support Vector</i> (Awad i Khanna, 2015)	

Taula 2.3: Diferents algorismes desenvolupats per obtenir una classificació final, tenint en compte si son de base simple o de base múltiple (veure explicació al text).

A continuació, es resumeixen els trets bàsics i característics de cada algorisme per tal de poder mostrar l'objectiu principal de cadascun, així com els fonaments teòrics bàsics i imprescindibles per a poder seguir i entendre els mètodes que es descriuran en capítols posteriors (veure Secció

3.3). Per aprofundir més en tots i cadascun d'aquests algorismes, així com per complementar aquesta informació bàsica que es recull aquí, el lector pot consultar la bibliografia citada.

El grup de models classificadors de **base simple** utilitzats en aquesta tesi són:

Decision tree: L'aprenentatge basat en arbres de decisió consisteix en la selecció automàtica amb una única estratègia representada en un arbre de decisió (com a model predictiu) (Hastie et al., 2001). Els nodes interns, que separen o uneixen les branques representen les propietats característiques de les dades (o aquells valors característics), i les branques, les regles de decisió (variables independents), que es representen per fulles com la classe classificada (variables dependents). El procés de decisió consisteix en començar a l'arrel i, basant-nos en el valor de les variables independents, anar prenent la branca corresponent. Els arbres de decisió on la variable dependent és discreta (pren un nombre finit de valors) s'anomenen arbres de classificació a diferència dels arbres de regressió, on la variable dependent és continua.

Gaussian Naive Bayes: És un algorisme que utilitza el teorema de Bayes, per obtenir una classificació de les dades basant-se en que les variables predictores són independents entre si. Considera les prediccions de valor continu i les agrupa com distribucions normals, també anomenada gaussianes (Hastie et al., 2001).

K-Nearest Neighbor. En aquest model, s'emmagatzema les dades i es classifiquen les noves dades en funció de la distància/semblances als seus veïns emmagatzemats, que poden variar en número (K). El paràmetre opcional Number of neighbors controla el nombre de veïns. (Hastie et al., 2001)

Linear Discriminant Analysis: L'algorisme busca agrupacions de característiques per diferenciar linealment les classes, s'utilitza sobretot per classificacions binàries (Hastie et al., 2001). Les classes són ajustades amb la funció de densitat Gaussiana amb la mateixa matriu de covariància. El paràmetre opcional més important és el referent al tipus d'estimador: SVD (Single Value Decomposition), LSQR (Least Square Solutions) o Eigen (Principal Component Analysis)

Quadratic Discriminant Analysis: L'algorisme treballa de la mateixa manera que el Linear Discriminant Analysis (Hastie et al., 2001), però sense assumir una mateixa matriu de covariància. Aquesta variació li dona més flexibilitat per diferenciar les classes. El paràmetre opcional Reg Param varia el valor de la covariància.

Multi-Layer Perceptron: Aquest classificador es basa en xarxes neuronals preventives, de capes connectades (Murtagh, 1991). La xarxa consta de capa d'entrada i una de sortida, amb capes ocultes on s'estimen els valors. Les dades d'aprenentatge passen per les capes definint límits de decisió lineal amb un pes i un biaix fins activar les funcions d'activació. La finalitat es aconseguir una capa de sortida amb resultats de classificació iguals a les capes d'entrada ja classificades. D'aquesta manera es defineixen els millors límits de separació de classe o hiperplans. Els paràmetres opcional més importants són: Solver, per definir el mètode estimador

(quasi-Newton method, Stochastic Gradient Descent, Stochastic gradient-based optimizer), la funció d'activació de les capes ocultes i el nombre de les capes ocultes.

Support Vector. L'objectiu de l'algorisme és crear línies o límits de decisió agregant múltiples dimensions a l'espai per poder classificar o definir millor els límits entre classes amb els hiperplans (Awad i Khanna, 2015). L'algorisme escull els casos extrems per definir vectors dels hiperplans. Els paràmetres opcional més importants són: Tipus de kernel per fer una classificació no lineal a lineal, Gamma per definir el coeficient del kernel i el paràmetre de regularització C del kernel.

Els mètodes classificadors que utilitzen **base múltiple**, són:

Adaptive Boost. Els algorismes que utilitzen la tecnologia Boost combinen models de baixa precisió per crear models d'alta precisió (Hastie et al., 2009). És un model iteratiu, on un model classificador dèbil inicia el procés i en les següents iteracions el model s'enfoca en les dades erròniament classificades, reassignant el pes (gra d'aprenentatge) de cada dada fins a obtenir la mètrica desitjada o assolint el nombre d'iteracions establert. El creixement iteratiu és seqüencial creant-se un node amb dues fulles en cada procés reduint el biaix i la covariància. Els paràmetres opcionals més importants són: El nombre d'estimadors i el grau d'aprenentatge.

Random Forest. És un model on un grup de models dèbils es combinen per crear un de fort. Diferents arbres de decisions classificant escissions de dades per acceptar el resultat en funció del més proposat. El classificador crea un nombre determinat d'arbres expandint els nodes, al no tenir una profunditat definida (Hastie et al., 2001). Els paràmetres més destacats són el nombre d'estimadors (nombre d'arbres aleatori) i la funció per mesurar la qualitat de l'escissió de les dades (Gini o Entropia).

Extra Tree. És un model per conjunts on s'agreguen els arbres de decisió per correlació semblant a Random Forest, però amb la diferència de com es construeix el bosc (Geurts et al., 2006). Cada arbre es construeix amb les dades d'aprenentatge, i a cada node se li proporcionen dades aleatòries. L'arbre escull la millor característica i escindeix les dades. Els paràmetres més importants són el nombre d'estimadors, el criteri d'escissió (Gini o Entropia) i la profunditat del bosc.

Extreme Gradient Boosting (XGBoost): Aquest mètode està basat en un arbre de decisions que utilitza un procediment d'impuls per millorar la tasca (Chen i Guestrin, 2016). Combina un nombre d'estimadors dèbils per crear classificació robusta. Les iteracions dels estimadors són ponderades amb la iteració anterior. Els resultats correctes són ponderats a la baixa i els erronis a l'alta. La combinació d'aquestes solucions fins a arribar a la mètrica escollida per donar la solució a la classificació. Els paràmetres més importants són: el nombre de sub-processos paral·lels, el tipus de reforç d'arbre o linears, el grau de ponderació, el nombre d'estimadors.

CAPÍTOL 3

Metodologies i resultats

3.1. Introducció

3.2. AVALUACIÓ DE LES TÈCNIQUES REMOTES EN MODELS DFN APLICATS A LA SIMULACIÓ DE FLUX

3.3. PROPOSTA METODOLÒGICA PER A LA IDENTIFICACIÓ DE BLOCS CAIGUTS AMB NÚVOLS DE PUNTS I *MACHINE LEARNING*

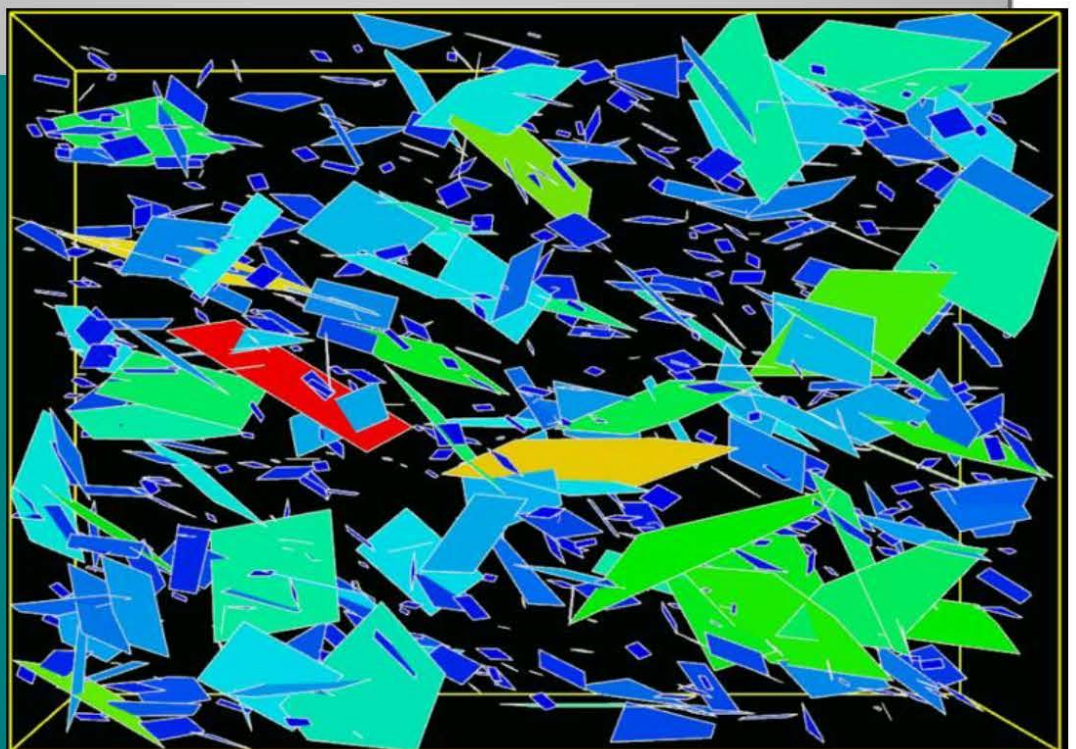
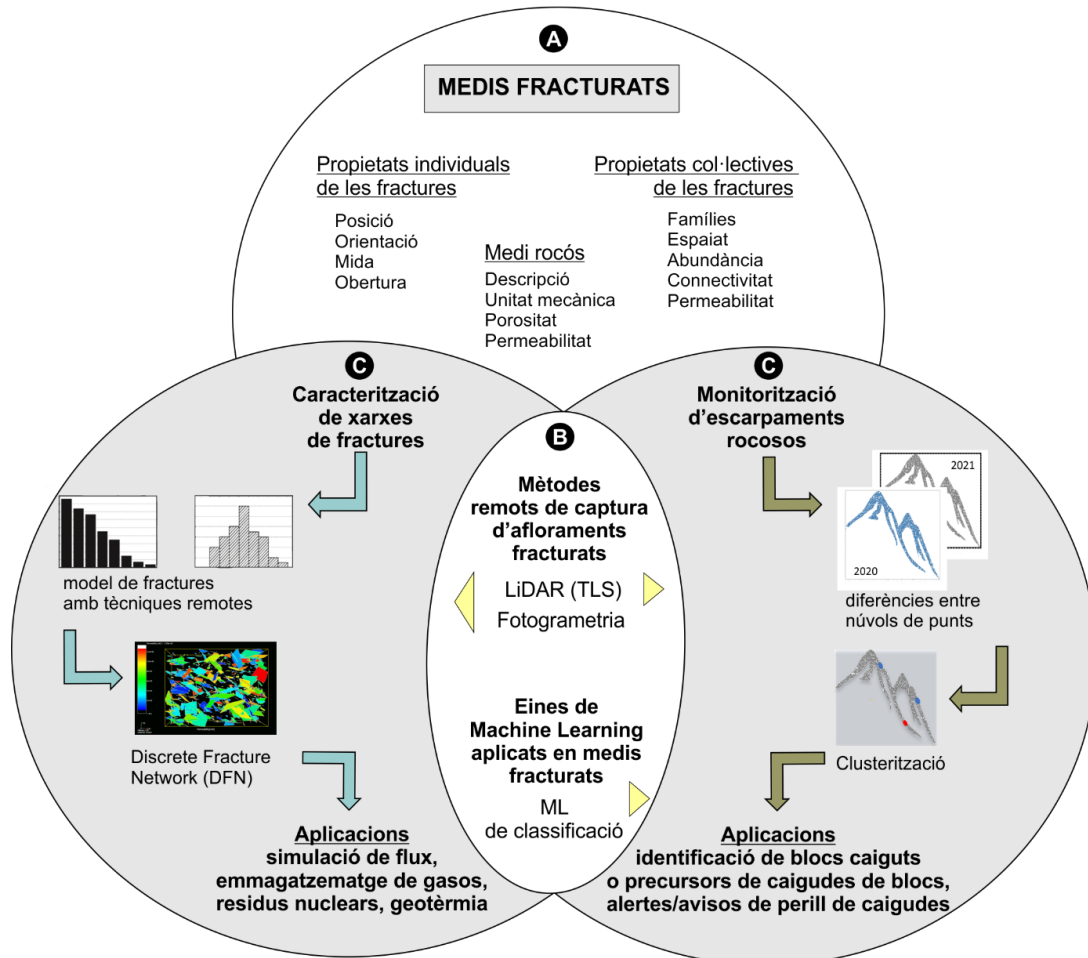


Figura 3.1: (veure pàgina següent) Esquema gràfic dels principals components de la present tesi, on (A) s'enumeren les característiques geomètriques dels medis fracturats (discontinuitats mecàniques i medi rocós) i (B) s'indiquen les tècniques digitals utilitzats per capturar i processar les dades dels afloraments fracturats i l'ús de machine learning. Finalment, (C) s'inclouen (marcats amb fons gris) dos fluxos de treballs resumits, on s'han utilitzat aquestes tècniques. En el cas de l'esquerra, s'han utilitzat les tècniques remotes per la caracterització de la xarxa de fractures, amb l'objectiu d'avaluar-les a l'hora de construir un model DFN, identificant les avantatges i desavantatges de cada tècnica segons la propietat representada. En el cas de la dreta, hem disposat de les dades de monitoratge d'un escarpament rocós escanejat amb TLS des de 2009, on a partir dels núvols de punts, s'identifiquen els blocs caiguts. S'ha desenvolupat una nova metodologia introduint eines de machine learning, per l'automatització de la classificació de blocs caiguts i precursors, substituint així, un procés d'identificació pràcticament manual. Aquest desenvolupament, està publicat a l'article (Blanco et al., 2022).

3.1. Introducció

La digitalització dels medis fracturats amb tècniques remotes permeten, per una banda, caracteritzar la xarxa de fractures presents en els afloraments i construir models discrets de fractures que es podran utilitzar per la simulació de flux i per altra banda, aquesta digitalització, repetida en el temps en un mateix aflorament (monitoratge), permet comparar i determinar diferències entre els escanejors e identificar els blocs caiguts o que estan a punt de caure (precursors). També el tractament digital de milions de punts, amb la introducció d'algorismes de *machine learning* s'ha utilitzat en aquesta tesi doctoral, per l'automatització de processos fins ara manuals (veure Figura 3.1).

Aquest capítol s'ha dividit en dos seccions. A la Secció (3.2), s'avaluen les tècniques remotes en la representació de xarxes de fractures en models DFN i a la Secció (3.3), es presenta una metodologia per a l'automatització de la classificació de blocs caiguts a partir de tècniques remotes i *machine learning*.



3.2. AVALUACIÓ DE LES TÈCNiques REMOTES EN MODELS DFN APLICATS A LA SIMULACIÓ DE FLUX

Per avaluar la construcció de models DFN amb dades capturades amb tècniques remotes, caldrà caracteritzar geomètricament la xarxa de fractures amb la digitalització dels afloraments amb tècniques remotes, TLS i fotogrametria. També, es recullen dades directament amb els *scan-lines* (tècniques manuals). A partir dels resultats prèviament processats, s'han construït 3 models DFN, un per a cada tècnica i s'han identificat les diferències de cada tècnica en la parametrització de les principals propietats geomètriques de les fractures capturades; posició, orientació, espaiat, alçada i longitud.

L'estudi de la zona, s'ha complementat amb un tall geològic d'orientació E-W i una columna litològica, a més de realitzar la presa de 5 mostres sobre l'aflorament. Mostres que han servit per identificar les principals textures i datar els materials amb làmines primes i obtenir la porositat i permeabilitat de la roca amb anàlisis de laboratori, ampliant els coneixements del medi que conté les fractures.

Finalment, la realització d'estudis complementaris, amb tècniques geofísiques i el recull de sondejos bibliogràfics, estan dirigits a obtenir informació de l'interior del volum estudiat, prèviament a les voladures. Per una banda, identificant el comportament del flux a través d'ell i per altra banda, ampliant el coneixement de la distribució i mida de les fractures. Aquesta última part, també és necessària per calibrar els models de fractures obtinguts.

L'àrea d'estudi és una zona extractiva aflorant dins de la Pedrera La Fou de Vallirana, on s'han pogut estudiar els fronts d'explotació que es succeïen a mida que l'explotació avançava, obtenint abundants dades de les fractures amb tècniques remotes, tècniques directes (*scan-line*) i altres tècniques o mostrejos.

3.2.1. Metodologia de treball

La metodologia de treball utilitzada per a la construcció dels models DFN mitjançant tècniques remotes i posterior avaluació, segueix el diagrama de flux de la Figura 3.2, consistent en 4 etapes que es descriuen, breument a continuació: la captura de dades, el processat de les dades, els resultats i l'avaluació o discussió dels resultats obtinguts.

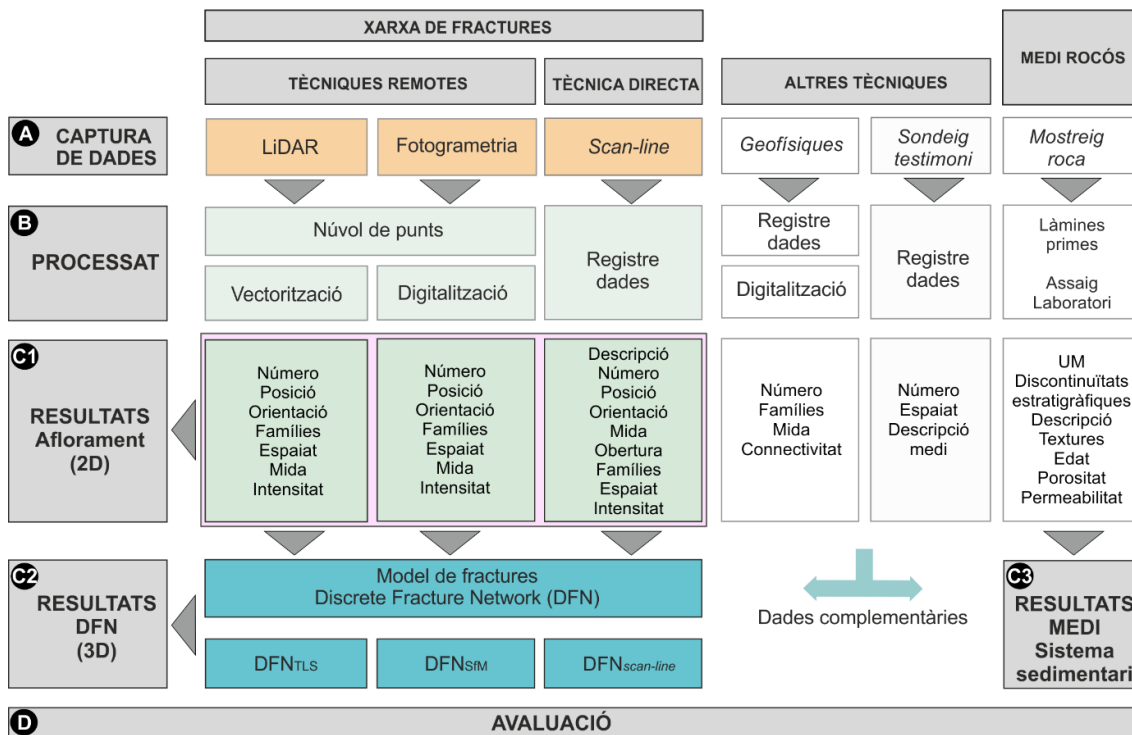


Figura 3.2: Etapes realitzades durant la investigació per avaluar les tècniques remotes aplicades en la modelització discreta de fractures (DFN) tridimensionalment. A la etapa A i B, es capturen i processen les dades per a ser analitzats. Hi ha tres grups de dades; (1) les recollides sobre els aflorament per caracteritzar la xarxa de fractures (TLS, Fotogrametria i scan-lines), (2) les obtingudes per a caracteritzar el medi rocós, un sistema sedimentari i (3) les complementàriament, geofísiques o sondejos bibliogràfics, per calibrar i corroborar les dades. A la etapa C, es mostren, en tres part, els resultats obtinguts; (C3) els que caracteritzen el sistema sedimentaris que conté les fractures, (C1) els que mostren les propietats de les fractures per a cada aflorament i tècnica i (C2) els que presenten els tres models DFN tridimensional, un per a cada tècnica aplicada. A la etapa D s'avaluen i discuteixen els resultats.

(A) Captura de dades i pre-processat de les dades

La digitalització dels afloraments es duu a terme mitjançant l'aplicació de les tècniques remotes TLS i fotogrametria. Apart, també es procedeix a la recollida de dades de les fractures mitjançant el desenvolupament de *scan-lines* (Figura 3.2A i B). Tal i com s'ha esmentat en apartats anteriors, mitjançant l'ús de *scan-lines* s'obtenen dades de les superfícies de les fractures preses directament sobre l'aflorament, és a dir, sense la necessitat de realitzar cap processat. En canvi, en el cas de les tècniques remotes, el procés de captura permet obtenir un núvol de punts que, tal i com es descriu a la Secció 2.4.1), cal processar individualment per tal de reconstruir les

superfícies de les fractures identificades. Aquest pre-processat cal aplicar-lo per cada tècnica i en cadascun dels afloraments subjecte a estudi.

(B) Resultats

En base al processat i anàlisi de les dades preses a camp, s'obtenen tres tipus de resultats (veure Figura 3.2C).

Per una banda (Figura 3.2C3), s'obté la caracterització del sistema sedimentari que conté les fractures, establerta en base al tall geològic i la columna lito-estratigràfica, la descripció detallada de les làmines primes, els assajos de laboratori realitzats per calcular la porositat i la permeabilitat de la matriu rocosa i la identificació de les unitats mecàniques.

En un segon apartat de resultats, (Figura 3.2C1) es processen les dades de les fractures preses sobre cada aflorament mitjançant les tècniques de TLS, Fotogrametria i *scan-line*, amb l'objectiu d'identificar, per cada tècnica i per cada aflorament, les principals propietats geomètriques de les fractures (veure Figura 3.2C). Finalment, a partir de les dades processades (Figura 3.2C2) s'elaboren tres models DFN tridimensionals, un model per tècnica aplicada; DFN_{TLS} , DFN_{SfM} i $DFN_{scanline}$.

En aquest darrer sentit, i una vegada generats els 3 models, es va poder apreciar com la informació de sortida que ofereixen és similar, presentant algunes dades comunes, com són les principals discontinuïtats estratigràfiques i les obertures, aquestes últimes verificades a partir dels *scan-lines*.

(C) Avaluació dels resultats o discussió

En aquesta última etapa s'analitzen, avaluen i es discuteixen els resultats obtinguts, amb l'objectiu d'identificar les avantatges o desavantatges de cada tècnica quan s'utilitzen per parametritzar les propietats geomètriques de les fractures capturades amb tècniques remotes, i valorar la seva representativitat en models DFN per simulacions de flux o per altres aplicacions en les quals, la investigació de la fracturació del massís pren rellevància. (veure Secció 3.2.7 Discussió).

3.2.2. Aplicació en la Pedrera La Fou a Vallirana

L'àrea d'estudi dins de la pedrera de La Fou, se situada al terme municipal de Vallirana (Barcelona) al nord del massís del Garraf (Picart et al., 2016), formant part del bloc de materials mesozoics amb què presenten estratificació tabular, encara que amb un suau basculament en direcció W-SW i una fracturació molt desenvolupada (Vega et al., 1973). (veure Figura 3.3).

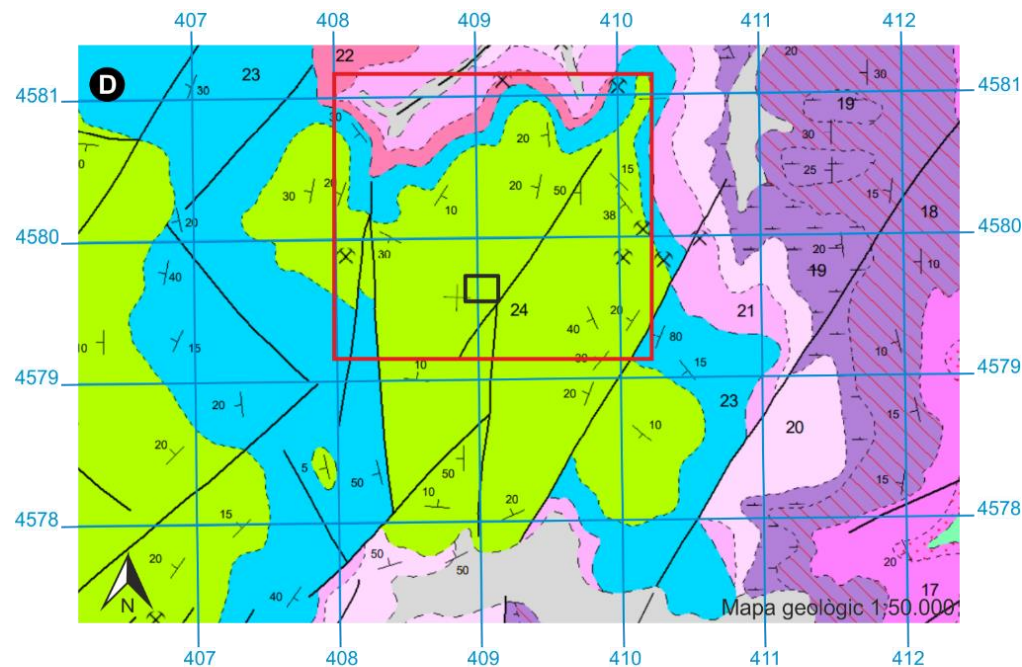
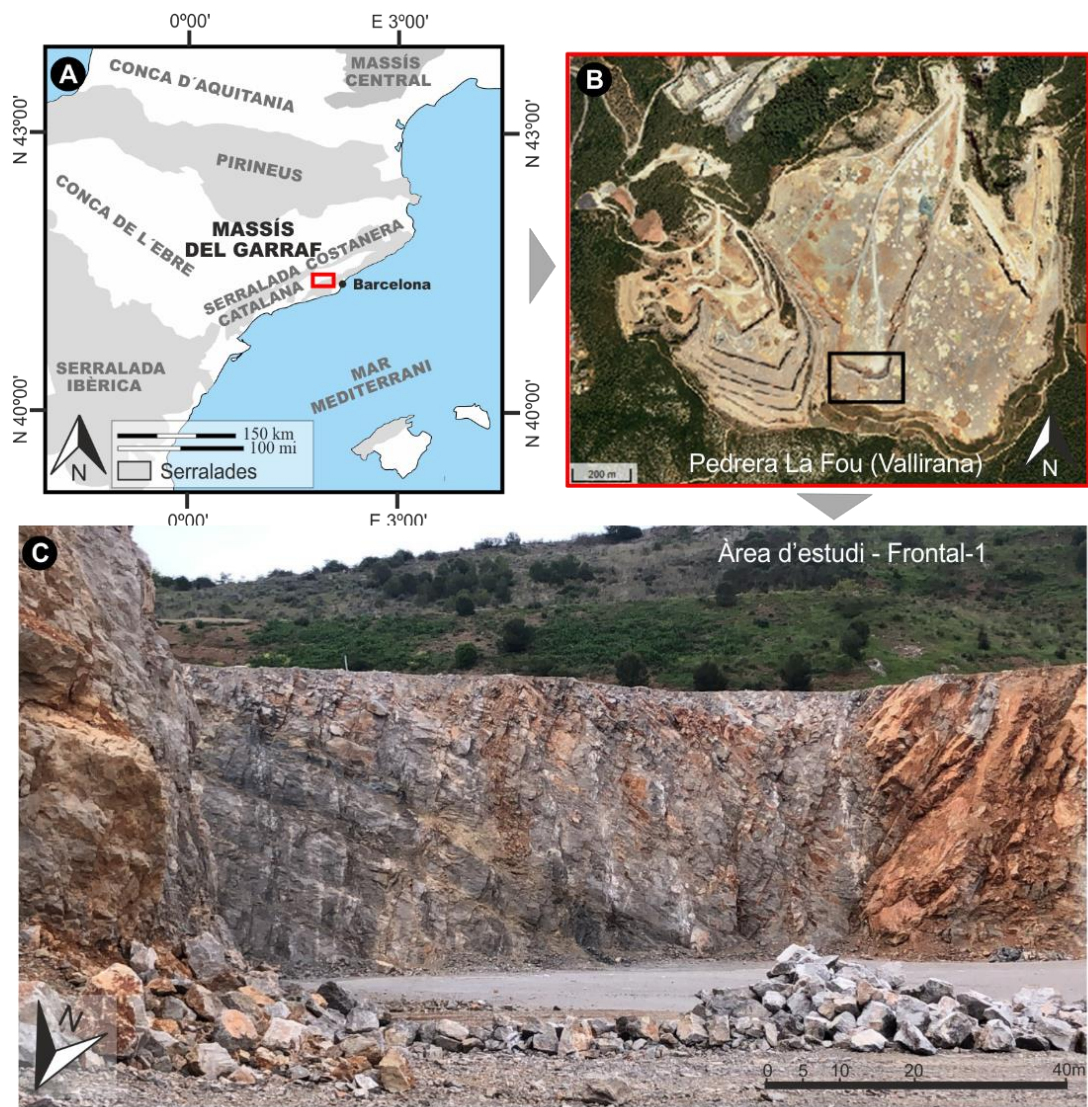
El conjunt de materials mesozoics aflorants a l'àmbit de la pedrera són principalment carbonats dipositats durant el trànsit del Juràssic superior-Cretaci inferior, caracteritzats pel desenvolupament de successius episodis transgressius i regressius, que donen lloc a la creació

de grans plataformes carbonatades. Durant aquests períodes es dipositen fonamentalment roques calcàries i dolomies.

Aquests materials, tot i que estan afectats per varis sistemes de fractures, es distribueixen a grans trets seguint la seva sèrie estratigràfica de NE a SW, situant-se els materials més antics al NE i els més moderns al SW, en consonància amb el seu cabussament vers el SW.

L'aflorament estudiat, d'edat cretàica, es troba integrat per unes capes potents i extenses de calcàries de color gris clar o blanquinós de potències decimètriques, que s'intercalen amb nivells de marges grises a groguenques de poca potència. També presenta superfícies estratigràfiques de potència decimètrica, falles normals de baix angle i de petit salt, juntament amb una fracturació vertical important. Aquests paquet d'estructures constitueixen les discontinuïtats mecàniques en part derivades de l'activitat tectònica de l'àrea.

La zona destaca per un alt grau de carstificació i la pedrera i la zona d'estudi no n'estan exemptes, encara que de forma molt localitzada (Daura et al., 2014). El front en explotació estudiat, amb una orientació ESE-WNW, rep el nom de Fr-3, en la toponímia de l'empresa Ciments Molins S. A. El front extractiu té una litologia calcària amb unes dimensions de 80 m d'amplada per 25 d'alçada.



EDAT			DESCRIPCIÓ	
CRETÀCI INFERIOR	Aptiense	25	25 Margues grises blavoses amb sostre de bicalcarenites d'orbitolines	
	Barremià	24	24 Calcàries massives de rudites, calcàries lacustres i calcarenites grises.	
	Hauteriviense	23	23 Dolomies grises fosques de gra groller	
	Valanginiense	22	22 Camioles i dolomies blanquinoses	
JURÀSSIC	Portlandiense	21	21 Margues irisades i guixoses.	
		20	20 Calcàries grises i dolomies.	
TRIÀSIC	MUSCHEL KALK	Keuper	19	19 Margues i arenisques vermelloses amb lenticions de guix.
		SUPERIOR	20	
		MEDIO	18	18 Calcàries massives i dolomies grises
	Buntsandstein	INFERIOR	17	
			16	

Pedrera La Fou
 Àrea d'estudi

Figura 3.3: Imatge A, B, C, plànols de situació de l'àrea d'estudi, situada al SE del massís del Garraf, a la Pedrera la Fou. Imatge D: Mapa geològic 1:50.000 (Vega et al., 1973). Font: IGME. Emmarcat en vermell, els límits de la Pedrera (imatge B) i emmarcat en negre l'àrea d'estudi, en materials d'edat Cretàcic inferior.

La fracturació de la cobertora mesozoica del Massís del Garraf en aquesta zona segons (Guimerà, 1984) és a conseqüència del que s'interpreta com cisalla fràgil originada per dues fases compressives alpines. Durant el Neogen es dona lloc a una fase distensiva que fa jugar com a falles normals la fracturació generada per les fases compressives. Més en detall Guimerà, (1982), proposa un primer esquema sintètic de l'evolució de la fracturació a la cobertura del Massís del Garraf, tal com es mostra a la Figura 3.4), on; (A) s'observa l'existència de dos falles conjugades; aproximadament NNW-SSE i E-W. Aquests dos sistemes de fractures conjugades es produïrien per una compressió NW-SE, que degué bisecar l'angle agut entre ambdues. (B) Després, una fase compressiva amb un escurçament NE-SW, que segons les dades, sembla aprofità les falles ja existents en comptes de formar-ne de noves. (C) Al final de la fase compressiva, podria haver-hi una petita rotació que faria actuar l'escurçament NW-SE, (les falles destres NW-SE quedarien orientades perpendicularment a la direcció d'escurçament, i se'n bloquejaria l'actuació com a falles de direcció, la qual cosa afavoreix la formació d'un nou sistema de falles destres conjugat de les sinistres de direcció E-W i que tindria una direcció NNE-SSW). Esquerdas de tensió direcció NW-SE i NE-SW. (D) Finalment, una fase distensiva on l'allargament principal $X=NW-SE$, l'eix intermedi $Y = NE-SW$ també és d'allargament; les falles E-W, NNE-SSW i NNW-SSE actuen com a falles normals.

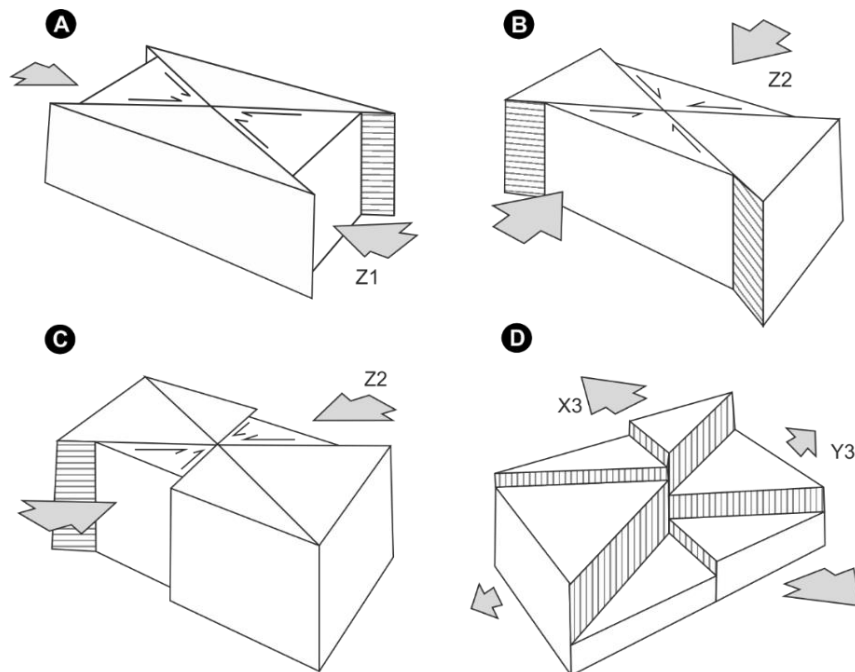


Figura 3.4: Esquema sintètic de l'evolució de la fracturació a la cobertura del Massís del Garraf, extret de Guimerà, (1982).

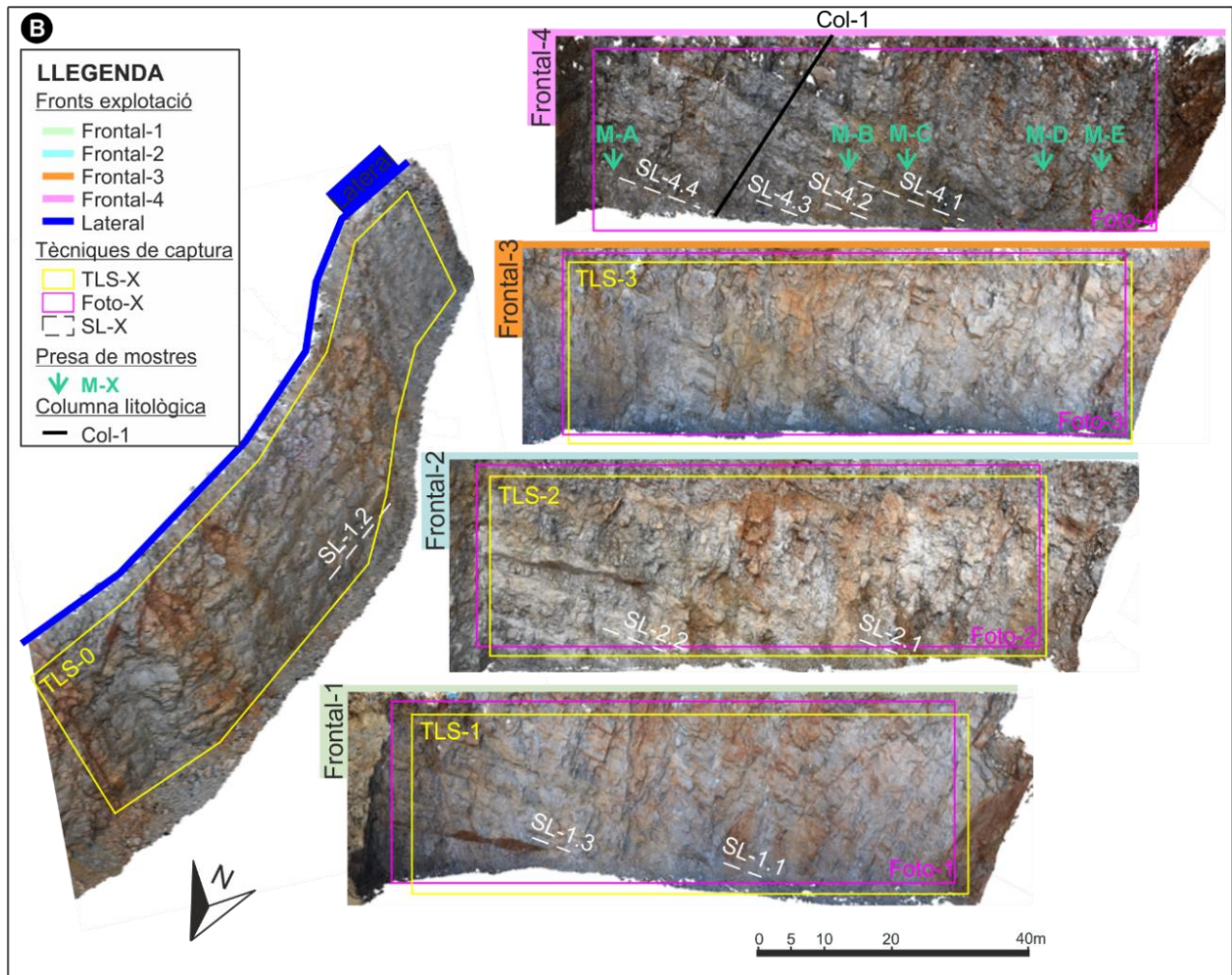
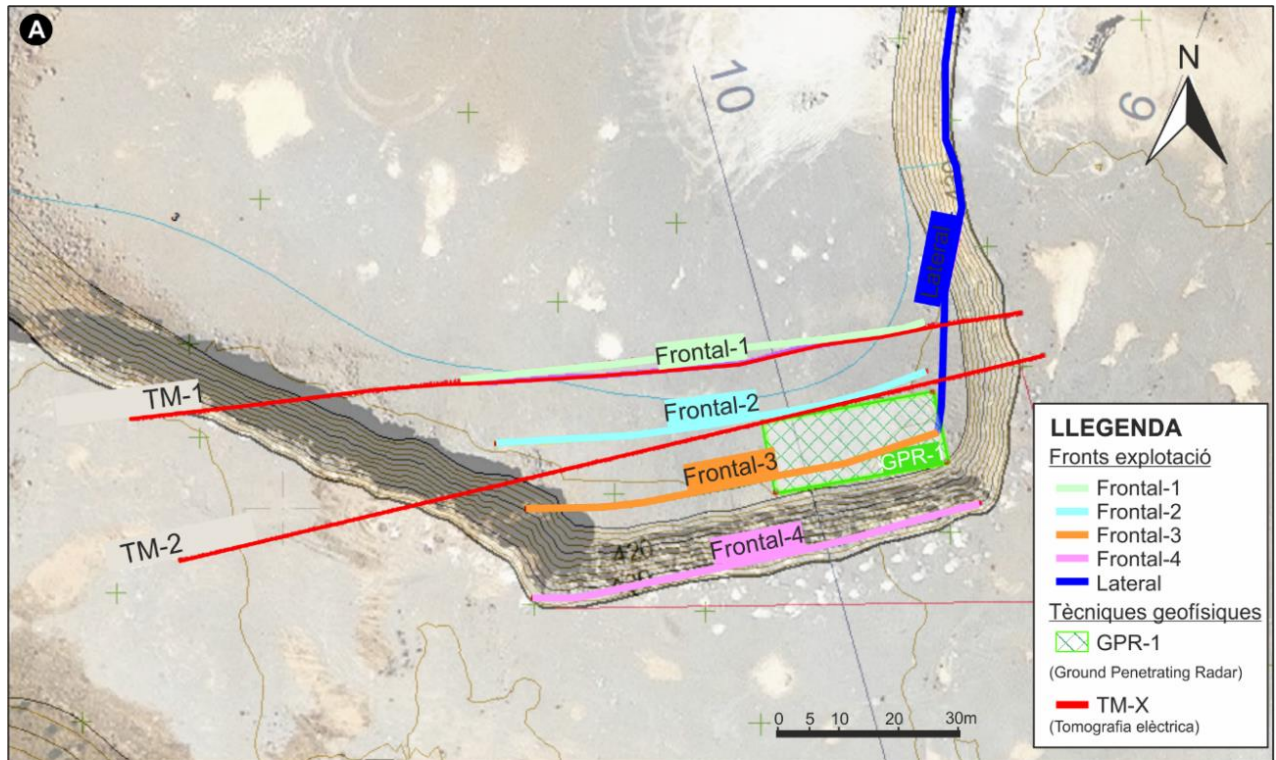
3.2.3. Adquisició de dades

Entre març de 2019 i març de 2020, Ciments Molins va realitzar tres voladures i es va tenir accés a 4 frontals consecutius d'orientació cap al nord i al lateral d'orientació cap a l'oest, que va quedar exposat al produir-se el retrocés dels frontals. La planificació de les campanyes de camp va estar condicionada a l'activitat extractiva. A la Taula 3.1, es presenta la relació cronològica dels principals treballs realitzats i que es descriuen en els següents apartats. Les situacions dels treballs realitzats durant les campanyes de camp es representen gràficament a la Figura 3.5.

	Gener	Febrer	Març	Abril	Maig	Juny	Juliol	Agost	Setembre	Octubre	Novembre	Desembre
2019	Frontal 1			Frontal 2								
	1r escaneig TLS 1r escaneig Foto 3 scan-lines			2 perfils PTM						Cub GPR		2n escaneig TLS 2n escaneig Foto 2 scan-lines
2020	Frontal 3			Frontal 4								
	3r i lateral escaneig TLS i Foto			4t escaneig Foto 4 scan-lines						1 Columna 1 Perfil geològic 11 mostres de roca		

Taula 3.1: Cronograma de les campanyes de camp realitzades a la Pedrera de La Fou de Vallirana. A la Taula, s'indiquen els trams temporals on queden exposats cadascun dels frontals mitjançant colors; verd (Frontal-N-1), blau (Frontal-N-2), groc (Frontal-N-3 i Lateral-W) i rosa (Frontal-N-4). Els límits responen a les dates de detonació: 27/03/2019, 7/1/2020 i 25/3/2020. Finalment, s'inclouen les tasques realitzades per mesos. Els acrònims utilitzats són els següents: TLS: Terrestrial Laser Scanner, Foto: Fotogrametria, PTM: Perfils de Tomografia elèctrica, GPR: Cub de Ground Penetrating Radar i Columna: Columna litoestratigràfica.

Figura 3.5: (veure pàgina següent) Situació en planta i en alçat de les captures realitzades a l'àrea d'estudi. En la part superior de la Figura, l'anomenada (A), es mostra el plànol en planta de l'àrea d'estudi, amb la situació dels fronts d'explotació, els perfils tomogràfics i el cub de GPR estudiat. La imatge es correspon a la situació final, després de totes les voladures, on s'observa el límit del frontal-N-4. Cal tenir en compte que les tomografies elèctriques i el cub GPR, es van fer prèviament a les voladures sobre el bancal superior. En la meitat inferior, la (B), es mostra la representació gràfica dels 4 frontals estudiats amb orientació N i el lateral-W, a més de situar els límits recollits en el escaneig amb TLS i la captura amb fotogrametria. També s'inclouen la situació dels scan-lines. Notar que la orientació en el apartat B és diferent que en el apartat A.



Dades capturades i processades amb tècniques remotes i scan-lines

Durant les diferents campanyes de camp, s'han capturat digitalment els afloraments amb tècniques remotes (TLS i fotogrametria) i recollit dades directament sobre l'aflorament mitjançant *scan-lines*. A diferència dels *scan-lines*, les dades adquirides amb tècniques remotes han necessitat de la reconstrucció de les superfícies de les fractures, per tal d'obtenir les dades de posició, orientació i dimensions de les superfícies identificades.

Digitalització amb Fotogrametria

La tècnica de captura es va aplicar sobre els quatre frontals d'orientació cap al nord i sobre el lateral d'orientació cap a l'oest de la pedrera. Tal com requereix la metodologia els afloraments es van fotografiar multitud de vegades des de diferents punts de vista i amb un alt grau de solapament entre elles. El processament per obtenir el núvol de punts es va realitzar amb el programari d'SfM (veure Subsecció 2.3.3) Agisoft Metashape. Les dades en detall estan mostrades a la Taula 3.2).

Fotogrametria					
Nom	Frontal	Número de punts	Núm. Imatges	Distància focal (mm)	Distàncies mitges (m)
Foto-1	1	36.449.223	58	18	11
Foto-2	2	58.443.038	43	36	20
Foto-3	3	18.740.583	54	24	15
Foto-4	4	140.009.144	49	24	8

Taula 3.2: Quadre resum de les dades capturades amb tècniques de Fotogrametria (remotes)

Un model de càmera DSLR (Digital Single Lens Reflex) Nikon D5100 de múltiples objectius va registrar el *Frontal-N-1* i la resta d'aflorament es va utilitzar una càmera DSLR de múltiples objectius Nikon D850. La resolució de la primera càmera es de 4928 x 3264 píxels i la segona càmera de 8256 x 5504 píxels. Les distàncies mitges al front varien entre 8 i 20 m depenent de la distància a la qual es podia accedir al front, distància focal i càmera utilitzada.

La configuració per els processos d'ajust fotogramètric amb el programari comercial *Agisoft Metashape* han estat de "High Accuracy" sense limitació de punts de clau i en el pas d'alineació de les fotografies, optimitzant la calibratge dels paràmetres. Durant el procés de reconstrucció de la superfície de l'aflorament els paràmetres per la qualitat de generació dels mapes de fondària varien entre "Medium" i "High" i els mode de filtració entre "Aggressive" i "Mild". El producte final d'aquest procés són els diferents núvols de punts de cada aflorament amb un nombre de punts mostrats a la Taula 3.2.

Les fractures contingudes en els núvols de punts fotogramètrics han estat identificades manualment durant la seva digitalització amb el programari *Metashape* (veure Subsecció 2.4.1). De la mateixa manera s'han digitalitzat les superfícies d'estratificació més destacades (veure figura 3.6). Les fractures s'han exportat en format dxf (*Drawing Exchange Format*) al programari

Afloraments fracturats digitalitzats.

Avaluació de les tècniques remotes en models *DFN* i aplicació de *Machine Learning*.

Microstation (Bentley) per mesurar la posició, orientació i dimensions per cada fractura identificada, amb un total de 1.094 superfícies caracteritzades.

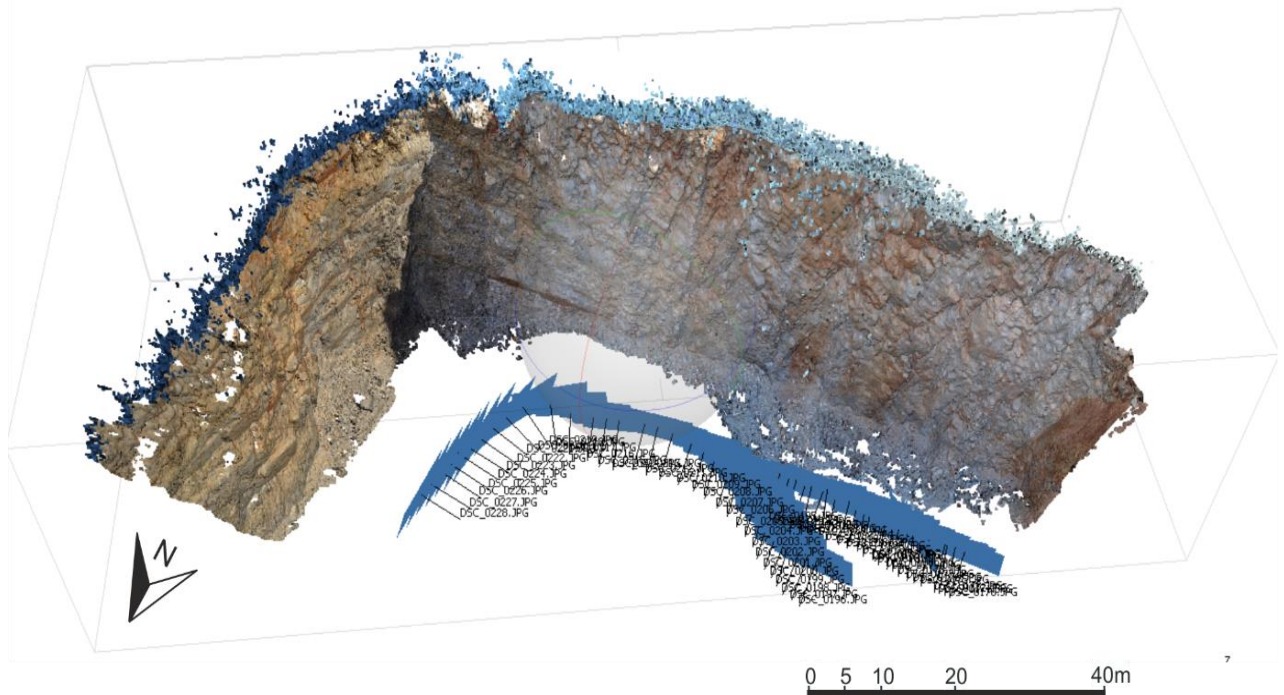


Figura 3.6: Imatge del Frontal-N-1, amb la posició de les fotografies captades.

Escaneig amb *Terrestrial Laser Scanner (TLS)*

Els tres primers Frontals d'orientació cap al nord i el Lateral d'orientació cap a l'oest es van capturar digitalment amb un model de làser escàner ILRIS-3D de la companyia *Teledyne Optech*. El número d'estacions per escanejar aflorament va ser de tres i les seves posició van ser establerta per GPS diferencial amb modalitat de post-procés, donant un error mig de cada posició de 0.05m. L'angle de cobertura de l'escàner és de 40° i per cobrir els afloraments amb unes densitats entre 1.2 i 0.3 punts per cm² es va haver d'anar rotant l'aparell en cada estació. El nombre d'imatges preses en cada estació, per tant varia de dos a quatre. Per cobrir tots l'àrea en cada posició s'anava rotant l'escàner prenent una imatge en cada rotació. El número final d'imatges varia entre 12 i 8 depenent de l'àrea de solapament i els laterals. Posteriorment amb el software *PolyWorks* les imatges es van alinear en un únic sistema de referència i posteriorment traslladades al sistema *Universal Transverse Mercator (UTM)*, fus 31T, datum WGS84.

L'objectiu de l'escaneig era registrar l'aflorament amb suficient resolució per poder identificar la superfície de les fractures posteriorment en la etapa de post-procés. La (Taula 3.3) mostra les propietats de cada escaneig.

TLS						
Nom	Frontal	Número punts	Núm. Estacions	Núm. Imatges	Longitud (m)	Densitat (Punts/cm ²)
TLS-1	Frontal-N-1	19.064.220	3	12	80	1.27
TLS-2	Frontal-N-2	19.243.951	3	12	75	1.2
TLS-3	Frontal-N-3	8.552.358	3	8	82	0.54
TLS-0	Lateral-W	8.484.001	3	5	113	0.3

Taula 3.3: Quadre resum de les dades capturades amb tècniques de TLS (remotes)

El resultat de la captura digital de cada aflorament són els núvols de punts amb un format de les coordenades cartesianes de cada punt i la intensitat del retorn del senyal emès per l'escàner. El pas següent consisteix a identificar els punts que representen les superfícies de les fractures, per reconstruir-les (veure Subsecció 2.3.2) i mesurar la posició, orientació i dimensions de cadascuna de les fractures reconstruïdes, que han sigut 52.507 pel conjunt de frontals orientats cap al nord i 50.554 pel lateral orientat cap a l'oest.

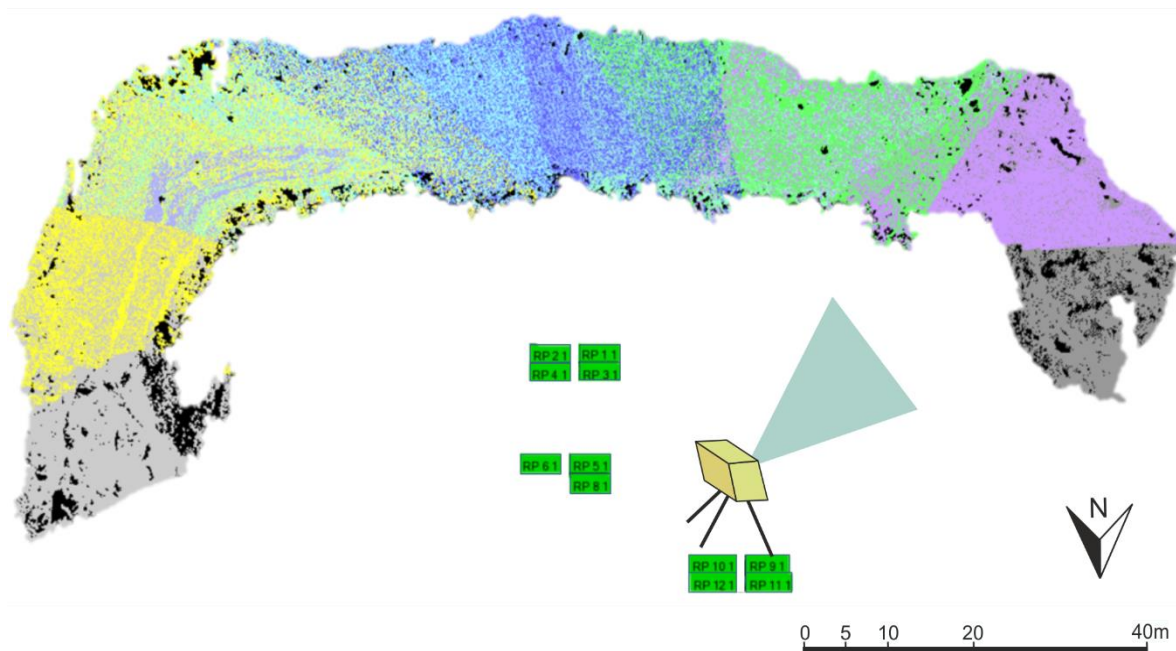


Figura 3.7: Imatge del Frontal-N-1, amb la posicions de les 3 estacions imatges i el número de imatges per estació. Els diferents colors del núvol de punts corresponen a les diferents imatges.

Scan-lines

Les dades registrades en els 9 *scan-lines* caracteritzen 398 fractures en un total de 89m.l, distribuïdes tal i com es mostra a Taula (3.4).

Scan-line			
Nom	Frontal	Longitud (m)	Núm. Fractures
SL-1.1	Frontal-N-1	16.80	15
SL-1.2	Lateral-W	2.70	6
SL-1.3	Frontal-N-1	6.80	22
SL-2.1	Frontal-N-2	10.40	34
SL-2.2	Frontal-N-2	8.90	28
SL-4.1	Frontal-N-3	18.40	61
SL-4.2	Frontal-N-3	5.10	68
SL-4.3	Frontal-N-3	5.50	68
SL-4.4	Frontal-N-3	12.40	96

Taula 3.4: Quadre resum de les dades capturades amb tècniques d'*scan-line* (directes)

S'han recollit dades de la posició, la orientació, la mida, les obertures, els reompliment i l'espaiat entre fractures. Els *scan-lines* s'han realitzat amb una inclinació d'entre 20-30° cap a l'oest dins d'una mateixa unitat mecànica. No s'han pogut realitzar *scan-lines* a totes les unitats identificades, ja que la disposició d'aquestes ha estat condicionada per els accessos i el perill de caiguda de blocs en el aflorament, de la mateixa manera, la longitud dels *scan-lines*, habitualment ha sigut curta perquè la inclinació de 20-30° de la línia d'*scan-line*, feia que en pocs metres no hi hagués accés a les fractures.

Preses de mostres, làmines primes i assaigs de laboratori

Les mostres per a l'anàlisi de les microfàcies, han estat preses sobre el Frontal-N-4 (situació a la Figura 3.5), a partir de les característiques petrològiques observades amb la lupa de mà i de la seva distribució, de tal manera que també són representatives de tota l'àrea analitzada. El total de làmines primes descrites fetes a partir de cada mostra es detalla a la Taula 3.5).

Mostra	Número de làmines primes
M-A	1
M-B	1
M-C	2
M-D	1
M-E	1

Taula 3.5: Taula resum amb la relació de mostres seleccionades per a realitzar làmines primes, segons les observacions fetes amb la lupa de mà.

Les **làmines primes** s'han estudiat amb un microscopi Zeiss Axiophot, amb l'objectiu de descriure les microfàcies de les unitats mecàniques identificades i determinar l'edat de la successió estudiada a partir de la fauna fòssil.

A partir de les mostra assajades i de les observacions a camp, s'ha confeccionat una columna litològica al fornal-4. La columna perpendicular al pla estratigràfic té 29m de longitud. (veure Figura 3.5)

S'ha **assajat la porositat** total (porositat connectada) i la **permeabilitat** de 4 mostres seleccionades entre en el *Frontal-N-4*, per a caracteritzar les relacions porositat-permeabilitat i identificar les característiques del medi que conté la xarxa de fractures. Les mostres han estat la M-A, M-B, M-C i M-D, buscant una bona distribució i recollint les diferents textures identificades amb la lupa i a partir també de les descriptives de les làmines primes. De cada mostra s'ha obtingut dos cilindres de mostreig, paral·lel i perpendicular al pla estratigràfic, amb mesures de 20 mm de diàmetre i 57 cm de longitud. De les 4 mostres s'han obtingut 7 cilindres per assajar, ja que un dels cilindres es va trencar en el procés de tall. La mostra M-C en direcció paral·lela al pla estratigràfic es va trencar mentre es tallava. A la taula següent (Taula 3.6), es presenta la relació de mostres assajades.

Mostra	Orientació mostra	Longitud (cm)	Diàmetre (cm)	Àrea (cm ²)	Geom. vol. (cm ³)
M-Aa	// S0	5.51	4.00	12.57	69.24
M-Ab	⊥ S0	5.22	4.00	12.57	65.60
M-Ba	// S0	5.51	4.00	12.57	69.24
M-Bb	⊥ S0	5.50	4.00	12.57	69.12
M-Ca	⊥ S0	5.50	4.00	12.57	69.12
M-Da	// S0	5.03	4.00	12.57	63.21
M-Db	⊥ S0	5.51	4.00	12.57	69.24

Taula 3.6: Taula resum amb la relació de mostres assajades al laboratori. S'indiquen les mides del cilindre i el volum que ocupa. Per indicar la orientació de la mostra s'utilitzen els símbols "// S0" quan el cilindre de la mostra s'orienta paral·lel al pla estratigràfic i "⊥ S0" quan la orientació és perpendiculars en aquest pla.

Els assajos s'han realitzat al laboratori Francès de Besancon, del Département de Géosciences UMR CNRS 6249 Chronoenvironnement, a partir dels cilindres tallats al servei de làmina prima que disposa la Facultat de Ciències de la Terra de la Universitat de Barcelona. La porositat total es va quantificar amb el mètode del porosímetre d'Arquímedes (Cavailhes et al., 2013; Monicard, 1982) amb una precisió analítica de $\pm 0,05\%$. La permeabilitat del gas es va mesurar mitjançant un permeàmetre de cèl·lules Hassler (Cavailhes et al., 2013; Mertz, 1991) mètode de flux en estat estacionari i nitrogen com a fluid, en les mateixes condicions. (pressió de confinament d'entre 1,5 i 4,2 bars i pressió de porus d'entre 1,5 i 4,2 bars). La precisió analítica de $\pm 0,5\%$ (kz 0,001 mD). Els resultats obtinguts es mostren a continuació a la Subsecció (3.2.3).

Estudis complementaris amb tècniques geofísiques

La realització de prospeccions geofísiques, prèviament als avanços d'explotació, a permès disposar de dades de l'interior del volum que s'ha de modelitzar. Amb el *Ground Penetrating*

Radar o GPR, s'estudia un cub de dimensions conegudes i amb la tomografia elèctrica es realitzen 2 perfils.

Cub estudiat amb GPR

Amb les tècniques de GPR, s'han realitzat dos experiments, la presa de dades en un cub definit, per obtenir les dimensions i distribució de les discontinuïtats detectades i la introducció d'aigua en el cub, per obtenir dades de transmissivitat de l'aigua.

Prèviament a les voladures, es delimita una superfície rectangular de dimensions conegudes, situada a la part superior del massís estudiat, com queda representat a la Figura 3.5. Amb el dispositiu GPT, d'Opera Duo de la companyia IDS GeoRadar, amb unes freqüències central de 250 i 700 MHz, es recobreix l'àrea per obtenir dades en fondària i construir un volum de dades, que responen a la resposta electromagnètica del terreny. El cub estudiat, té unes dimensions de 30 m de longitud per 10 m d'amplada i amb una previsió de penetració de 10m de fondària. Les traces digitalitzades de les discontinuïtats registrades en el cub estudiat, podrien mostrar la distribució i dimensions d'aquestes. (veure Figura 3.8).

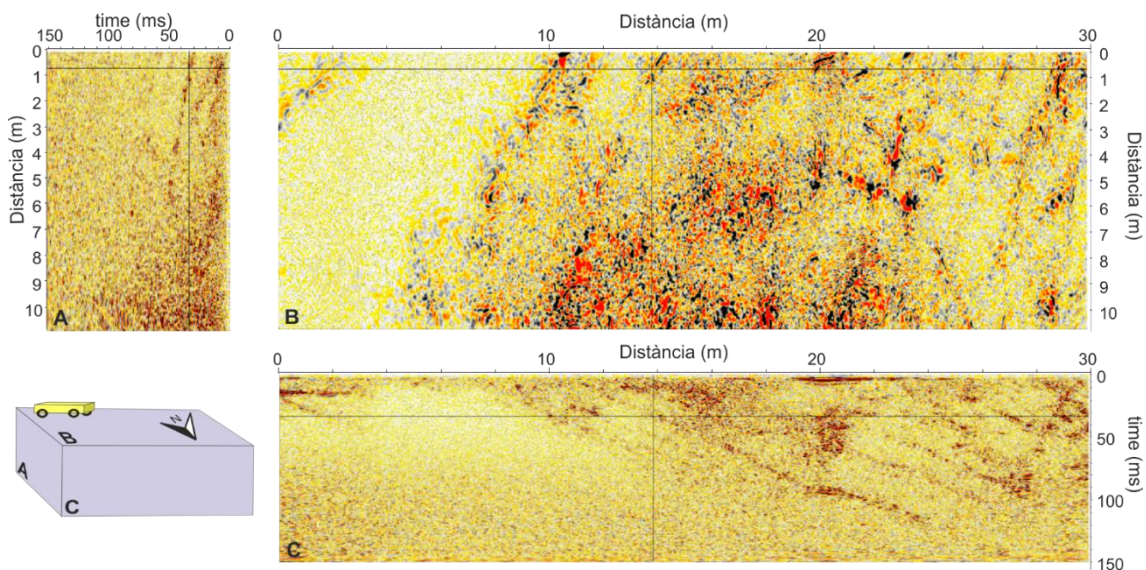


Figura 3.8: Imatges processades del cub estudiat amb GPR. La imatge B, es correspon amb la superfície situada per sobre dels afloraments estudiats, abans del avanço d'explotació, situant-se entre el Frontal-N-2 i 3. La imatge A i C mostren la fondària en funció del temps de resposta dels polsos d'energia electromagnètica transmesos en el terreny.

El segon experiment, es fa sobre la mateixa superfície definida, però amb l'aportació controlada de 12 m³ d'aigua (durant 30 minuts). En aquest ocasió, es defineix un transsecte on es passa el dispositiu GPR de forma repetida, per obtenir perfils en el transcurs de 2 hores, amb l'objectiu d'identificar la circulació de el flux a través de les fractures i poder calcular la seva velocitat. En aquest segon cas, els successiu perfils recollits, no van presentar canvis.

Ground Penetrating Radar

Nom	Frontal	Longitud (m)	Amplada (m)	Rang time (ms)	Volum estudiat (m ³)
GPR-1	2 - 3	30	10	0-150	3.000

Taula 3.7: Quadre resum de les dimensions del cub estudiat amb tècniques de GPR (tècnica geofísica).

Perfils de tomografia elèctrica

En el mes d'abril, es realitzen 2 perfils de tomografia elèctrica (TM) de 140 metres de longitud, paral·lels a la capçalera, a 4 i 10 metres d'aquesta, sobre la superfície superior de l'àrea d'estudi (veure la Figura 3.5). La fondàries de penetració va arribar a 32,10m a la part central. El perfil TM-1, es superposa aproximadament amb el Frontal-N-1 i el perfil TM-2 amb el Frontal-N-2. Les dades es recullen amb l'objectiu d'identificar heterogeneïtats atribuïbles a zones preferents de circulació de fluids, com carstificacions o falles de major ordre que alterin la homogeneïtat de la distribució de les fractures.

Perfils de tomografia elèctrica					
Nom	Frontal	Espaiat entre elèctrodes (m)	Iteració	Longitud (m)	Fondària (m)
TM-1	1	2	59	140	32,1
TM-2	2	2	59	140	32,1

Taula 3.8: Quadre resum dels perfils realitzats amb tomografia elèctrica (tècnica geofísica)

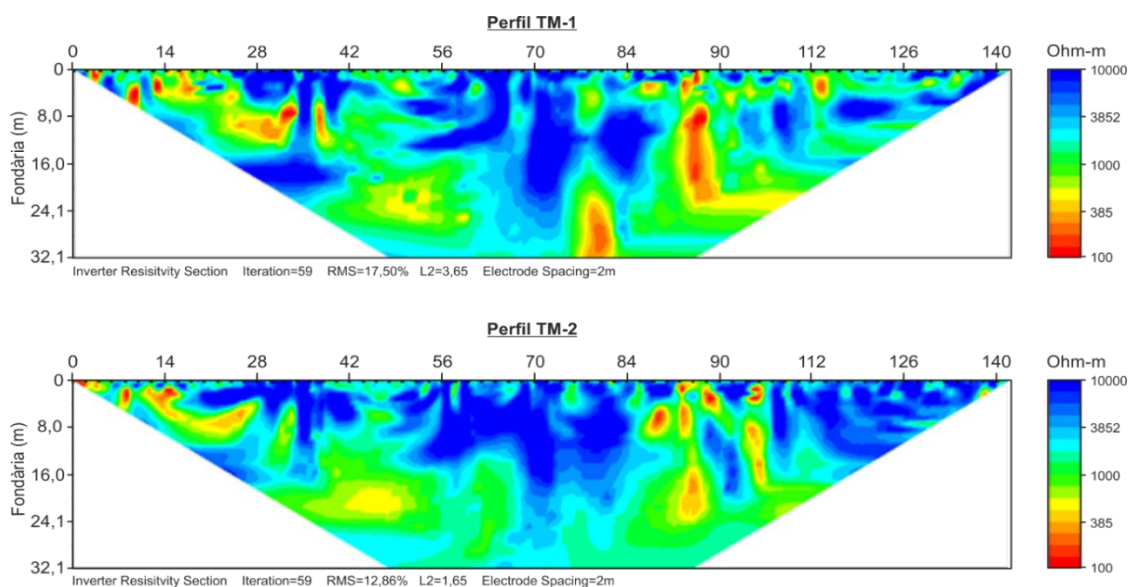


Figura 3.9: Perfils tomogràfics realitzats a l'àrea d'estudi. La longitud dels perfils és de 140 metres i assolixen una fondària màxima de 32,1m a la part central. El perfil TM-1 se situa pròxim amb el Frontal-N-1 i el perfil TM-2 està més pròxim al Frontal-N-2.

Sondejos mecànics amb testimoni continu (bibliogràfics)

S'ha tingut accés a treballs interns de Ciments Molins S.A. com l'*Etude de la cerriere calcarie des Cementos Molins*, (Bayle, 1973). En aquest treball s'analitza el volum potencialment explotable a la pedrera, amb 5 sondejos d'entre 57 i 155 m de fondària on es pot reconèixer una falla de direcció NNW-SSE que travessa tota la pedrera. En el bloc de l'est de la falla, el sondeig S-3, detecta el contacte de les calcàries amb les dolomies a una fondària de 68m, en canvi en el bloc

de l'oest, no s'arriba a detectar, perforant fins a 155m de fondària en el sondeig S-1. El treball de Sanchez (2009), *Estudi Geològic de la cantera en explotació "La Fou"*, realitzat per caracteritzar la litologia de la zona amb estudis químics de les calcàries aporta tres nous sondejos d'entre 44,6 i 60m de fondària. El sondeig S-1, es troba a 223m de distància dels fronts estudiats i s'utilitzarà en el desenvolupament de la tesi doctoral. La situació de tots els sondejos realitzats a la Pedrera la Fou, se situen a la Figura 3.10.

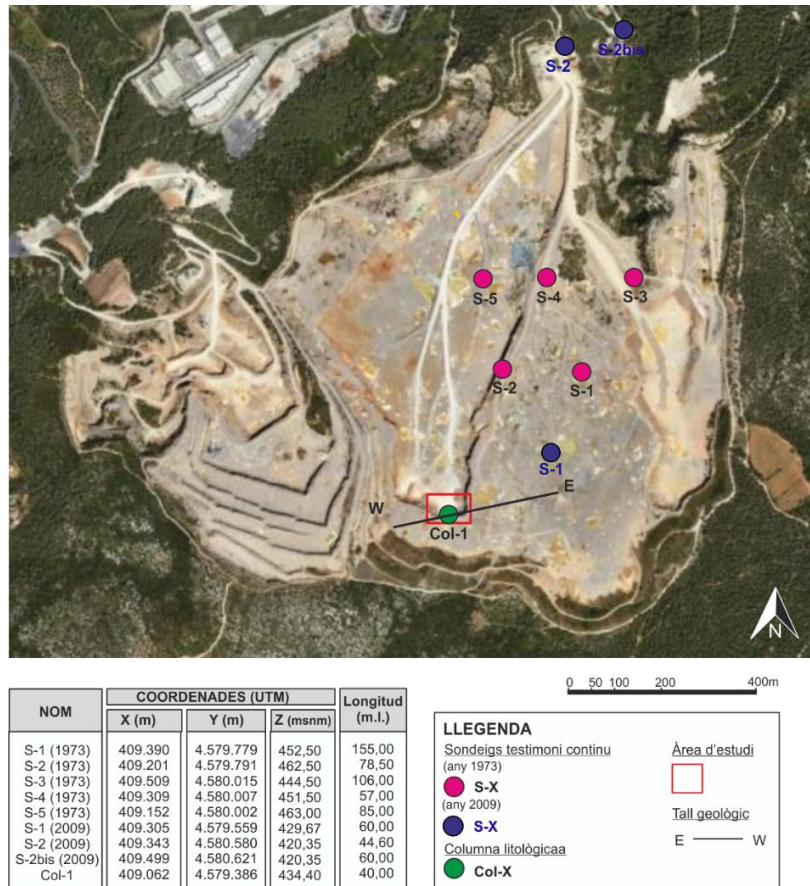


Figura 3.10: Plànol de situació dels sondejos recopilats a l'entorn de la pedrera la Fou de Vallirana i situació en planta de la columna litològica confeccionada. També es situa el tall geològic realitzat, representatiu de l'àrea d'estudi.

3.2.4. Resultats

El resultat s'organitzen en tres apartats tal com s'ha descrit al Subapartat 3.2.1; l'anàlisi previ del medi rocós, la caracterització de la xarxa de fractures i la modelització discreta de fractures.

Caracterització prèvia del medi rocós estudiat

La successió estratigràfica estudiada es caracteritza per una successió de nivells carbonats amb un lleuger basculament amb una orientació mitja de 280/23°, distribuït en paquets de gruixos centimètrics a decamètrics amb intercalació de nivells margosos de poca entitat i la presència de fauna fòssil. Destaquen 4 plans estratigràfics amb major penetració que els altres, on s'observa humitat i degoteig de l'aigua després de pluges. També queden representats a la Figura 3.13.

La successió sedimentària analitzada amb làmina prima, està formada per roques carbonàtiques marines amb textures mudstone (Figura 3.11A), wackestone (Figura 3.11B) i packstone (Figura 3.11E). Els components esquelètics més abundants que presenten aquestes textures són foraminífers bentònics com l'orbitolínid *Palorbitolina lenticularis*, miliòlids, i *Choffatella decipiens* i *Nautiloculina*. Aquestes roques també inclouen foraminífers bentònics monoseriats, biseriats, i d'altres foraminífers sense identificar. Les mostres presenten fragments de mol·luscs, equinoderms, espícules d'esponja, serpúlids i *Chondrodonta*, i de forma més puntual, fragments de braquiòpodes. També s'hi troben fragments de *Permocalculus dasycladals*. Localment també presenten bioturbació.

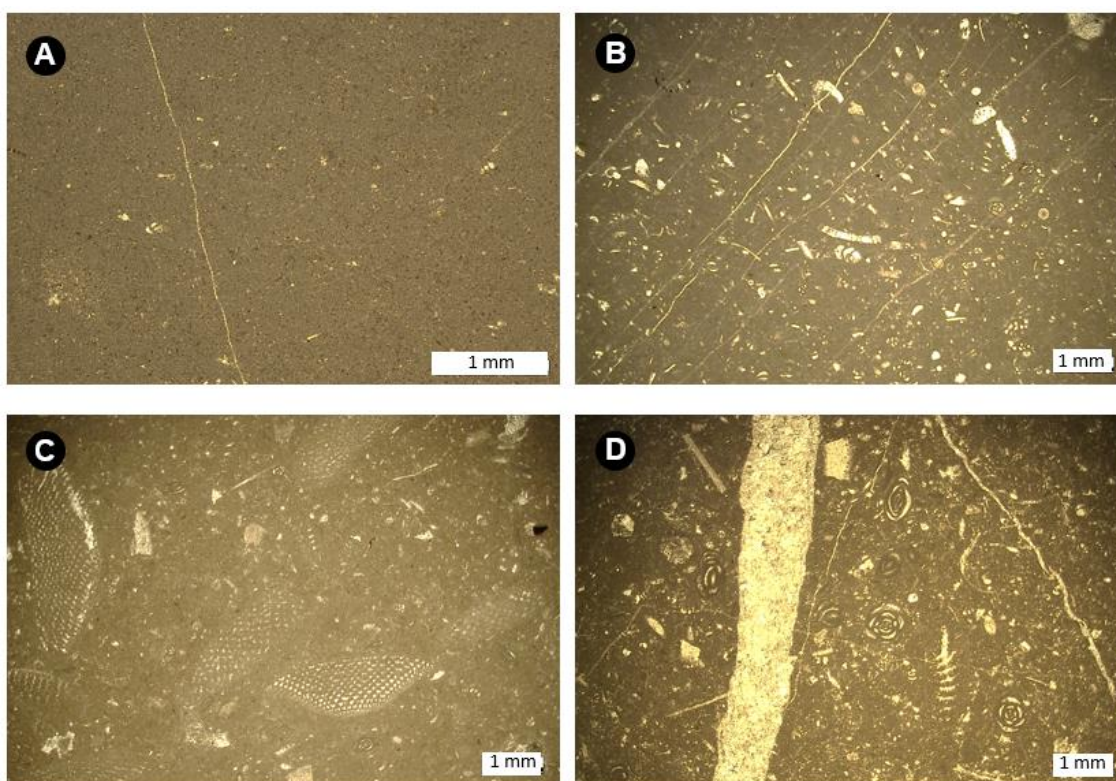


Figura 3.11: Microfàcies identificades i representatives de l'àrea d'estudi. (A) Fotomicrografia realitzada en llum paral·lela que mostra una textura mudstone. Mostra M-A. (B) Fotomicrografia en llum paral·lela, que mostra una textura wackestone amb fragments de mol·luscs i foraminífers. Mostra M-B. (C) Fotomicrografia en llum paral·lela d'una textura wackestone-packstone d'orbitolínids que també inclou *Choffatella decipiens* i altres foraminífers bentònics. Mostra M-C (D) Fotomicrografia en llum paral·lela, que mostra una textura packstone dominada per miliòlids i *Choffatella decipiens*. Mostra M-D.

La successió estudiada (Figura 3.12A) conté *Choffatella decipiens* (Figura 3.12A-C) que té un rang d'edat Barremià-Aptià inferior (Neagu i Cîrnaru, 2004). A les mostres estudiades, la presència de *Choffatella* és abundant. La presència de *Choffatella* en abundància es correlaciona amb una edat Barremià (Granier et al., 2003). El fet que les roques estudiades també presentin *Palorbitolina lenticularis* (Figura 3.12D), que té un rang d'edat Barremià superior-Aptià inferior (Schroeder et al., 2010), fa que l'edat de la successió es pugui determinar amb precisió i adscriure al Barremià superior. Altres treballs anteriors a la zona havien datat aquest registre sedimentari al Valanginià - Barremià (Vega et al., 1973) o Valanginià – Hauterivià (Picart et al., 2016).

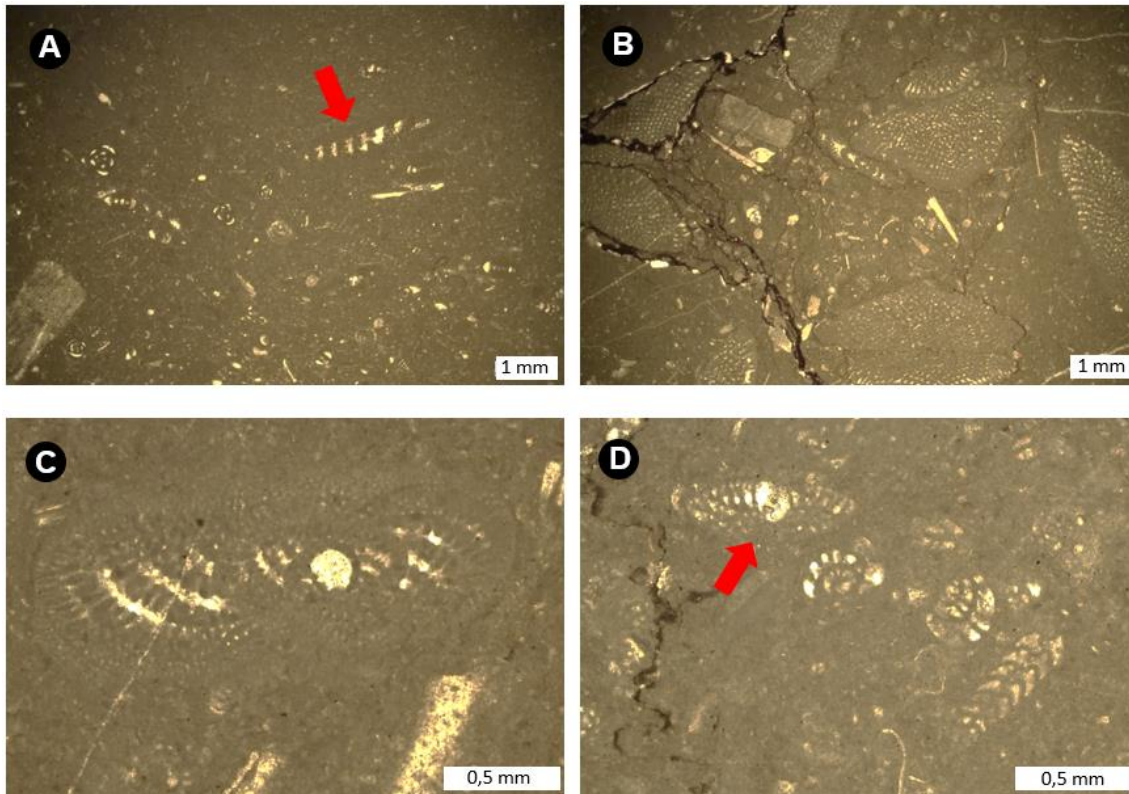


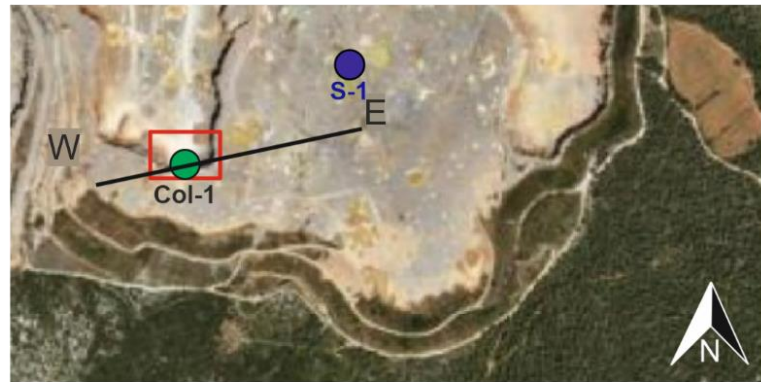
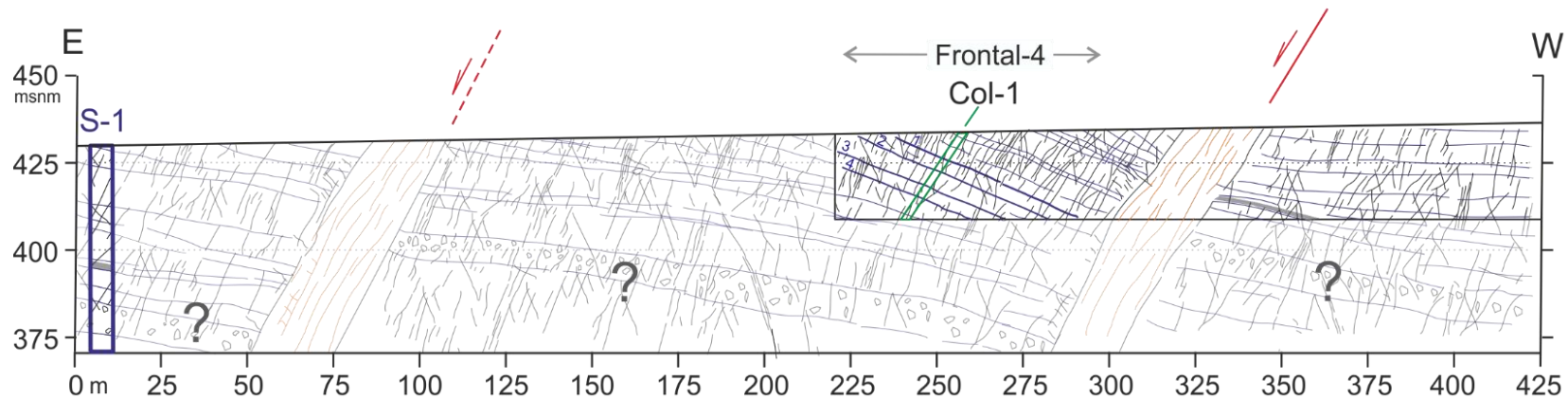
Figura 3.12: Imatges del contingut fòssil que ha permès datar la successió estudiada a La Fou de Vallirana. (A) Microfàcies wackestone-packstone de foraminífers bentònics on s'hi observa un espècimen de *Choffatella decipiens* (fletxa vermella), a més de miliòlids, orbitolínids i *Nautiloculina*, entre altres components esquelètics. Fotomicrografia en llum paral·lela. Mostra M-B. (B) Wackestone-packstone d'orbitolínids que també inclou altres foraminífers bentònics com *Choffatella decipiens* o miliòlids, a més de fragments de mol·luscs, equinoderms i espícules d'esponja. Mostra M-C. (C) Exempler de *Choffatella decipiens*. Fotomicrografia en llum paral·lela. Mostra M-C. (D) Exempler de *Palorbitolina lenticularis* (fletxa vermella) que mostra la cambra embrionària i permet fer un diagnòstic a nivell d'espècie. Fotomicrografia en llum paral·lela. Mostra M-C.

Els resultats dels assajos de porositat i permeabilitat realitzats a les mostres recollides es detalla a continuació, a la Taula 3.9.

Mostra	Porositat total*						P Ones de so		Permeabilitat**	
	M _{dry} (g)	M1 (g)	M2 (g)	Plug Vol. (cm ³)	D _{Bulk}	D _{min}	N _t (%)	T _{dry} (μs)	V _{p dry} (m.s ⁻¹)	K _{klink} (mD)
M-Aa	183.86	115.78	184.85	69.07	2.66	2.701	1.43	8.8	6261	4.30E-04
M-Ab	174.50	109.83	175.42	65.59	2.66	2.698	1.40	8.4	6214	8.50E-05
M-Ba	184.74	116.65	185.52	68.87	2.68	2.713	1.13	8.9	6191	9.13E-04
M-Bb	185.78	117.24	186.28	69.04	2.69	2.711	0.72	8.6	6395	9.76E-05
M-Ca	185.58	116.90	185.91	69.01	2.69	2.702	0.48	8.8	6250	4.22E-05
M-Da	169.89	107.16	170.20	63.04	2.69	2.708	0.49	7.7	6532	7.87E-06
M-Db	185.89	117.30	186.28	68.98	2.69	2.710	0.57	8.5	6482	2.84E-04

Taula 3.9: Resultats dels assajos de porositat i permeabilitat en les mostres assajades. Les mostres amb codi M-Xa, corresponen a les mostres realitzades perpendicularment al pla d'estratificació i les mostres amb codi M-Xb, es corresponen a les mostres realitzades paral·lelament al pla d'estratificació. * by Archimedes' method porosimetry. ** Klinkenberg-corrected gas permeability. En la columna on s'indica el nom de la mostra, la lletra a, representa la mostra tallada perpendicular al pla de referència, en canvi la lletra b, és el contrari, representa la mostra tallada paral·lela en aquest pla.

A Perfil geològic E-W



- Col-1 Columna litoestratigràfica
- S-1 Sondeig amb testimoni continu

Àrea d'estudi



LLEGENDA PERFIL/COLUMNA

Litologia

- Calcàries grises molt fracutrades i amb fòssils
- Margo-calcàries massives negres
- Calcàries bretxoides amb fòssils

- Zona de rubefacció associada a falla
- Falla identificada Falla intepretada
- Xarxa de fractures sintetitzades
- Superfícies estratigràfiques sintetitzades
- Superfícies estratigràfiques (destacades)

Principal fauna fòssil observada

- Miliòlids
- Orbitolínids
- Fragments equínids
- Fragments molúscs

B Columna litològica Col-1 (Frontal-4)

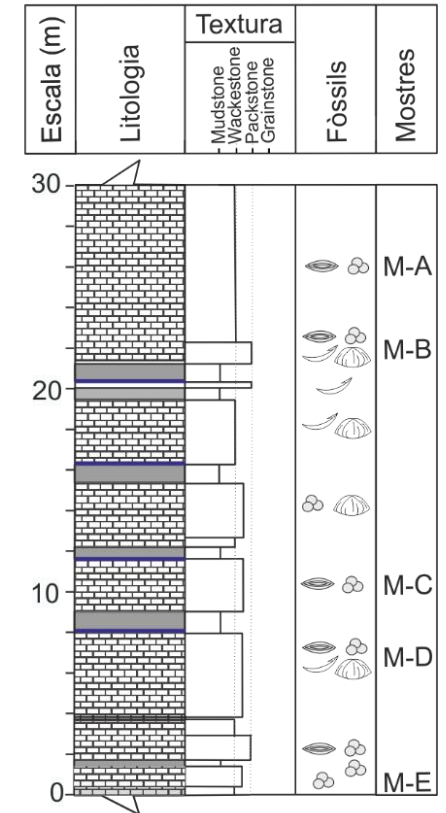


Figura 3.13: (A) Tall geològic representatiu de l'àrea d'estudi, de 425m de longitud, on es correlacionen les dades de l'aflorament (Frontal-N-1) amb el sondeig S-1. Sobre la secció s'identifica el sondeig S-1, el frontal-N-1 i la columna litològica projectada, Col-1. Les fractures i els plans estratigràfics són una representació reduïda d'aquestes. Es destaquen 4 plans de discontinuïtat estratigràfics, com els més penetratius, i es numeren. Les unitats estratigràfiques (B) Representació de la columna sintètica litoestratigràfica realitzada en el Frontal-N-4, representativa de l'àrea d'estudi. S'inclou la textura i la fauna fòssil identificada.

Caracterització de la xarxa de fractures

Les dades capturades mitjançant l'ús de tècniques remotes i les mesures amb els *scan-lines*, s'han processat mitjançant el programa SEFL, seguint la metodologia descrita a García-Sellés et al., (2018), per caracteritzar geomètricament les fractures. El processat s'ha realitzat individualment per a cada aflorament i tècnica.

Per tals efectes, en funció de les famílies de fractures i les unitats mecàniques identificades, obtenim els valors mitjos representats amb gràfics estadístics. Les propietats individuals representades són el nombre de fractures, l'alçada i la longitud i les propietats col·lectives, l'espaiat, els valors d'intensitat, P_{10} , P_{20} , P_{21} i P_{32} , el coeficient de variació i el Ratio d'espaiament de les fractures (FSR).

A continuació, es mostren la relació d'afloraments i tècniques utilitzades en el processat SEFL (veure Taula 3.10).

SITUACIÓ	TÈCNICA
Frontal 1 (EW)	Terrestrial Laser Scan Fotogrametria Scan-line
Frontal 2 (EW)	Terrestrial Laser Scan Fotogrametria Scan-line
Frontal 3 (EW)	Terrestrial Laser Scan
Frontal 4 (EW)	Fotogrametria Scan-line
Lateral (NS)	Terrestrial Laser Scan

Taula 3.10: Taula resum amb la relació d'afloraments estudiats i les tècniques aplicades.

A l'Apèndix 1 estan disponibles els gràfics de totes les característiques calculades, per família i unitat mecànica dels fronts. A continuació es presenta un recull de les dades més rellevants per propietats individuals i col·lectives, presentades per famílies i unitats mecàniques.

Famílies de fractures

Les orientacions de les fractures han estat representades en projeccions estereogràfiques amb els pols de les fractures per cada tècnica d'exploració com es pot veure en la Figura 3.14. La tècnica que ha permès identificar amb major facilitat les 4 famílies de fractures ha estat la *scan-line*, ja que les tècniques remotes destaquen les famílies amb major densitat de punts i emmascaren la resta (veure Figura 3.15). A les projeccions, es pot apreciar els màxims de densitat dels pols d'orientacions i les diferències en funció de les capacitats de cada tècnica. En total, s'han identificat quatre famílies de fractures, amb orientacions preferents ESE-WNW per la

família de fractures I. Les famílies de fractures II i III són molt properes entre elles però clarament diferenciada amb la tècnica de l'Scan-line amb unes orientació NNW-SSE per la família II i N-S per la família III. La família de fractures IV té una orientació NE-SW.

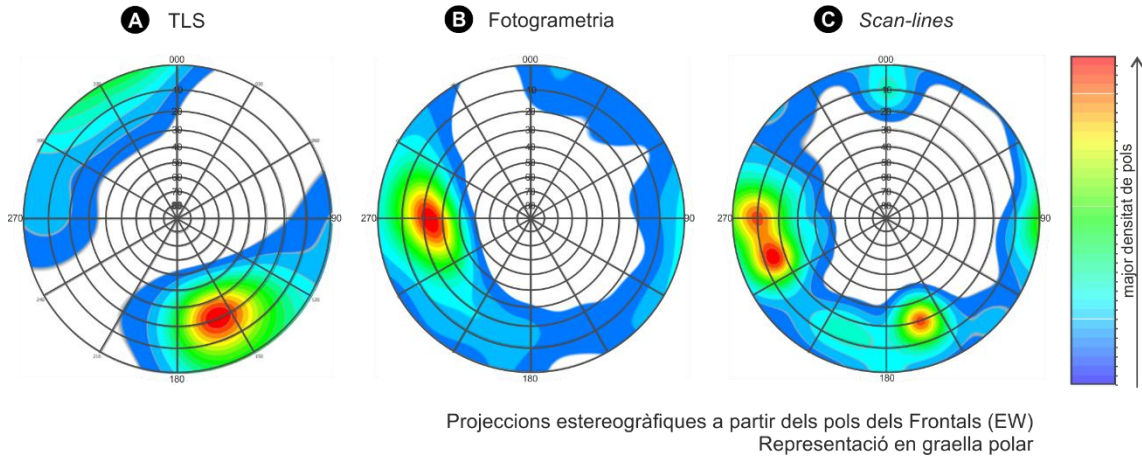


Figura 3.14: Projeccions estereogràfiques i identificació de les distribucions de les famílies, amb les dades de tots els fronts estudiats a l'àrea d'estudi. (A) TLS (B) Fotogrametria (C) Scan-lines

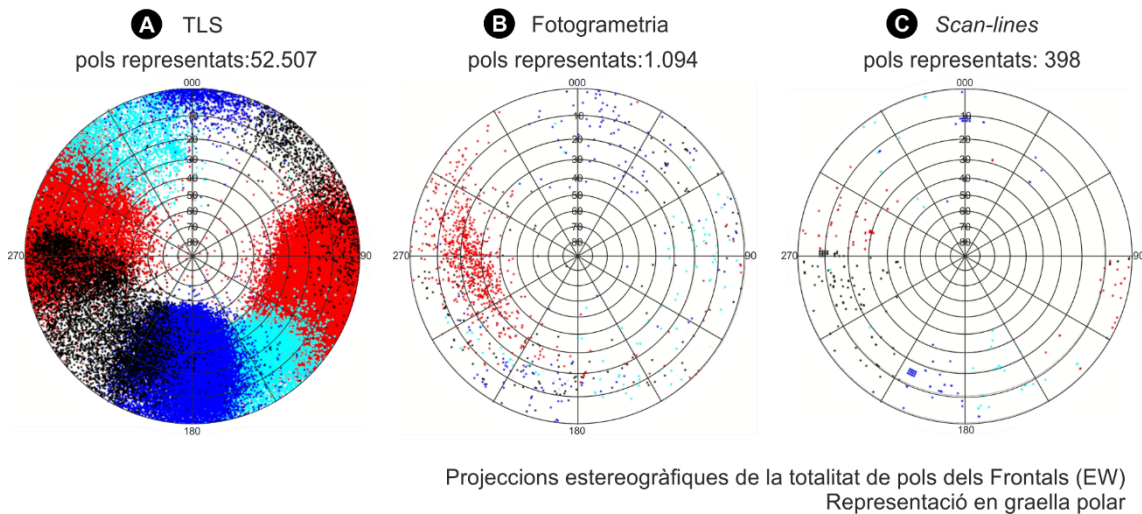


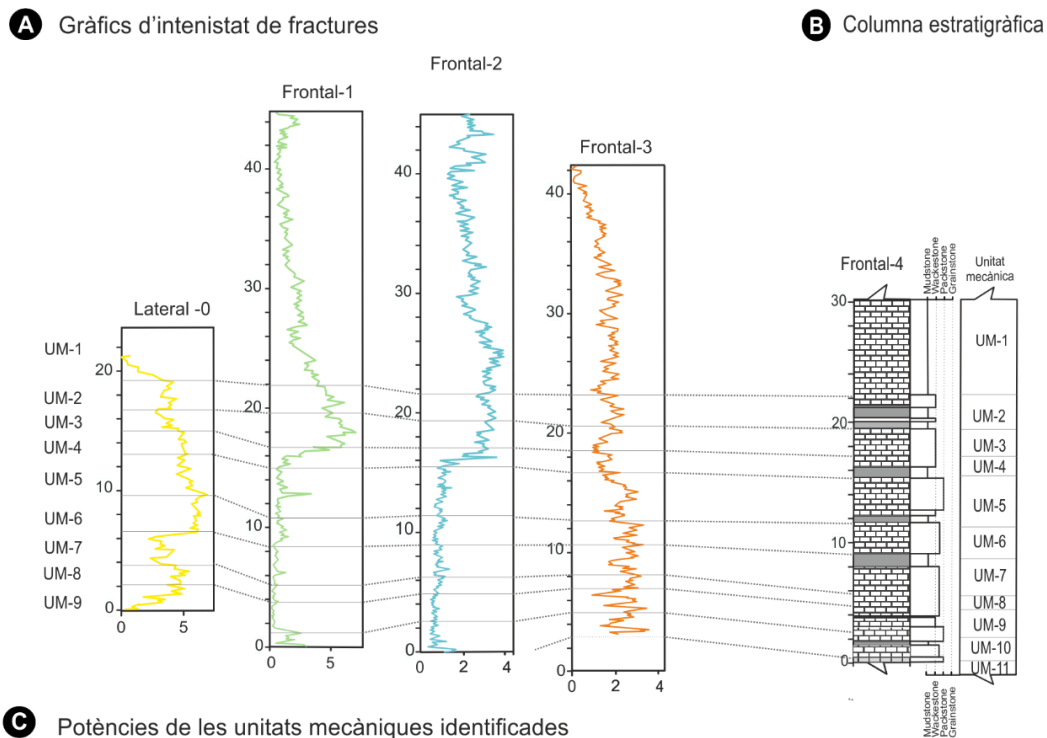
Figura 3.15: Projeccions estereogràfiques i identificació de les distribucions de les famílies, amb les dades de tots els fronts estudiats a l'àrea d'estudi. (A) TLS (B) Fotogrametria (C) Scan-lines

Unitats mecàniques identificades

La identificació de les Unitats Mecàniques s'ha realitzat amb la metodologia *Fracture Stratigraphy* descrita per Laubach et al (2009) i utilitzant l'atribut de la intensitat de les fractures. La intensitat de les fractures s'ha calculat amb el processament de les dades obtingudes amb TLS, al considerar-se la més completa, ja que explora la totalitat dels afloraments. Per mesurar la intensitat de les fractures el programari SEFL realitza comptatges de fractures en paral·lel a l'estratificació i cada 10 cm d'espaiament horitzontal, els resultats es mostren en la Figura 3.16,

juntament amb el perfil estratigràfic que ha servit de guia per correlacionar els límits. En total s'han identificat 11 Unitats Mecàniques.

El número de fractures per metre o intensitat de fractures varia segons la unitat mecànica. En la UM-1, aquest valors està al voltant de 2, amb un gruix de la unitat que supera els 25m, en canvi les unitats UM-2, UM-3, UM-4 d'entre 1,7 i 2,5m de gruix, arriben a 7 fractures per metre. La unitat UM-8, entre 1,2 i 1,8m de gruix, té intensitats de fractures que arriben a 4 i la resta d'unitats, d'entre 2 i 4,3m de gruix, presenten dades d'intensitat variades, en funció del perfil analitzat.



C Potències de les unitats mecàniques identificades

Potència	UM-1	UM-2	UM-3	UM-4	UM-5	UM-6	UM-7	UM-8	UM-9	UM-10	UM-11
Gruix (m)	>25,0	2,5-2,4	2,3-1,9	1,7-1,9	4,3-4,1	2,8-2,3	3,1-2,8	1,8-1,2	2,1-2,0	3,1-2,0	>1,0
Gruix mig (m)	>25,0	2,45	2,1	1,8	4,2	2,5	3,0	1,6	2,15	2,6	>1,0

Figura 3.16: (A) Correlació de les UM per a al Frontal-N 1, 2, 3 i el lateral-W. (B) Es projecten les unitats mecàniques identificades en els Frontals-N 1, 2, 3 i Lateral-W, sobre la columna litològica sintètica del quart Frontal-N. (C) Es mostren els gruixos de les unitats mecàniques identificades, la mitja.

Alçada i longitud de les fractures

Els valors mitjans de les alçades de les fractures mostren valors mitjans variats, però es pot destacar que la família de fractures I en totes les tècniques pràcticament és inferior als 2m i les altres tres famílies són de major mida, sobre els 8–12m. També és de destacar la diferència de valors en funció de la unitat mecànica. Respecte a les longituds, cada tècnica presenta valors dispars, però es pot generalitzar, que menys les fractures obtingudes amb TLS, que són inferiors al metre, la resta de tècniques varien entre 6–10 m fins els 18m. (veure Figura 3.17)

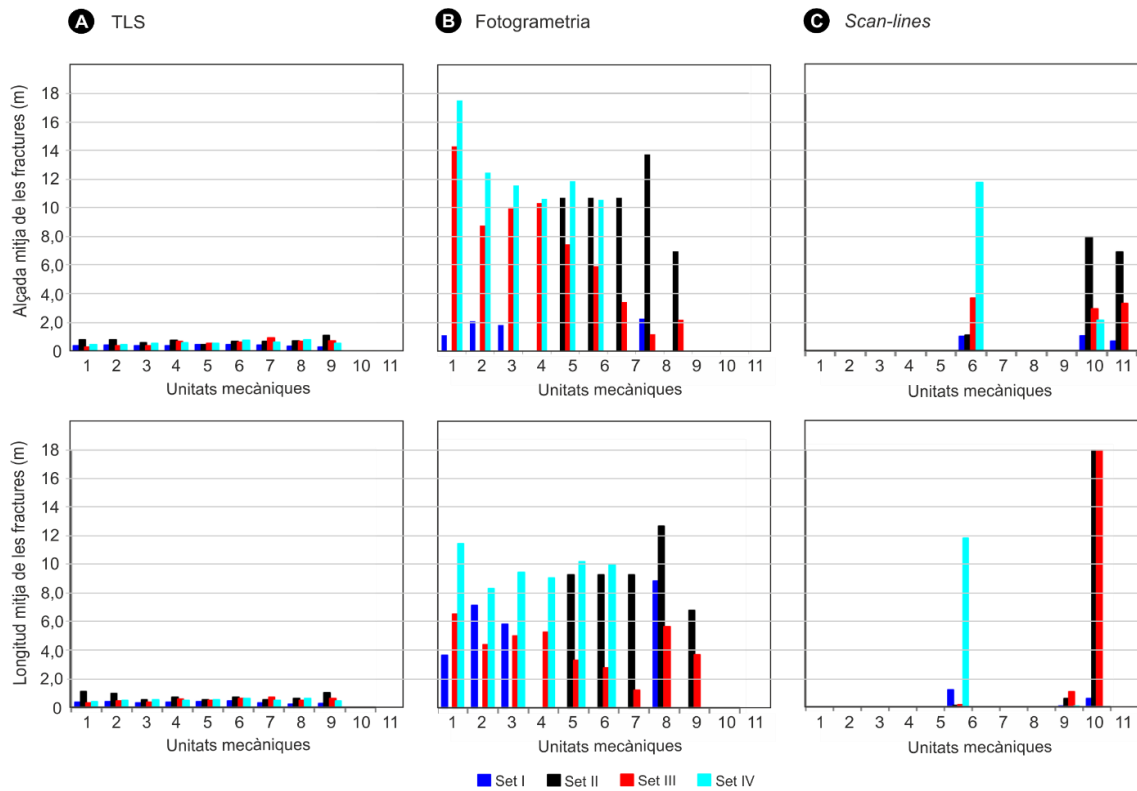


Figura 3.17: Alçada i longitud entre les fractures mesurades (Frontal-N-1), separades per famílies segons cada tècnica. (A) TLS (B) Fotogrametria (C) Scan-lines

Espaiat entre fractures

Els resultats mostren que els valors calculats d'espaiat amb TLS per la família de fractures I destaquen sobre la resta de famílies amb uns valors centimètrics que en les altres tècniques d'exploració no es repeteix. També, la família de fractures IV, en moltes unitats presenta espaiats centimètrics. En general, per la resta de tècniques els valors calculats d'espaiament es poden considerar dispers (veure figura 3.18).

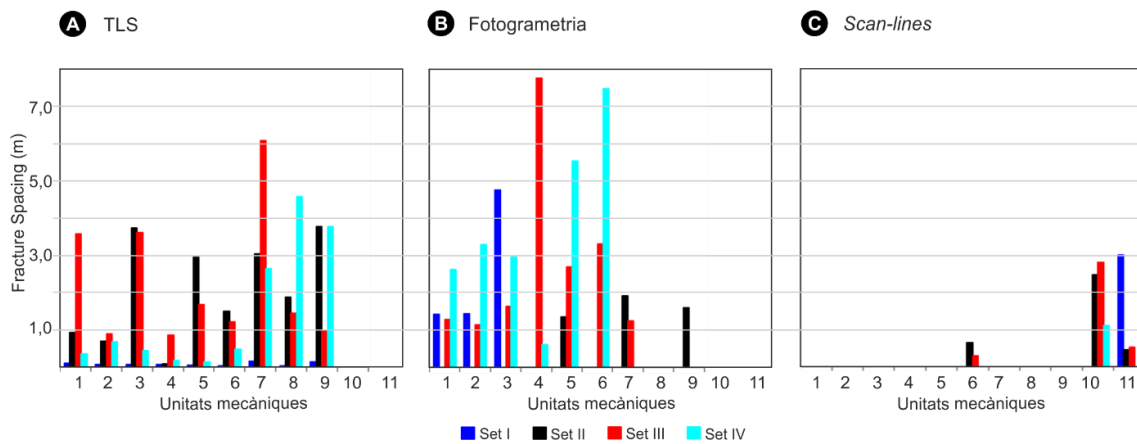


Figura 3.18: Espaiat entre les fractures mesurades (frontal-N-1), separades per famílies segons cada tècnica. (A) Lidar (B) Fotogrametria (C) Scan-lines

Els coeficients de variació obtinguts presenten certa dispersió en funció de de la família i la unitat mecànica estudiada (veure Figura 2.6). En general, el rang de valors obtinguts entre 0,5 i 1,4 en el cas de la fotogrametria és similar als valors obtinguts amb la tècnica directa. En canvi, amb el TLS els coeficients de variació es troben entre els valors 1 i 2, indicant una major agrupació de les fractures. (veure Figura 3.19).

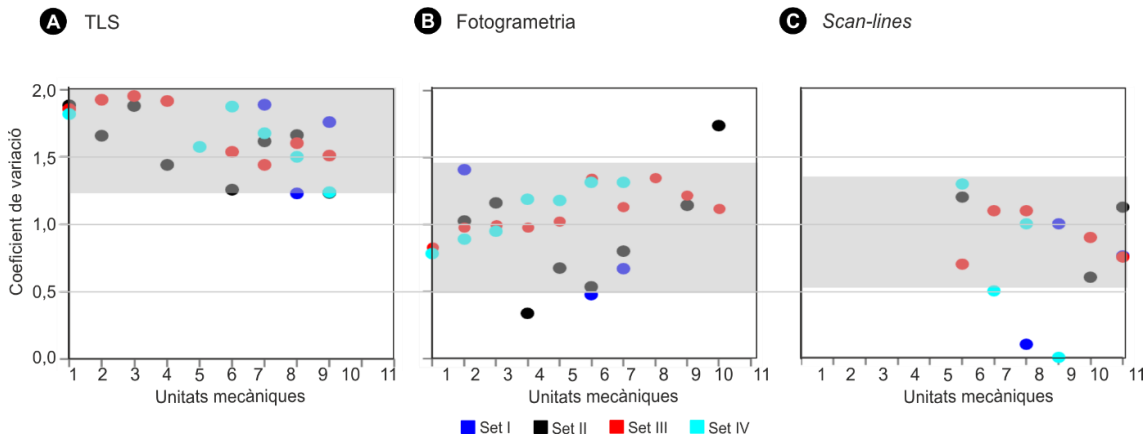


Figura 3.19: Gràfics del Coeficient de variació de l'espaiat (nivell d'agrupament de fractures (Gillespie et al., 1999)) de tots els afloraments estudiats. (A) Lidar (B) Fotogrametria (C) Scan-lines.

Intensitat de fractures

La intensitat linear de les fractures, o P10 en la nomenclatura de Dershowitz i Einstein (1984), calculades amb les dades TLS mostren que la família I en els fronts d'orientació Nord, és la família de fractures amb els valors més elevats en la majoria de les unitats mecàniques. Els valors de la intensitat varien entre 1 i 3 fr/m. El predomini no és hegemònic, sent rellevada normalment per la família de fractures IV amb valors de 1fr/m en algunes unitats mecàniques. En alguna unitat mecànica, la família de fractures III també arriba a presentar les intensitats més altes. En el front d'orientació E-W els valors s'inverteixen i la família III és la família més dominant en quant a valors d'intensitats més elevades, entre 1–3 fr/m. La família de fractures I amb aquesta orientació del front es qui presenta el valors més reduïts, inferiors als 0.1 fr/m.

La intensitat calculada amb les dades fotogramètriques en el primer Frontal-N, també mostren l'hegemonia de la família de fractures I però amb valors més baixos, 1 fr/m i inferiors. Però en el segon Frontal-N, la intensitat de les fractures en la família II arriba a valors hegemònics entre 1 i 8 fr/m, mantenint-se la resta de famílies amb valors inferiors a 1 fr/m. En el tercer frontal d'orientació Nord les famílies de fractures III i IV son les més destacades però amb valors d'intensitat inferiors al 1fr/m amb excepció de la unitat mecànica 1 i 3. Les dades d'*scan-line* també mostren intensitat superiors a 1fr/m per les famílies de fractures I, III i IV.

Els valors d'intensitat mesurat en una àrea (P21) i en un volum (P32) confirmen les dades obtingudes en la mesura de la intensitat lineal (P10).

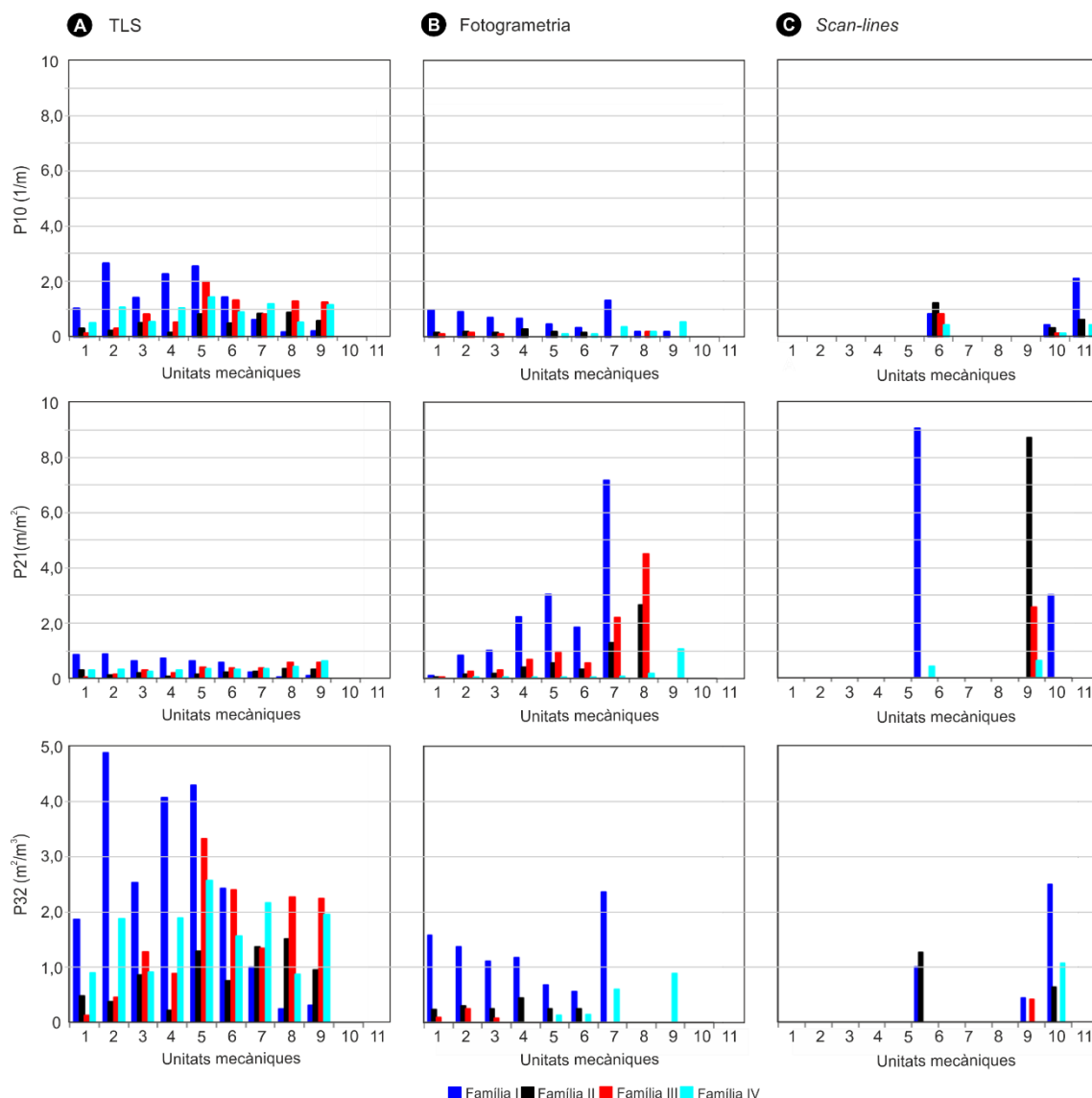


Figura 3.20: Valors mitjos d'intensitat (P_{21} i P_{32}), per a cada unitat mecànica i família de fractures al Frontal-N-1. (A) TLS, (B) Fotogrametria i (C) scan-line.

Obertures

L'obertura de les fractures solament ha estat mesurada amb la tècnica de l'*scan-line*, identificant l'obertura buida amb un rang de valors mesurats entre 0,01 i 50mm. A la Figura 3.21, es representen els valors mitjos obtinguts, agrupats per famílies de fractures. No s'han mesurat les obertures cinemàtiques però si que s'han descrit els materials de reompliments en cada fractura, obtenint la distribució dels reompliments, en funció de les famílies i les fractures mesurades.

D'aquestes dades se'n desprèn que les famílies amb menor presència de reompliment és la família I i II i en canvi, la família III, contindria major quantitat de reompliment, destacant la calcita i les argiles de descalcificació.

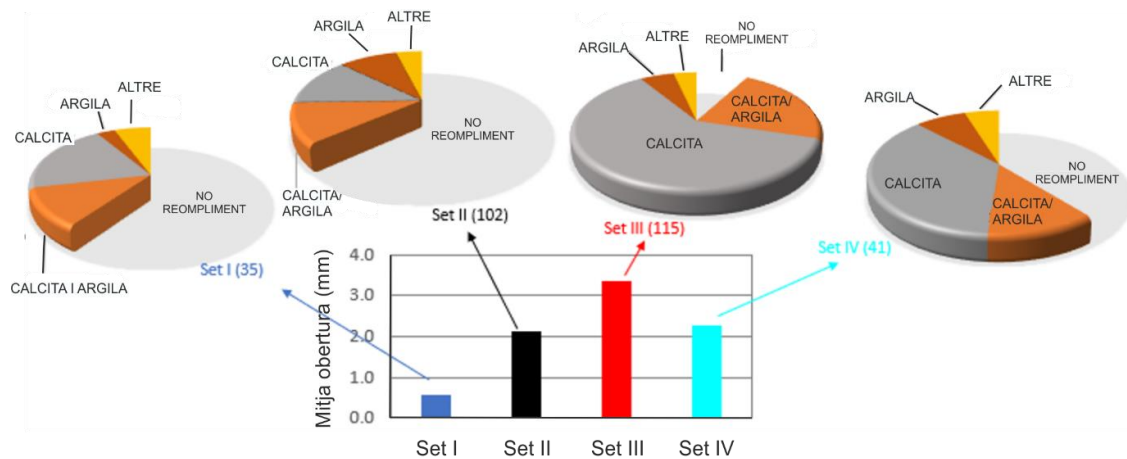


Figura 3.21: Dades mitges de les obertures obtingudes i agrupades per famílies de fractures mitjançant la tècnica d'scan-line. També es representa la distribució del reompliment descrit, segons la fractura i la família.

Corbes de distribució per la propietat de l'espaiat, l'alçada i longitud de les fractures

Altres maneres de representar les dades, poden ser les corbes de distribució, on obtenim una distribució dels valors de la propietat analitzada en funció de la freqüència que es repeteix aquell valor. El gran nombre de dades obtingut amb els mètodes remots permeten representar les dades de l'espaiat, l'alçada i la longitud. (veure Apendix 1)

En aquest apartat, els resultats obtinguts mostren certes diferències segons la tècnica i la propietat analitzada. En la representació estocàstica d'aquestes dades en els models DFN, caldrà analitzar com queden representades, per avaluar-les en un model tridimensional.

Modelització discreta de fractures amb tècniques remotes

En un model matemàtic discret de xarxes de fractures (DFN), les dades introduïdes per modelitzar-lo dependran directament de les dades capturades o recollides a la zona i de la manera de representar-les. En models realitzats, una per cada tècnica de captura (TLS, Fotogrametria i *scan-lines*), s'han introduït estocàsticament les principals característiques geomètriques de les fractures; orientacions, espaiats i mida, per poblar el model tridimensional.

Per a realitzar el model, s'ha delimitat un cub, on els límits incorporassin els afloraments estudiats, tal com es mostra a la 3.22. La dimensió total del cub és de 120.000m³, amb una alçada de 25m, una amplada de 80m i el fons de 60m. Per a la representació de les dades capturades en el model DFN, s'han introduït totes les fractures identificades en els Frontals orientats cap al nord, separades per famílies. També, s'han introduït els plans estratigràfics més penetratius. (veure Figura 3.23 i 3.24)

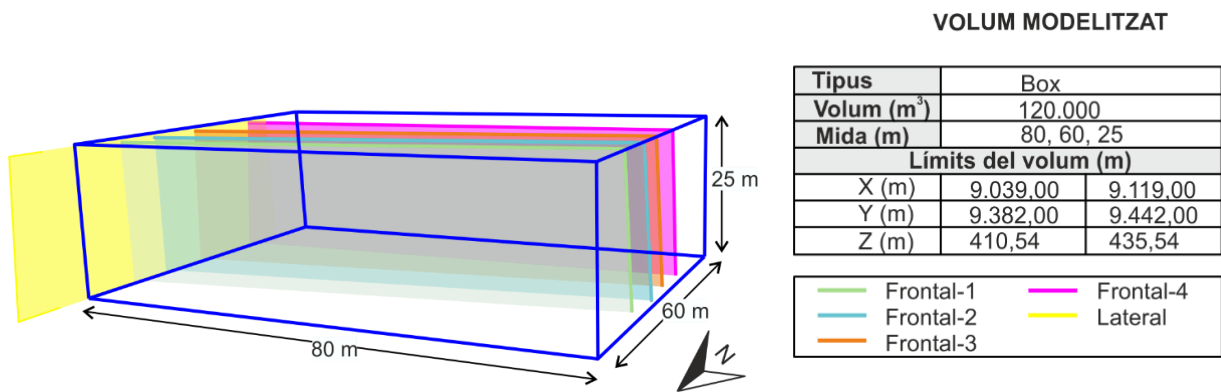


Figura 3.22: Volum modelitzat de 120.000m³, de 80m d'ample, per 60m de fons i 25m d'alçada.

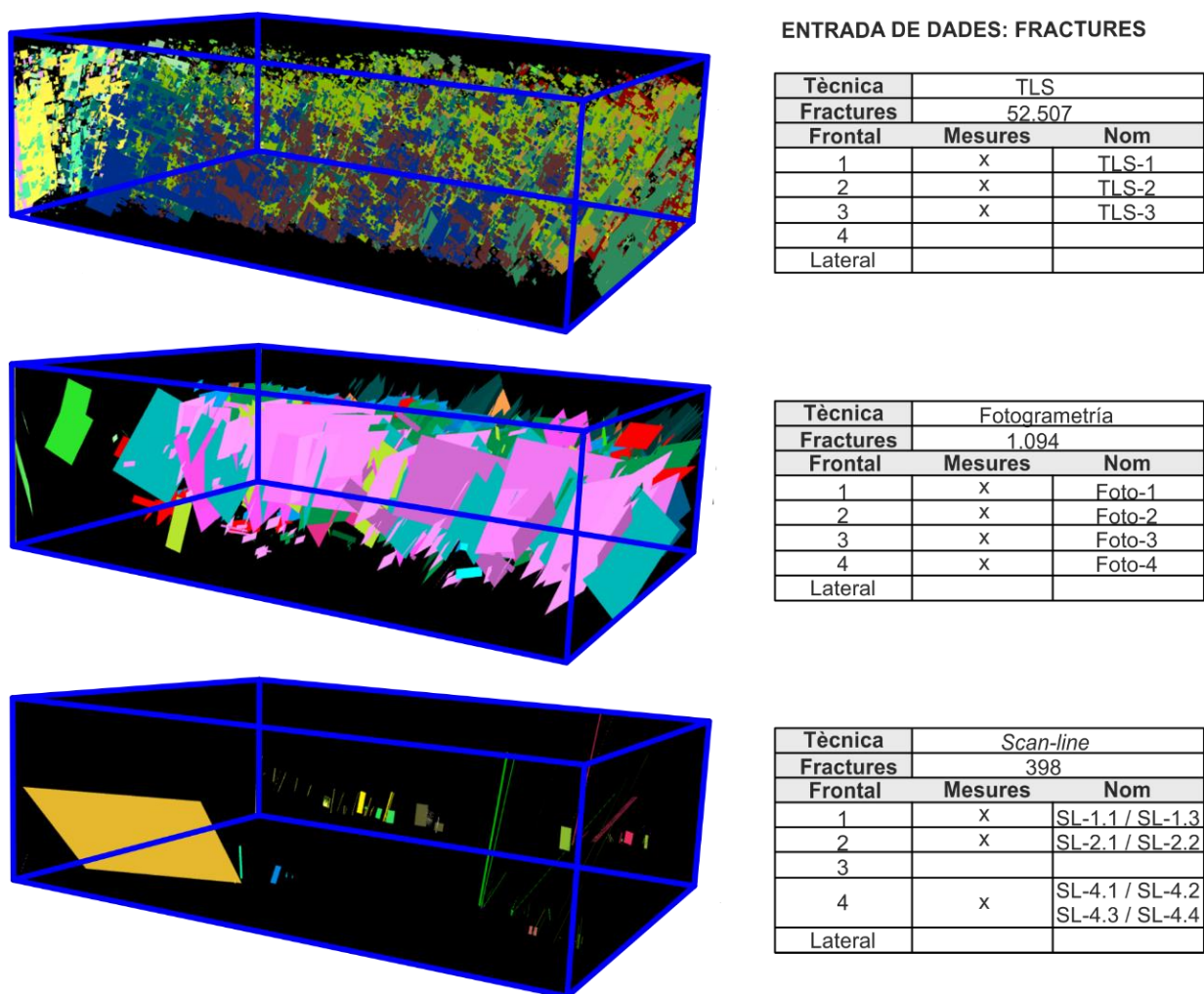


Figura 3.23: Visualització de les superfícies de fractures introduïdes per a cada tècnica; TLS, Fotogrametria i scan-line.

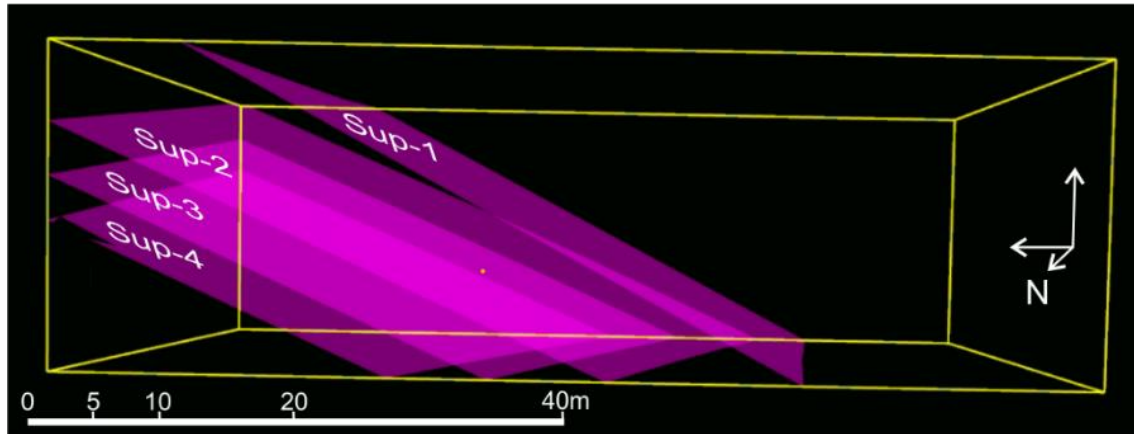


Figura 3.24: Projecció de la superfície dels plans de discontinuïtats estratigràfics més penetratius que s'han identificat com superfícies que poden influenciar en el transport de flux a través del massís fracturat.

De les diferents opcions que proposa el programari *FracMan* per generar fractures s'ha optat per la recreació estadística, ja que l'objectiu és avaluar les tècniques remotes i els scan-lines i d'aquesta manera, controlem millor els valors introduïts. L'àrea d'estudi s'ha adaptat a les dimensions dels afloraments tal i com es mostra a la Figura 3.26.

Els tres models creats s'han basat en el model de càlcul, en aquest cas Baecher millorat (Baecher et al., 1978) i s'han considerat els centres de les fractures com a referència de càlcul.

Obtenim 3 models DFN tridimensionals, un per cada tècnica de captura, que anomenarem DFN_{TLS} , DFN_{SfM} i $DFN_{Scan-line}$, tal com es representen a la Figura 3.25.

L'espaiat de les fractures en els models presenta mitges de 0,1m per els models DFN_{TLS} i $DFN_{Scan-line}$. En canvi en model DFN_{SfM} , l'espaiat mig és de 0,3m. Aquests valors van amb consonància amb els valors de P10 obtinguts. Observant amb més detall la Figura 3.25, s'identifiquen també diferències en la representació de les fractures per a cada família. Per exemple la família I, mostra majors similituds en la seva representació entre el DFN_{TLS} i $DFN_{Scan-line}$, però en canvi, en el cas de la família III, les semblances són entre el model DFN_{SfM} i $DFN_{Scan-line}$.

En els tres models s'han introduït els mateixos valors d'obertura obtinguts únicament amb *scan-lines*. Posteriorment, una de les funcionalitats del programari ha permès generar el model de traces. El mapa de traces es crea a partir de la intersecció entre el conjunt de fractures representades i la regió definida. L'objectiu és calcular el valor de la porositat (P33) i la permeabilitat (veure figura 3.26).

Després d'introduir les obertures i generar el mapa de traces s'obtenen diferents resultats. Per una banda, els valors d'obertura mitjos entre 3,1 i 4,6mm, són valors acceptables en funció de les dades introduïdes. Els valors mitjos de P32 del model analitzat que representen les superfícies de fractures per unitat de volum, mostren valors de 1,5 en el DFN_{TLS} , 2,5 per el DFN_{SfM} i 3,6 per el model $DFN_{Scan-line}$.

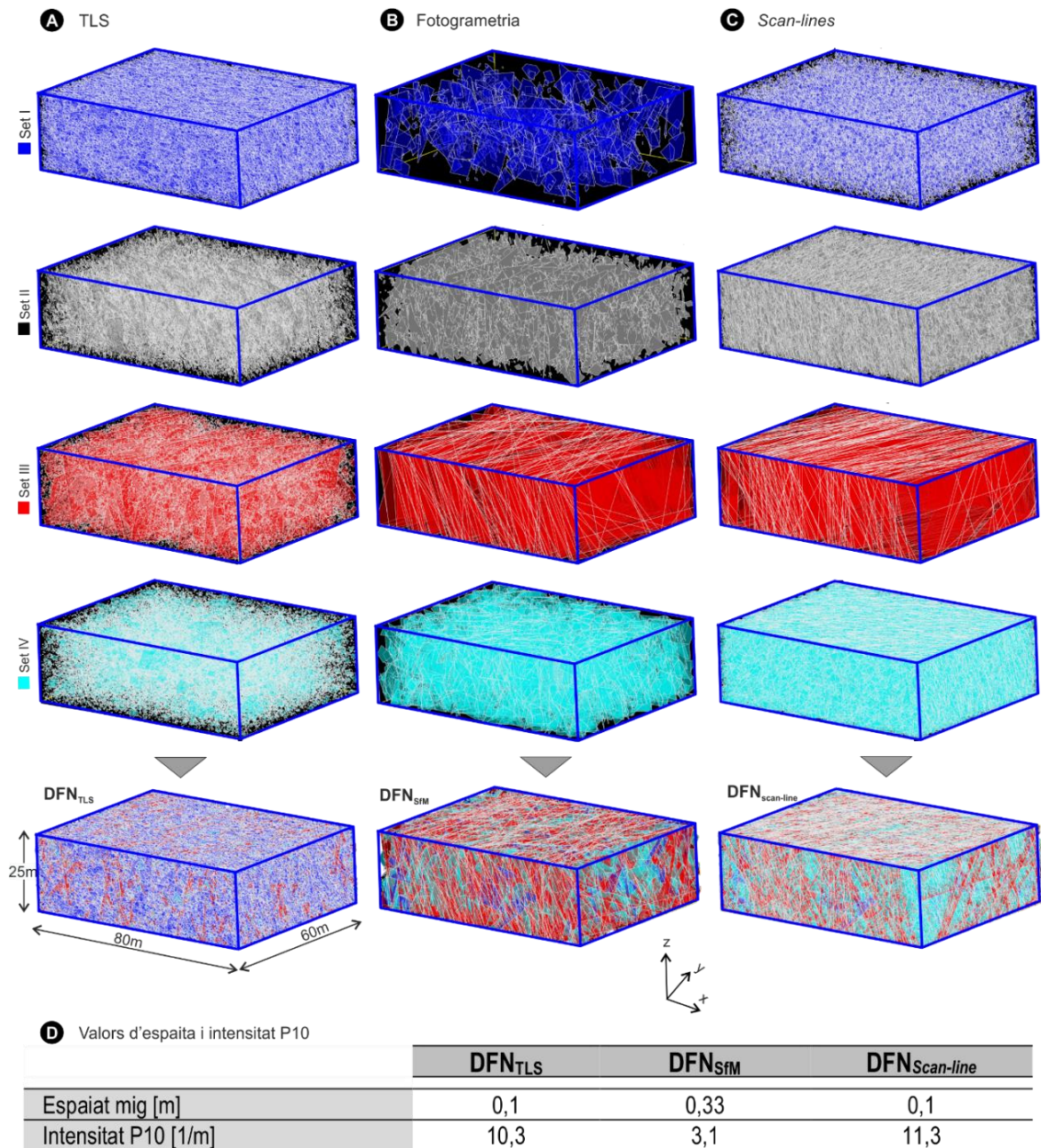
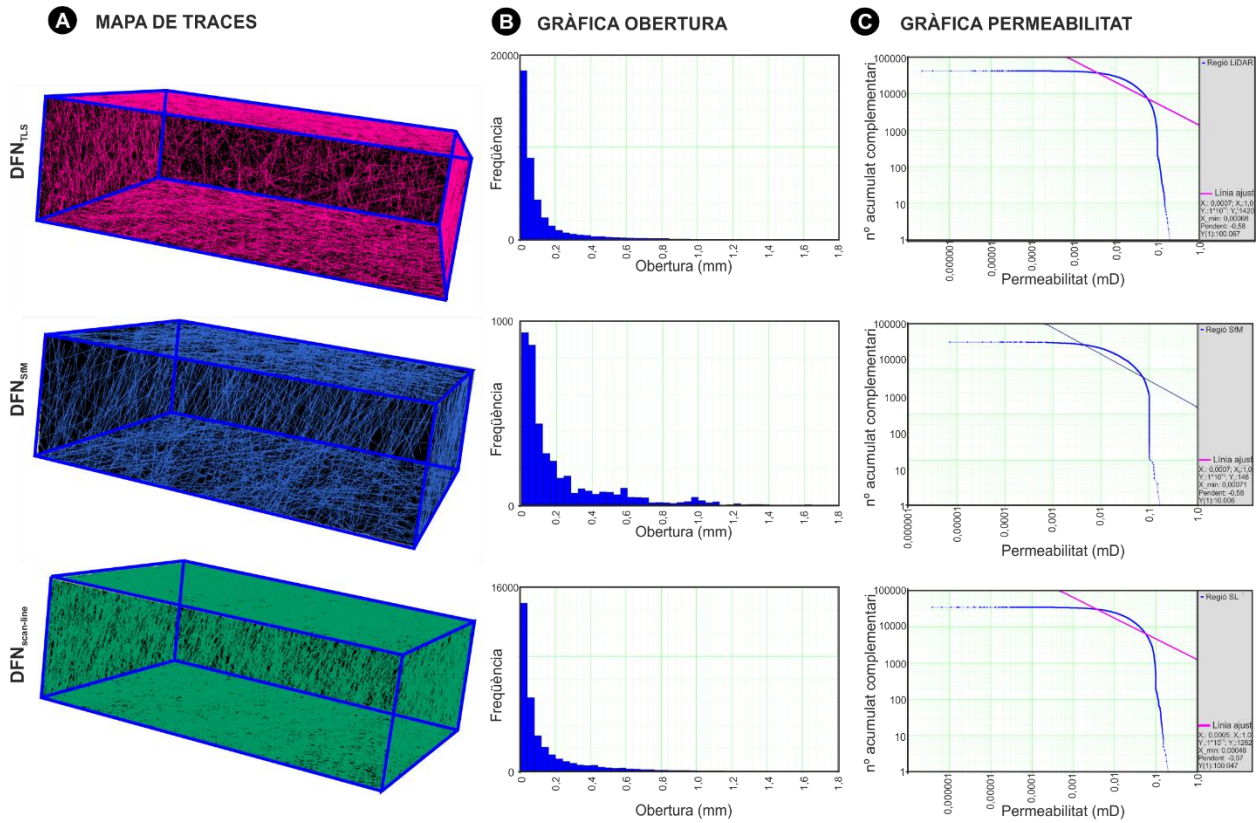


Figura 3.25: A la part superior de la Figura, s'observa la representació gràfica del volum estudiat, separats per famílies i per tècniques. Extret del programa FracMan (GOLDER). (D) La Taula presenta els espaiats mitjos i el P10, per cada model DFN construït.

Finalment, obtenim els valors de porositat a partir del P33 i els valors de permeabilitat en mD o cm^2 . Els resultats mostren valors semblants per a tots els models. Com a resultat obtenim que la porositat mitja en el DFN_{SfM} i DFN_{scan-line} és de 5,7% i en el DFN_{TLS} és de 5,4% i que l'estimació de la permeabilitat mitja de la xarxa de fractures en el model DFN_{SfM}, és de 4×10^5 mD, lleugerament superiors als models DFN_{TLS} o DFN_{scan-line}, que ofereixen valors mitjos de $6,2 \times 10^5$ mD.



D	DFN _{TLs}	DFN _{SfM}	DFN _{Scan-line}
Obertura mitja (m)	0,0038	0,0046	0,0031
Obertura – Desviació estàndard	0,0027	0,0027	0,0024
P ₃₂ (fractures per àrea / volum) [1/m]	1.524	2.474	3.568
P ₃₃ (fracVolume/volume) [%]	5,4	5,7	5,7
Permeabilitat mitja (mildarcy)	616.431,14	402.951,42	642.512,15
Permeabilitat – Desviació estàndard	366.099,87	414.031,9	370.238,28
Permeabilitat mitja (cm ²)	6,04 * 10 ⁻⁶	3,98 * 10 ⁻⁶	6,34 * 10 ⁻⁶

Figura 3.26: (A) Representació gràfica de les traces, (B) gràfica de la distribució de les obertures i (C) gràfica de tendència de la permeabilitat. (D) Dades mitges i desviació estàndard de l'obertura, porositat i permeabilitat a partir del mapa de traces, extreus del programa FracMan, per els models DFN construïts.

3.2.5. Discussions

Es discuteixen i s'avaluen els resultats específics obtinguts en aquest experiment a través dels següents apartats: (1) Consideracions de l'anàlisi previ del medi rocós, (2) Representativitat de la caracterització de la xarxa de fractures (3) Avaluació de les tècniques remotes en models DFN.

Consideracions de l'anàlisi del medi rocós

Estudiar el medi que conté les fractures, permetrà identificar les discontinuïtats estratigràfiques que puguin intervenir en les simulacions de flux i determinar si la porositat primària del medi cal considerar-la en el model. Addicionalment en aquest cas concret, hem pogut identificar a partir de làmines primes l'edat de la successió estratigràfica estudiada.

Els assajos de porositat i permeabilitat realitzats amb testimonis de roca, representen la matriu i les microfractura presents. Els valors obtinguts suggereixen que les matriu rocosa és pràcticament impermeable. Aquests resultats han resolt una de les qüestions inicials on calia decidir quin model utilitzaríem per a la representació de les fractures per posteriors simulacions de fluxos; un model de doble porositat o un que únicament consideri la porositat generada per les fractures. Amb les dades obtingudes podem afirmar que el tipus de model per representar les fractures no necessitarà considerar doble porositat (Secció 2.4.1), en conseqüència, el model adequat i el que s'ha utilitzat en aquesta tesi doctoral ha estat un model discret de fractures (DFN) on únicament es contempla la porositat i permeabilitat generada per les fractures.

Per altra banda, s'observa en el aflorament que algunes superfícies estratigràfiques més penetratives, poden tenir una forta influència en la circulació del flux. Aquest fet també s'identifica en les anomalies detectades en els perfils tomogràfics. Aquestes anomalies electromagnètiques podrien ser degudes a zones preferents de flux o zones reomplertes d'argiles. Per discernir entre les dues possibilitats, s'ha superposat la imatge obtinguda amb la tomografia sobre les imatges de l'aflorament i s'han buscat relacions amb les fractures. En aquest sentit, si observem la figura 3.27, es marquen amb fletxes vermelles, les discontinuïtats coincidents amb les anomalies identificades en els perfils de les tomografies elèctriques.

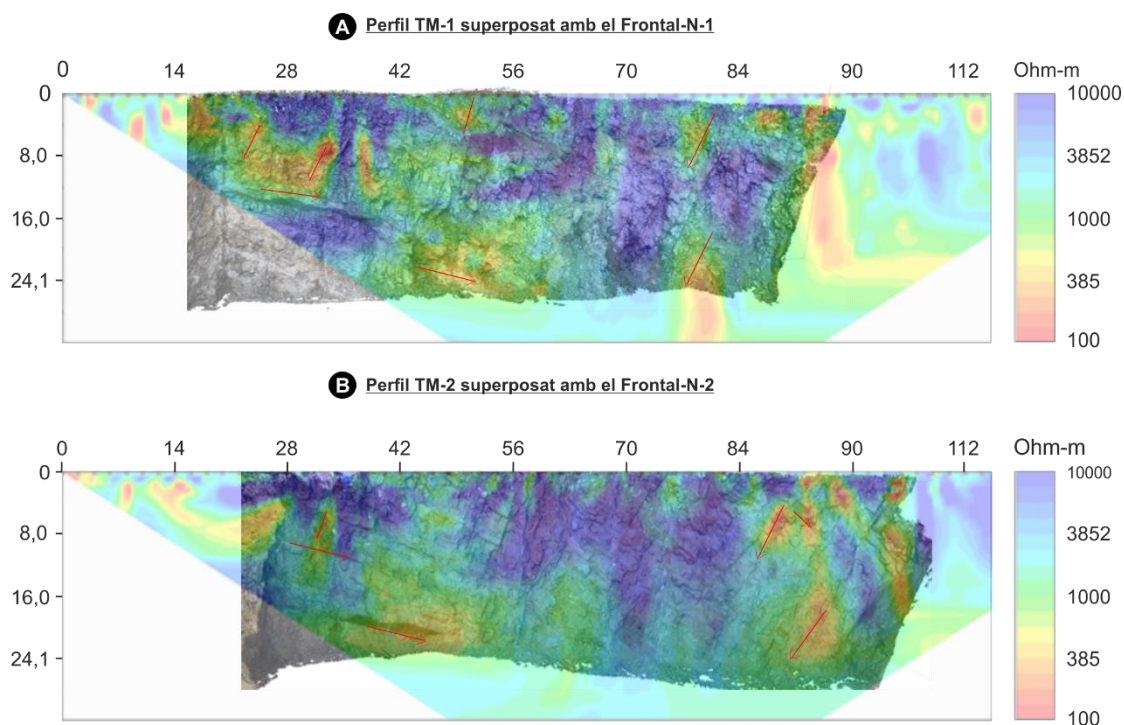


Figura 3.27: (A) Superposició de la tomografia elèctrica TM-1 amb el Frontal-N-1 (B) Superposició de la tomografia elèctrica TM-2 amb el Frontal-N-2. Amb fletxes vermelles, s'identifiquen les zones humides identificades després de la pluja. S'observen coincidències amb les coloracions groguenques i vermelles de la tomografia. Aquestes coloracions corresponen a les zones amb menor resistivitat.

Les discontinuïtats, s'atribueixen a fractures i a superfícies estratigràfiques. Per comprovar si les anomalies detectades poden ser zones preferents de flux, s'ha visitat l'àrea d'estudi els dies posteriors a la pluja i s'observa com l'aigua brolla o humiteja per aquestes discontinuïtats marcades, confirmant així, certa connectivitat entre elles. Les superfícies estratigràfiques penetratives, s'han incorporat en el model, per considera la rellevància d'aquestes discontinuïtats en el medi fracturat. Addicionalment, a partir del contingut fòssil reconegut i de les textures descrites a l'àrea d'estudi amb làmina prima, s'interpreta l'ambient deposicional de la successió sedimentària estudiada, com a una plataforma carbonatada d'aigües poc profundes i baixa energia, del Barremià i l'Aptià. Ambients sedimentaris de plataforma carbonàtica de gran extensió dominats per *Palorbitolina lenticularis*, *Choffatella decipiens* i miliòlids que van desenvolupar-se als marges nord i sud del Mar de Tetis. (Albrich, et al., 2006; Pittet et al., 2002; Vilas et al., 1995). Les principals textures identificades són mudstone, wackestone i packstone. La identificació de les textures i la confecció de la columna litològica, tot i no ser un dels objectius de la tesi han aportat un major coneixement de la zona.

Representativitat de la caracterització de la xarxa de fractures

L'avaluació dels diferents models de DFN creats segons cada tècnica, requereix d'uns processaments previs per caracteritzar la xarxa de fracturació del medi. En comú per totes les tècniques utilitzades en aquesta tesi doctoral, els resultats de les propietats individual de les fractures comprèn la seva orientació, posició i mida i posteriorment les propietats col·lectives, les famílies de fractures, espaiament, intensitat i d'altres descrites anteriorment.

Famílies de fractures

Les màximes densitats dels pols d'orientació de les fractures varia segons cada tècnica, però mantenen unes proximitats inferiors de 15°, inclús inferiors. Les particularitats de cada tècnica explicant en part aquestes diferències. Primer cal recordar que totes tres tècniques son dependent de la orientació de l'aflorament i en el cas de l'*scan-line* de la línia d'exploració. I en segon lloc, que la metodologia aplicada a les dades TLS permet identificar solament superfícies, la fotogrametria traces i l'*scan-line* explora un sector limitat. Cal destacar que la tècnica que mostra les 4 famílies sobre la projecció estereogràfica amb més claredat ha estat l'*scan-line*. En canvi les tècniques remotes, amb major nombre de pols representats, acostumen a emmascarar algunes famílies i destacar d'altres, necessitant estudiar amb major detall les dades. En tot cas, totes les tècniques obtenen dades de les quatre famílies, sense contradir-se.

Les orientacions de les famílies identificades són coherents amb les principals orientacions establertes en estudis previs (Guimerà, 1982), centrats en l'evolució del procés de fracturació de la cobertura del Massís del Garraf (veure Subsecció 3.2.2). Tal i com descriu Guimerà, (1982), semblaria que el procés de fracturació de la família IV és posterior al de les famílies I, II i III.

Unitats Mecàniques

Les 11 unitats mecàniques identificades pretenen reconèixer les heterogeneïtats mecàniques del medi fracturat. Algunes unitats han pogut ser identificades amb certa facilitat, com en el front d'orientació Est i els fronts 1 i 2 d'orientació Nord, encara que d'altres poden ser discutibles. En general s'observa que el medi conté heterogeneïtats en el seu comportament mecànic i que es reflexa en la intensitat de les fractures (Figura 3.20 i Apèndix 1). A la figura 3.28, es representen les unitats mecàniques incorporant altra informació; la columna litològica confeccionada, les textures identificades i la porositat i permeabilitat obtingudes. Treballs previs com el de Barbier et al., (2012), presenten estudis que relacionen aquestes característiques. Amb aquestes dades observem com els límits de les unitats mecàniques identificades coincideixen amb canvis de litologia o porositat i permeabilitat, confirmant aquesta delimitació de les unitats.

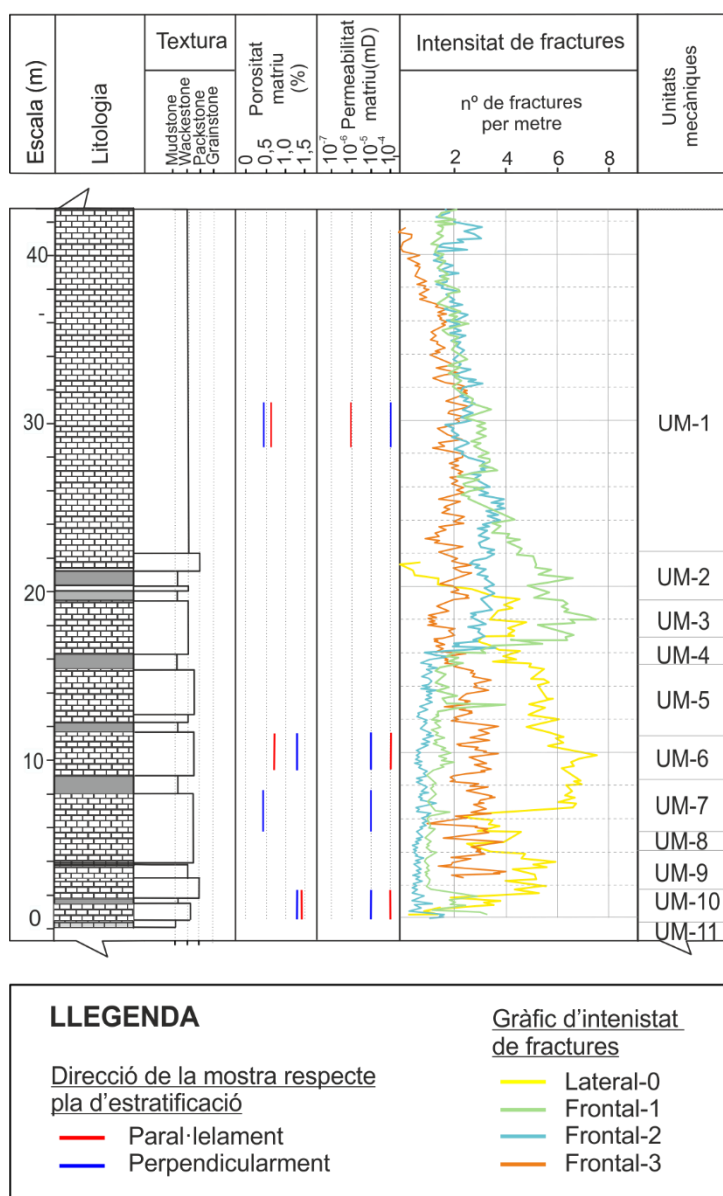


Figura 3.28: Correlació de la porositat i permeabilitat amb les unitats mecàniques identificades amb la intensitat de fractures i les propietats petrogràfiques de la columna estratigràfica aixecada al Frontal-N-4.

Nombre de fractures

El total de fractures identificades i processades per el programa SEFL, en funció de les famílies i unitats mecàniques identificades, han estat de 6.414 fractures amb *TLS*, 1.088 amb tècniques de fotogrametria i 398 amb les tècniques clàssiques d'*scan-line*. Tal i com ja hem comentat, la orientació de l'aflorament observat condiona la recollida de dades, i això es trasllada també en el nombre de fractures recollides i en conseqüència a les altres propietats com la intensitat. Per exemple, en el cas del Lateral-W, únicament processat per *TLS*, mostra diferències si el comparem amb la resta de frontals d'orientació nord analitzats amb aquesta tècnica (veure Figura 3.29).

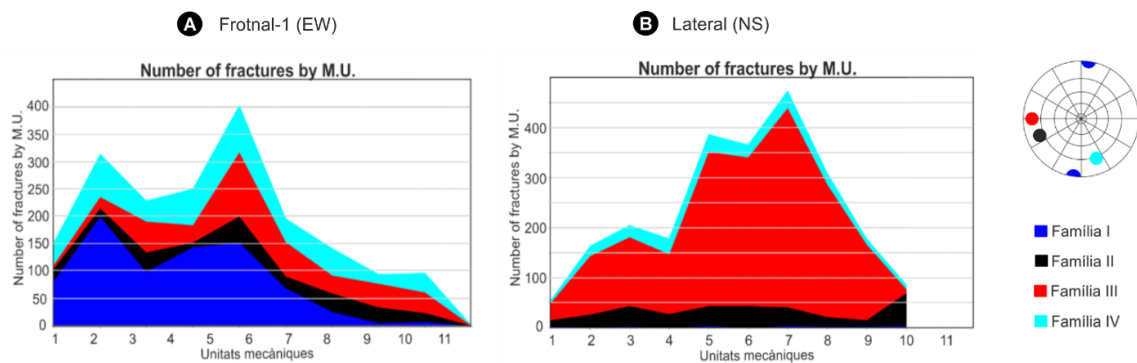


Figura 3.29: Nombre de fractures identificades i mesurades amb *TLS*. (A) Frontal-N-1 i (B) Lateral-W

Alçada i longitud de les fractures

La mitja dels valors d'alçada i longitud per els fronts, mostren un valors per unitats mecàniques força dispersos com reflexa la Figura 3.17. Els valors calculats amb dades *TLS* mostren unes alçades i longituds de fractures mitjans inferiors al metre. En el cas de les alçades i en part les longituds, aquest fet pot ser explicat pel tractament semi-automàtic al reconstruir les fractures, ja que una intersecció amb el pla d'estratificació segmenta la fractura i es computa com final de fractura, sense interpretar la seva possible prolongació. D'aquesta manera, s'incrementa el número de fractures i es retalla les dimensions.

L'alçada de les fractures quan són digitalitzades amb dades de fotogrametria presenten uns valors al voltant dels 10 m per les famílies amb una orientació perpendicular als fronts de la pedrera i molt més baixes a las paral·leles al front. En fotogrametria, el que es digitalitza són les traces de les fractures, és a dir, la intersecció de la fractura amb la superfície de l'aflorament i és un tractament manual. És tendència de la persona digitalitzadora a identificar i digitalitzar les fractures més ben definides i visibles en les imatges per la seva facilitat. Les fractures més curtes i de traça quasi paral·lela al front de l'aflorament presenten una dificultat que l'usuari a efectes de filtre, deixa sense digitalitzar, possiblement sobredimensionant a la família de fractures. En

aquest sentit, també la qualitat del conjunt de fotografies adquirides pot influenciar, especialment quan s'observen zones d'ombres o sobrexposicions, que dificulten la digitalització.

La tècnica de l'*scan-line* segurament és la més fiable, al poder dedicar més atenció a cada fractura, però al realitzar les mesures de manera manual la productivitat no és molt elevada. És una tècnica més costosa en temps, més limitada en l'accés i queda reservada a l'exploració d'una reduïda secció de l'aflorament, com en el cas de la Fou. Els valors d'alçada i longitud, en tot cas, són força heterogenis.

A més, cal tenir present que la majoria de les longituds de les fractures queden censurades pel propi aflorament, i en menor mesura l'alçada. Això s'observa sobre tot en les unitats mecàniques situades en els extrems, la 1 i la 11. És a dir, el percentatge de fractures complertes identificades és baix. El fragment visible i mesurable de fractura no se sap quin percentatge realment representa de la fractura. Es coneix el valor mínim, però no es pot identificar el valor màxim.

Es partia de la hipòtesis de que les dades de GPR podrien donar més informació respecte a les dimensions, sobre tot a la longitud de les fractures. La visualització dels seus resultats va revelar la impossibilitat de identificar aquestes propietats.

Les consideracions abans comentades per mesurar directa i indirectament les propietats individuals de les fractures condicionen els resultats de les propietats col·lectives més importants, l'espaiament i la intensitat.

Espaiat i intensitat de fractures

Els valors mitjos d'espaiat identificats en la tècnica TLS, mostren rangs variables entre els pocs centímetres i els 6m, excepte en la Família I, que els espaiats mitjos són centimètrics. En els scan-lines, aquestes dades són més restrictives arribant a valors mitjos d'espaiat màxim de 3m. En la fotogrametria, el rang de valors es troba entre els pocs centímetres i els 8 metres. En general són valors majors que els obtinguts en les altres tècniques. Especialment en la fotogramètrica, els espaiats acostumen a ser majors que els obtinguts amb la resta de tècniques, pel propi procés de digitalització. La digitalització per fotogrametria, és molt dependent de de l'usuari a l'hora de digitalitzar les traces de les fractures, tal i com hem comentat. L'espaiat entre les fractures també estarà condicionat per l'orientació de l'aflorament captat, on per exemple, en els frontals d'orientació cap al nord, les tècniques TLS captaran més fractures de les famílies I i IV i la fotogrametria de les famílies II i III.

La distribució d'aquestes fractures també és una qüestió important. En els afloraments de la Fou, s'observen corredors o zones clusteritzades. Són zones amb major intensitat de fractures, que arriben a tenir espaiats mil·limètrics i agrupacions de gruixos totals d'entre 20 i 30cm. Amb la tècnica d'*scan-line*, s'han mesurat intencionadament, per observar si quedaven representats en els anàlisis. Segons les dades obtingudes dels coeficients de variació (Figura 3.19), només la

tècnica TLS, amb valors d'entre 1,2 i 2, recull aquestes agrupacions. La resta de tècniques presenten certa dispersió, però no una agrupació de fractures, amb valors entre 0,5 i 1,2. Ja em observat que tot i haver agafat les dades dels corredors en els *scan-lines*, els gràfics de coeficient de variació no recullen aquestes agrupacions de fractures que s'associen als corredors, en canvi, les tècniques TLS si que les han captat. La importància de representar aquests corredors en la simulació de flux pot ser rellevant, ja que acostumen a ser zones preferents de flux (Sanderson i Peacock, 2019) i de fet, s'observa com es dies després de pluja, l'aigua brolla especialment en aquestes agrupacions de fractures. Aquestes distribucions influenciaran en les generades en el model DFN i caldrà veure si queden representades.

Per altra banda, els gràfics FSR que ens poden mostrar una certa relació entre la distribució de les fractures i les unitats mecàniques en aquest cas no mostra dades rellevants.

Intensitat de les fractures

Les dades de P10, en general es troben entre 0,10 i 2fr/m, però poden arribar a valors que superen les 8fr/m. Les dades obtingudes a partir de les tècniques TLS, mostren valors constants i similars en el diferents fronts estudiats. S'observen certes diferències en el Lateral-W suggerint la dependència respecte la orientació de l'aflorament tal i com hem comentat anteriorment. Tot i així, les dades no superen les 4fr/m d'intensitat. En canvi, en la resta de tècniques, es mostra major variabilitat de les dades P10, en funció del frontal i la unitat mecànica estudiada. En el cas de la tècnica de la fotogrametria, aquestes diferències s'observen especialment en les unitats mecàniques 1, 2 i 3. Aquests valors, mostren una major densitat de fractures en aquelles unitats per a les famílies II i III, que podria ser conseqüència del processat d'aquestes dades, que combinada amb el capbussament aparent d'aquestes, tendeix a agrupar-les en les primeres unitats mecàniques. En tot cas, aquestes dades solen ser puntuals. Els valors obtinguts del P21 o P32, van en consonància amb el P10 tal i com cabria esperar. S'utilitzen dues metodologies per el càlcul del P32. Els resultats mostren major variabilitat de les dades en el primer cas P32(C13) i certa constància dels resultats en el segon cas amb el P32(C23).

L'anàlisi s'ha realitzat amb l'estudi dels afloraments individualment i per cada tècnica de captura. Per altra banda, si s'analitzen les dades obtingudes per a cada tècnica de captura, comparant el primer frontal amb els successius frontals orientats cap al nord, s'observa certa homogeneïtat de les dades representades. Les unitats mecàniques es traslladen amb un lleuger capbussament cap al oest i els gràfics estadístics de les mitges representades mostren valors similars quan analitzem individualment cada tècnica. En aquest sentit, observem que les dades en conjunt, tot i ser diferents entre les tècniques, són coherents amb els afloraments analitzats.

Les fractures del Lateral-W s'han estudiat únicament amb les tècniques TLS. El perill de caigudes de blocs ha limitat les zones d'accés per realitzar els *scan-lines* i l'adaptació al calendari de l'explotació, en algunes ocasions no ha permès la realització d'alguns escanejos. (p.e. no s'ha

pogut realitzar l'escaneig amb TLS en el Frontal-N-4 ni la captura amb fotogrametria del Lateral-W).

En el cas de la tècnica TLS, si es comparen les dades del Frontal 1 orientat cap al nord, amb les obtingudes en el Lateral orientat cap al oest, observarem noves diferències. Tot i que les unitats mecàniques es traslladen lateralment confirmant la homogeneïtat lateral, els gràfics estadístics ens presenten diferències entre les famílies, demostrant de nou, la dependència d'aquesta tècnica amb la orientació de l'aflorament i les fractures. S'observa com la família orientada paral·lela al frontal estudiat, és la que destaca en nombre de mesures captades. En el cas del frontal 1, la família-I i en el cas del Lateral-W, és la família-III i en menor mesura, la família-II.

Els models DFN tridimensionals s'han construït amb les dades únicament dels Frontals orientats cap al nord, i no s'ha tingut en compte les del Lateral-W. Aquesta decisió que semblaria banal té les seves repercussions com ja hem comentat anteriorment.

Avaluació de les tècniques remotes en la construcció d'un model DFN

La construcció dels tres models DFN tridimensional, pretén identificar les diferències geomètriques entre la representació de la xarxa de fractures, també analitzades en els afloraments 2D, però sobretot està encaminat a obtenir la seva representativitat en termes de connectivitat, porositat i permeabilitat.

Amb l'objectiu de disposar de les dades equivalents a les captades, la representació d'aquestes fractures s'ha realitzat estocàsticament, a partir de les dades de les superfícies de les fractures agrupades per famílies i introduïdes en cada model. S'han representat també, els plans de les 4 discontinuïtats estratigràfiques més penetratives. La decisió de no introduir les unitats mecàniques identificades pretén focalitzar l'avaluació sobre la representativitat de les tècniques remotes en la construcció dels models DFN, que és l'objectiu definit.

Finalment, l'obertura, necessària per determinar la porositat i permeabilitat secundària del model, únicament s'ha pogut mesurar amb les tècniques d'*scan-line*. Aquestes mesures s'han realitzat sobre els afloraments estudiats, on s'identifica visualment una sobre-dilatació de les fractures per l'efecte dels explosius utilitzats per la pedrera en la seva activitat extractiva.

A continuació es detallen les discussions dels aspectes més rellevants, dels models DFN obtinguts.

Mida de les fractures

La **mida de les fractures** és un altra propietat rellevant en la representació de les fractures i dependrà de la longitud i alçada mesurada. En aquests models els valors mitjos de les àrees de les fractures obtinguts, que es presenten a la Taula 3.13, mostren que la tècnica de la fotogrametria s'ajusta millor als valors obtinguts amb els *scan-lines*. Del resultat obtingut amb

TLS, s'observen valors excessivament petits i poc representatius de les mides de fractures presents en aflorament. Tal com es comenta en el apartat anterior aquestes dades són conseqüència del processat aplicat al extreure alçades i longituds.

	DFN _{TLS}	DFN _{SfM}	DFN _{Scan-line}
Àrea mitja [m ²]	0,813	414,204	385,554

Taula 3.11: Valors d'espaiat mig i intensitat (P_{10}) obtinguts amb el *FracMan*.

Espaiat entre les fractures

Els valors mitjos **d'espaiat o d'intensitat de les fractures** obtinguts en els models DFN, suggereixen una major correspondència entre el model construït a partir de les dades de TLS i els *scan-line* (Taula 3.11). En el cas de la tècnica de fotogrametria, s'obtenen espaiats 3 vegades majors, mostrant de nou la dependència d'aquestes dades a la digitalització representada per l'usuari.

	DFN _{TLS}	DFN _{SfM}	DFN _{Scan-line}
Espaiat mig [m]	0,1	0,33	0,1
Intensitat P10 [1/m]	10,3	3,1	11,3

Taula 3.12: Valors d'espaiat mig i intensitat (P_{10}) obtinguts amb el *FracMan*.

Per demostrar quines de les dades obtingudes són representatives del medi fracturat, s'ha realitzat la comparació dels valors d'espaiat obtinguts amb els valors d'espaiat mesurats en sondejos bibliogràfics, a partir del *Rock Quality Designation (RQD)*, definit per Deere, (1969) i, Deere i Deere, (1989). La mesura de l'RQD es presenta com el percentatge de trossos de testimonis recuperats majors de 10 cm en relació a la longitud total i la recuperació de la mostra d'un sondeig (veure Subsecció 2.3.1). A través del programa *FraMan* s'han generat pous virtuals (veure Figura 3.31) i s'ha calculat el valor d'RQD. Concretament, el sondeig S-1 (Figura 3.30) es correlaciona amb l'àrea d'estudi tal i com s'observa en el tall geològic representat a la Figura 3.13 i és el que s'ha utilitzat per obtenir l'RQD en el tram entre 6 i 30m. A la Taula 3.12, es resumeixen els resultats obtinguts dels pous simulats en els models DFN i del sondeig mecànic obtingut bibliogràficament.

	ESTUDIS GEOTÈCNICS	DISCRETE FRACTURE NETWORK		
	S-1 Sondeig	DFN _{TLS}	DFN _{SfM}	DFN _{Scan-line}
RQD(10cm)	76.77%	73.76%	96.74%	68%

Taula 3.13: Valors RQD extrets d'estudis geotècnics i dels models obtinguts amb el *FracMan*.

Els valors d'RQD generats en els models DFN_{TLS} i el DFN_{scan-line} s'aproximen al valor obtingut en el sondeig S-1. En canvi, l'RQD obtingut en el model DFN_{SfM}, és un 20% superior al obtingut amb el S-1, mostrant espaiats majors, és a dir, una menor intensitat de fractures. Podríem afirmar així,

que els models de fractures que s'aproximen millor a la realitat, pel que respecta a la propietat d'espaiat, són els obtinguts de les tècniques TLS i *scan-lines*, encara que mostren una intensitat de fractures lleugerament superiors. Aquesta diferència podria ser conseqüència de la descompressió dels materials en aflorament i de les voladures de la pròpia activitat extractiva, incorporant noves fractures "artefacte".

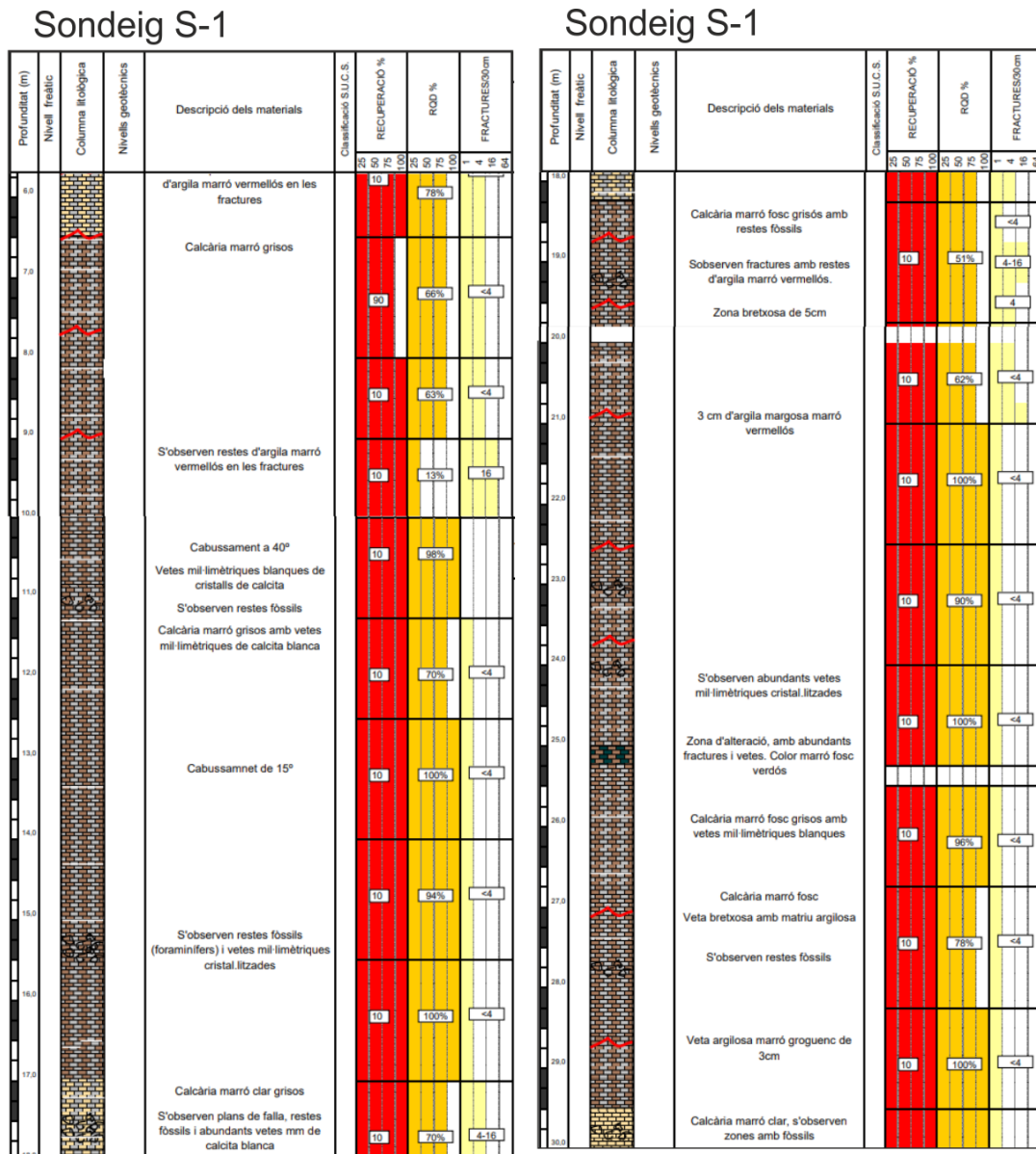


Figura 3.30: Testificació del sondeig bibliogràfic S-1 on es recull el registre d'RQD.

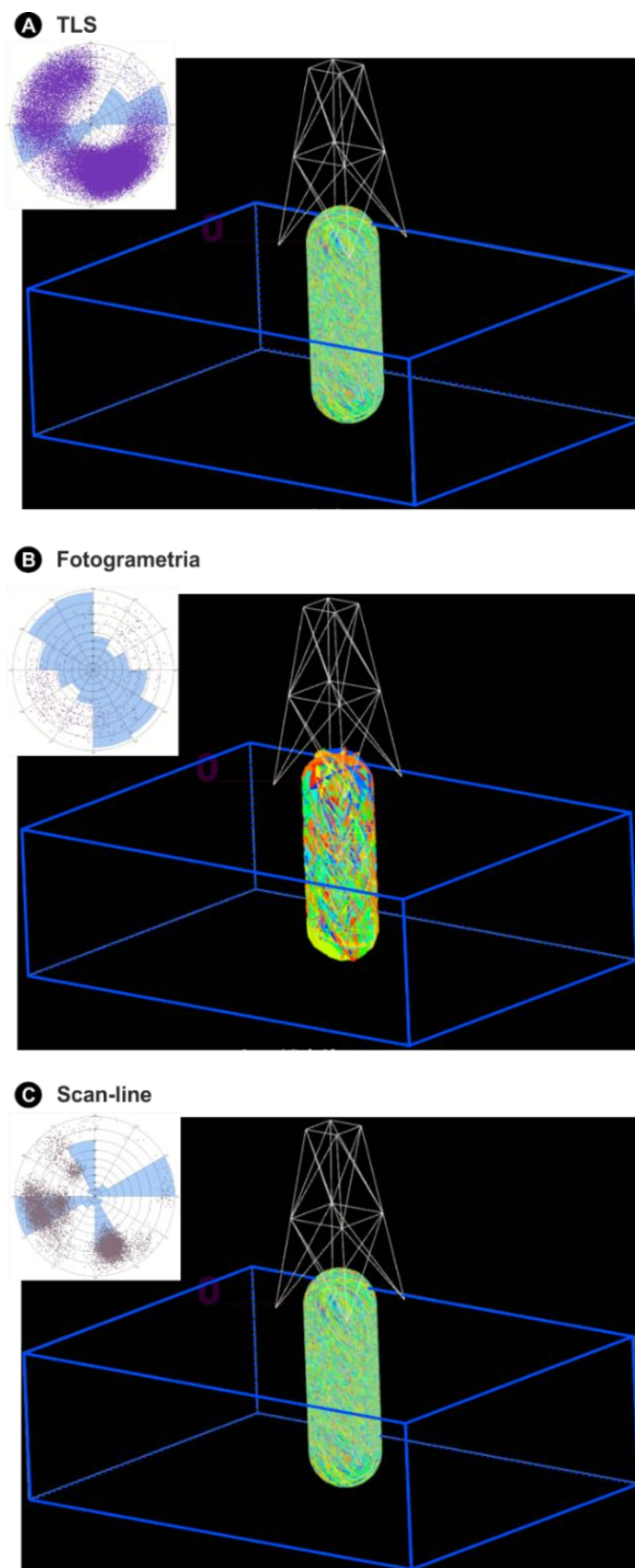


Figura 3.31: Pous virtuals generats en *FracMan*. De 25 m de fondària i un diàmetre de 200mm mitjançant les diferents tècniques. (A) TLS (B) Fotogrametria i (C) scan-line

Porositat i permeabilitat

Les obertures generades per el programa *FracMan* (*FracMan® User's Manual*, 2020), un cop processat el model, mostren valors d'obertura mitjos molt similars. Aquest fet, corrobora que les obertures introduïdes han sigut les mateixes. (veure Taula 3.14 i Figura 3.32). El valor mig d'obertura més gran el té la fotogrametria.

	DFN _{TLS}	DFN _{SfM}	DFN _{Scan-line}
Obertura mitja (m)	0,0038	0,0046	0,0031
Obertura – Desviació estàndard	0,0027	0,0027	0,0024

Taula 3.14: Valors mitjos obtinguts d'obertura, extretes del programa *FracMan*, per els models DFN construïts.

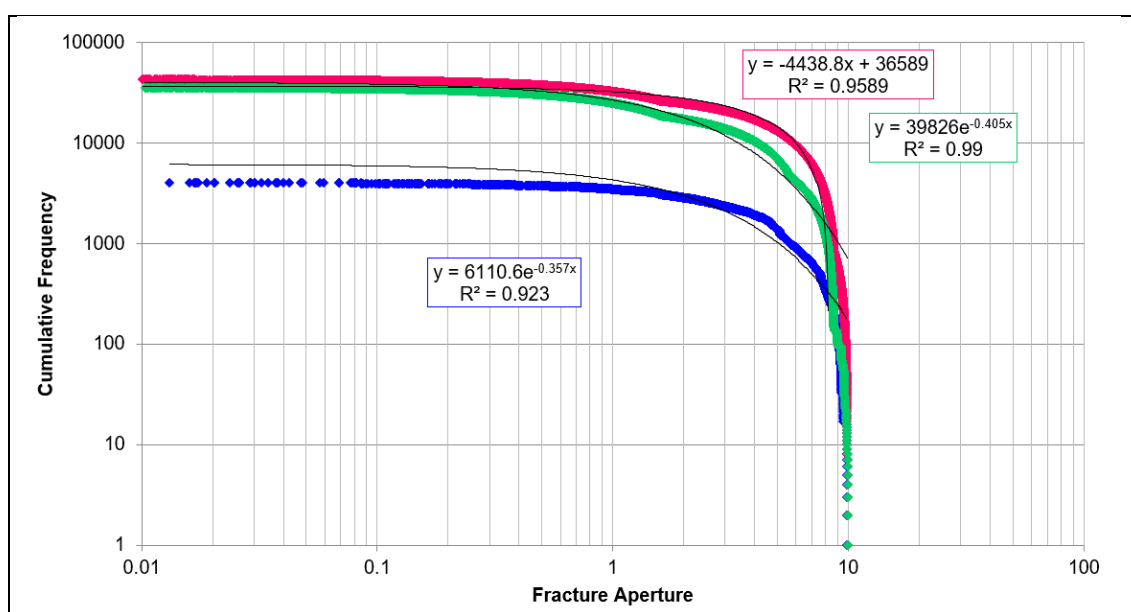


Figura 3.32: Corbes de freqüència acumulativa de les obertures de les fractures per a cada tècnica de captura extretes del programa *FracMan*. Llegendes: color Blau: Fotogrametria, Vermell: TLS i Verd: Scan-line.

Cal recordar que la pedrera realitza els avanços amb voladures controlades (Persson, 1994). En aflorament, s'identifiquen fractures al voltant de les barrines perforades sobre-dilatades i la creació de noves petites fractures derivades de la ona expansiva de l'explosió. Les conseqüències poden ser, una major intensitat de fractures derivat de la creació de noves fractures i la sobre-dilatació de les obertures. Si s'afegeix, que les obertures en aflorament habitualment estan afectades per la descompressió o pels efectes de l'erosió. Per tot això, els valors d'obertura mesurats es considera que estan amplificats i no són representatius del medi fracturat en fondària.

A partir dels models de traces generats amb les 3 tècniques de captura, s'obtenen els resultats mitjos de la porositat (P_{33}) i de la permeabilitat de la xarxa de fractures (veure Taula 3.15 i Figura 3.33).

	DFN _{TLS}	DFN _{SfM}	DFN _{Scan-line}
P33 (fracVolume/volume)[%]	5,4	5,7	5,7
Permeabilitat mitja (mildarcy)	616.431,14	402.951,42	642.512,15
Permeabilitat – Desviació estàndard	366.099,87	414.031,9	370.238,28
Permeabilitat mitja (cm ²)	6,04 * 10 ⁻⁶	3,98 * 10 ⁻⁶	6,34 * 10 ⁻⁶

Taula 3.15: Característiques de la xarxa de fractures, extretes del programa FracMan, per els models DFN construïts.

Encara que s'han obtingut diferències en la representació de les principals propietats geomètriques, els valors obtinguts de porositat (P_{33}) són valors molt similars tal com es recull a la Taula 3.15. La porositat en el DFN_{SfM} i DFN_{scan-line} és de 5,7% i en el DFN_{TLS} de 5,4%. Aquests valors, tal i com cabria esperar, són majors als obtinguts per a la porositat primària (roca matriu més micro-fractures) entre 0.5-1.4%. Els valors bibliogràfics, com els definits per Sanderson, 1998, presenten una porositat eficaç entre 0,1 i 5% en el cas de les calcàries sense carstificacions, com seria el nostre cas. Amb els valors obtinguts, pel que es refereix a la porositat secundària, ja s'observa que ens trobem lleugerament per sobre dels rangs habituals, reforçant la hipòtesis de que les obertures recollides estan sobre-dilatades.

L'estimació de la permeabilitat mitja de la xarxa de fractures en el model DFN_{SfM}, és de 4×10^5 mD amb una desviació estàndard de $4,1 \times 10^5$ mD, lleugerament superiors als models DFN_{TLS} o DFN_{scan-line}, amb valors mitjos de de $6,2 \times 10^5$ mD i una desviació estàndard de $3,7 \times 10^5$ mD en tots dos casos (veure Taula 3.15). Els valors de permeabilitat obtinguts de la roca matriu són molt inferiors, per sota de 0.0005mD, pràcticament impermeables.

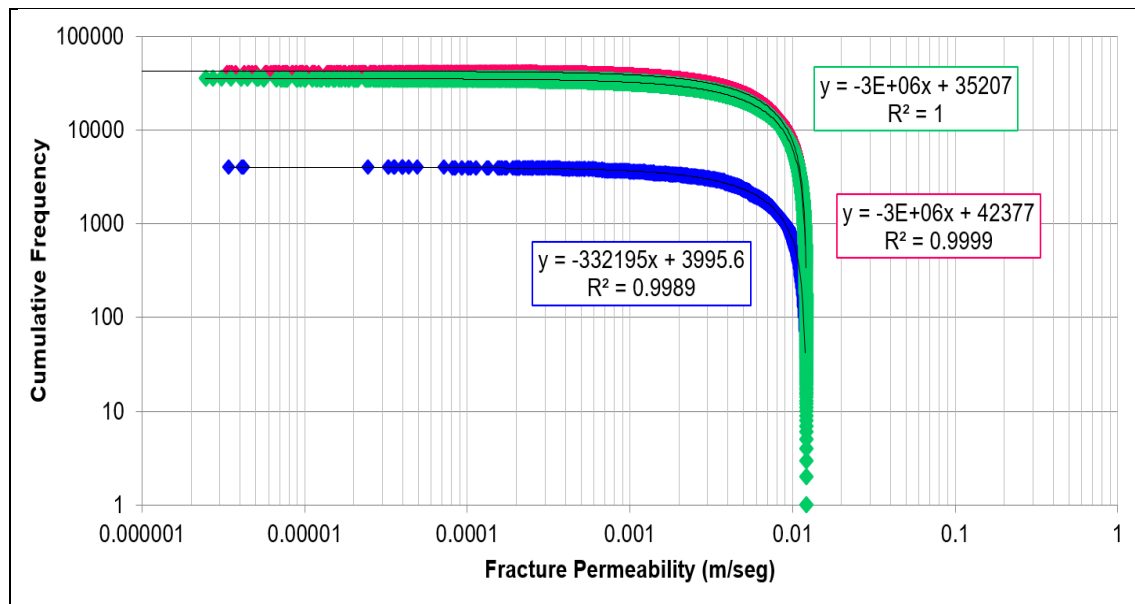


Figura 3.33: Corbes de freqüència de la permeabilitat per a les tècniques de captura extretes del programa FracMan. Llegendes: color Blau: Fotogrametria, Vermell: TLS i Verd: Scan-line.

Hoek, 1981, proposa mesurar la permeabilitat a partir de les obertures i les juntes. Amb aquesta metodologia, en el cas de la Fou s'identifica un rang de permeabilitat entre 10^6 i 10^3 mD. Els valors obtinguts de permeabilitat en el model DFN estarien dins del rang. Però si tenim en compte els valors habituals de permeabilitat en aquests materials segons Freeze i Cherry, (1979), les permeabilitats són més baixes, entre 10^{-1} i 10^3 mD i per tant els valors obtinguts de nou confirmarien la influència de la sobre-dilatació en la mesura de les obertures i la dependència de la permeabilitat en aquesta propietat.

Per obtenir una validació més objectiva de les dades i calibrar els models, seria necessari la realització d'assajos hidràulics amb l'objectiu d'estimar els valors d'obertura equivalent (veure Subsecció 2.2.2). En aquest experiment no s'han pogut realitzar. Si que es va realitzar un assaig amb el GPR a partir de l'aportació controlada de 12 m^3 d'aigua (durant 30 minuts) per obtenir un valor de permeabilitat hidràulica (veure Subsecció 3.2.4). Es van plantejar aquests estudis complementaris ja que s'havien realitzat experiments semblants amb bons resultats a treball com el de Marchesini et al., (2013) i Guo et al., (2019). En el cas d'estudi de la Fou, no s'han obtingut resultats aplicables. La falta de la resposta en fondària d'aquesta tècnica aplicada al subsol de la pedrera, ha estat la principal limitació.

Altres propietats que no s'han estudiat en aquesta tesi doctoral, també poden ser determinants en el comportament del flux, com les rugositats, les terminacions de les fractures o les relacions entre elles. Aquesta tesi, s'ha centrat en l'estudi de les principals propietats geomètriques que es poden registrar i caracteritzar amb les tècniques remotes.

3.2.6. Conclusions

En medis fracturats amb una permeabilitat matricial insignificant com és el cas de La Fou, la resposta hidràulica de les xarxes de fractura és d'especial importància per entendre les propietats globals de flux i transport, ja que les vies de flux principals es regeixen per fractures connectades (Neuman, 2005). Els medis naturals fracturats solen mostrar una forta complexitat hidràulica procedent de la topografia interna de les fractures (és a dir, la geometria dels espais buits dins d'una sola fractura).

L'aplicació de tècniques remotes (TLS i fotogrametria) i els *scan-lines*, per a la reconstrucció de models discrets de fractures (DFN) en tres dimensions, en l'experiment realitzat a la Pedrera de la Fou de Vallirana, permet obtenir unes conclusions: (A) sobre el medi rocós estudiat, (B) sobre les tècniques de captura per la caracterització de xarxes de fractures i (C) sobre els models DFN tridimensionals de xarxes de fractures reconstruïts a partir de dades obtingudes amb tècniques remotes.

(A) Sobre el medi rocós estudiat

L'àrea d'estudi, es correspon amb una successió de materials carbonatats que s'interpreta com a una plataforma carbonatada d'aigües poc profundes i baixa energia, del Barremià i l'Aptià (Cretaci Inferior).

Els assajos de porositat i permeabilitat realitzats amb mostres de roca presenten valors de porositat entre 0.5-1.4% i valors de permeabilitat per sota de 0.0005mD. Aquests resultats suggereixen que la matriu rocosa és pràcticament impermeable i per tant, no serà necessari utilitzar un model de doble porositat per simular el flux. En el cas d'estudi l'avaluació de les tècniques s'ha utilitzat el model discret de fractures (DFN) on únicament es contempla la porositat i permeabilitat generada per les fractures.

S'identifiquen 4 superfícies estratigràfiques penetratives amb una orientació mitja de 280/23°, que intervenen en el transport del flux i que s'han introduït en els models. La seva rellevància es demostra a partir de les observacions a camp i les dades extretes dels perfils tomogràfics.

(B) Sobre les tècniques de captura per la caracterització de xarxes de fractures

Les tècniques de captura utilitzades durant la tesi doctoral, han permès identificar les principals propietats de les fractures estudiades. Les propietats individuals representades són el nombre de fractures, l'alçada i la longitud i les propietats col·lectives, l'espaiat, els valors d'intensitat, P_{10} , P_{20} , P_{21} i P_{32} , el coeficient de variació i el Ratio d'espaiament de les fractures (FSR). L'única propietat que s'ha mesurat només amb tècniques directes d'*scan-lines*, és l'obertura, la qual no es pot mesurar a partir de tècniques remotes.

La fracturació estudiada, es presenta molt verticalitzada i intensa, on s'identifiquen fins a 4 famílies de fractures amb característiques pròpies. Família I WNW- ESE, Família II NNW/SSE, Família III, N-S i Família IV NE-SW. Totes les tècniques obtenen dades de les quatre famílies, sense contradir-se. La gran quantitat de dades recollides en les tècniques remotes i la distribució desbalancejada per famílies captades, dificulten la delimitació de les **famílies de fractures** amb aquestes tècniques.

La tècnica amb **major quantitat de fractures caracteritzades** ha estat el TLS amb més de 52.000, després la tècnica de fotogrametria amb 1.094. En aquests casos la informació s'obté de tota la superfície de l'aflorament. En canvi, amb els *scan-lines* s'han recollit 398 fractures en 89 metres lineals d'*scan-lines* en sectors reduïts i en zones on s'ha pogut accedir. En aquest sentit, les tècniques remotes, permeten caracteritzar superfícies majors que amb els *scan-lines*. Com a conseqüència i especialment en estudis regionals, les remotes permeten obtenir major informació de les heterogeneïtats de l'àrea estudiada i reduir incerteses en les interpretacions.

En totes les tècniques remotes, **la orientació de l'aflorament i de les fractures** influenciarà en els resultats. En el cas de les tècniques TLS s'identifiquen les superfícies i per tant, les famílies

amb orientacions ortogonals a l'aflorament seran les millor captades. En canvi, la tècnica de la fotogrametria, digitalitza les traces de les fractures i en conseqüència identificarà amb major detall les fractures amb les orientacions més perpendiculars a l'aflorament. Els *scan-lines* exploren sectors limitats on la disposició aparent de les fractures pot dificultar algunes mesures puntuals.

La mida de les fractures es mesura a partir de l'**alçada i la longitud** de les fractures capturades. La majoria de les longituds i en menor mesura les alçades, queden censurades pel propi aflorament. El percentatge de fractures complertes identificades és baix. S'observen majors correspondències entre la tècnica de fotogrametria i els *scan-lines*, amb rangs de valors molt variables de centímetres fins a 18m. En les Famílies II, III i IV, és on s'obtenen majors valors mitjos d'alçades i en canvi, en la Família I, els valors estan més continguts arribant als 6 metres d'alçades mitges màximes. Cal tenir en compte que en aflorament, la família III IV, superen en alguns casos els 25m d'alçada (alçada dels afloraments), valors que no podran ser captats. En canvi, els valors obtinguts amb dades TLS mostren unes alçades i longituds de fractures mitjanes inferiors al metre, mesures que no representen els afloraments estudiats. Aquesta tècnica, per tant, ha comportat una limitació respecte les altres tècniques a l'hora de identificar les alçades i longituds.

En aquesta tesi doctoral, la distribució de la xarxa de fractures es representa amb els valors de l'**espaiat o de la intensitat de fractures** (P_{10}). Els resultats mostren valors mitjos obtinguts amb tècniques TLS i els *scan-lines* que es corresponen millor amb la realitat de l'aflorament. En canvi, la tècnica de fotogrametria presenta espaiats que poden arribar a triplicar els valors obtinguts amb les altres tècniques. Aquests resultats, mostren una limitació de la tècnica davant de la resta, conseqüència del procés de digitalització on l'usuari tendeix a destacar les fractures més rellevants i a filtrar les més petites o imperceptibles. En general s'observa que els espaiats de la Família I són centimètrics i bastant regulars, en canvi la resta de famílies té valors més variables.

En la representació de la distribució de les fractures també s'observen algunes diferències. El cas més clar ha estat la dels **corredors de fractures** identificats en els afloraments. Aquests corredors, són zones amb major intensitat de fractures, que arriben a tenir espaiats mil·limètrics. Els gruixos totals d'aquests corredors són de 20-30cm i de pocs metres d'alçada, que acostumen a ser zones preferents de flux. Amb la tècnica d'*scan-line*, s'han mesurat alguns d'ells intencionadament. Els resultats mostren que tot i haver agafat les dades dels corredors en els *scan-lines*, els gràfics de coeficient de variació no recullen aquestes agrupacions de fractures que s'associen als corredors, en canvi, les tècniques TLS sí que les han representat. Aquestes distribucions influenciaran en els models DFN. En aquest sentit i en treballs futurs, caldria estudiar-los amb més detall com es proposa a l'article de Sanderson i Peacock, (2019) mitjançant els *scan-lines* o també es poden identificar intencionadament amb la digitalització per fotogrametria. D'aquesta manera es podrien representar en el model posteriorment i de forma manual, per recollir-los en totes les tècniques.

(C) Sobre els models DFN tridimensionals de xarxes de fractures reconstruïts a partir de dades obtingudes amb tècniques remotes

La xarxa de fractures ha quedat representada en els 3 models tridimensionalment, un per a cada tècnica de captura: DFN_{TLS}, DFN_{SfM} i DFN_{scan-line} (veure Figura 3.25). Les diferències identificades i analitzades a partir dels gràfics representatius de cada aflorament, han quedat també incorporades en els model DFN.

- La representació de les **orientacions de les famílies** dels models, són analitzades a partir del centre de densitats dels pols d'orientacions. Es presenten petites divergències en funció de la tècnica, tal i com ja hem comentat, conseqüència de la captura però sense rellevància. Els valors mitjos del centre de distribució dels pols per totes les tècniques i per a cada família són: Família I 190/12, Família II 234/21, Família III, 282/21 i Família IV 142/19. Aquestes dades s'ajusten a les orientacions inicialment identificades amb les dades capturades.
- En el cas de la mida de **les fractures** s'obtenen valors mitjos al voltant de 400m² en el model DFN_{SfM} i DFN_{scan-line}. En canvi, els valors representats de les dades TLS, mostren superfícies que no arriben al metre quadrat. Del resultat obtingut amb TLS, observem aquests valors excessivament petits i que no representen les dades reals. Cal recordar, que el pre-processat de la tècnica ha influenciat en aquests resultats.
- En el cas de **l'espaiat entre fractures**, presenten valors mitjos de 0,10m en el model DFN_{TLS} i en el DFN_{scan-line}. En canvi, de nou, la tècnica de la fotogrametria ens mostra valors d'espaiat mitjos majors, en aquest cas 0,3m. Aquests valors també s'expressen amb **la intensitat de les fractures (P10)** va amb consonància amb les dades obtingudes de l'espaiat.
- Amb els **valors d'RQD** extrets d'un sondeig pròxim (S-1) i la simulació de pous en els models per obtenir l'RQD d'aquests, es demostra que les distribucions de fractures que s'ajusten millor al medi fracturat del subsòl, són les obtingudes en els models de TLS i *scan-line*. Tot i les similituds dels valors, les dades dels models TLS i *scan-line*, presenten una intensitat lleugerament superior als obtinguts en el S-1, que podrien ser conseqüència de la captura involuntària de "fractures artefactes" generades pels explosius de l'activitat extractiva.

A partir del mapa de traces, obtenim que la porositat mitja en el DFN_{SfM} i DFN_{scan-line} és de 5,7% i en el DFN_{TLS} és de 5,4%. L'estimació de la permeabilitat mitja de la xarxa de fractures en el model DFN_{SfM}, és de 4x10⁵ mD, lleugerament superiors als models DFN_{TLS} o DFN_{scan-line}, que ofereixen valors mitjos de 6,2x10⁵mD. Comparant la informació de sortida dels models, s'observa que, per a la construcció de models DFN, s'obtenen majors similituds en els resultats de

permeabilitat quan s'apliquen les tècniques de TLS i *scan-lines*. En canvi, quan es fa ús de la fotogrametria, resulten valors de porositat i permeabilitat lleugerament més baixos. Aquest valors, es poden considerar equivalents.

En aquest cas concret, les obertures mesurades presenten certa sobre-dilatació conseqüència de la descompressió, però sobretot per l'ús d'explosius en l'activitat minera, valors que podrien haver influenciat augmentant lleugerament els valors de porositat i permeabilitat de les fractures.

En definitiva, els resultats obtinguts suggereixen, que els tres models permeten obtenir la porositat i permeabilitat del mapa de traces i que els valors obtinguts són similars. La tècnica de fotogrametria, presenta valors lleugerament més baixos de permeabilitat i la tècnica TLS de porositat. Per tant, les diferències geomètriques representades en el conjunt de dades, no presenta importants conseqüències en l'estimació dels valors de porositat i permeabilitats de la xarxa de fractures en el cas d'estudi de La Fou.

3.3. PROPOSTA METODOLÒGICA PER A LA IDENTIFICACIÓ DE BLOCS CAIGUTS AMB NÚVOLS DE PUNTS I *MACHINE LEARNING*

En aquesta secció, es presenta la metodologia desenvolupada durant la tesi i publicada a la revista *Remote Sensing* (Blanco et al., 2022) per la detecció de blocs caiguts, a partir dels monitoratges amb TLS i l'ús de tècniques de *machine learning*.

La finalitat no es altra que millorar la metodologia estàndard existent (Secció 2.4.2) on la identificació i classificació dels blocs caiguts és pràcticament manual. La proposta metodològica es concreta en una eina digital per augmentar l'automatisme del procés d'identificació i classificació, i reduir els criteris subjectius de l'usuari.

Al proporcionar una eina més eficient, s'aconsegueix també, confeccionar inventaris històrics dels vessants rocosos on es donen caigudes de blocs. El monitoratge d'aquests escarpaments, basades en la comparació de núvols de punts per identificar els blocs caiguts donen una informació geomètrica d'aquests blocs com pot ser la situació, morfologia i el volum així com la data del despeniment. D'aquesta manera, les dades recollides en els inventaris històrics caracteritzen la dinàmica de la vessant permeten així l'assignació de la infraestructura de mitigació del perill més adient (Williams et al., 2019).

La nostra proposta es basa en l'adaptació de les metodologies existents per la creació de nova informació geomètrica continguda a les dades digitals. Així, tota la informació nodreix a les eines d'aprenentatge automàtic per poder classificar els blocs caiguts (Angluin, 1988; Cohn, 1996).

Les següents seccions descriuen l'adaptació de la tècnica estàndard acceptada, prenent com exemple l'àrea d'estudi de Degotalls a Montserrat (Barcelona). Aquests escarpament fracturat té un patró de fractures descrit per Alsaker et al., (1996) que afavoreix les caigudes de blocs de dimensions decamètriques esdevingudes, pròximes a zones d'infraestructures del Monestir de Montserrat (Janeres et al., 2017).

3.3.1. Proposta metodològica

Aquesta secció exposa la proposta metodològica desenvolupada durant la tesi per detectar blocs caiguts en monitoratges de vessants i amb núvols de punts. La principal aportació d'aquesta proposta és l'ús de tècniques de *machine learning* en l'àmbit de les caigudes de blocs per aconseguir els objectius proposats. Part d'aquesta nova metodologia està basada en la metodologia estàndard existent (veure Subsecció 2.4.2) però evolucionada per aplicar tècniques de classificació amb *machine learning*. Per validar la metodologia s'ha realitzat un estudi d'inestabilitat de vessants al massís de Montserrat (Barcelona), aprofitant els treballs i ja existents en aquesta àrea (Royan 2915, Janeras et al., 2017).

La proposta metodològica d'aquesta tesi respecte la metodologia estàndard habitual presentada a la Subsecció 2.4.2, consisteix en la modificació de les etapes de: C) Càlcul de les diferències entre núvols de punts; D) Agrupament de punts i; E) Classificació de grups de punts candidats a pertànyer a blocs caiguts, tal i com queda representada a la Figura 3.34.

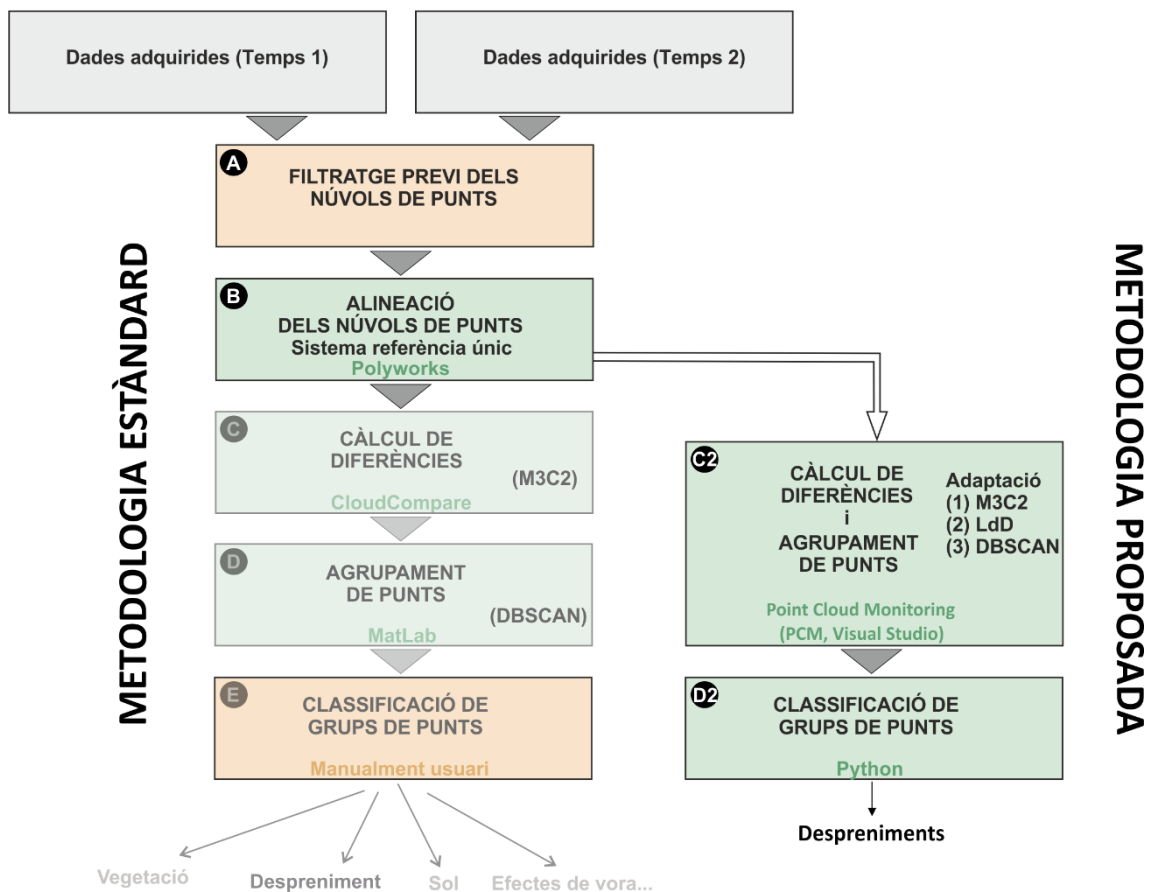


Figura 3.34: Esquema del flux de treball del mètode actual per identificar caigudes de roques a partir de dades del núvol de punts i el flux de treball de la nova metodologia proposada en el cas exemple de la Muntanya de Montserrat desenvolupat. S'incorporen la etapa (C2) per mesurar les diferències i agrupar els punts i la etapa (D2), per classificar els blocs caiguts amb *machine learning*.

Igual que la metodologia estàndard, els dos núvols de punts adquirits durant el monitoratge, han de ser capturats des de les mateixes estacions, condicions similars (posició, densitat de punts o orientació). I les dos primeres passes de la metodologia són compartides amb la metodologia estàndard. Aquestes, passes comprenen una lleugera classificació dels objectes en els núvols de punts (per eliminar vegetació, elements antròpics o altres classes que poden provocar desalineació), i l'alineació en un mateix sistema de referència entre el núvol de punts de referència i el comparat. Els programaris comercials com *Polyworks* o lliures com *CloudCompare* poden realitzar aquestes funcions.

Les contribucions de la nova metodologia permeten implementar nous atributs que caracteritzin les agrupacions de punts, amb la finalitat de possibilitar la classificació de les agrupacions amb mètodes de *machine learning* i incrementar l'automatització del procés. D'aquesta manera, la proposta de la metodologia es concreten en quatre contribucions:

(1) La adaptació de l'algorisme M3C2 desenvolupat per Lague et al., 2013, passant a ser de tipus de pla a pla a ser de tipus de punt-a-punt i calcular nous atributs per cada punt necessaris pel procés de *machine learning*. Aquests atributs són d'ordre geomètric, com la **diferència entre punts**, les orientacions dels vectors calculats de cada punt, així com els índexs de coplanarietat i col·linealitat dels punts utilitzats per calcular el vector de cada punt.

(2) Desenvolupament d'un mètode d' auto-calibratge per definir automàticament el **límit de detecció (LdD)** del sistema de captura digital. Aquest paràmetre permet atribuir les diferències entre núvols de punts a canvis reals en la superfície de l'aflorament, de l'error del sistema de detecció, és a dir, TLS o sistema fotogramètric utilitzat.

(3) Adaptació de l'algorisme DBSCAN. La finalitat és crear nous atributs estadístics associats a la fase **d'agrupament de punts**. Els atributs guarden relació amb la predominança de diferències positives o negatives (avanç o retrocés) calculades entre els núvols de referència i el comparat.

(4) L'anàlisi de diferents models de *machine learning* per classificar agrupacions de punts associades a caigudes de blocs.

En la nova etapa (C2), s'incorporen tres de les contribucions (1) la mesura de diferències, (2) el calibratge automàtic i (3) l'agrupament de punts. Aquestes han estat implementades en el programari lliure *Point Cloud Monitoring* o PCM i es desenvolupa en el llenguatge de programació BASIC en l'entorn Visual Studio 2019 (Studio, 2019).

En la nova etapa (D2), s'incorpora (4) la classificació d'agrupament de punts, que s'ha implementat en el llenguatge de programació Python amb el nom de *Cluster Classification*. Els mètodes implementats en aquesta tesi estan disponibles en el repositori GitHub de l'Institut de Recerca Geomodels de la U.B. (https://github.com/Geomodels-UB/Risknat_Detection) (accedit el 12 de juny de 2022).

Adaptació de l'algorisme M3C2

Essencialment, l'adaptació de l'algorisme M3C2 per mesurar diferències entre els punts dels núvols de punts de referència i comparats consisteix en calcular les diferències, més un afegit de nous atributs associats a l'esmentat càlcul. Per aquest motiu, l'algorisme M3C2 ha estat modificat passant de tipus model a model a un tipus punt a punt, canviant el mètode de recerca dels punts comparats.

En el processament, per cada punt del núvol de referència (P_{REF} , en la Figura 3.35) es calcula un vector normal (V_N , en la Figura 3.35) amb el pla definit pels punts inclosos dins del radi de cerca (R_s , en la Figura 3.35) fixat per l'usuari. El mètode per fixar el pla és el d'anàlisi dels components principals (Jolliffe, 2011) i el mètode dels valors propis i vectors propis (Woodcock, 1977). Des de cada punt de referència (P_{REF}) i en la direcció del vector calculat (V_N), l'algorisme busca amb el criteri del punt més proper el punt dins el núvol de punts de la vessant comparada (P_{COM} , en la Figura 3.35). Igualment, el cilindre assignat per Lague et al. (2013) és substituït per un doble conus truncat, amb la finalitat d'augmentar la flexibilitat de la cerca de punts. L'usuari, per tant, defineix la màxima i la mínima distància horitzontal del conus truncat ($MaxHd$ i $MinHd$ en la Figura 3.35), i la màxima distància vertical ($MaxVd$ en la Figura 3.35) de recerca al llarg del vector normal (V_N). Durant el procés, un vector és calculat en cada punt del núvol de punts de referència i també en el comparat. Els vectors són descompostos per conèixer l'orientació azimutal i pendent del vector per caracteritzar les vessants. A més, el mètode dels valors propis calcula els índexs de coplanarietat i col·linealitat dels punts escollits per caracteritzar la seva distribució.

En aquest procés són registrats 35 atributs que queden associats a cada punt del núvol de referència com mostra la Taula 3.16. Els nous atributs són els relacionats amb els components geomètrics de la distància (distància, distància horitzontal i distància vertical). També són considerats nous atributs els components dels vectors (i , j , k), angle entre els vectors de referència i comparat, les orientacions respecte al Nord i l'horitzontal i els índexs de coplanarietat i col·linealitat. Els paràmetres associats a la intensitat del pols làser del TLS i les textures RGB es cas d'obtenir-les i el número de punts seleccionats són igualment considerats com atributs i registrats.

Calibratge automàtic

Prèviament a l'agrupament de punts, és necessari quantificar quin és el LdD. Aquest paràmetre marca el valor límit que ens permetrà discernir de; si les diferències entre punts monitorats són diferències associades a un canvi real en l'aflorament o si són atribuïbles a la limitació de precisió del nostre sistema de registre (veure Secció 2.4.2 apartat Agrupació de diferències).

La proposta d'adaptació del algorisme DBSCAN per calcular automàticament el paràmetre LdD es En realitzar una segona captura digital de la superfície de l'aflorament en el mínim interval possible de temps (T_0-T_1) i calcular les diferències. D'aquesta manera es minimitza la possibilitat de canvis.

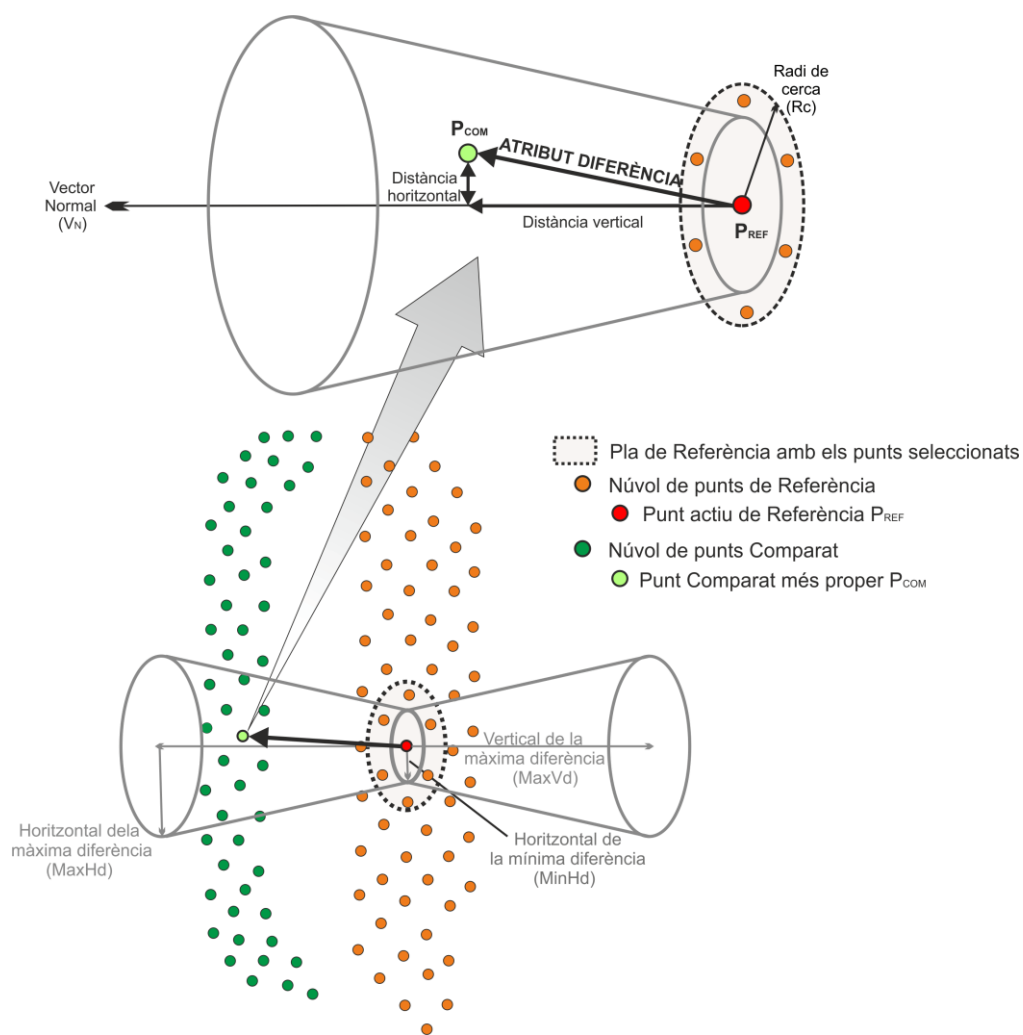


Figura 3.35: Representació de l'algorisme M3C2. A partir de cada punt del núvol de punts 3D de referència, es calcula un vector normal (VN) amb els punts veïns en un radi de cerca definit per l'usuari (R_s). El vector normal (VN) defineix la direcció en què es busca el punt més proper (P_{COM}) en el núvol de punts 3D comparat. La distància horitzontal màxima i mínima i les distàncies verticals màximes ($MaxHd$, $MinHd$ i $MaxVd$) són paràmetres assignats per l'usuari per definir un doble conus.

Atributs	Importància
Distància	Distància total entre el punt de referència i el comparat
Distància Vertical	Distància al llarg del vector normal
Distància horitzontal	Distància perpendicular al vector normal
Angle entre normals	Angularitat entre el vector de referència i el comparada
Sentit	La direcció del vector normal respecte la superfície
Vector	Vector normal (i, j, k) per a cada punt i veïns
Azimut	Orientació de cada vector respecte al nord geogràfic
Pendent	Orientació vertical de cada vector respecte a la horitzontal
Colinealitat	Grau de distribució dels punts veïns al llarg d'una línia
Coplanarietat	Grau de distribució dels punts veïns al llarg d'un pla

Taula 3.16: Resum dels atributs calculats amb l'adaptació de l'algorisme M3C2 (Lague et al., 2013). Els vectors comentats són referits tant en el núvol de punts de referència com en el comparat.

En una segona fase, la distribució dels valors de diferència segons la seva freqüència es fixat a un model de distribució gaussiana, on la funció es definida per la mitja i la desviació estàndard, establint la precisió dels sistema (Figura 3.36). Posteriorment, durant la fase de monitoratge del vessant i en un espaiat de temps major (T_1-T_2), la funció del valor diferències es fixat de nou a una funció de distribució gaussiana. Comparant les dues distribucions, els punts d'intersecció de les funcions indica els valors que superen els errors del nostre sistema i, per tant, s'assignen a canvis reals en el vessant (Figura 3.36). Aquests valors són el llindar entre el soroll i els canvis detectats.

Els factors que defineixen el soroll del nostre sistema són variats (Kromer et al., 2015) i poden variar en cada escaneig de la vessant. La distància entre l'aparell i l'angle d'incidència del pols làser emès i el seu retorn varia en funció de l'alçada de la vessant, la distància no es constant i per tant hi ha un rang. Aquests factors afecten a la precisió dels valors de les diferències entre núvols de punts. Amb la calibratge del LdD s'intenta minimitzar la pèrdua de precisió durant el calibratge utilitzant només els punts pertanyents a superfícies perpendiculars al TLS i a una distància mitjana representativa de l'aflorament.

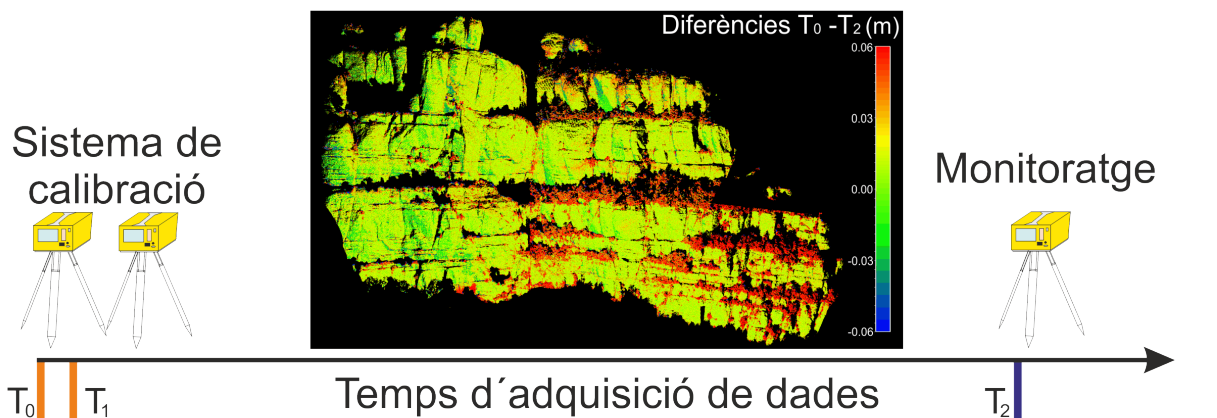
Adaptació DBSCAN

L'adaptació de l'algorisme DBSCAN consisteix a inspeccionar els valors de diferència dels seus punts veïns durant la agrupació, creant quatre nous atributs (Taula 3.18). Els quatre nous atributs calculats a cada punt de referència són paràmetres de caràcter estadístic per representar els valors de diferència predominants dels punts veïns. El predomini de les diferències durant el càlcul d'agrupament de punts es realitza en funció de la seva relació amb els valors del LdD. D'aquesta manera es discriminen quatre noves situacions en funció de la seva estadística: (1) **Indeterminació**, corresponent a l'error de l'aparell, (2) **avanç**, que es correspon a la deformació del massís, (3) **retrocés**, conseqüència d'un despreniment i (4) **sense predomini**, atribuïble a la vegetació, com es mostra a la Figura 3.37.

El càlcul dels predominis es realitza entre els punts veïns de cada punt del núvol de referència i el valor estadístic es incorporat com un atribut. Els següents quatre nous atributs representen el percentatge en els punts veïns de cada classe. D'aquesta manera es tracta de caracteritzar cada punt de l'agrupament amb en els seus punts veïns (veure Taula 3.17).

Atribut	Significat
Predomini	Valor de diferència predominant (avanç, retirada o soroll)
Predomini del soroll	Percentatge de punt classificat com a soroll
Predomini avançat	Percentatge de punt classificat com a avançat
Predomini de la retirada	Percentatge de punt classificat com a retirada

Taula 3.17: Nous quatre atributs calculats amb l'adaptació de l'algorisme DBSCAN. Els resultats s'incorporen a l'esdeveniment i als atributs de l'agrupació de punts.



(a) Sistema de calibració (T_0-T_1)

(b) Monitoratge (T_0-T_2)

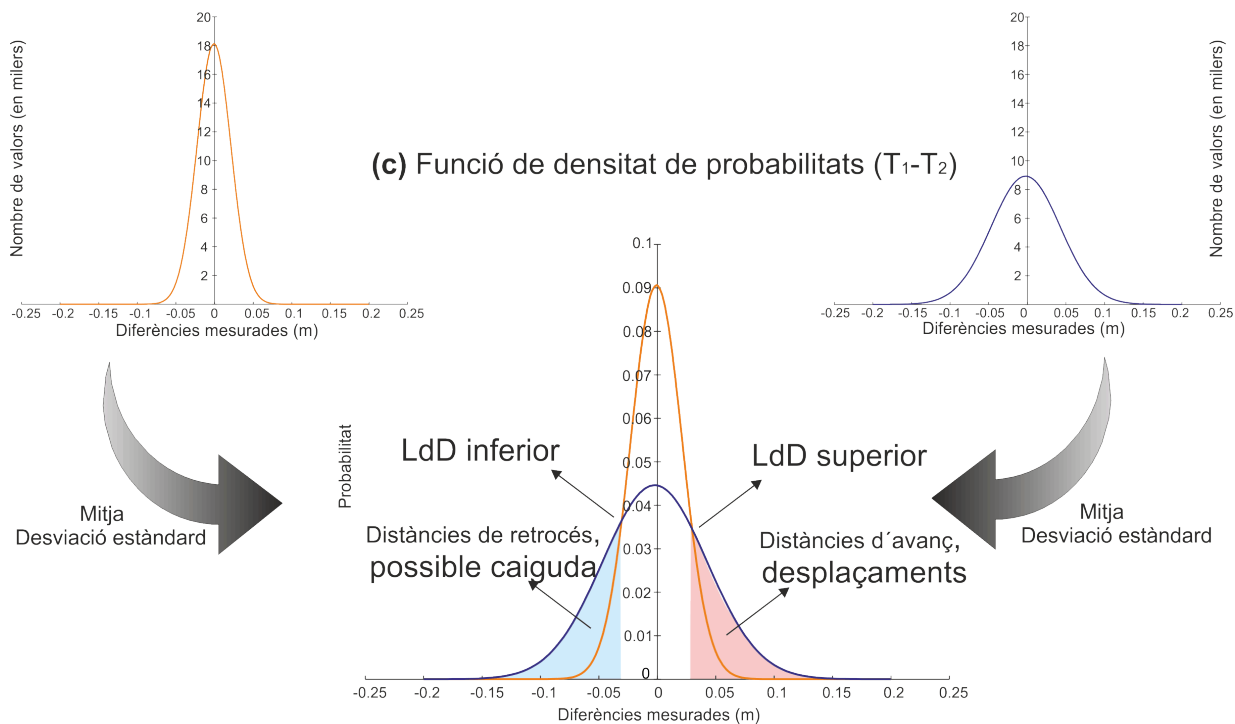


Figura 3.36: a) Distribució de les diferències entre l'adquisició de dades inicial i la repetició en l'interval de temps mínim (T_0-T_1), intentant repetir les mateixes condicions per calcular el sistema TLS-vessant. Les mesures se centren en zones perpendiculars i distàncies mitjanes representatives per optimitzar l'error. b) Distribució de les diferències calculades durant el monitoratge en un període de temps determinat. c) Superposició de distribucions per calcular el soroll del sistema (àrea en color blau). Les zones en color vermell marquen els valors assignats als processos d'avanç i els assignats en color blau com disminució en una pèrdua de volum.

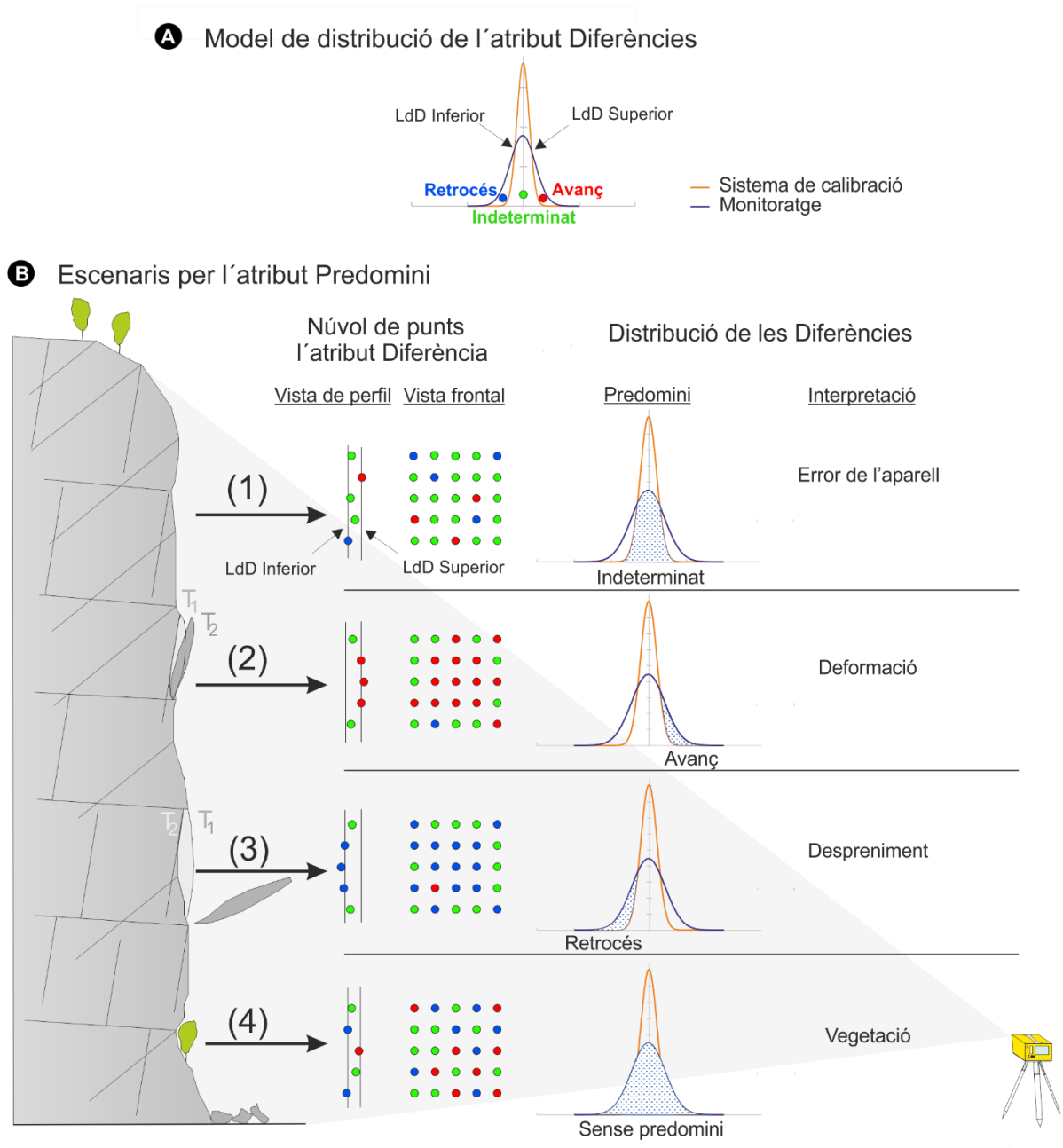


Figura 3.37: (A) Gràfic de distribució de referència de la presa de dades del dispositiu TLS (veure Figura 3.40). S'observen les diferències de distàncies entre punts, entre el sistema de calibratge i el monitoratge en la presa de dades. (B) Esquema gràfic dels possibles escenaris detectats en un escarpament de roca durant el monitoratge. Es relaciona amb la distribució de diferències recollides en els núvols de punts i la seva interpretació (a la dreta). (1) Soroll del sistema, les distàncies entre punts es troben dins de la corba gaussiana de calibratge. (2) Deformació. Quan els valors de les diferències es classifiquen de manera predominant com a avanços. (3) Caiguda de roques. Quan els valors de les diferències es classifiquen predominantment com a retirades. (4) Vegetació. Quan les classificacions d'avanç i retrocés tenen una distribució aleatòria i superen el LdD, matemàticament similar al predomini del soroll.

Els atributs que caracteritzen la resta de cada agrupació de punts (veure Taula 3.16) es donen estadísticament (mitjanes i desviacions estàndard) pel conjunt d'atributs de cada punt que formen les agrupacions. El centroide de cada agrupació es representa pel centre de massa de tots els punts, calculat amb el mètode de l'anàlisi de components principals. El volum de cada agrupació

es calcula per triangulació respecte una base plana, tant els punts de referència com els punts comparats es triangulen per separat respecte aquesta base plana comuna (volum de referència i comparat). El volum total correspon a la suma d'aquests volums, i sempre conservant el sentit positiu i negatiu respecte la posició TLS. El número d'atributs que caracteritzen cada agrupació de punts és de 56, encara que per el procés de *machine learning* solament es computen 33 atributs (Taula 3.18). Resolta l'agrupació de punts i els seus atributs, el programari PCM finalitza la seva tasca.

ID.	ATRIBUT	DESCRIPCIÓ
1	n *	Reference point index
2	m	Compared point index
3	Coordinate X	
4	Coordinate Y	Point reference coordinates
5	Coordinate Z	
6	Code n *	Reference index texture (0 n/a, 1 Intensity, 2 RGB, 3 RGB+ Int)
7	R ¹	
8	G ¹	Texture reference points RGB format
9	B ¹	
10	Intensity ²	Reference intensity texture
11	Vector i	
12	Vector j	Reference vector normal vector components
13	Vector k	
14	Orientation	Reference strike azimuth (degree)
15	Dip	Reference strike slope (degree)
16	Collinearity	Reference point index of collinearity
17	Coplanarity	Reference point index of coplanarity
18	Selected	Number of points to calculate the normal vector
19	Distance	Distance selected between closest and average
20	Vertical Distance	Vertical distance along vector with direction (+ or -)
21	Horizontal Distance	Horizontal distance component between points
22	Distance closest	Shorter distance between Refer. and Comp. point
23	Coordinate X	
24	Coordinate Y	Point compared coordinates
25	Coordinate Z	
26	Code m *	Compared index texture (0 n/a, 1 Intensity, 2 RGB, 3 RGB + Int)
27	R ¹	
28	G ¹	Texture compared points RGB format
29	B ¹	
30	Intensity 2	Compared intensity texture
31	vector i	
32	vector j	Compared vector normal vector components
33	vector k	
34	Orientation	Compared strike azimuth (degree)
35	Dip	Compared strike slope (degree)
36	Collinearity	Compared point index of collinearity
37	Coplanarity	Compared point index of coplanarity
38	Selected	Number of points to calculate the normal vector
39	Angle	Angle between Ref. and Comp. normal vectors
40	Angle Direction	Angle with direction
41	Minimal distance	Shortest distance between those inscribed in the geometric figure
42	Average distance	Average distance between those inscribed in the geometric figure
43	Maxima distance	Longest distance between those inscribed in the geometric figure
44	Dev. Stand distance	Dev.Stand distance between those inscribed in the geometric figure
45	Selected points	Number of points inscribed in the geometric figure

Taula 3.18: Atributs associats al núvol de punts després de calcular les diferències. * No es calcula com a atributs. ¹ Només s'utilitza un ordinador amb format RGB. ² Només s'utilitza el format Intensity calculat

Classificació automàtica de l'agrupació de punts

L'objectiu del darrer programari desenvolupat és la determinació de la naturalesa de cada agrupació de punts mitjançant els seus atributs, utilitzant mètodes de *machine learning*. Caiguda de blocs, vegetació, efectes de vora, deformacions del terreny o neu són les classes més freqüents que poden crear agrupaments de punts en el monitoratge, depenent de cada aflorament. Perseguint la identificació dels blocs caiguts, la classificació de cada agrupació de punts en la proposta de metodologia s'ha concretat de forma binària, és a dir: candidat de pertany a una caiguda de blocs o no, independentment si és vegetació, efecte de vora o altres. Amb aquesta finalitat, el flux de treball es divideix en diferents etapes com il·lustra la Figura 3.38 i també es pot consultar la secció 2.5.1: (1) Una etapa d'entrenament, on un número determinat de d'agrupacions classificada visualment com a caiguda de bloc serveix per caracteritzar les caigudes de bloc amb els seus atributs al model de *machine learning*, y (2) una etapa predictiva, per classificar agrupacions de punts de naturalesa desconeguda. Les agrupacions de punts etiquetades com a "Candidat", en la darrera etapa es validada manualment amb una comprovació sobre el terreny amb imatges d'alta resolució. La resta d'agrupacions etiquetades com a "No candidats" pertany a la resta d'esdeveniments amb el nom de "Desconegut".

La primera etapa del flux de treball comença amb la normalització de la col·lecció de dades que constitueixen les dades de les agrupacions de punts, és a dir convertir-les dins del rang [0–1] per evitar que certs atributs tinguin més pes que altres..

Probablement, alguns escenaris de classificació presenten un desequilibri entre el nombre d'agrupacions de cada classe a causa del nombre reduït de caigudes de blocs, com el cas pràctic del Massís de Montserrat (Royán et al., 2014; Janeras et al., 2017). En el cas dels escarpaments de Degotalls, els esdeveniments de caiguda de blocs "Candidats" estan molt desequilibrats respecte la resta d'agrupacions, podent assolir proporcions de centenars de vegades més baixes que els etiquetats com a "Desconegut". En els diferents exercicis de monitoratge, el càlcul de diferències amb M3C2 i agrupament de punts es poden crear de l'ordre de 6.000 agrupaments, per identificar un nombre inferior a la desena de "Candidats".

Per corregir aquest desequilibri, tal com proposa Zoumpikas et al., (2021), s'han implementat estratègies de remostreig, ja sigui reduint les classes majoritàries (mètode de sub-mostreig) o augmentant sintèticament el nombre de classes minoritàries (mètode de sobre-mostreig). Els mètodes de mostreig implementats en la metodologia proposada es descriuen a la secció 2.5.2 i s'exposen a la Taula 3.19.

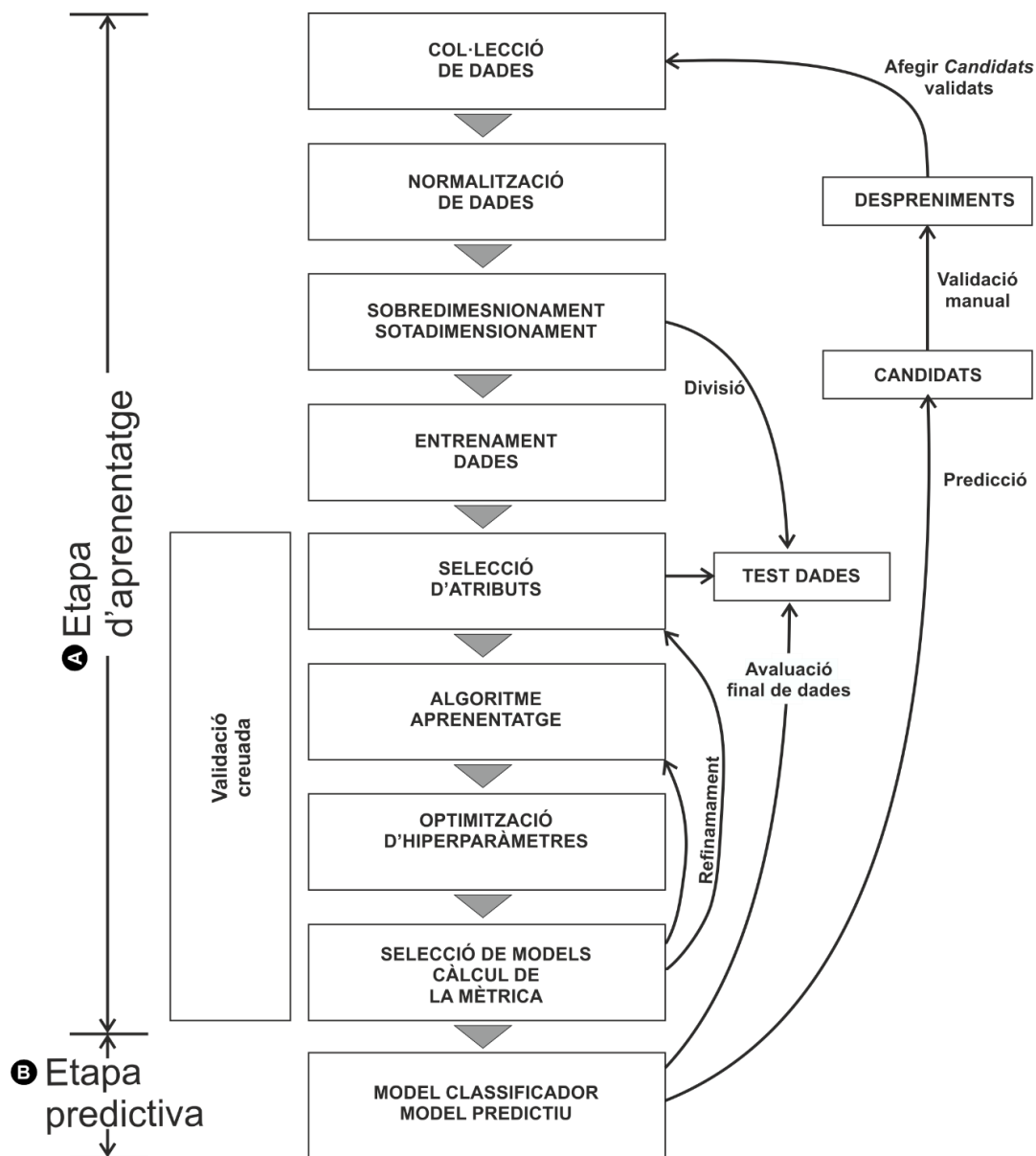


Figura 3.38: Diagrama de flux del model supervisat d'aprenentatge automàtic per classificar grups de blocs caiguts. La recollida de dades correspon a grups de punts creats anteriorment.

Subdimensionament	Sobredimensionament (Oversampling)	
Cluster Centroid	SMOTE	SMOTE-IPF
Cluster Representative	ADASYN	LEE
	SPIDER	SMOBD
	SWIM	G-SMOTE
	Polynom-Fit SMOTE	LVQ-SMOTE
	ProWsyn	Assembled-SMOTE
		TomekLinks- SMOTE

Taula 3.19: Recull de diferents tècniques de remostreig per corregir el desequilibri existent entre grups, agrupades segons si permeten un sotamostreig o sobremostreig.

Els models classificadors de *machine learning* utilitzats en el procés de refinament pertanyen a dues famílies; de base simple o múltiple descrits a la Secció 2.5.3. La Taula 3.20 mostra els mètodes contemplats en aquest estudi.

Base simple	Base múltiple
<i>Decision tree</i>	<i>Adaptative Boost</i>
<i>Gaussian Naive Bayes</i>	<i>Random Forest</i>
<i>K-Nearest Neighbor</i>	<i>Extra Tree</i>
<i>Linear Discriminant Analysis</i>	<i>Extreme Gradient Boosting (XGBoost)</i>
<i>Quadratic Discriminant Analysis</i>	
<i>Multi-Layer Perceptron</i>	
<i>Support Vector</i>	

Taula 3.20: Diferents algorismes desenvolupats per obtenir una classificació final, tenint en compte si son de base simple o de base múltiple (veure explicació al text).

En el processament dels models classificadors s'ha utilitzat un valor de 10 per els subconjunts de validació creuada per avaluar l'efectivitat del paràmetres que controlen els model. S'ha de tenir en compte que dels 10 subconjunts utilitzats en la validació creuada, l'entrenament del sistema utilitza solament 9 i la resta és per realitzar l'avaluació del resultat. El disseny de l'arquitectura de cada model és inferit per l'afinació dels models d'híper-parametrització, optimitzant la puntuació de la precisió (*Recall*). La configuració està mostrada a la Taula 3.21.

Models classificadors	Híper-Paràmetres
Linear Discriminant Analysis	Solver: svd, lsqr, eigen
Quadratic Discriminant Analysis	Reg param: 0.1, 0.3, 0.5
K-Nearest Neighbors	Number of neighbors: 1, 17
Gaussian Naive Bayes	Var smoothing: logspace (0, -9, num = 100)
Decision Tree	Criterion: gini, entropy; Maximum depth: 3–15
Support Vector	C: 0.1, 1, 10; Gamma: 1, 0.01; Kernel: rbf
Multi-Layer Perceptrons	Solver: lbfgs, SGD, ADAM; Activation: relu; Hidden layer sizes: 50, 100, 150
AdaBoost	Number of estimators: 1–50; Learning rate: 0.2
Random Forest	Number of estimators: 1–20
Extra Trees	Criterion: gini, entropy; Maximum depth: 3–15; Number of estimators: 1–20
XGBoost	Criterion: gini, entropy; Maximum depth: 3–15; Number of estimators: 1–20; Criterion: gini, entropy; Maximum depth: 3–15; Nthread: 4; Booster: gblinear, gbtrees; Missing: -999; Learning rate: 0.1, 0.2, 0.3; Number of estimators: 50, 100, 500; Seed: 1337; Disable default metric: True

Taula 3.21: La taula mostra la configuració per una exhaustiva cerca dels millors paràmetres per cada model classificador amb la tècnica del híper-paràmetre. El millor paràmetre està implementat d'acord al criteri de la millor puntuació de la precisió.

Finalitzada l'etapa d'entrenament dels 11 models classificadors segons el criteri de l'òptima precisió en els resultats (percentatge de resultats que són identificats correctament), comença

l'etapa de predicció. En aquest punt, cada un dels 11 models classificadors, basant-se en les dades de l'etapa d'entrenament que incloent la combinació dels 15 mètodes de remostreig, està configurat per classificar noves agrupacions de punts en funció dels seus 35 atributs. Per aquesta nova etapa, les dades utilitzades pertanyen a nous monitoratges que estan sense classificar. El resultat de la combinació de models classificadors i re-mostrejadors a l'etapa de predicció és un llistat de models amb l'assignació d'una etiqueta de "Candidat" i "Desconegut". La bondat de la predicció es defineix per dos desafiaments: (a) aconseguir un baix nombre de falsos positius (FP), és a dir, candidats que realment no corresponguin a desprendiments o caigudes de blocs reals; i (b) evitar falsos negatius (FN), és a dir, agrupacions de punts que corresponguin a desprendiments reals, però etiquetats com a "Desconeguts". El primer desafiament és degut al temps necessari per la validació dels agrupaments de punts etiquetats com a "Candidats" i una vegada realitzada la tasca de validació es comprova que realment no ho són. Del segon desafiament, se'n deriva el fet de que caigudes reals són classificades com no caigudes sense poder-les validar, produint-se un error. Per tant, l'objectiu és que el nombre real de caigudes de blocs sigui igual als veritables positius (VP) classificat pels models. I els valors de FP i FN siguin 0 o el més proper possible al 0.

La metodologia es completa amb la validació manual dels agrupaments "Candidat" amb imatges d'alta resolució de l'aflorament. En cas positiu, la nova agrupació de punts s'afegeix a les dades d'entrenament, per incrementar la classe "candidats". La intenció és reduir el desequilibri amb la classe "Desconegut" i augmentar l'eficàcia de l'etapa d'entrenament del procés de *machine learning* en els següents monitoratges.

3.3.2. Aplicació en els escarpaments de Degotalls-N i E a la Muntanya de Montserrat

El massís de Montserrat es troba a uns 40 km al NW de Barcelona (Espanya) amb una extensió aproximadament de 7,5 km per 2,5 km amb alçades al voltant dels 1.000 m de pinacles i agulles que perfilen innumbrables escarpaments (veure Figura 3.39). El massís és un Parc Natural amb un santuari que atreu un gran nombre de turistes i pelegrins. Així, es requereixen grans infraestructures, algunes de les quals es troben a prop dels perills de caiguda de roques, com, al costat dels escarpaments.

Esdeveniments recents han ocorregut a Montserrat, destacant els successius desprendiments que van tenir lloc entre l'any 2007 i 2008, a l'àrea anomenada Degotalls, desprendiment de 1200m³, que van afectar a la carretera d'accés i al pàrquing del Monestir de Montserrat, obligant a tenir tancat l'accés al monestir, durant diverses setmanes, amb un gran perjudici econòmic. Aquests esdeveniments mostren la necessitat d'iniciar un pla de mitigació del risc. Des d'aleshores, l'Institut Cartogràfic i Geològic de Catalunya conjuntament amb l'I.R. Geomodels de la UB, ha estat estudiant la zona i recollint dades TLS amb l'objectiu d'analitzar el comportament dels desprendiments mitjançant l'obtenció de corbes de magnitud i freqüència i l'anàlisi dels seus precursors (Vilajosana et al., 2008; Abellán et al., 2014; Janeras et al., 2015; Royan et al., 2016).

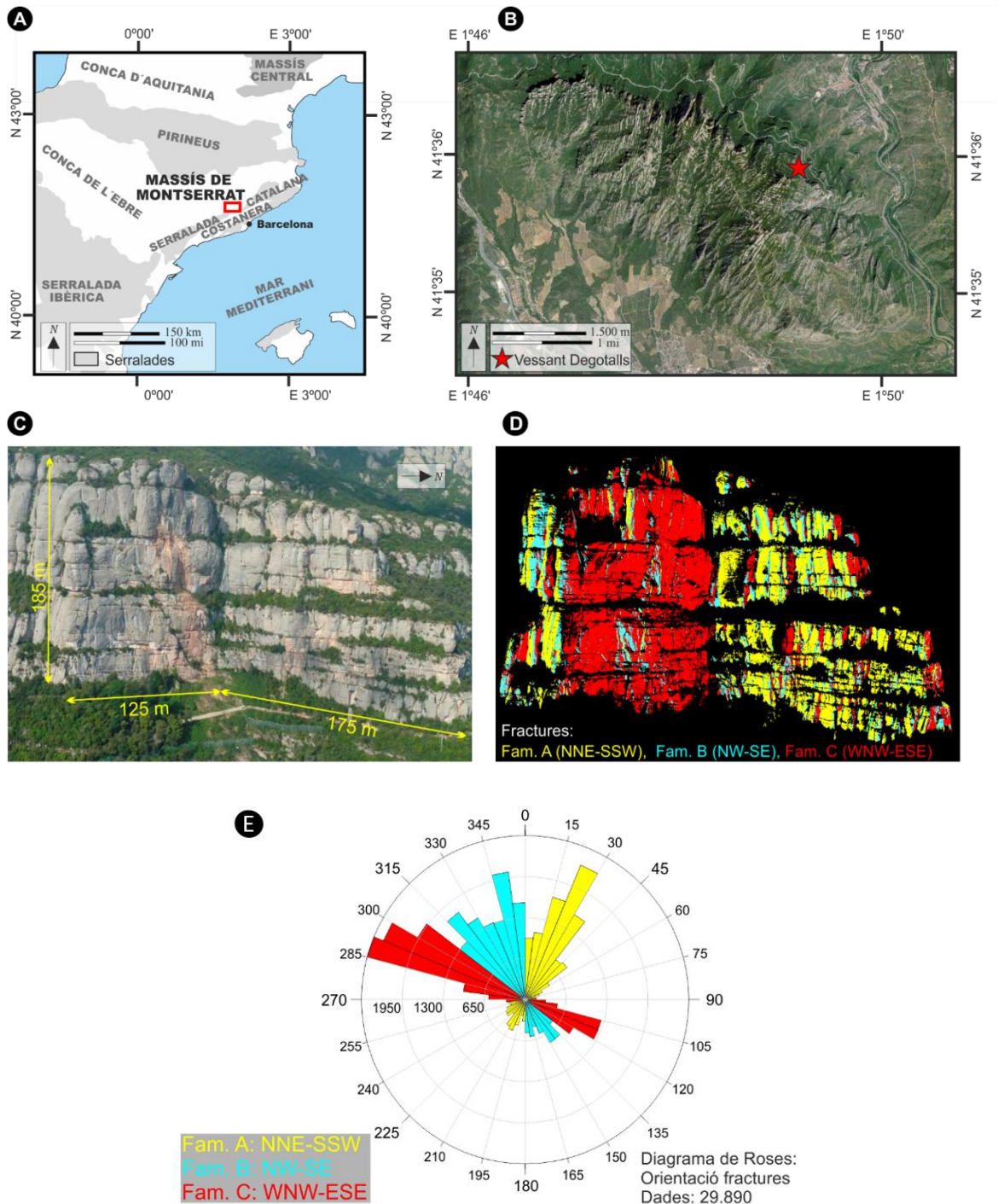


Figura 3.39: Diagrama de roses d'azimut respecte la línia de capa de les fractures modelades amb dades TLS (García-Sellés et al., 2018)

Geològicament, ens trobem en un ventall costaner (López-Blanco, 2007) on s'hi dipositaren centenars de metres de materials grollers, que posteriorment ens han donat als conglomerats montserratins del Bartonià-Priabonià (Gómez-Paccard et al., 2012). Una successió de conglomerats de més de 1000 m de gruix (conglomerats de Montserrat) entrelaçats per argiles

vermells i gresos (formació La Salut i Artés) amb llit subhoritzontal. El sistema deposicional correspon a una acumulació del complex ventall-delta a l'època de l'Eocè tardà al llarg del marge sud-est de la Conca d'Avantpaís de l'Ebre i adjacent a les Serralades Costaneres Catalanes (Anadón et al., 1985; López-Blanco et al., 2010; Gómez-Paccard et al., 2012).

La fracturació de Montserrat queda descrita a (Alsaker et al., 1996). El patró de fractura controla la morfologia del massís, amb dues orientacions preferents de fractures d'orientacions NNE-SSW (família A) y WNW-ESE (família C) *Figura 3.39*. Aquestes fractures, subverticals, penetratives, i amb una alta freqüència tallen el massís en blocs de mida decamètrica, que juntament amb l'acció de la meteorització caracteritzen el peculiar paisatge del massís. La traça superficial de les fractures pot ser seguida durant algun kilòmetre amb fotos aèries. La família de fractures B, amb una orientació NW-SE contribueix a la inestabilitat del vessant, encara que té una baixa freqüència, d'igual manera que la conjugada d'aquesta família de fractures, amb una orientació NE-SW i amb una presència residual a l'àrea de Degotalls. L'ordre alfabètic de les famílies de fractures defineix la cronologia de més antiga a més moderna (Alsaker et al., 1996). Les fractures representades a la *Figura 3.39E* han estat modelades amb dades TLS segons la metodologia Garcia-Selles et al. (2018).



Figura 3.40: Fotografia realitzada des del pàrquing de Montserrat, enfrontats amb la vessant de Degotalls. Al fons de la imatge es troba la vessant Est de Degotalls i al davant de la fotografia i a la dreta, veiem la recollida de dades mitjançant Terrestrial Laser Scanner (TLS). A partir d'aquestes dades recollides de forma freqüent i les dades històriques existents, hem pogut desenvolupar eines d'aprenentatge automàtic específiques amb la mateixa línia d'investigació que Zoumpikas, et al (2021). Fotografia pròpia realitzada el maig de 2019.

Els precedents de caigudes de blocs són nombrosos als voltants del monestir i del cingle de Degotalls. L'esdeveniment de caigudes de blocs durant el període 2001-2008 en la façana d'orientació N-S de Degotalls va totalitzar un volum de caigudes de blocs superior als 1000 m³ (Janeras et al., 2017). La interacció entre les fractures d'alta persistència vertical de la família de

fractures C i les superfícies d'estratificació, juntament amb altres famílies de fractures (A i C) va produir el desenganxament d'un gran bloc en fragments de dimensions mètriques i decamètriques. Aquest fet va motivar a les autoritats del Parc Natural del disseny d'un pla de mitigació de riscos juntament amb l'Institut Cartogràfic i Geològic de Catalunya on va col·laborar el Grup de Recerca RISKNAT de la Universitat de Barcelona. Inclòs en dit pla de mitigació, diferents escarpaments del massís són monitorats des de l'any 2007 amb TLS i amb la metodologia estàndard per la detecció de blocs caiguts.

Per la validació de la metodologia desenvolupada en aquesta tesi s'ha seleccionat el període de monitoratge 2007-2020 efectuat amb el model de TLS Iiris 3D de *Teledyne Optech Co.* en dos vessants ortogonals de Degotalls. Els vessants reben la denominació de Degotalls-N pel vessant d'orientació N-S i Degotalls-E pel vessant d'orientació E-W i es disposa dels resultats previs de l'estudi del seu monitoratge (Royan, 2015; Janeras et al, 2017). En aquests estudis previs es quantifiquen un nombre determinat de caigudes de blocs i els classifiquen en tres categories,

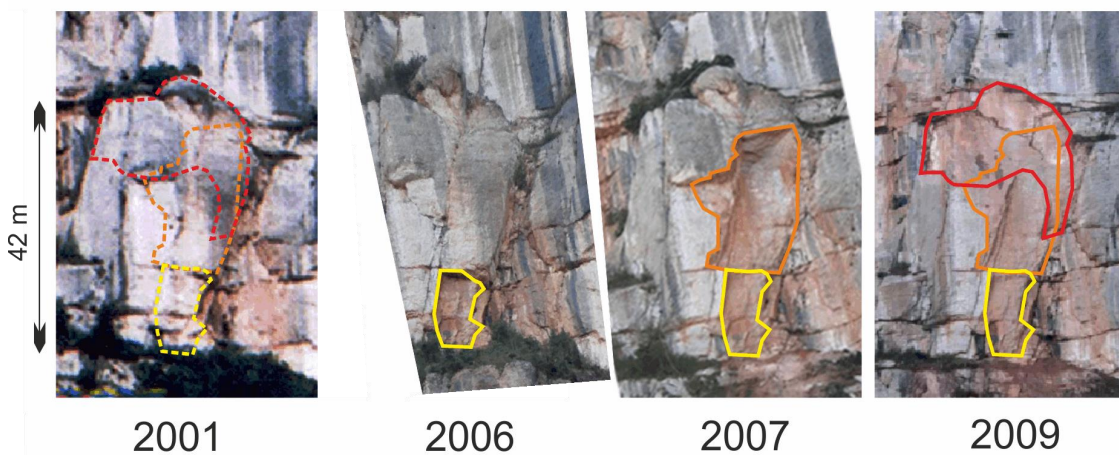


Figura 3.41: Exemple de caiguda de roques de gran volum al cingle de Degotalls-N com a conseqüència de fractures i discontinuïtats estratigràfiques. Imatge 2001: Zona degotalls abans de la caiguda de roques (en línies discontinües). L'any 2008, la superfície del penya-segat s'ha estabilitzat després de diferents episodis de desprendiments i va totalitzar un volum de caiguda de roques superior als 1000 m³.

Els estudis previs quantifiquen el nombre de blocs caiguts i els classifiquen en tres categories (Janeras et al., 2017) segons els seus processos generadors d'inestabilitat i els seus volums:

- a) **Grans blocs**, controlats per discontinuïtats mecàniques com poden ser fractures i superfícies d'estratificació produint desenganxaments d'una àmplia gamma de volums segons la geometria de les discontinuïtats, trobant-se entre el metre cúbic i alguns milers.
- b) **Còdols o agregats de còdols** despresos per la matriu meteoritzada o secundàriament influenciada per petites fractures locals, generalment de mig a petit volum, de 10 m³ a inferiors al 1 m³.

- c) **Plaques**, corresponents a escates de meteorització o exfoliació tèrmica (Martel , 2006; Guerin et al., 2020) en lloses amb volums normalment inferiors al metre cúbic.

3.3.3. Adquisició de dades

El monitoratge va consistir en l'adquisició de 12 núvols de punts amb TLS models ILRIS-3D de la companyia Optech Teledyne Co. (precisió $\sigma = 0.7$ cm als 100m) ((Optech, 2014)) durant els 13 anys d'observacions, com mostra la Taula 3.22. La longitud d'ona del làser utilitzat és de 1.535 nm, dins de la regió espectral de l'infraroig.

	Gener	Febrer	Març	Abril	Maig	Juny	Juliol	Agost	Setembre	Octubre	Novembre	Desembre
2007					N/E							
2008												
2009												N/E
2010												N/E
2011					N/E							
2012											N/E	
2013												N/E
2014							N/E					
2015												N/E
2016										N/E		
2017												N/E
2018												
2019											N/E	
2020		E								N/E		

Taula 3.22: Cronograma indicant en vermell, els mesos on es hi ha registre de dades dels afloraments de Degotalls-E i Degotalls-N, mitjançant *Terrestrial Laser Scanner* (TLS). Destaca en verd, les visites que s'han realitzat per recollir dades, únicament a l'aflorament de Degotalls-E, durant la tesi doctoral.

Amb aquest objectiu van ser necessaris dues estacions de TLS per capturar els dos vessants ortogonals (Degotalls-N i Degotalls S). L'aflorament de Degotalls-N va requerir d'una sola imatge d'escaneig (veure Figura 3.42) per cobrir tot el vessant mentre que Degotalls-E va requerir de dues imatges d'escaneig des de la segona estació. La denominació dels dos escanejors de Degotalls-E són primera secció (Nord) i segona secció (Sud) respectivament com mostra la Figura 3.42. El rang entre les estacions i la superfície dels vessants és d'uns 175 m, i les seves alçades per sota de 185 m. La densitat dels escanejors són d'un punt cada 7 cm aproximadament.

La pressa de dades durant els escanejors han estat complementades amb fotografies d'alta resolució dels afloraments per validar visualment les caigudes de blocs dins de la fase final del procés de classificació amb els models de *machine learning*.

La geometria dels 2 escarpaments amb unes dimensions aproximades de 400 metres d'alçada i 150 metres d'amplada, s'han estudiat des de dues estacions definides a la carretera que es troba a sota, obtenint dades de tota la superfície de l'escarpament.

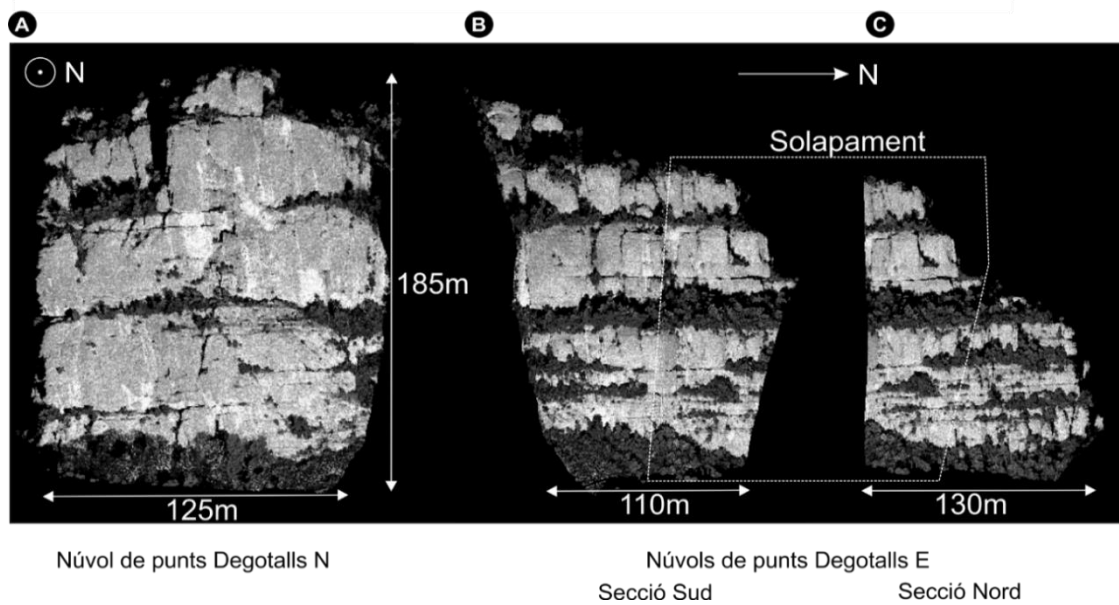


Figura 3.42: (A) Núvol de punts de la cinglera dels Degotalls N: 2.370.000 punts. Estació 1 (B) Núvol de punts del tram sud de la cinglera dels Degotalls E: 2.860.000 punts. Estació 2, orientació 1 (C) Núvol de punts de la secció nord de la cinglera dels Degotalls E: 2.060.000 punts. Estació 2, orientació 2 (reorientació de l'escàner). Els núvols de punts i les figures corresponen a la intensitat de textura del senyal retornat TLS (1.530 nm).

3.3.4. Processat de les dades i resultats

Les dades recollides en el camp amb TLS s'incorporen al flux de treball tal i com queda presentat a la figura 3.38. En una primera etapa (Figura 3.38A), es realitza un filtratge manual de la vegetació més exuberant, per aconseguir òptimes alineacions i alleugerar la quantitat de punts per reduir els temps de càlcul (veure figura 3.43).

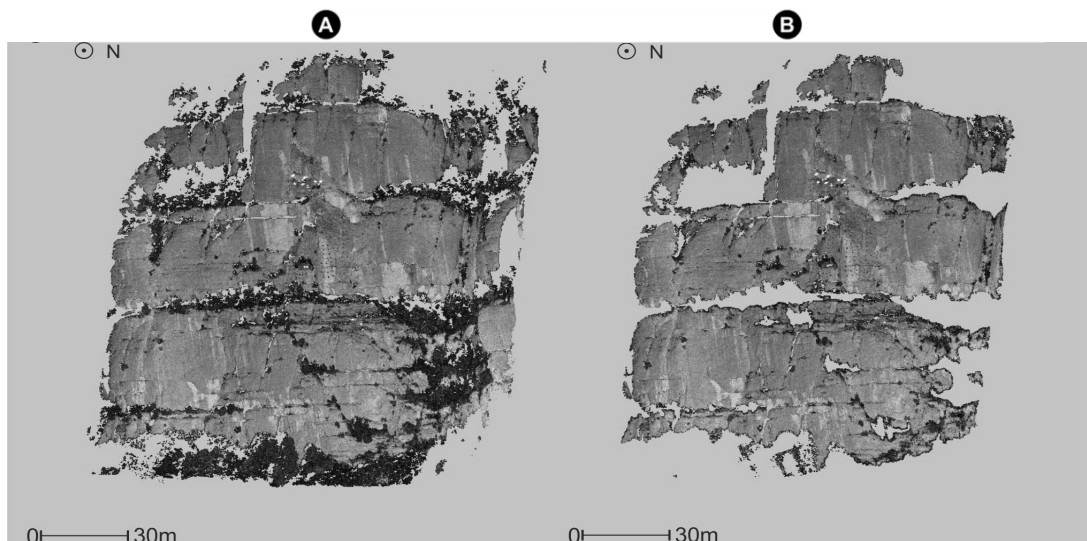


Figura 3.43: Exemple a Degotalls-N de filtratge previ del núvol de punts per eliminar la informació innecessària i aconseguir una òptima alineació. (A) Núvol de punts capturat amb TLS amb la intensitat de la senyal retornada. (B) Núvol de punts posterior a l'eliminació de la vegetació més exuberant.

(C2) Càlcul de diferències i agrupament de punts

Els valors dels paràmetres necessaris en els passos següents com càlcul de diferències i agrupament de punts són mostrats a la Taula 3.23 1) i 2). A la Taula 3.23 a l'apartat 3) es mostra el resultat del calibratge del LdD per l'àrea d'estudi de Degotalls.

El processament de dades de Degotalls continua amb l'alineació dels núvols de punts (Figura 3.38B) obtinguts durant els monitoratges respecte als núvols de punts de referència. En el següent pas i amb el programari PCM s'inicia el càlcul de les diferències (veure figura 3.38C2 i figura 3.45), amb els paràmetres mostrats a l'apartat 1) de la Taula 3.23. Aquests valors s'obtenen després de diferents exercicis per intentar ajustar-se a diferències conegudes de referència a l'aflorament.

L'agrupament de punts (Figura 3.38C2) és la continuació del procés amb el programari PCM amb els valors dels paràmetres del punt 2) de la Taula 3.23.

1) CÀLCUL DE DIFERÈNCIES	AFLORAMENTS DE DEGOTALLS (m)
Radi cerca punts veïns	0.2
Màxima vertical (MaxVd)	0.5
Mínim horitzontal (MinHd)	0.08
Màxima horitzontal (MaxHd)	0.10
2) AGRUPACIÓ DE PUNTS	
Distància llinar entre punts (eps)	0.15
Nombre mínim de punts (minPts)	10 punts
Diferència	Atribut diferència de cada punt
3) CÀLCUL LdD	
Mitja	-0.000268
Desviació estàndard	0.019547

Taula 3.23: Paràmetres per establir els processos de càlcul de diferències, agrupament de punts: 1) Paràmetres utilitzats per definir la geometria del doble con truncat en la mesura de les diferències; 2) Agrupació de paràmetres amb la nomenclatura equivalent a l'algorisme DBSCAN. La distribució 3) Mitja i desviació estàndard per definir la funció de distribució dels valors de diferència per TLS ILRIS-3D a l'àrea de Degotalls per a distàncies de 175 m.

Previ al procés d'agrupament de punts, es calcula el LdD (Figura 3.38C2) amb dos escanejos consecutius en una diferència de 30 min respecte el núvol de referència, i amb els mateixos valors dels paràmetres. La funció de distribució del valor de diferència entre sectors escollits són mostrats a la Taula 3.38 3). Els resultat de la diferència de punts es presenten a la Figura 3.44

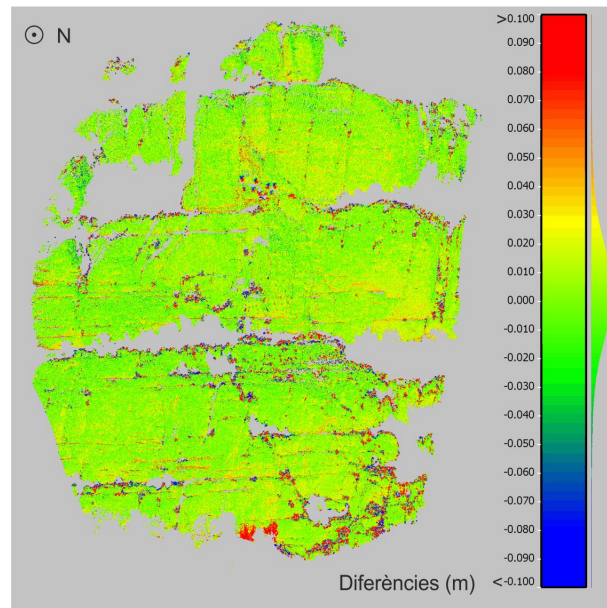


Figura 3.44: Mapa de les diferències calculades a mode de exemple a Degotalls-N amb el programari PCM. L'histograma mostra com la majoria dels valors estan compresos entre 0.035 i -0.035. Els valors més destacats corresponen a diferències registrades.

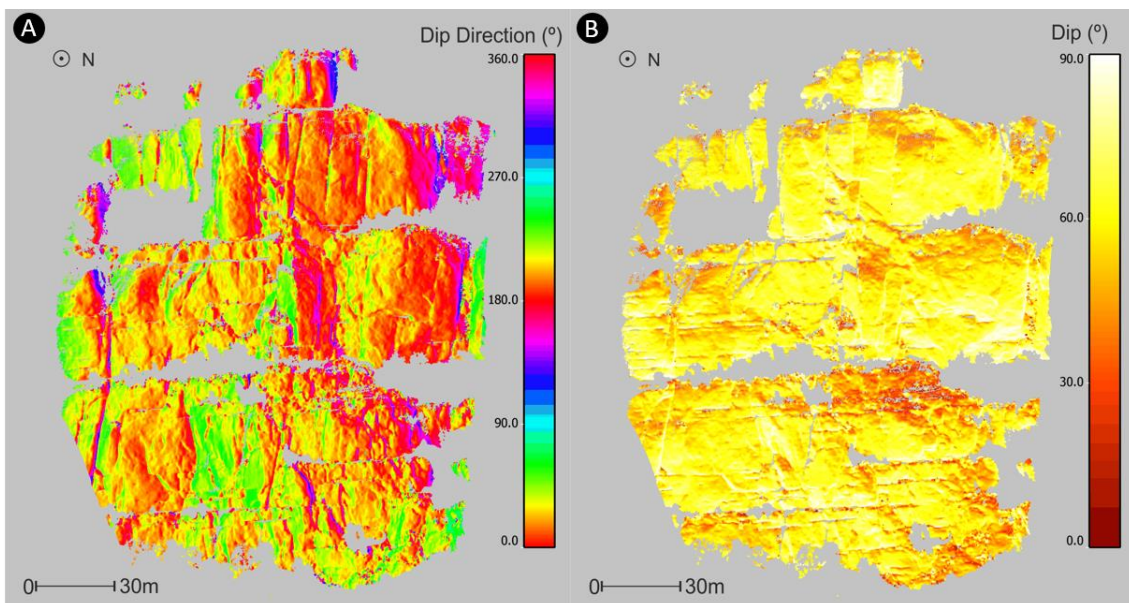


Figura 3.45: Mapa de les orientacions calculades amb el programari PCM com a atribut de cada punt. Les orientacions són mostrades en (A) direcció de màxim pendent o Dip Direction i; (B) Dip format.

Calculades les diferències (Figura 3.38C2), s'identifica el valor d'LdD avaluant que la mètrica de les diferències s'ajusta a una distribució gaussiana per estimar els paràmetres estadístics de mitja i desviació estàndard. Aquests resultats són comparats amb els valors obtinguts en el calibratge per obtenir els valors de LdD. La Taula 3.24 mostra els LdD mitjos en cada sector de Degotalls.

SECTORS		LdD:	MITGES (m)	STD (m)
Degotalls E:	Secció Sud	Superior	0.03242	0.00336
		Inferior	-0.03189	0.00453
	Secció Nord	Superior	0.03430	0.00590
		Inferior	-0.03511	0.00385
Degotalls N:		Superior	0.03928	0.01685
		Inferior	-0.04026	0.01626

Taula 3.24: Sumari estadístic dels LdD a Degotalls en el període estudiat 2007–2021. El valor de Degotalls E estan calculats amb els 12 monitoratges i a Degotalls N amb els dos monitoratges (2007–2017 i 2017–2021) que s'han realitzat per cobrir el període.

Finalitzat el càlcul de les diferències i establerts els valors del LdD (resumits a la taula 3.24), el següent pas és agrupar els punts amb els paràmetres de la taula 3.23 2) i amb el programari PCM. El format dels resultats està a la taula 3.25.

ID.	ATRIBUT	DESCRIPCIÓ
1	Identification	
2	Coordinate X*	
3	Coordinate Y*	Coordinates of cluster centroid
4	Coordinate Z*	
5	Item number *	Cluster differences number
6	Points number	Number of points
7	TotalVolume	Cluster total volume
8	PositiveVolume	Volume behind the reference surface and TLS
9	NegativeVolume	Volume in front the reference surface and TLS
10	Area	Planimetric cluster area 2D, perpendicular to TLS
11	Code *	Cluster classification (Unknown, Candidate)
12	Confidence *	Confidence index
13	Predominance_Mean	Mean predominance (noise 0, advance 1, retreat 2)
14	Predominance_Sigma	STD predominance classification (0, 1, 2)
15	Percentage_1_Mean	Mean advance predominance (1)
16	Percentage_1_Sigma	STD advance predominance (1) classification
17	Percentage_0_Mean	Mean noise predominance (0)
18	Percentage_0_Sigma	STD noise predominance (0) classification
19	Percentage_2_Mean	Mean retreat predominance (2)
20	Percentage_2_Sigma	STD retreat predominance (2) classification
21	OrientationSetsRef	Reference cluster strike azimuth
22	OrientationSetsCom	Compared cluster strike azimuth
23	IndexTextureRef *	Reference texture index (0, 1 Int, 2 RGB, 3 RGB + Int)
24	R_Mean_Ref ¹	
25	R_Sigma_Ref ¹	
26	G_mean_Ref ¹	
27	G_Sigma_Ref ¹	
28	B_mean_Ref ¹	Texture. Mean & Std of reference clusters.
29	B_Sigma_Ref ¹	
30	I_Mean_Ref ²	
31	I_Sigma_Ref ²	
32	IndexTextureCom *	Compared texture index (0, 1 Int, 2 RGB, 3 RGB + Int)
33	R_mean_Com ¹	
34	R_Sigma_Com ¹	
35	G_mean_Com ¹	
36	G_Sigma_Com ¹	
37	B_mean_Com ¹	Texture. Mean & Std dev. of compared clusters
38	B_Sigma_Com ¹	
39	I_Mean_Com ²	
40	I_Sigma_Com ²	
41	AziRef_Mean	Mean strike azimuth of Reference points
42	SloRef_Mean	Mean strike slope of Reference points
43	AziCom_Mean	Mean strike azimuth of Compared points
44	SloCom_Mean	Mean strike slope of Compared points
45	CopRef_Mean	Mean coplanarity of Reference points
46	CopRef_Sigma	STD coplanarity of Reference points
47	ColRef_Mean	Mean collinearity of Reference points
48	ColRef_Sigma	STD collinearity of Reference points

49	CopCom_Mean	Mean coplanarity of Compared points
50	CopCom_Sigma	STD coplanarity of Compared points
51	ColCom_Mean	Mean collinearity of Compared points
52	ColCom_Sigma	STD collinearity of Compared points
53	Ang_Mean	Mean angularity between normal vectors
54	Ang_Sigma	STD angularity between normal vectors
55	Reference File *	String
56	Compared File *	String

Taula 3.25: Funcions associades als agrupaments de punts, ordre de format. * No s'utilitza en la classificació *ML*. ¹ Només s'utilitza un ordinador amb format *RGB*. ² Només s'utilitza el format *Intensity* calculat.

El nombre d'agrupacions de punts establertes en cada monitoratge són d'aproximadament 5.800 al Sector Sud de Degotalls-E, 2.600 al Sector Nord de Degotalls-E i de 3.700 a Degotalls-N. Els diferents atributs de cada punt són els que contribueixen a caracteritzar estadísticament l'agrupació de punts. La Figura 3.46 mostra les agrupacions amb els atributs de cada punt.

(D2) Classificació de grups de punts

La nova etapa de classificació d'agrupaments amb el programari *Cluster Classification* (Figura 3.38D2) s'inicia amb la selecció i validació manual de 10 candidats necessaris per la fase d'entrenament del primer monitoratge de Degotalls-E, Secció Sud. Entre les primeres 2.000 agrupacions de punts d'un total de 5.957 del monitoratge 2007–2009 es troben els 10 candidats per completar la fase d'entrenament. Completada la fase d'entrenament es continua amb la fase de predicció, entre les 3.957 agrupacions restants.

La combinació dels 11 models classificadors de *machine learning*, amb els 15 mètodes de remostreig totalitzen un valor de 165 configuracions amb 10 subconjunts de validació creuada.

Els resultats van ser validats manualment per identificar les corresponents a caigudes reals.

D'aquest període, són identificats 8 nous casos de caigudes de blocs i s'identifica el millor mètode classificador (*Quadratic Discriminant Analysis*) i millor mètode de remostreig (*Polynom Fit-SMOTE*), segons la seva mètrica (VP+FP). Seguint amb el procediment, aquests 8 nous casos són afegits a la base de dades de candidats per entrenar el següent monitoratge (2009–2010) i classificar en la fase de predicció les 5.100 agrupacions de punts. El procediment es va repetir fins a completar tota la sèrie a Degotalls S, Secció Sud.

A Degotalls-E, Secció Nord, el procés de classificació s'inicia aprofitant els candidats validats a la Secció Sud, però en aquest cas s'ha optat per fer una única comparació entre l'adquisició TLS de l'any 2007 amb la darrera adquisició TLS de l'any 2019. La motivació d'aquest allargament del període de comparació és intentar detectar deformacions suficientment grans per poder ser detectades, ja que amb períodes més reduïts com els de la Secció Sud, la recerca ha estat negativa. En aquest cas, també la recerca d'aquestes deformacions que es podrien donar abans de la caiguda de bloc, van donar un resultat negatiu. Per datar els anys de les caigudes de blocs s'utilitza la seva identificació amb les imatge d'alta resolució.

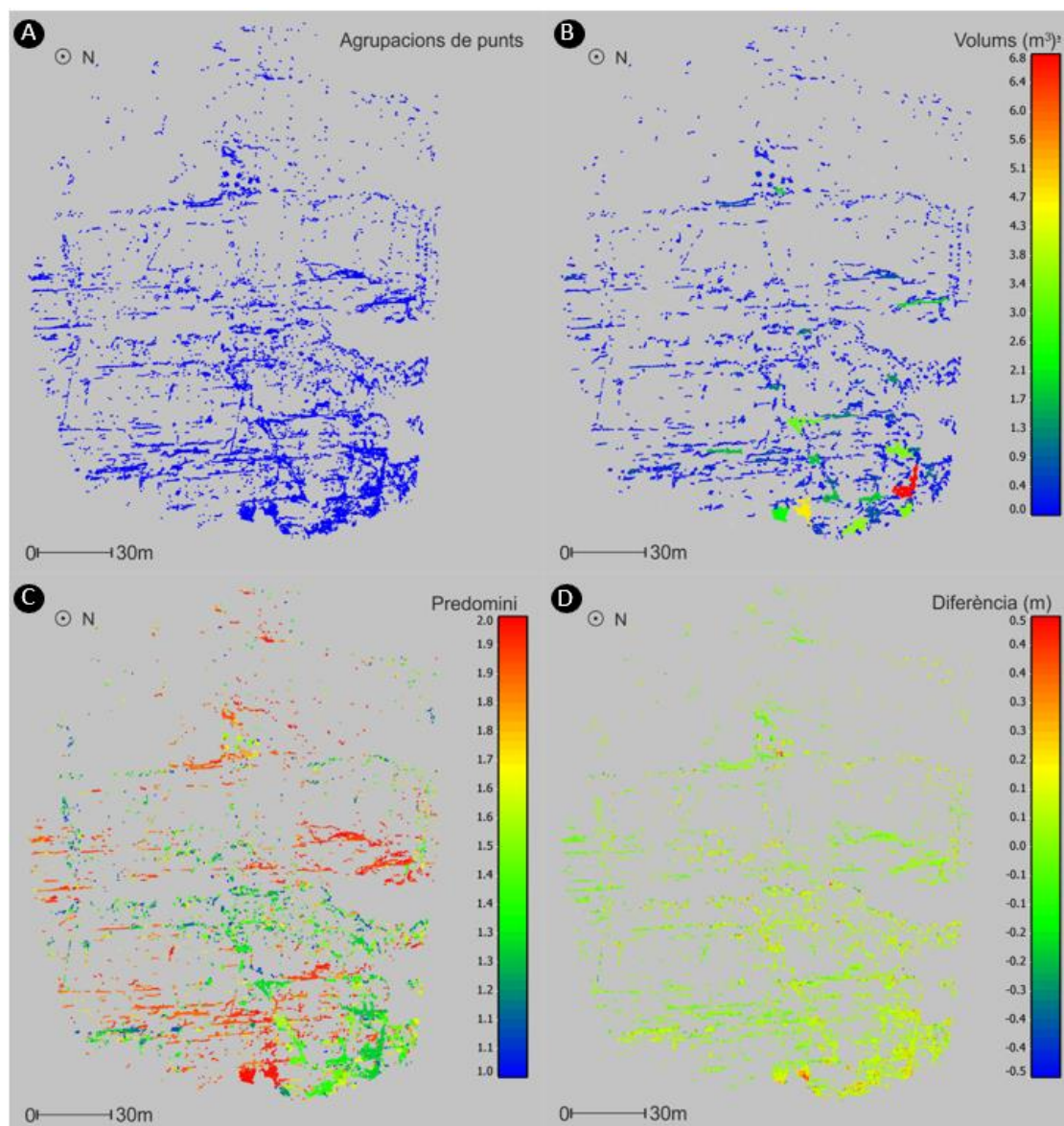


Figura 3.46: Resultats de les “agrupacions de punts” de la comparació 2007-2017 a Degotalls N. (A) Punts que conformen les 3.712 agrupacions de punts. (B) Distribució de l’atribut “Volum” calculat pel programari PCM de les diferents agrupacions de punts en funció de les diferències calculades entre els diferents núvols de punts. (C) “Predomini” i (D) “Diferències”

Degut a la diferència d’orientació de l’escarpament de Degotalls-N, s’ha optat per desestimar els candidats validats obtinguts a Degotalls-E i començar una nova col·lecció de candidats validats, específica per Degotalls-N.

El motiu d’aquesta decisió ha estat que els atributs relatius a l’orientació de les superfícies de les agrupacions candidates validades podrien condicionar la fase d’entrenament de la classificació. Igualment, com a Degotalls-E, Secció Nord, s’ha reduït el nombre de comparacions per allargar el temps entre monitoratges amb l’objectiu de detectar deformacions. La comparació es realitza solament entre els escanejos dels anys 2007–2017 i 2017–2020. La classificació del primer període s’inicia amb 10 candidats validats manualment per completar la fase d’entrenament o

aprenentatge. En la fase de predicció, els resultats reporten que la mètrica de millor model de classificació, amb un total de 1.211 de Falsos Negatius (FN), valor massa elevat per un total de 3.700 agrupacions de punts.

Com alternativa per ajustar els resultats, s'ha condicionat els resultats a un nombre molt reduït de Falsos Negatius (FN), per identificar el millor model classificador i mètode de remostreig. Els candidats validats s'han incorporat en el següent procés d'aprenentatge. En aquest tercer cas, també la recerca de deformacions ha estat negativa, sense identificar-se agrupaments de punts associats a deformacions.

Els resultats tant a Degotalls-E com a Degotalls-N han estat negatius pel que fa a la identificació d'agrupaments de punts representatius de deformacions prèvies a caigudes de blocs. Al no poder-se validar cap agrupació de punts, no s'ha pogut entrenar al sistema per la classificació amb *machine learning*.

Els resultats obtinguts de la classificació amb *machine learning* mostren que els millors models classificadors i el millor model de remostreig en una comparació, no té perquè ser la millor en la següent comparació. Per altra banda, també s'observa a Degotalls-E, Secció Sud que el nombre de FP té una tendència a la reducció (Taula 3.26).

AFLORAMENT	CAIGUDES DE BLOCS PER L'ENTRENAMENT	MILLOR MODEL CLASSIFICADOR	MILLOR MÈTODE DE REMOSTREIG	CAIGUDES REALS	VP	FP	FN
Degotalls-E							
Primera secció							
2007-2009	10*	Quadratic Discr.	Pol. Fit-SMOTE	8	8	91	0
2009-2010	18	Linear Discr. A.	Cluster Centr.	5	5	7	0
2010-2011	23	KNN C.	Cluster Centr.	4	4	139	0
2011-2012	27	XGBoost C.	S. TomekLinks	4	4	48	0
2012-2013	31	Extra Trees C.	Cluster Centr.	2	2	10	0
2013-2014	33	XGBoost C.	Cluster Centr.	1	1	1	0
2014-2015	34	SVC	Cluster Centr.	1	1	0	0
2015-2016	35	Linear Discr. A.	Stefanowsky	3	3	53	0
2016-2017	38	-	-	0			
2017-2019	38	Linear Discr. A.	Cluster Centr.	3	3	9	0
2019-2020	41	-	-	0			
2020-2020	41	Extra Trees C.	Stefanowsky	2	2	2	0
Segona secció							
2007-2019	43	Quadratic Discr.	LVQ-SMOTE	22	22	97	0
Degotalls-North							
2007-2017 ^a	10*	Linear Discr. A.	Cluster Repres.	107	107	1211	0
2007-2017 ^b	10	Decision Tree C.	Cluster Repres.	107	104	296	3
2017-2020 ^a	117	Quadratic Discr.	SWIM	16	16	455	0
2017-2020 ^b	117	Quadratic Discr.	Pro WSyn	16	15	256	1

Taula 3.26: Resum dels resultats de l'etapa de predicció a Degotalls durant la classificació dels agrupaments de punts. A la taula s'indica el número d'agrupaments de punts utilitzats com a candidats validats en la fase d'entrenament. Els millors model de classificació, així com el de remostreig i les mètriques d'aquests. Degotalls N presenta dos solucions: a) La millor solució amb la premissa de FN = 0; i b) la premissa de FN ≠ 0, acceptant un número reduït de FN. * Candidats identificats manualment, al ser inici de procés.

La mètrica dels millors models de classificació són considerats com resultats molt favorables, especialment a Degotalls-E. També cal remarcar que el nombre inicial dels agrupaments de punts en cada monitoratge és d'entre 5.000 i 6.000 i el nombre de candidats per validar dels models de classificació Fals Positiu (FP) més Verdader Positiu (VP) té una significativa reducció. Malgrat aquest fet, els resultats a Degotalls-N presenten un elevat nombre de Falsos Positius (FP), especialment quan la condició imposada per la total identificació de Verdaders Positius (VP) és imposada.

Aquests resultats van encaminats a buscar la predicció amb la millor combinació de model classificador i remostrejador.

A la Taula 3.2 es mostra l'exemple complert del procés de classificació si utilitzem el primer millor model classificador (*Quadratic Discriminant Analysis*) juntament amb el primer millor mètode de remostreig (*Polynom Fit-SMOTE*), obtingut en el període 2007–2009. Els resultats obtinguts per la resta de períodes, no serien acceptables degut al gran nombre de Falsos Positius (FP) i per tant, el gran nombre d'agrupacions a validar.

SECTOR I PERÍODE	CAIGUDES REALS	VP	FP	FN
Degotalls E				
Secció Sud				
2007–2009	8	8	91	0
2009–2010	5	5	148	0
2010–2011	4	4	461	0
2011–2012	4	4	258	0
2012–2013	2	2	235	0
2013–2014	1	1	224	0
2014–2015	1	1	188	0
2015–2016	3	3	315	0
2016–2017	0	-	-	-
2017–2019	3	3	517	1
2019–2020	0	-	-	-
2020–2020	2	2	111	0

Taula 3.27: Resum dels resultats de l'etapa de predicció a Degotalls durant la classificació dels agrupaments de punts. A la taula s'indica el número d'agrupaments de punts utilitzats com a candidats validats en la fase d'entrenament. Els millors model de classificació, així com el de remostreig i les mètriques d'aquests. Degotalls N presenta dos solucions: a) La millor solució amb la premissa de FN = 0; i b) la premissa de FN ≠ 0, acceptant un número reduït de FN. * Candidats identificats manualment, al ser inici de procés.

Per tant, la dificultat resideix em com identificar d'entre tots els models classificadors i de remostreig correctes la millor combinació de models per cada període estudiat, ja que el percentatge de models que prediuen correctament agrupacions de punts de caigudes de blocs reals és significativament elevat. També han de complir la condició de tenir el menor nombre possible de Falsos Positius (FP), per facilitar la validació. D'aquesta manera, serà complex, identificar el millor model de predicció quan el nombre i posició de les caigudes de blocs reals no és conegut. Per la validació de la metodologia, aquestes dades eren conegudes, però no tenim un coneixement previ per futures prediccions.

Per resoldre aquesta problemàtica, el programari *Cluster Classification* calcula el percentatge de vegades que cada agrupació de punts es etiquetada com a “Candidats” per qualsevol model de predicció. Per tant, si intentem validar només les agrupacions de punts més proposades pels models predictius, és possible millorar la mètrica de la classificació.

Amb aquesta premissa, les caigudes de blocs reals a Degotalls-E sempre es troben entre les primeres 115 agrupacions més proposades pels models (entre VP i FP). El total inicial analitzat és de 5.800 agrupacions de punts per un monitoratge a Degotalls-E. Això significa una reducció del 98% d'agrupacions per validar (VP + FP). A Degotalls-N, la reducció seria del 80.16% per a una identificació completa, i per identificar el 96% de les caigudes de blocs reals, la reducció seria del 90%, partint d'una població de 3.700 agrupacions de punts.

La validació de les agrupacions etiquetades com a “Candidats” és realitzada amb imatges d'alta resolució. A la Figura 3.47 es mostra un exemple de validació, comparant les imatges d'una agrupació de punts abans i després del monitoratge. En aquest sentit, a la Taula 2.28 es mostren els atributs de l'agrupació de punts.

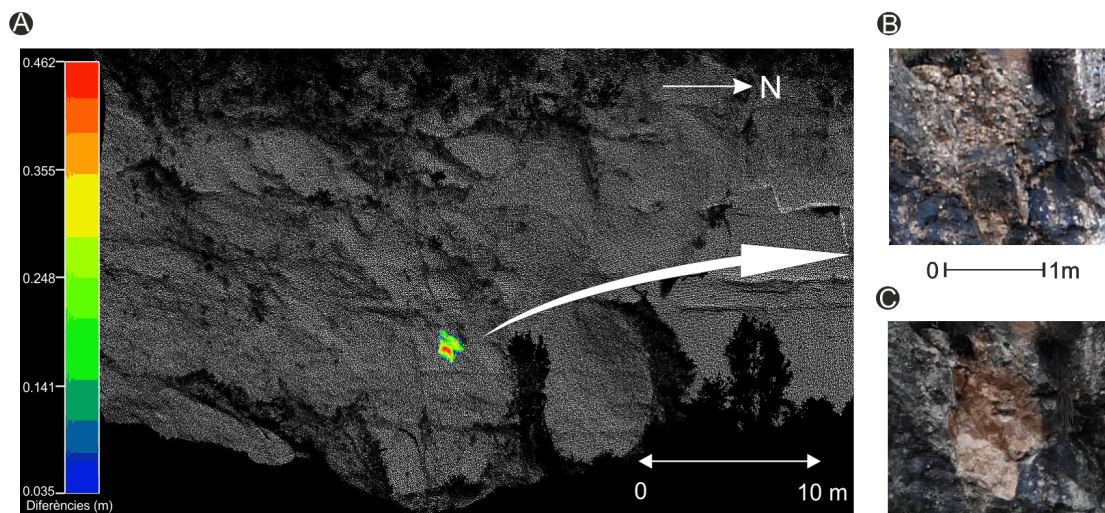


Figura 3.47: A) La imatge correspon al núvol de punts amb textura d'Intensitat a Degotall E amb l'atribut diferències del clúster #1326 (Monitoratge 2017-2019), les dimensions d'aquest exemple són: 1,4 m (alçada) i 1 m (amplada). B) Imatge del clúster abans del monitoratge. C) Imatge post caiguda de bloc on és visible la falca i la superfície de les fractures que controlen el desprendiment. Aquest desprendiment de blocs està sub-classificat com a gran bloc.

ATRIBUT DE L'AGRUPAMENT	VALOR	ATRIBUT DE L'AGRUPAMENT	VALOR
Número de l'agrupament de punts	1326	Pendent Núvol Comparat	70,91°
Centroide Coord. X	-23.169 m	Desviació estàndard	0.07
Coord. Y	192.847 m	Punts % Soroll	10.48%
Coord. Z	-10.063 m	Avanç	0.21%
Nombre de punts	428	Retrocés	89.31%
Volum positiu	0.27164 m ³	Intensitat Núvol Referència	210.32
Àrea	1.45 m ²	Intensitat Núvol Comparat	228.43
Predomini: Mitjana	2 (Retrocés)	Azimut Núvol Referència	165,10°
Pendent Núvol Referència	71,11°	Azimut Núvol Comparat	166,20°

Taula 3.28: Resum del atributs de l'agrupament de punts utilitzat per caracteritzar l'exemple de la Figura 3.51.

El nombre total d'agrupacions de blocs classificats com a "Candidats" i validats com a esdeveniment de caigudes de blocs durant el període de monitoratge a Degotall-E és de 65 i 95 a Degotall-N (Figura 3.48). El nombre més important d'esdeveniments a Degotall-N correspon al període 2007-2009, i estan associats al gran esdeveniment de l'any 2001-2007 (Figura 3.41) i a les activitats de mitigació posteriors. Dels 160 agrupaments de punts identificats, 43 a Degotalls-E i 2 a Degotalls-N es corresponen a 2 grans esdeveniments associats a les activitats aplicades per mitigar el risc de caigudes. En una segona etapa, els dos escarpaments entren en un període d'estabilitat. El nombre i mida reduït de les caigudes s'associa en part, a les actuacions de contenció incorporades als escarpaments com el bulons, malles de triple torsió o ancoratges. S'observa una lenta reactivació de l'activitat a Degotalls-N, a partir del 2018, encara que no significativa en volum.

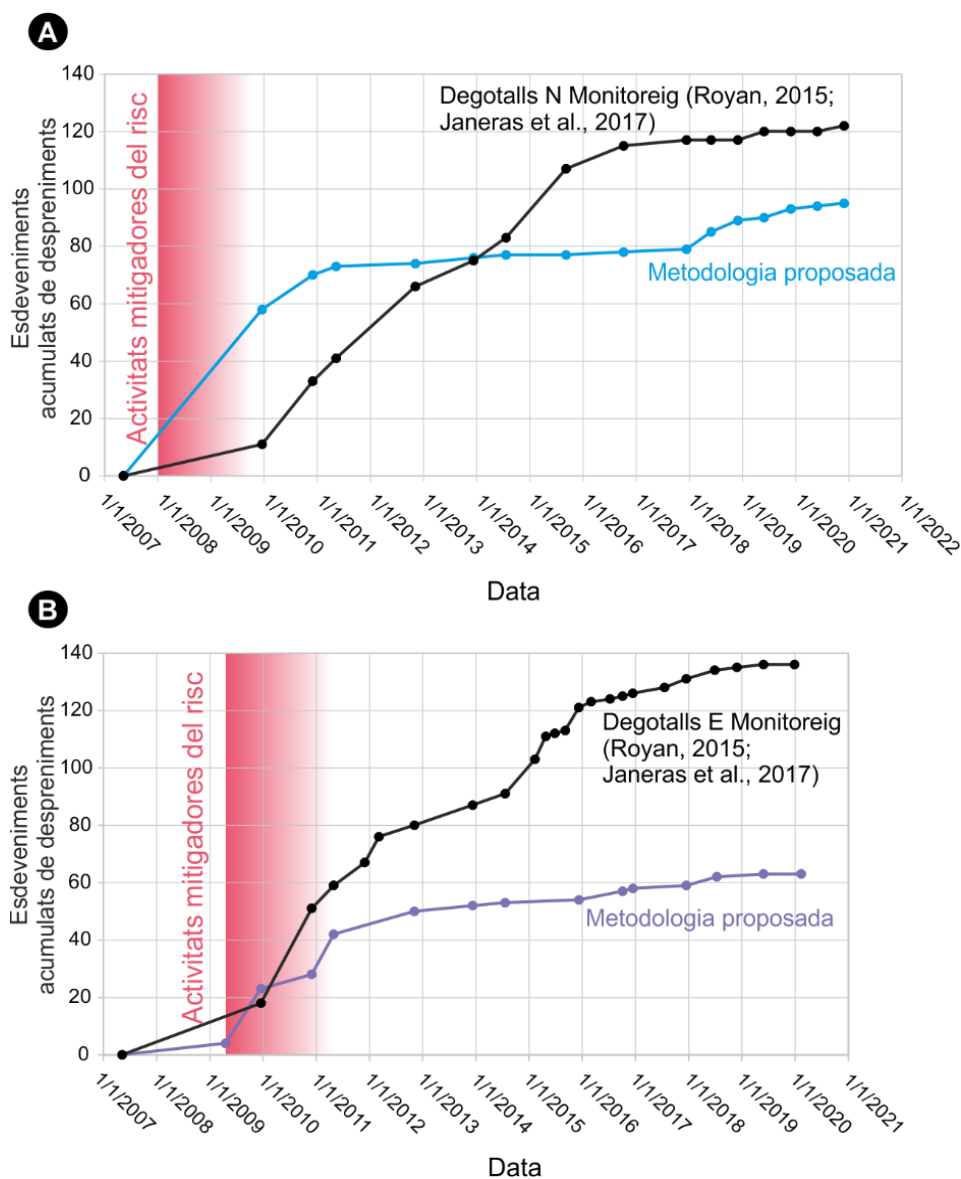


Figura 3.48: Esdeveniments de caiguda de roques a Degotalls-N (a) i Est (b) color porpra). La línia marró representa el nombre d'esdeveniments registrats amb la metodologia actual en comparació amb la metodologia proposada (color blau i morat).

3.3.5. Discussions

Es discuteixen i s'avaluen els principals resultats de la proposta metodològica per a l'automatització de la classificació de blocs caiguts.

En aquest estudi, es construeix un inventari de caigudes de roques a partir de núvols de punts amb dos nous mètodes (PCM i *Cluster Classification*). El programari PCM s'implementa per caracteritzar i identificar clústers de diferències durant els processos de monitoratge, i el programari *Cluster Classification* classifica la naturalesa dels agrupaments de punts. Concretament, està entrenat per classificar agrupacions de punts associats a despreniments amb tècniques de *machine learning*. L'inventari s'ha realitzat per al període 2007-2020 a la zona d'escarpaments de Degotalls (massís de Montserrat, Espanya). L'escarpament, està dividit en dos; Degotalls-N i Degotalls-E, en funció de la seva orientació, d'aquesta manera, es construeix un inventari de caigudes de roques a partir de núvols de punts amb dos nous mètodes (PCM i *Cluster Classification*), per cadascun dels escarpaments. Degotalls-E i Degotalls-N, presenten un inventari de 65 i 95 despreniments, respectivament, i aquests resultats s'ajusten als valors esperats i coneguts dels estudis anteriors a la zona (Royan, 2015; Janeras et al., 2017).

El programari PCM calcula el LdD (veure Taula 3.24), i l'ordre de magnitud dels resultats és similar als utilitzats anteriorment en el seguiment dels Degotalls i recomanats en treballs anteriors (Royan, 2015; Janeras et al., 2017). En general, els valors de la LdD (estimació de l'error de l'aparell) a Degotalls-N són un 20% més alts que els obtinguts a Degotalls-E, però això és atribuïble a la major alçada de l'escarpament de Degotalls-N i, per tant, amb una distància TLS vs. escarpament una mica més gran. Les diferències de mesura i els punts d'agrupació són processos integrats al programari PCM i ofereixen resultats similars als estudis anteriors realitzats a l'àrea amb els algorismes M3C2 i DBSCAN per Royan, (2015) i Janeras et al., (2017).

El programari PCM complementa els resultats amb atributs que caracteritzen primer els punts i, posteriorment, els agrupaments (per exemple, el predomini, la coplanarietat o el vector normal) per alimentar el procés de classificació de *machine learning*.

Els resultats de la classificació d'agrupaments de punts proporcionen diferències en el nombre de despreniments comptats respecte a treballs anteriors, però aquest fet és atribuïble als diferents mètodes de validació. En aquest treball, només s'han acceptat aquells clústers etiquetats com a "Candidats" validats en imatges d'alta resolució. Els dispars resultats dels models predictius per a ambdós escenaris s'han interpretat a Degotalls-E com a prometedors, la qual cosa obre la possibilitat d'estudiar amb més detall la importància de cada atribut en l'aportació dels models. A més, això es pot estendre a l'estudi de l'eficiència de cada tècnica de remostreig i cada model de classificació per fer evolucionar la identificació de clústers de deformació corresponent als episodis previs a la caiguda del bloc, on comença el desenganxament de la roca (veure l'apartat de Treballs futurs).

A Degotalls-E, la validació manual dels primers 16 agrupaments de punts més proposats pels 165 models predictius aconseguen identificar el 65% dels desprendiments coneguts. Si la validació manual s'estén als 115 primers agrupaments de punts més proposats pels models, la identificació arriba al 100% dels desprendiments. Això implica una reducció important dels agrupaments inicials per validar-los com a desprendiments. L'eficiència dels models predictius també tendeix a augmentar quan la base de dades de caiguda de roques augmenta en nombre d'exemples.

A Degotalls-N, la validació manual dels agrupaments de punts presenta uns resultats més desfavorables, requerint la validació de les 370 primeres agrupacions per identificar el 90% de les caigudes de blocs. Però es aquí on es mostra el potencial de millora del mètode per reduir el número de candidats a validar i augmentar el percentatge d'èxit.

L'anàlisi dels agrupaments validats revela una relació entre l'atribut volum i el percentatge de models predictius que els proposen. El percentatge de models predictius que proposen blocs es baix amb els agrupaments de gran volum (veure Figura 3.49A), mentre que Degotalls-E (veure Figura 3.49B) no presenta volums tan grans.

Degotalls-N presenta els agrupaments etiquetats com a "Candidats" amb menor percentatge de predicció pels models pertanyents principalment a la categoria de **grans blocs** amb volums superiors a 0,1 m³. D'altra banda, els agrupaments corresponents a la categoria **plaques** amb volums més petits obtenen els percentatges més alts d'identificacions reals de verdader Positiu (VP) en la predicció. Això és degut a que presenten atributs més homogènies i hi ha més exemples, per tant, faciliten l'entrenament. Els agrupaments de punts classificats com a desprendiments de roques a Degotalls-N es poden considerar heterogenis (en termes d'atributs: per exemple, volum, orientació, intensitat) i, per tant, són més difícils d'aprendre en l'etapa d'entrenament, ja que els atributs dels agrupaments no estan tan polaritzades com en Degotalls-E. Els dos agrupaments de major volum a Degotalls-N tenen el percentatge més baix de prediccions del model perquè tenen un alt grau de singularitat, la qual cosa no contribueix a definir una classe homogènia per a l'aprenentatge del model. Un augment d'exemples de caigudes de roques de categoria **grans blocs** en la recollida de dades, pot corregir aquest problema.

La categoria de **grans blocs** (>1 m³) va ser predominant a Degotalls-N durant els dos primers anys de seguiment, i es va associar amb el gran esdeveniment controlat per fractures de l'episodi 2001-2007 i les seves posteriors activitats de mitigació de riscos, com es pot veure a la Figura 3.48. En la segona etapa, ambdós escarpaments entren en un període d'estabilitat en el qual predomina un nombre reduït i constant de desprendiments pertanyents a la "Categoria" **plaques**, vinculats al persistent procés de meteorització. A Degotalls-E s'han observat alguns desprendiments de petits volums amb alguns esdeveniments controlats per fractures al voltant del metre cúbic de volum. L'augment dels esdeveniments registrats en ambdós escarpaments (veure

Figura 3.48) des del 2018 correspon a les plaques de categoria amb volums petits (vegeu Figura 3.49) de menys de mig metre cúbic.

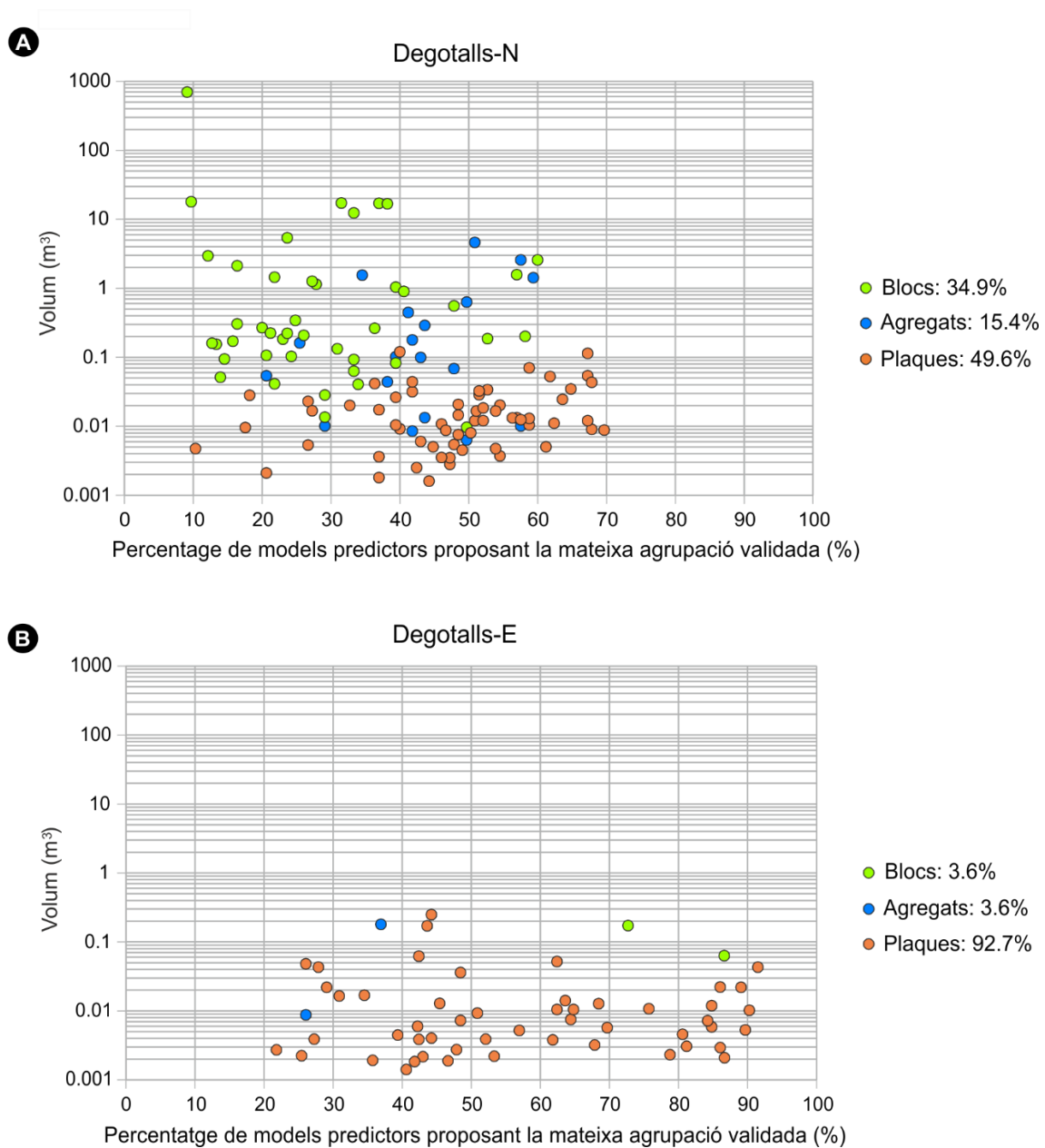


Figura 3.49: Relació entre el volum dels agrupaments de caiguda de blocs, les classes de caiguda de blocs i el percentatge de models que prediuen els mateixos agrupaments de caiguda de blocs validats als escarpaments de Degotalls. (A) Les classes de caiguda de roques de Degotalls-N són plaques majoritàries, normalment associades a processos de meteorització amb volums petits, i, blocs grans, a causa del gran desprendiment durant el període 2007-2009. Ambdues classes defineixen un escenari heterogeni que dona lloc a més dificultats en l'etapa d'entrenament dels models predictius. (B) Degotalls-E presenta una classe homogènia amb volums petits que faciliten la identificació de models predictius.

La distribució acumulada de volums registrada als Degotalls presenta exponents de la llei de potència lleugerament diferents en ambdós escarpaments, malgrat que les condicions estructurals de la massa rocosa són les mateixes, només variant la orientació de la paret rocosa.

En el cas de Degotalls-N, la distribució cobreix un rang més ampli d'ordres de magnitud a causa del gran esdeveniment de caiguda de roques del període 2007-2009 (veure Figura 3.50). En conseqüència, el mostreig comú en el rang de 0,01 a 10 m³ és la més representativa de l'activitat de caiguda de blocs. Aquests resultats haurien de determinar els escenaris a considerar, en una avaluació addicional del perill per a l'àrea d'aparcament del santuari (Mineo et al., 2018).

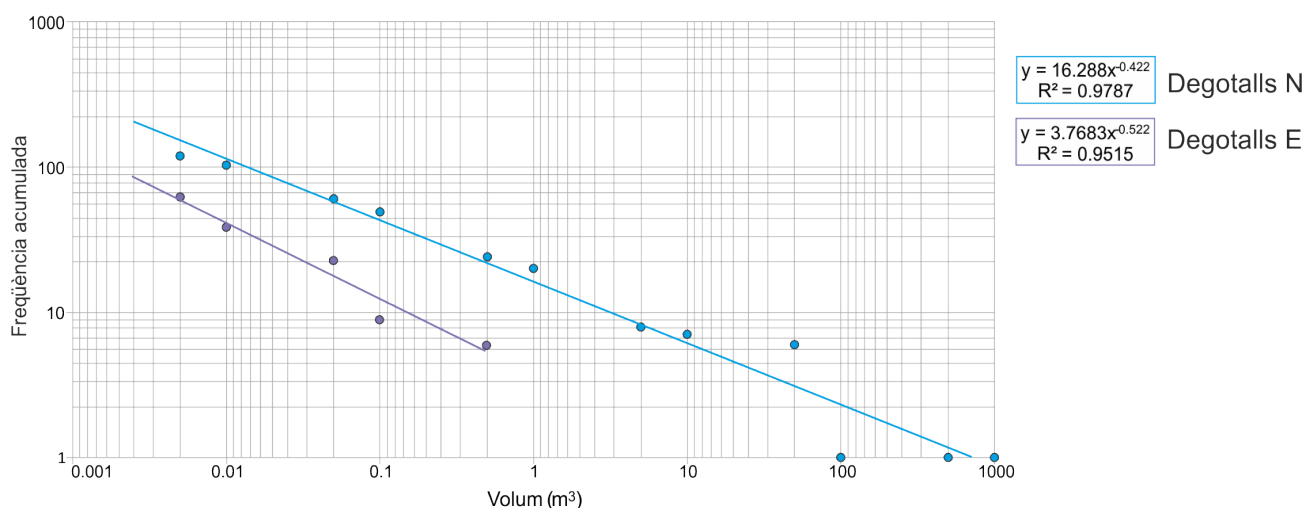


Figura 3.50: Magnitud-volum als dos escarpaments rocosos de Degotalls. Degotalls-N en color blau i Degotalls-E en color violeta.

Una possible nova estratègia per optimitzar els resultats a Degotalls-N és alimentar l'etapa d'entrenament amb cada Sub-categories dels candidats (grans blocs, agregats i plaques), sempre que estigui disponible un nombre necessari d'exemples d'entrenament.

En comparació amb els resultats de la metodologia estàndard aplicada a Degotalls (Janeras et al., 2017; Royan, 2015), es troba que hi ha una variació en la quantitat i la distribució temporal dels esdeveniments. Això, sens dubte, es pot atribuir a la reducció del nombre d'agrupaments de punts a validar que permet augmentar la qualitat de la validació i en conseqüència, redueix el nombre de casos dubtosos, apart de la validació amb imatge. A més, l'avaluació objectiva de cada agrupament com a candidat a una caiguda de roques millora el procés, i contribueix a millorar la interpretació de l'evolució dels escarpaments i la valoració de les activitats de mitigació del risc.

L'anàlisi de la naturalesa dels agrupaments no ha permès identificar cap procés de deformació prèvia a una caiguda de bloc, pel que no s'ha pogut crear una nova classe. L'estudi de la zona en anys previs a la caiguda de bloc, amb la intenció de detectar aquestes deformacions ha estat improductiu. El motiu d'aquesta no identificació pot ser degut a l'ús d'una resolució temporal massa baixa per registrar aquest procés, és a dir, que la deformació prèvia a la caiguda, s'esdevé en un temps inferior a la freqüència d'escanejos. Possiblement, la modificació de la resolució temporal i espacial durant el monitoratge permetin la identificació d'agrupaments d'avanç o de

deformació prèvia a la caiguda, i per tant possibilita la seva classificació amb la metodologia proposada.

3.3.6. Conclusions

El seguiment de l'evolució dels escarpaments fracturats per identificar les caigudes de blocs de manera ràpida i eficient amb núvols de punts, requereix mètodes de processament de grans conjunts de dades, que sovint són desenvolupats per la comunitat científica.

La captura periòdica en format digital de superfícies d'escarpaments amb instruments com el TLS no proporciona directament un inventari de les caigudes de roques que s'hagin produït. Per a aquest motiu, és necessari processar les dades capturades i arribar a identificar, la freqüència de mesures i la resolució espacial del TLS d'acord amb la dinàmica de les caigudes de blocs de l'escarpament estudiat. La captura de núvols de punts, la seva alineació, la mesura de les diferències i l'agrupament de punts són els aspectes més investigats i evolucionats, en termes d'automatitzacions dels processats. Tanmateix, la classificació dels agrupaments de punts és un procediment que no s'aborda amb tanta freqüència. El present treball proposa una solució en aquest problema basada en la tècnica del *machine learning* i els models predictius.

La metodologia proposada està basada en els programes *Point Cloud Monitoring* (PCM) i *Cluster Classification*, per identificar automàticament els agrupaments de punts mitjançant tècniques de *machine learning* i classificar-los com a caigudes de blocs. Que els agrupaments de punts tinguin atributs similars facilita l'etapa d'aprenentatge, fet que s'observa en els exemples estudiats als escarpaments de Degotalls.

Per realitzar la etapa del *Cluster Classification* es combina 11 models o algorismes de classificació amb 15 mètodes de remostreig (13 de sotamostreig i 2 de sobremostreig). Els models, han estat refinats i validats mitjançant tècniques de validació creuada i hiperparametrització.

A Degotalls-E, amb les diferents combinacions de models predictius i mètodes de remostreig s'aconsegueix identificar, entre els 115 primers agrupaments de punts més proposats pels models, el 100% dels desprendiments. Això implica una reducció important dels agrupaments inicials (més de 5.800) per validar-los com a desprendiments. A Degotalls-N, la aquests resultats són més desfavorables, requerint la validació de les 370 primeres agrupacions per identificar el 90% de les caigudes de blocs. L'eficiència dels models predictius també tendeix a augmentar quan la base de dades de caiguda de blocs també ho fa en nombre d'exemples.

En el cas d'estudi, l'èxit dels models de predicció de caiguda de blocs depèn de l'homogeneïtat o heterogeneïtat de les característiques dels diferents atributs dels agrupaments classificats com a caiguda de blocs. Aquesta diferència la observem clarament en els escarpaments de Degotalls. En el cas de Degotalls-E, els exemples de caigudes de blocs són classe **plaques**, en general

amb atributs més homogenis i polaritzats (veure figura 3.49). En el cas Degotalls-N presenta un menor percentatge de predicció pels models, quan ha d'identificar la categoria de **grans blocs** amb volums superiors a 0,1 m³. En canvi, la categoria **plaques** amb volums més petits classe **plaques**, obtenen els percentatges més alts d'identificacions ja que hi ha més exemples i de major homogeneïtat.

S'ha observat que la combinació de models o algorismes de classificació i mètodes de remostreig més eficients, canvia segons el període estudiat. Si tenim en compte que els models classificadors són 11 i els de remostreig 15, les combinacions són altes. Per aquest motiu, no ha estat possible identificar una combinació única per a la predicció dels blocs caiguts als escarpaments de Degotalls. Per resoldre aquesta problemàtica, el programari *Cluster Classification* calcula el percentatge de vegades que cada agrupació de punts es etiquetada com a "Candidats" per qualsevol model de predicció. Per tant, si intentem validar només les agrupacions de punts més proposades pels models predictius, aconseguim millorar la mètrica de la classificació.

Amb l'anàlisi realitzat des del 2007 fins al 2020, substituint la metodologia estàndard aplicada a Degotalls (Janeras et al., 2017; Royan, 2015), per la metodologia proposada s'observen noves dades, tant en la quantitat com en la distribució temporal dels esdeveniments. Això, sens dubte, es pot atribuir a la reducció del nombre d'agrupaments de punts a validar que permet augmentar la qualitat de la validació i en conseqüència, redueix el nombre de casos dubtosos, a part de la validació amb imatge (veure Figura 3.48).

Tot que s'han estudiat diferents períodes de temps per identificar les deformacions prèvies a caigudes, no ha estat possible. El motiu, es pot atribuir a l'anàlisi de freqüències temporal massa baixa per registrar aquest procés, és a dir, que la deformació prèvia a la caiguda, s'esdevé en un temps inferior a la freqüència d'escanejos. La modificació de la resolució temporal i espacial durant el monitoratge podria permetre la identificació d'agrupaments de punts de deformació prèvia a la caiguda, i per tant possibilitaria la introducció d'una nova classe en la metodologia proposada.

Els desprendiments dels escarpaments de Degotalls a la Muntanya de Montserrat, es troben actualment en fase d'estabilització, en part també, gràcies a les mesures de contenció i fixació aplicades pel Santuari per mitigar el risc de caiguda. Tot i que els desprendiments registrats en el període d'estudi (2007-2020), han superat la centena, aquests, es corresponen a caigudes categoritzades com a **plaques**. Els processos que provoquen la caiguda d'aquestes **plaques** s'associen a la meteorització i no a les fractures del massís. En aquesta direcció, per aconseguir associar caigudes de blocs amb les famílies de fractures presents, caldrà obtenir nous exemples de caigudes de **grans blocs** que es puguin associar a les fractures.

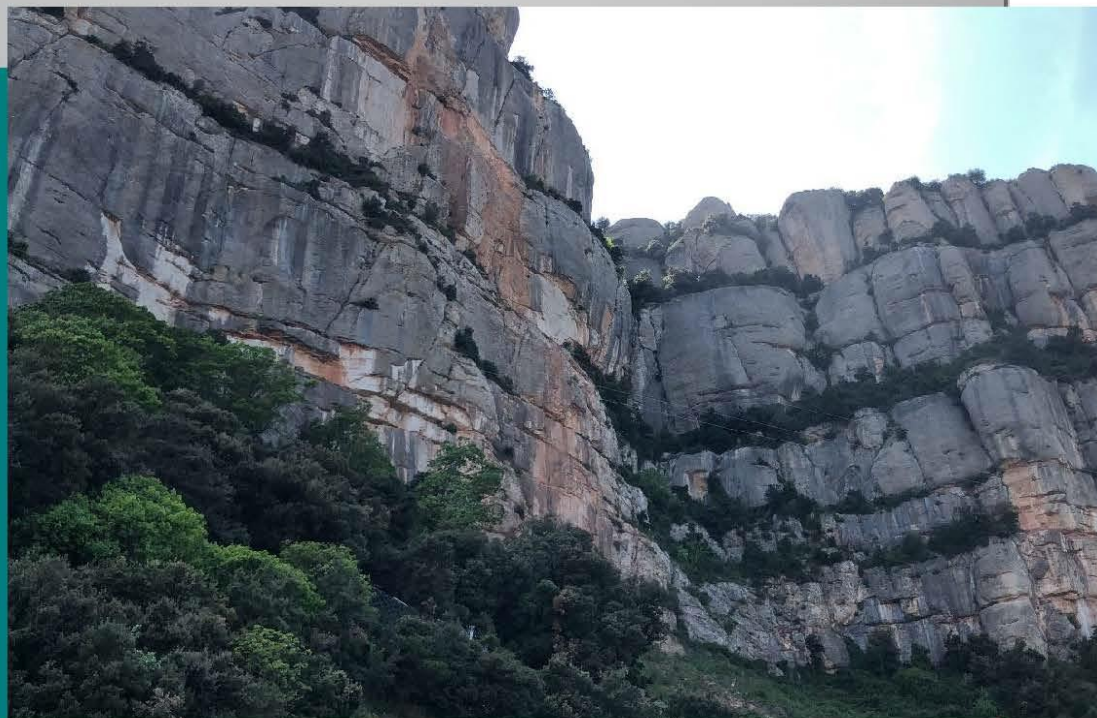
El seguiment de les caigudes de blocs en escarpaments amb núvols de punts, és una tasca amb processos sofisticats, que es beneficia de les estratègies de *machine learning*, en les etapes

d'identificació i classificació de blocs caiguts. Aquest fet és més beneficiós quan, com en el cas dels escarpament de Montserrat, els blocs caiguts són uns pocs a identificar i classificar entre milers d'agrupaments de punts (fins a 5.800 en el cas de Degotalls-E). En aquesta tesi doctoral s'ha validat aquesta hipòtesi amb l'experiment realitzat als escarpaments de Degotalls situats al massís de Montserrat amb resultats satisfactoris.

CAPÍTOL 4

Discussions i conclusions finals i perspectiva d'avanç

- 4.1. Introducció
- 4.2. Resum dels principals resultats
- 4.3. Discussions finals
- 4.4. Conclusions finals
- 4.5. Treballs futurs



4.1. Introducció

Partint de la hipòtesis que, la incorporació tant de noves tècniques remotes com de les eines de *machine learning* augmenten exponencialment el coneixement de medis fracturats, però que a la vegada, dificulten el seu tractament i control de les dades, en aquest doctorat industrial, s'avalua l'abast d'aquestes tècniques en el reconeixement de discontinuïtats, bàsicament fractures, i s'incorporen algorismes de *machine learning* en exemples aplicats.

La importància de les discontinuïtats presents en els medis fracturats i el paper cabdal que hi juguen per a la societat ho és també per a les empreses que es dediquen a la geotècnia o hidrogeologia. En aquest context, es defineixen dos objectius específics. Per una banda, aplicar tècniques remotes o indirectes per la digitalització d'afloraments i avaluar la representativitat de les dades captades en la construcció de models discrets de fractures tridimensionals. Per altra banda, desenvolupar un model predictiu de *machine learning* en la monitorització digital d'un aflorament, per capturar, identificar i classificar les caigudes de blocs que s'hi desencadenen.

4.2. Resum dels principals resultats

En el primer cas d'estudi, realitzat a la pedrera La Fou a Vallirana, els resultats obtinguts mostren que les tècniques remotes obtenen un alt nombre de superfícies de fractures, especialment quan s'aplica el TLS (més de 52.000 fractures) i que les dades representades mostren diferències entre les tècniques (TLS, fotogrametria i *scan-lines*), en funció de quina sigui la propietat geomètrica estudiada (Priest, 1993); orientacions, espaiat, alçades i longituds de fractures. Per tals efectes, i considerant les mateixes obertures per tots tres models (atès que aquest paràmetre només es pot obtenir mitjançant la pràctica dels *scan-lines*), s'han generat els mapes de traces i s'ha calculat la porositat i permeabilitat secundària. Com a resultat obtenim que la porositat mitja en el DFN_{SfM} i DFN_{scan-line} és de 5,7% i en el DFN_{TLS} és de 5,4% i que l'estimació de la permeabilitat mitja de la xarxa de fractures en el model DFN_{SfM}, és de 4×10^5 mD, lleugerament superiors als models DFN_{TLS} o DFN_{scan-line}, que ofereixen valors mitjos de $6,2 \times 10^5$ mD. Comparant la informació de sortida dels models, s'observa que, per a la construcció de models DFN, s'obtenen majors similituds en els resultats de porositat i permeabilitat quan s'apliquen les tècniques de TLS i *scan-lines*, que quan es fa ús de la fotogrametria, de la qual resulten valors de porositat i permeabilitat lleugerament més baixos.

En el segon cas d'estudi, realitzat als escarpaments de Degotalls a Montserrat, els resultats generats a partir de la combinació de 11 algorismes classificadors amb 15 mètodes de remostreig, han permès identificar la combinació predictiva amb major efectivitat per classificar les agrupacions de punts associades a caigudes de blocs, per cadascun dels exemples

analitzats. Aquest aspecte és rellevant, atès que, pel cas estudiat, a Degotalls-E, a partir d'aproximadament 5.800 agrupaments identificats per cada monitoratge, s'assoleix una reducció del 98% del nombre d'agrupaments a validar, entre Falsos positius (FP) i Veritables Positius (VP), entenent (VP) com les caigudes reals de blocs. A Degotalls-N la reducció ha estat del 80,16% per a una identificació completa, mentre que per identificar el 96% dels caigudes de blocs reals, la reducció ha estat del 90%, a partir d'una població d'uns 3.700 agrupaments inicials. És important també assenyalar, que la validació de la tasca mitjançant la comparació d'imatges d'alta resolució, ha permès augmentar notablement la fiabilitat dels resultats obtinguts.

Els resultats ampliat per a cada cas d'estudis es poden consultar en la subsecció 3.2.4 per el cas on s'apliquen les diferents tècniques remotes en la construcció de models DFN i en la subsecció 3.3.4 quan es detallen els resultats obtinguts d'aplicar la proposta metodològica de classificació automàtica de blocs caiguts als escarpaments de Degotalls.

4.3. Discussió general

La introducció de tècniques remotes per identificar, caracteritzar i analitzar els medis fracturats, permet estudiar en detall les estructures aflorants. Al tractar-se de tècniques indirectes, la captura de les dades es realitza des de certa distància a l'aflorament, fet que comporta grans avantatges i alguns inconvenients.

Per una banda, s'aconsegueix capturar superfícies de medis aflorants amb grans alçades o distàncies inaccessibles, que d'altra forma comportaria grans dificultats per accedir-hi. En aquest sentit, al no ser necessari un contacte directe amb l'aflorament, evita també situacions de risc per caiguda de blocs. A més, els afloraments digitalitzats són perdurables en el temps i podran ser tractats en el futur per assolir altres objectius, donant-li un valor afegit considerable.

Per altra banda, són tècniques sofisticades amb un alt grau de complexitat en el seu tractament. El tractament de les dades i del seu anàlisi augmenta considerablement respecte altres tècniques manuals, convertint-se en ocasions, en tot un repte. Algunes propietats, no queden recollides en la captura indirecta dels afloraments i cal utilitzar tècniques directes per a recollir les dades, com és el cas de les obertures. També tenen algunes limitacions pròpies de cada tècnica com hem estat veient al llarg de la tesi doctoral (Subsecció 3.2.5). La orientació de l'aflorament i de les fractures, per exemple, permetrà capturar amb major precisió algunes famílies de fractures que no pas unes altres. També la mesura d'algunes propietats, com per exemple; l'alçada de les fractures que pot quedar fraccionades per la configuració del dispositiu TLS, com ha succeït a la tesi en el cas de la Fou. O l'espaiat, que amb la tècnica de fotogrametria, tindrà una forta dependència dels resultats obtinguts en la digitalització manual de les traces de les fractures, que a la vegada dependrà de l'usuari digitalitzador i de la qualitat de les fotografies preses.

En definitiva, l'ús de tècniques remotes en medis fracturats, quan el que es pretén és estudiar en detall les xarxes de fractures, permet obtenir dades que amb altres metodologies no es podrien

obtenir i reduir incerteses en la caracterització de les fractures. La combinació de les diferents tècniques aprofitant les avantatges de cadascuna d'elles per obtenir les propietats de les fractures més aproximades a la realitat, permetrà obtenir models amb menors incerteses. Gràcies en aquesta investigació, s'obtenen noves dades, que serviran com a eines de decisió per a la utilització de les tècniques remotes i/o manuals en l'anàlisi de xarxes de fractures encaminades a la modelització discreta de fractures (DFN) aplicats a simulacions de flux.

Per altra banda, l'ús de tècniques remotes per monitorar un aflorament fracturat és una eina fonamental i no té competència amb altres tècniques. La identificació i classificació de caigudes de blocs, es podrà realitzar a partir de la digitalització de l'aflorament monitorat en espais de temps diferents. Els processos són complexos i la gran quantitat de dades a tractar comporta temps molt elevats quan alguna de les etapes és manual. En aquests casos, si aquests processos manuals pretenen classificar algun atribut com són les caigudes de blocs, la utilització de les eines de *machine learning*, són la millor opció, tal i com es demostra en aquesta tesi doctoral. Per a tals efectes, s'ha desenvolupat una metodologia de classificació automàtica de blocs caiguts per substituir un procés de classificació manual que comportava subjectivitat en els resultats i grans períodes de temps en la classificació. Al substituir la metodologia i comparar els resultats s'observen noves dades amb major fiabilitat. També s'aconsegueix reduir el temps d'identificació i classificació i s'observa que amb nous exemples, l'aprenentatge és major i la identificació i classificació és més eficient. Els resultats han estat molt satisfactoris i també han quedat publicats en l'article Blanco et al., (2022)

En aquesta secció s'han discutit de forma generalista, els grans resultats de la tesi doctoral. En canvi, les discussions específiques de cada cas d'estudis es troben a la Subsecció 3.2.5, quan s'avaluen les tècniques remotes aplicades en models DFN i en la subsecció 3.3.5, quan es detalla la proposta metodològica de classificació automàtica de blocs caiguts.

4.4. Conclusions finals

En termes generals s'ha assolit l'objectiu principal de la tesi doctoral, on es pretenia incorporar noves tècniques remotes i eines de *machine learning* en la caracterització i parametrització de les discontinuïtats presents en el medi fracturat, per avaluar, optimitzar i/o millorar el seu ús en l'àmbit de l'obra civil o per al control i l'explotació sostenible de recursos naturals.

Més específicament, la investigació s'ha centrat en dos casos d'estudi per donar resposta als objectius específics definits a la tesi doctoral. S'han assolit també aquests objectius específics amb resultats molt interessants i aplicables al món industrial. També han sorgit noves incerteses que a la vegada ens presenten noves necessitats d'estudi de futur.

- Enmarcat en el primer objectiu específic, en el cas d'estudi de la Pedrera la Fou, s'ha aconseguit identificar, caracteritzar, parametritzar i analitzar les discontinuïtats presents

al medi rocós fracturat mitjançant diferents tècniques remotes avançades, per tal de valorar la seva aplicació en hidrogeologia a partir de models DFN. En aquest sentit, els resultats finals de l'avaluació han estat exitosos, obtenint les avantatges i inconvenients de cada tècnica. Aquests resultats permeten discernir entre les tècniques per decidir en cada cas d'estudi individual, quina o quines tècniques poden ser més convenients per caracteritzar la xarxa de fractures. Durant la investigació també han sorgit canvis, en ocasions han estat canvis que no s'esperaven i en d'altres ocasions noves oportunitats d'aprofundir en les investigacions. En aquest sentit, alguns dels treballs que s'havien plantejat inicialment, com la recollida de dades en 3D de les fractures amb GPR no han mostrat els resultats esperats, segurament limitat per la incompatibilitat de la tècnica amb el terreny. En canvi, altres noves oportunitats d'estudi han aparegut, com la possibilitat de disposar de les testificacions d'alguns sondejos bibliogràfics fet que ens ha permès estudiar alguns aspectes rellevants en termes de distribució de fractures.

En el cas d'estudi de la Fou, es conclou que els models DFN construïts a partir de les tècniques remotes són una bona combinació per estudiar la xarxa de fractures d'un medi fracturat aplicats a la simulació del flux. Tal i com es mostra en aquesta tesi doctoral, les dades capturades i tractades amb tècniques remotes queden representades en models DFN amb resultats molt acceptables. Les avantatges d'aquestes tècniques en l'estudi de La Fou, són varies però destaquen (1) L'obtenció d'un gran volum de dades representatives (2) La captura d'aquestes dades sobre tota la superfície aflorant (3) La major seguretat que comporta recollir les dades a certa distància de l'aflorament amb perill de desprendiments, (4) la durabilitat de les dades, que poden ser processades en diferents objectius al llarg del temps, (5) la obtenció de les propietats geomètriques de les fractures a partir de processats. Cal recordar que segons la propietat geomètrica de la fractura estudiada, la orientació de l'aflorament i de les fractures, s'obtidran resultats més precisos amb una tècnica o una altra. Les principals limitacions que ha comportat treballar amb aquestes tècniques en l'estudi de La Fou han estat; (1) aquesta dependència a la orientació de l'aflorament i les fractures, (2) la propietat de l'espaiat identificada amb la tècnica de la fotogrametria sobredimensionada, conseqüència del procés de digitalització on l'usuari actua de filtre involuntari obviant algunes fractures menys destacades i (3) la propietat de l'alçada i de la longitud de les fractures identificada amb TLS sotadimensionada, conseqüència del pre-processat en la reconstrucció de les fractures, la tècnica TLS tendeix a fragmentar les alçades de les fractures, obtenint mesures que no arribaven al metre.

En definitiva, els resultats obtinguts suggereixen, que els tres models permeten obtenir la porositat i permeabilitat del mapa de traces i que els valors obtinguts són similars. La tècnica de fotogrametria, presenta valors lleugerament més baixos de permeabilitat i la tècnica TLS de porositat, però en diferències insignificants en termes de flux de fluid regional. Per tant, les diferències geomètriques representades en el conjunt de dades,

no presenta importants conseqüències en l'estimació dels valors de porositat i permeabilitats de la xarxa de fractures en el cas d'estudi de La Fou. En conseqüència, tots els models serviran per obtenir els paràmetres equivalents geomètrics de les fractures i realitzar simulacions de flux, en un programari específic, tenint en compte les limitacions comentades.

- El segon objectiu específic centrat en els escarpaments de Degotalls a Montserrat, pretenia incorporar eines de machine learning per l'automatització de la classificació de blocs caiguts. En aquest sentit, s'ha aconseguit obtenir una proposta de metodologia que permet identificar i classificar els blocs caiguts amb un molt alt percentatge d'encert. Aquesta metodologia aplicada a Degotalls, millora a mida que es nodreix de nous exemples de caigudes. Per tot això, aquesta metodologia s'ha establert com a eina de treball, en el monitoratge dels escarpaments de Degotalls que es realitzen entre l'ICGC i la Facultat de Ciències de la Terra. Aquesta proposta metodològica identifica i classifica la caiguda de blocs en afloraments fracturats que són digitalitzats amb tècnica LiDAR amb el dispositiu TLS. A contra, una de les hipòtesis d'aquest estudi on es pretenia relacionar els blocs caiguts amb les famílies de fractures del medi estudiat no s'ha assolit. Tot i que els desprendiments registrats en el període d'estudi (2007-2020), han superat la centena, aquests, es corresponen a caigudes categoritzades com a **plaques**. Els processos que provoquen la caiguda d'aquestes **plaques** s'associen a la meteorització i no a les fractures del massís. En aquesta direcció, per aconseguir associar caigudes de blocs amb les famílies de fractures presents, caldrà obtenir nous exemples de caigudes de **grans blocs** que es puguin associar a les fractures. El Patronat de Montserrat ha incorporat mesures i construït infraestructures aplicades en la estabilització dels escarpaments en els últims anys són impecables, amb l'aplicació de bulons en els sectors més inestables, malles de triple torsió i d'altres mesures de contenció. Per aquesta raó trobar nous exemples de **grans blocs** a Degotalls, ha estat difícil.

En tots els casos, les tècniques remotes aplicades a l'escaneig i estudi dels medis fracturats augmenta la seguretat del tècnic geòleg quan ha d'estudiar un aflorament fracturat. Ja sigui en una pedrera activa, en un túnel que avança, en un escarpament inestable, la perillositat de caiguda de blocs treballant directament sobre l'aflorament augmenten amb el temps d'exposició. Per aquesta raó, les avantatges que implica utilitzar tècniques remotes per l'anàlisi d'afloraments fracturats, en termes de seguretat, són incomparables amb altres tècniques directes.

Les conclusions específiques de cada cas d'estudi i que amplien el detall de les conclusions finals, es troben a la Subsecció 3.26 en l'avaluació de tècniques en models DFN i en la Subsecció 3.3.6 en la proposta metodològica utilitzant machine learning, per la classificació de blocs caiguts.

4.5. Perspectives d'avanç

Les perspectives d'avanç d'aquesta tesi doctoral van en consonància amb el que representa un doctorat industrial, a través de projectes amb la Universitat i la empresa Anufra. En aquest sentit projectes com el *Digital transition in Geology: New challenges from field to laboratory datasets (GEODIGIT)* adjudicat en el 2022, que incorpora l'àrea d'estudi de la Pedrera La Fou. Aquest ja és un exemple real de la continuïtat i la perspectiva d'avanç.

Pel que fa a l'avaluació de les tècniques remotes i per completar algunes incerteses de **l'experiment dirigit a l'avaluació de les tècniques remotes en la construcció de models DFN en tres dimensions**, es proposen els següents passos;

- La construcció d'un nou model discret de fractures combinant les tècniques remotes o indirectes, per obtenir les propietats de les fractures, en funció de la tècnica més adequada, permetria obtenir un model més pròxim a la realitat. En aquest nou model, caldria també, incorporar alguns aspectes manualment, com són els corredors de fractures, ja que són un pas preferent del flux.
- A partir d'aquesta modelització caldria anar un pas més enllà i estudiar la simulació de flux (veure Secció 2.4.1) a partir de programes com FracMan, Modflow o Petrel, i calibrar el model a partir d'assajos hidràulics, per corroborar les dades obtingudes de connectivitat i permeabilitat hidràulica de les fractures. Aquests assajos s'haurien de realitzar mitjançant assajos Lugeon o assajos de bombeig. Amb aquestes dades obtindríem les dades d'obertura del medi fracturat.

També, la incorporació d'aquests models DFN, en projectes de l'empresa d'Anufra, a partir de dades reals adquirides amb tècniques remotes, on la intenció és la modelització de flux ens presenten oportunitats d'anàlisi que es poden aplicar a l'estudi de la circulació i comportament de fluids en medis rocosos, l'estudi de reservoris d'hidrocarburs o el transport tèrmic en geotèrnia. També tindrà aplicabilitat en d'altres àmbits com l'anàlisi de les conseqüències de la propagació d'explosius en pedreres o l'anàlisi de la estabilitat de massissos rocosos associats a túnels o mines, entre d'altres.

Per altra banda, la **implantació de la metodologia desenvolupada amb *machine learning* per la identificació de blocs caiguts**, en els escarpaments de Degotalls ha donat resultats satisfactoris. Aquesta metodologia, es continua utilitzant en el monitoratge de Degotalls-N i Degotalls-S. No obstant això, sorgeixen noves perspectives de futur com l'aplicabilitat en altres escarpaments, la comptabilitat del mètode amb altres dispositius TLS, com els de petit format, l'aplicabilitat en la detecció de deformacions de la roca, prèvies a les caigudes de blocs per implementar sistemes d'alerta i les possibles millores de la metodologia proposada.

- La implantació de la metodologia en nous exemples d'escarpaments també és un projecte que ja és realitat. Aquests nous exemples comporten beneficis clars en la identificació i classificació de blocs caiguts i permeten validar una metodologia experimental. També la identificació d'exemples de caigudes de blocs de major mida "blocs", identificaria quins models i mètodes s'adaptarien millor a la nova configuració. Es milloraria així, el coneixement de l'eina i les possibilitats reals, en altres tipus d'escarpaments o de classificacions, podent ser la llavor d'un software comercial.
- Els dispositius TLS de petit format, de baix cost i de fàcil mobilització però, amb menor resolució i precisió, tenen una gran avantatge que és el temps de captura. En pocs segons captura l'escarpament i s'obté un núvol de punts. Actualment, la Facultat de Ciències de la Terra, ja n'ha adquirit un per instal·lar-lo a Puigcerçós o Montserrat. L'objectiu és intentar captar les possibles deformacions de la roca, prèviament a caure, elevat la freqüència de captura. La incorporació de la nova metodologia, permetria un grau elevat d'automatització en la classificació dels agrupaments i podríem arribar a identificar dos classes; caigudes de blocs i deformacions. L'adaptació del mètode a les característiques de l'aparell TLS serà bàsica per obtenir l'èxit. Incorporant un sistema de telecomunicacions, es podrien obtenir els núvols de punts remotament i treballar sobre les dades. L'automatisme dels processos són rellevants si els que es pretén és incorporar un sistema d'alerta.
- Algunes qüestions sense resoldre durant la tesi també són objecte de millora, com per exemple, provar la discriminació de cada atribut introduït en el procés d'aprenentatge. Les conseqüències d'aquest treball, permetrien fer processats més eficients, incorporant els atributs amb major rellevància en el model de predicció.

Finalment, queda pendent, la publicació de dos articles en els que s'està treballant. En un primer cas, juntament amb l'ICGC, l'article té com objectiu final, publicar l'obtenció de la nova corba de magnitud i freqüència dels escarpaments de Degotalls amb la utilització de la metodologia proposada i descrita en aquesta tesi doctoral i a l'article (Blanco et al., 2022). En un segon cas, l'article a publicar correspondria a l'avaluació de les tècniques remotes com a eina per a construir models discrets de fractures i les seves avantatges o desavantatges per a la simulació de flux.

CAPÍTOL 5

Bibliografia



- Abellán, A., Jaboyedoff, M., Oppikofer, T., & Vilaplana, J. M. (2009). Detection of millimetric deformation using a terrestrial laser scanner: experiment and application to a rockfall event. *Natural Hazards and Earth System Sciences*, 9(2), 365–372. <https://doi.org/10.5194/nhess-9-365-2009>
- Abellán, Antonio, Oppikofer, T., Jaboyedoff, M., Rosser, N. J., Lim, M., & Lato, M. J. (2014). Terrestrial laser scanning of rock slope instabilities. *Earth Surface Processes and Landforms*, 39(1), 80–97. <https://doi.org/10.1002/esp.3493>
- Albrich, S., Bernaus, J., Boix Fayos, C., Caus, E., Martín-Closas, C., Salas, R., Vicedo, V., & Villalonga, R. (2006). Caracterización bioestratigráfica y paleoambiental del Cretácico Inferior (Berriasiense-Barremiense) del Macizo de Garraf (Cadena Costera Catalana). *Revista Española de Micropaleontología*, 38(2), 429–451.
- Alsaker, E., Gabrielsen, R. H., & Roca, E. (1996). The significance of the fracture pattern of the Late-Eocene Montserrat fan-delta, Catalan Coastal Ranges (NE Spain). *Tectonophysics*, 266(1–4), 465–491. [https://doi.org/10.1016/S0040-1951\(96\)00239-9](https://doi.org/10.1016/S0040-1951(96)00239-9)
- Anadón, P., Marzo, M., & Puigdefàbregas, C. (1985). *The Eocene fan-delta of Montserrat (southeastern Ebro basin, Spain)*.
- Andersson, J., Shapiro, A. M., & Bear, J. (1984). A stochastic model of a fractured rock conditioned by measured information. *Water Resources Research*, 20(1), 79–88.
- Angluin, D. (1988). *Queries and Concept Learning*.
- Annan, A. P. (2001). GPR—trends, history, and future developments. *Proceedings of the EAGE 2001 Conference, Delft, Netherlands*.
- Assali, P., Grussenmeyer, P., Villemain, T., Pollet, N., & Viguier, F. (2014a). Surveying and modeling of rock discontinuities by terrestrial laser scanning and photogrammetry: Semi-automatic approaches for linear outcrop inspection. *Journal of Structural Geology*, 66, 102–114. <https://doi.org/10.1016/j.jsg.2014.05.014>
- Assali, P., Grussenmeyer, P., Villemain, T., Pollet, N., & Viguier, F. (2014b). Surveying and modeling of rock discontinuities by terrestrial laser scanning and photogrammetry: Semi-automatic approaches for linear outcrop inspection. *Journal of Structural Geology*, 66, 102–114. <https://doi.org/10.1016/J.JSG.2014.05.014>
- Atkinson, K. B. (1996). *Close range photogrammetry and machine vision*. Whittles.
- Awad, M., & Khanna, R. (2015). Support vector machines for classification. In *Efficient learning machines* (pp. 39–66). Springer.
- Ayala Carcedo, F. J. (2006). *Manual de ingeniería de taludes* (I. geológico y minero de España (ed.); 1a ed.).
- B., I., G., R. P., R., P. J., & W., G. (2003). Two-Phase (Air and Water) Flow through Rock Joints: Analytical and Experimental Study. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(10), 918–928. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2003\)129:10\(918\)](https://doi.org/10.1061/(ASCE)1090-0241(2003)129:10(918))
- Baca, R. G., Arnett, R. C., & Langford, D. W. (1984). Modelling fluid flow in fractured-porous rock masses by finite-element techniques. *International Journal for Numerical Methods in Fluids*, 4(4), 337–348.
- Bachman, C. G. (1979). Laser radar systems and techniques. *Dedham*.
- Baecher, G B, Lanney, N. A., & Einstein, H. H. (1977). *Statistical Description Of Rock Properties And Sampling*.
- Baecher, G B, Lanney, N. A., & Einstein, H. H. (1978). *Proceedings of the 18th US Symposium on Rock Mechanics*.
- Baecher, Gregory B. (1983). Statistical analysis of rock mass fracturing. *Journal of the International Association for Mathematical Geology*, 15(2), 329–348. <https://doi.org/10.1007/BF01036074>
- Baghbanan, A., & Jing, L. (2007). Hydraulic properties of fractured rock masses with correlated fracture length and aperture. *International Journal of Rock Mechanics and Mining Sciences*, 44(5), 704–719. <https://doi.org/10.1016/j.ijrmms.2006.11.001>
- Bahat, D. (1983). SURFACE MORPHOLOGY ON CROSS-FOLD JOINTS OF THE APPALACHIAN PLATEAU, NEW YORK AND PENNSYLVANIA. *Tectonophysics*, 104 (1984) 299–313, 104(9), 299–313. <https://doi.org/10.1017/CBO9781107415324.004>
- Bahat, Dov. (1999). Single-layer burial joints vs single-layer uplift joints in Eocene chalk from the Beer Sheva syncline in Israel. *Journal of Structural Geology*, 21(3), 293–303. [https://doi.org/10.1016/S0191-8141\(98\)00118-7](https://doi.org/10.1016/S0191-8141(98)00118-7)

- Bai, T., & Pollard, D. D. (2000). Closely spaced fractures in layered rocks: Initiation mechanism and propagation kinematics. *Journal of Structural Geology*, 22(10), 1409–1425. [https://doi.org/10.1016/S0191-8141\(00\)00062-6](https://doi.org/10.1016/S0191-8141(00)00062-6)
- Baltsavias, E. P., Favey, E., Bauder, A., Bosch, H., & Pateraki, M. (2001). Digital surface modelling by airborne laser scanning and digital photogrammetry for glacier monitoring. *The Photogrammetric Record*, 17(98), 243–273.
- Barbier, M., Leprêtre, R., Callot, J. P., Gasparrini, M., Daniel, J. M., Hamon, Y., Lacombe, O., & Floquet, M. (2012). Impact of fracture stratigraphy on the paleo-hydrogeology of the Madison Limestone in two basement-involved folds in the Bighorn basin, (Wyoming, USA). *Tectonophysics*, 576–577, 116–132. <https://doi.org/10.1016/j.tecto.2012.06.048>
- Barenblatt, G. I., Zheltov, I. P., & Kochina, I. N. (1960). Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks [strata]. *Journal of Applied Mathematics and Mechanics*, 24(5), 1286–1303.
- Barker, J. A., & Black, J. H. (1983). Slug tests in fissured aquifers. *Water Resources Research*, 19(6), 1558–1564. <https://doi.org/https://doi.org/10.1029/WR019i006p01558>
- Barnhart, T. B., & Crosby, B. T. (2013). Comparing two methods of surface change detection on an evolving thermokarst using high-temporal-frequency terrestrial laser scanning, Selawik River, Alaska. *Remote Sensing*, 5(6), 2813–2837.
- Barton, N. (2002). Some new Q-value correlations to assist in site characterisation and tunnel design. *International Journal of Rock Mechanics and Mining Sciences*, 39(2), 185–216. [https://doi.org/10.1016/S1365-1609\(02\)00011-4](https://doi.org/10.1016/S1365-1609(02)00011-4)
- Barton, N., Løset, F., Lien, R., & Lunde, J. (1981). APPLICATION OF Q-SYSTEM IN DESIGN DECISIONS CONCERNING DIMENSIONS AND APPROPRIATE SUPPORT FOR UNDERGROUND INSTALLATIONS. *Subsurface Space: Environmental Protection Low Cost Storage Energy Savings*, 2, 553–561. <https://doi.org/10.1016/B978-1-4832-8421-7.50080-6>
- Barua, S., Islam, M., & Murase, K. (2013). ProWSyn: Proximity weighted synthetic oversampling technique for imbalanced data set learning. *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 317–328.
- Batista, G. E. A. P. A., Prati, R. C., & Monard, M. C. (2004). A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *SIGKDD Explor. Newsl.*, 6(1), 20–29. <https://doi.org/10.1145/1007730.1007735>
- Battulwar, R., Zare-Naghadehi, M., Emami, E., & Sattarvand, J. (2021). A state-of-the-art review of automated extraction of rock mass discontinuity characteristics using three-dimensional surface models. *Journal of Rock Mechanics and Geotechnical Engineering*, 13(4), 920–936.
- Bauer, M., & Tóth, T. M. (2017). CHARACTERIZATION AND DFN MODELLING OF THE FRACTURE NETWORK IN A MESOZOIC KARST RESERVOIR: GOMBA OILFIELD, PALEOGENE BASIN, CENTRAL HUNGARY. *Journal of Petroleum Geology*, 40(3), 319–334. <https://doi.org/https://doi.org/10.1111/jpg.12678>
- Bayle. (1973). *Etude de la carrière calcaire des Cementos Molins (no publicat)*.
- Bech, N., Bourguin, B., Castaing, C., Chilés, J. P., Christensen, N. P., Frykman, P., Genter, A., Gillespie, P. A., Høier, C., & Klinkby, L. (2001). Fracture interpretation and flow modelling in fractured reservoirs. *European Commission Joule Programme Report. European Commission, Brussels. No EUR, 18946*.
- Belayneh, M. W., Matthai, S. K., Blunt, M. J., & Rogers, S. F. (2009). Comparison of deterministic with stochastic fracture models in water-flooding numerical simulations. *AAPG Bulletin*, 93(11), 1633–1648. <https://doi.org/10.1306/07220909031>
- Benjamin, J., Rosser, N. J., & Brain, M. J. (2020). Emergent characteristics of rockfall inventories captured at a regional scale. *Earth Surface Processes and Landforms*, 45(12), 2773–2787.
- Berg, C. R. (2012). THE EFFECT OF FRACTURE AND BOREHOLE ORIENTATION ON FRACTURE Borehole versus Fracture Orientation — the Scan Line. *Report, 1965*.
- Berkowitz, B. (2002). Characterizing flow and transport in fractured geological media: A review. *Advances in Water Resources*, 25(8–12), 861–884. [https://doi.org/10.1016/S0309-1708\(02\)00042-8](https://doi.org/10.1016/S0309-1708(02)00042-8)
- Bertotti, G., Hardebol, N., Taal-van Kroppen, J. K., & Luthi, S. M. (2007). Toward a quantitative definition of mechanical units: New techniques and results from an outcropping deep-water turbidite succession (Tanqua-Karoo Basin, South Africa). *AAPG Bulletin*, 91(8), 1085–1098. <https://doi.org/10.1306/03060706074>
- Bertrand, L., Géraud, Y., Le Garzic, E., Place, J., Diraison, M., Walter, B., & Haffen, S. (2015). A multiscale analysis of a fracture pattern in granite: A case study of the Tamariu granite,

- Catalunya, Spain. *Journal of Structural Geology*, 78, 52–66. <https://doi.org/10.1016/j.jsg.2015.05.013>
- Besl, P. J., & McKay, N. D. (1992). A Method for Registration of 3-D Shapes, 1992. *IEEE Trans Pattern, 1067*, 239–256.
- Bieniawski, Z. T. (1973). Engineering Classification of Jointed Rock Masses. *Civ Eng S Afr*, 15(12), 335–343. [https://doi.org/10.1016/0148-9062\(74\)90924-3](https://doi.org/10.1016/0148-9062(74)90924-3)
- Blanch, X., Eltner, A., Guinau, M., & Abellan, A. (2021). Workflow for Enhanced Change Detection Using Time-Lapse Cameras. *Remote Sensing*, 13, 1460.
- Blanch, X., & Guinau, M. (2017). 10 YEARS OF ROCKFALL ANALYSIS IN MONTSERRAT (NE SPAIN). May.
- Blanco, L., García-Sellés, D., Guinau, M., Zoumpikas, T., Puig, A., Salamó, M., Gratacós, O., Muñoz, J. A., Janeras, M., & Pedraza, O. (2022). Machine Learning-Based Rockfalls Detection with 3D Point Clouds, Example in the Montserrat Massif (Spain). In *Remote Sensing* (Vol. 14, Issue 17). <https://doi.org/10.3390/rs14174306>
- Board, M. (1989). *UDEC (Universal Distinct Element Code) Version ICG1*. 5. <https://www.osti.gov/biblio/5532817>
- Bonneau, D. A., & Hutchinson, D. J. (2019). The use of terrestrial laser scanning for the characterization of a cliff-talus system in the Thompson River Valley, British Columbia, Canada. *Geomorphology*, 327, 598–609.
- Boro, H., Rosero, E., & Bertotti, G. (2014). Fracture-network analysis of the Latemar Platform (northern Italy): Integrating outcrop studies to constrain the hydraulic properties of fractures in reservoir models. *Petroleum Geoscience*, 20(1), 79–92. <https://doi.org/10.1144/petgeo2013-007>
- Bouaziz, S., Tagliasacchi, A., & Pauly, M. (2013). Sparse iterative closest point. *Computer Graphics Forum*, 32(5), 113–123.
- Brodu, N., & Lague, D. (2012). 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68, 121–134.
- Brown, D. C. (1976). The Bundle Adjustment-Progress and Prospect Invited Paper to ISP. *Corn. III*.
- Cao, Q., & Wang, S. (2011). Applying over-sampling technique based on data density and cost-sensitive SVM to imbalanced learning. *2011 International Conference on Information Management, Innovation Management and Industrial Engineering*, 2, 543–548.
- Carrea, D., Abellan, A., Derron, M.-H., Gauvin, N., & Jaboyedoff, M. (2021). MATLAB Virtual Toolbox for Retrospective Rockfall Source Detection and Volume Estimation Using 3D Point Clouds: A Case Study of a Subalpine Molasse Cliff. *Geosciences*, 11(2), 75.
- Carrera, J. (1988). State of the art of the inverse problem applied to the flow and solute transport equations. *Groundwater Flow and Quality Modelling*, 549–583.
- Cavailles, T., Sizun, J.-P., Labaume, P., Chauvet, A., Buatier, M., Soliva, R., Mezri, L., Charpentier, D., Leclère, H., Travé, A., & Gout, C. (2013). Influence of fault rock foliation on fault zone permeability: The case of deeply buried arkosic sandstones (Grès d'Annot, southeastern France). *AAPG Bulletin*, 97(9), 1521–1543. <https://doi.org/10.1306/03071312127>
- Chandler, J. (1999). Effective application of automated digital photogrammetry for geomorphological research. *Earth Surface Processes and Landforms*, 24(1), 51–63.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Chesnaux, R., Allen, D. M., & Jenni, S. (2009). Regional fracture network permeability using outcrop scale measurements. *Engineering Geology*, 108(3–4), 259–271. <https://doi.org/10.1016/J.ENGCEO.2009.06.024>
- Cignoni, P., Rocchini, C., & Scopigno, R. (1998). Metro: measuring error on simplified surfaces. *Computer Graphics Forum*, 17(2), 167–174.
- Cohn, D. A. (1996). Active Learning with Statistical Models. *Journal of Artificial Intelligence Research* 4 (1996) 129-145, 3, 868–887.
- Conyers, L. B., & Cameron, C. M. (1998). Ground-penetrating radar techniques and three-dimensional computer mapping in the American Southwest. *Journal of Field Archaeology*,

- 25(4), 417–430.
- Corbett, K., Friedman, M., & Spang, J. (1987). Fracture Development and Mechanical Stratigraphy of Austin Chalk, Texas. *American Association of Petroleum Geologists Bulletin*, 71(1), 17–28. <https://doi.org/10.1306/94886d35-1704-11d7-8645000102c1865d>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27.
- Cruden, D. M. (1977). Describing the size of discontinuities. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 14(3), 133–137. [https://doi.org/10.1016/0148-9062\(77\)90004-3](https://doi.org/10.1016/0148-9062(77)90004-3)
- Dahlin, T., & Zhou, B. (2004). A numerical comparison of 2D resistivity imaging with 10 electrode arrays. *Geophysical Prospecting*, 52(5), 379–398. <https://doi.org/https://doi.org/10.1111/j.1365-2478.2004.00423.x>
- Dale, T. N., & Gregory, H. E. (1911). *The granites of Connecticut* (Issue 484). US Government Printing Office.
- Daura, J., Sanz, M., Fornós, J. J., Asensio, A., & Julià, A. R. (2014). Karst evolution of the Garraf Massif (Barcelona, Spain): Doline formation, chronology and archaeopalaeontological archives. *Journal of Cave and Karst Studies*, 76(2), 69–87. <https://doi.org/10.4311/2011ES0254>
- Deere, D. U., & Deere, D. W. (1989). Rock Quality Designation (RQD) after Twenty Years. US Army Corps of Engineers Report GL-89-1. *Department of The Army - US Army Corps of Engineers*, 100. <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA207597>
- Deere, U. (1969). *ENGINEERING CLASSIFICATION OF IN-SITU ROCK*. May 1960.
- Dershowitz, W. S. i Einstein, H. H. (1988). *Characterizing Tock Joint Geometry with Joint System Models*. 1(2), 12–17.
- Dershowitz, W., & Doe, T. (1997). Analysis of heterogeneously connected rock masses by forward modeling of fractional dimension flow behavior. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 34(3–4), 61.e1. [https://doi.org/10.1016/S1365-1609\(97\)00237-2](https://doi.org/10.1016/S1365-1609(97)00237-2)
- Dershowitz, W., Lapointe, P., & Cladouhos, T. (1998). Derivation of fracture spatial pattern parameters from borehole data. *International Journal of Rock Mechanics and Mining Sciences*, 35(4–5), 508. [https://doi.org/10.1016/s0148-9062\(98\)00149-1](https://doi.org/10.1016/s0148-9062(98)00149-1)
- Dershowitz, W S. (1998). FracMan interactive discrete feature data analysis, geometric modeling and exploration simulation. *User Documentation*.
- Dershowitz, W S, & Einstein, H. H. (1988). Characterizing rock joint geometry with joint system models. *Rock Mechanics and Rock Engineering*, 21(1), 21–51. <https://doi.org/10.1007/BF01019674>
- Dershowitz, William S, & Einstein, H. H. (1984). Application Of Artificial Intelligence To Problems Of Rock Mechanics. In *The 25th U.S. Symposium on Rock Mechanics (USRMS)*.
- Dershowitz, William S, & Einstein, H. H. (1987). Three dimensional flow modeling in jointed rock masses. *6th ISRM Congress*.
- Dershowitz, Ws, Pointe, P. La, & Doe, T. (2004). Advances in discrete fracture network modeling. ... *the US EPA/NGWA Fractured ...*, 882–894. https://www.clu-in.org/products/siteprof/2004fracrockconf/cdr_pdfs/indexed/group1/882.pdf
- Difrancesco, P.-M., Bonneau, D., & Hutchinson, D. J. (2020). *remote sensing The Implications of M3C2 Projection Diameter on 3D Semi-Automated Rockfall Extraction from Sequential Terrestrial Laser Scanning Point Clouds*. <https://doi.org/10.3390/rs12111885>
- Doe, T., McLaren, R., & Dershowitz, W. (2014). Discrete Fracture Network Simulations of Enhanced Geothermal Systems. *Proceedings, 39th Workshop on Geothermal Reservoir Engineering*, 1–11. <https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2014/Doe.pdf>
- Douzas, G., & Bacao, F. (2017a). Geometric SMOTE: Effective oversampling for imbalanced learning through a geometric extension of SMOTE. *ArXiv Preprint ArXiv:1709.07377*.
- Douzas, G., & Bacao, F. (2017b). *Geometric SMOTE: Effective oversampling for imbalanced learning through a geometric extension of SMOTE*. <http://arxiv.org/abs/1709.07377>
- Dramschi, J. S. (2020). 70 years of machine learning in geoscience in review. *Advances in Geophysics*, 61, 1–55. <https://doi.org/10.1016/bs.agph.2020.08.002>
- Dueholm, K. S., & Olsen, T. (1993). Reservoir Analog Studies Using Multimodel Photogrammetry: A New Tool for the Petroleum Industry1. *AAPG Bulletin*, 77(12), 2023–2031. <https://doi.org/10.1306/BDF8FBA-1718-11D7-8645000102C1865D>
- Eberhardt, E., Stead, D., & Coggan, J. S. (2004). Numerical analysis of initiation and progressive failure in natural rock slopes—the 1991 Randa rockslide. *International Journal of Rock*

- Mechanics and Mining Sciences*, 41(1), 69–87.
- Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., Glennie, C. L., Joerg, P. C., LeWinter, A. L., & Magney, T. S. (2016). Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences. *Remote Sensing of Environment*, 186, 372–392.
- El-Naqa, A. (2001). The hydraulic conductivity of the fractures intersecting Cambrian sandstone rock masses, central Jordan. *Environmental Geology*, 40(8), 973–982. <https://doi.org/10.1007/s002540100266>
- Elmo, D., Rogers, S., Stead, D., & Eberhardt, E. (2014). Discrete fracture network approach to characterise rock mass fragmentation and implications for geomechanical upscaling. *Transactions of the Institutions of Mining and Metallurgy, Section A: Mining Technology*, 123(3), 149–161. <https://doi.org/10.1179/1743286314Y.0000000064>
- Engelder, T. (1987). Joints and some fractures in rocks. In *Fracture Mechanics of Rocks* (pp. 27–69).
- Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, 96(34), 226–231.
- Fanos, A. M., Pradhan, B., Alamri, A., & Lee, C.-W. (2020). Machine Learning-Based and 3D Kinematic Models for Rockfall Hazard Assessment Using LiDAR Data and GIS. *Remote Sensing*, 12(11). <https://doi.org/10.3390/rs12111755>
- Fanos, A. M., Pradhan, B., Mansor, S., Yusoff, Z. M., & Abdullah, A. F. bin. (2018). A hybrid model using machine learning methods and GIS for potential rockfall source identification from airborne laser scanning data. *Landslides*, 15(9), 1833–1850. <https://doi.org/10.1007/s10346-018-0990-4>
- Fisher, N. I. (1985). Spherical Medians. *Journal of the Royal Statistical Society: Series B (Methodological)*, 47(2), 342–348. <https://doi.org/10.1111/j.2517-6161.1985.tb01362.x>
- FracMan® User's Manual*. (2020). © 1999-2020 Golder Associates Inc.
- Freeze, R. A., & Cherry, J. A. (1979). *Groundwater*. Prentice-hall.
- Gale, J. F. W., Laubach, S. E., Olson, J. E., Eichhubl, P., & Fall, A. (2017). Natural fractures in shale: A review and new observations. *AAPG Bulletin*, 101(8), 2165–2216. <https://doi.org/10.1306/08121413151>
- García-Sellés, D., Falivene, O., Arbués, P., Gratacos, O., Tavani, S., & Muñoz, J. A. (2011). Supervised identification and reconstruction of near-planar geological surfaces from terrestrial laser scanning. *Computers and Geosciences*, 37(10), 1584–1594. <https://doi.org/10.1016/j.cageo.2011.03.007>
- García-Sellés, David, Sarmiento, S., Gratacós, O., Granado, P., Carrera, N., Lakshmikantha, M. R., Cordova, J. C., & Muñoz, J. A. (2018). *Fracture Analog of the Sub-Andean Devonian of Southern Bolivia: Lidar Applied to Abra Del Condor*.
- García Badal, M. (2018). *Millora metodològica per a la detecció i caracterització de despreniments amb dades de LiDAR Terrestre a la Muntanya de Montserrat*. <http://diposit.ub.edu/dspace/handle/2445/125130#YSYDXignu7g.mendeley>
- Gazzah, S., & Amara, N. E. Ben. (2008). New oversampling approaches based on polynomial fitting for imbalanced data sets. *2008 the Eighth IAPR International Workshop on Document Analysis Systems*, 677–684.
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42.
- Geyer, A., García-Sellés, D., Pedrazzi, D., Barde-Cabusson, S., Marti, J., & Muñoz, J. A. (2015). Studying monogenetic volcanoes with a terrestrial laser scanner: case study at Crosat volcano (Garrotxa Volcanic Field, Spain). *Bulletin of Volcanology*, 77(3), 1–14. <https://doi.org/10.1007/s00445-015-0909-z>
- Gigli, G., & Casagli, N. (2011). Semi-automatic extraction of rock mass structural data from high resolution LIDAR point clouds. *International Journal of Rock Mechanics and Mining Sciences*, 48(2), 187–198. <https://doi.org/10.1016/J.IJRMMS.2010.11.009>
- Gillespie, P. A., Howard, C. B., Walsh, J. J., & Watterson, J. (1993). Measurement and characterisation of spatial distributions of fractures. *Tectonophysics*, 226(1–4), 113–141. [https://doi.org/10.1016/0040-1951\(93\)90114-Y](https://doi.org/10.1016/0040-1951(93)90114-Y)
- Gillespie, P. A., Johnston, J. D., Loriga, M. A., McCaffrey, K. J. W., Walsh, J. J., & Watterson, J. (1999). Influence of layering on vein systematics in line samples. *Geological Society Special Publication*, 155(December), 35–56. <https://doi.org/10.1144/GSL.SP.1999.155.01.05>
- Girardeau-Montaut, D. (2021). *CloudCompare—3d Point Cloud and Mesh Processing Software*

(version 2.10. 1). Windows.

- Girardeau-Montaut, Daniel, Roux, M., Marc, R., & Thibault, G. (2005). Change detection on points cloud data acquired with a ground laser scanner. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(part 3), W19.
- Gómez-Paccard, M., López-Blanco, M., Costa, E., Garcés, M., Beamud, E., & Larrasoaña, J. C. (2012). Tectonic and climatic controls on the sequential arrangement of an alluvial fan/fan-delta complex (Montserrat, Eocene, Ebro Basin, NE Spain). *Basin Research*, 24(4), 437–455. <https://doi.org/10.1111/j.1365-2117.2011.00532.x>
- Granet, S., Fabrie, P., Lemonnier, P., & Quintard, M. (1998). A single-phase flow simulation of fractured reservoir using a discrete representation of fractures. *ECMOR VI-6th European Conference on the Mathematics of Oil Recovery*, cp-102.
- Granier, B., Al Suwaidi, A. S., Busnardo, R., Aziz, S. K., & Schroeder, R. (2003). New insight on the stratigraphy of the “Upper Thamama” in offshore Abu Dhabi (U.A.E.). *Carnets de Géologie, CG2003(A05)*, 1–17. <https://hal.archives-ouvertes.fr/hal-00005742>
- Gross, M. R. (1993). The origin and spacing of cross joints: examples from the Monterey Formation, Santa Barbara Coastline, California. *Journal of Structural Geology*, 15(6), 737–751. [https://doi.org/https://doi.org/10.1016/0191-8141\(93\)90059-J](https://doi.org/https://doi.org/10.1016/0191-8141(93)90059-J)
- Gross, M. R., Fischer, M. P., Engelder, T., & Greenfield, R. J. (1995). Factors controlling joint spacing in interbedded sedimentary rocks: Integrating numerical models with field observations from the Monterey Formation, USA. *Geological Society Special Publication*, 92, 215–233. <https://doi.org/10.1144/GSL.SP.1995.092.01.12>
- Gruen, A., & Akca, D. (2005). Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3), 151–174.
- Guerin, A., Stock, G. M., Radue, M. J., Jaboyedoff, M., Collins, B. D., Matasci, B., Avdievitch, N., & Derron, M.-H. (2020). Quantifying 40 years of rockfall activity in Yosemite Valley with historical Structure-from-Motion photogrammetry and terrestrial laser scanning. *Geomorphology*, 356, 107069.
- Guimerá, J. (1984). Palaeogene evolution of deformation in the northeastern Iberian Peninsula. *Geological Magazine*, 121(5), 413–420. <https://doi.org/10.1017/S0016756800029940>
- Guimerà, J. (1982). *Estudi estructural de les zones de fractura de Garraf i de Vallcarca (Massís de Garraf)* (Institut d'Estudis Catalans (Ed.); 1982nd ed.). Institut d'Estudis Catalans.
- Guo, J., Liu, Y., Wu, L., Liu, S., Yang, T., Zhu, W., & Zhang, Z. (2019). A geometry-and texture-based automatic discontinuity trace extraction method for rock mass point cloud. *International Journal of Rock Mechanics and Mining Sciences*, 124, 104132.
- Guo, L., Lin, H., Fan, B., Nyquist, J., Toran, L., & Mount, G. J. (2019). Preferential flow through shallow fractured bedrock and a 3D fill-and-spill model of hillslope subsurface hydrology. *Journal of Hydrology*, 576(June), 430–442. <https://doi.org/10.1016/j.jhydrol.2019.06.070>
- Hagan, T. O. (1980). A case for terrestrial photogrammetry in deep-mine rock structure studies. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 17(4), 191–198.
- Hancock, P. L. (1985). Brittle microtectonics: principles and practice. *Journal of Structural Geology*, 7(3–4), 437–457. [https://doi.org/10.1016/0191-8141\(85\)90048-3](https://doi.org/10.1016/0191-8141(85)90048-3)
- Hancock, Paul Lewis, & Kadhi, A. (1978). Analysis of mesoscopic fractures in the dhurma-nisah segment of the central arabian graben system. *Journal of the Geological Society*, 135(3), 339–347. <https://doi.org/10.1144/gsjgs.135.3.0339>
- Hardebol, N. J., & Bertotti, G. (2013a). DigiFract: A software and data model implementation for flexible acquisition and processing of fracture data from outcrops. *Computers and Geosciences*, 54, 326–336. <https://doi.org/10.1016/j.cageo.2012.10.021>
- Hardebol, N. J., & Bertotti, G. (2013b). DigiFract: A software and data model implementation for flexible acquisition and processing of fracture data from outcrops. *Computers & Geosciences*, 54, 326–336. <https://doi.org/https://doi.org/10.1016/j.cageo.2012.10.021>
- Hastie, T.; Friedman, J.; Tibshirani, R. (2001). The Elements of Statistical Learning. *Springer New York: New York, NY*.
- Hastie, T., Rosset, S., Zhu, J., & Zou, H. (2009). Multi-class adaboost. *Statistics and Its Interface*, 2(3), 349–360.
- He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 1322–1328. <https://doi.org/10.1109/IJCNN.2008.4633969>
- Healy, D., Jones, R. R., & Holdsworth, R. E. (2006). New insights into the development of brittle

- shear fractures from a 3-D numerical model of microcrack interaction. *Earth and Planetary Science Letters*, 249(1–2), 14–28. <https://doi.org/10.1016/j.epsl.2006.06.041>
- Hemalatha, T., Ramesh, M. V., & Rangan, V. P. (2019). Effective and Accelerated Forewarning of Landslides Using Wireless Sensor Networks and Machine Learning. *IEEE Sensors Journal*, 19(21), 9964–9975. <https://doi.org/10.1109/JSEN.2019.2928358>
- Hendrickx, H., Le Roy, G., Helmstetter, A., Pointner, E., Larose, E., Braillard, L., Nyssen, J., Delaloye, R., & Frankl, A. (2022). Timing, volume and precursory indicators of rock-and cliff fall on a permafrost mountain ridge (Mattertal, Switzerland). *Earth Surface Processes and Landforms*.
- Hodgson, R. A. (1961). Classification of structures on joint surfaces. In *American Journal of Science* (Vol. 259, Issue 7, pp. 493–502). <https://doi.org/10.2475/ajs.259.7.493>
- Hoek, E. (1964). Fracture of anisotropic rock. *Journal of the Southern African Institute of Mining and Metallurgy*, 64(10), 501–518.
- Hoek, E. (1981). Rock slope engineering: Civil and mining, 4th edition. *Rock Slope Engineering: Fourth Edition*, 1–432. <https://doi.org/10.1201/9781315274980>
- Hoek, E., & Bieniawski, Z. T. (1965). Brittle fracture propagation in rock under compression. *International Journal of Fracture Mechanics*, 1(3), 137–155. <https://doi.org/10.1007/BF00186851>
- Hoek, E., & Brown, E. T. (1997). Practical estimates of rock mass strength. *International Journal of Rock Mechanics and Mining Sciences*, 34(8), 1165–1186. [https://doi.org/10.1016/S1365-1609\(97\)80069-X](https://doi.org/10.1016/S1365-1609(97)80069-X)
- Hooker, J. N., Gale, J. F. W., Gomez, L. A., Laubach, S. E., Marrett, R., & Reed, R. M. (2009). Aperture-size scaling variations in a low-strain opening-mode fracture set, Cozzette Sandstone, Colorado. *Journal of Structural Geology*, 31(7), 707–718. <https://doi.org/10.1016/j.jsg.2009.04.001>
- Hooker, J. N., Laubach, S. E., & Marrett, R. (2013). Fracture-aperture sizedfrequency, spatial distribution, and growth processes in strata-bounded and non-strata-bounded fractures, cambrian mesón group, NW argentina. *Journal of Structural Geology*, 54, 54–71. <https://doi.org/10.1016/j.jsg.2013.06.011>
- Hoteit, H., & Firoozabadi, A. (2005). Multicomponent fluid flow by discontinuous Galerkin and mixed methods in unfractured and fractured media. *Water Resources Research*, 41(11).
- Huang, N., Jiang, Y., Liu, R., & Li, B. (2017). Estimation of permeability of 3-D discrete fracture networks: An alternative possibility based on trace map analysis. *Engineering Geology*, 226(April), 12–19. <https://doi.org/10.1016/j.enggeo.2017.05.005>
- Hudson, J. A. (2013). *A Review of Rock Engineering Systems (RES) Applications over the last 20 Years*.
- Hudson, J. A., & Priest, S. D. (1979). Discontinuities and rock mass geometry. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 16(6), 339–362. [https://doi.org/10.1016/0148-9062\(79\)90001-9](https://doi.org/10.1016/0148-9062(79)90001-9)
- Hwang, S. G., Guevarra, I. F., & Yu, B. O. (2009). Slope failure prediction using a decision tree: A case of engineered slopes in South Korea. *Engineering Geology*, 104(1–2), 126–134. <https://doi.org/10.1016/J.ENGCEO.2008.09.004>
- ISRM, I. (1978). Suggested methods for the quantitative description of discontinuities in rock masses. *Int J Rock Mech Min Sci Geomech Abstr*, 15(6), 319–368.
- Jaboyedoff, M., Baillifard, F., Philipposian, F., & Rouiller, J.-D. (2004). Assessing fracture occurrence using “weighted fracturing density”: a step towards estimating rock instability hazard. *Natural Hazards and Earth System Science*, 4(1), 83–93. <https://doi.org/10.5194/nhess-4-83-2004>
- Jaboyedoff, Michel, Oppikofer, T., Minoia, R., Locat, J., & Turmel, D. (2008). Terrestrial LIDAR investigation of the 2004 rockslide along Petit Champlain street, Québec City (Québec, Canada). *4th Canadian Conference on Geohazards: From Causes to Management*, 295–301.
- Jambayev, A. S. (2013). Discrete Fracture Network Modeling for a Carbonate Reservoir. *Doctoral Dissertation, Colorado School of Mines. Arthur Lakes Library*, 1–87. <https://doi.org/10.1016/j.jcjo.2015.09.011>
- Janeras, M., Jara, J. A., Royán, M. J., Vilaplana, J. M., Aguasca, A., Fàbregas, X., Gili, J. A., & Buxó, P. (2017a). Multi-technique approach to rockfall monitoring in the Montserrat massif (Catalonia, NE Spain). *Engineering Geology*, 219, 4–20. <https://doi.org/10.1016/j.enggeo.2016.12.010>

- Janeras, M., Jara, J. A., Royán, M. J., Vilaplana, J. M., Aguasca, A., Fàbregas, X., Gili, J. A., & Buxó, P. (2017b). Multi-technique approach to rockfall monitoring in the Montserrat massif (Catalonia, NE Spain). *Engineering Geology*, 219, 4–20. <https://doi.org/10.1016/J.ENGCEO.2016.12.010>
- Jeannin, M., Garambois, S., Grégoire, C., & Jongmans, D. (2006). Multiconfiguration GPR measurements for geometric fracture characterization in limestone cliffs (Alps). *Geophysics*, 71(3), B85–B92.
- Jiang, X. W., Wan, L., Wang, X. S., Wu, X., & Zhang, X. (2009). Estimation of rock mass deformation modulus using variations in transmissivity and RQD with depth. *International Journal of Rock Mechanics and Mining Sciences*, 46(8), 1370–1377. <https://doi.org/10.1016/j.ijrmms.2009.05.004>
- Jolliffe, I. (2011). *Principal Component Analysis BT - International Encyclopedia of Statistical Science* (M. Lovric (Ed.); pp. 1094–1096). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_455
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2019). Machine Learning for the Geosciences: Challenges and Opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8), 1544–1554. <https://doi.org/10.1109/TKDE.2018.2861006>
- Kazhdan, M., Bolitho, M., & Hoppe, H. (2006). Poisson surface reconstruction. *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, 7.
- Kemeny, J., & Donovan, J. (2005). *Rock mass characterisation using LIDAR and automated point cloud processing*.
- Kemeny, J., & Post, R. (2003). Estimating three-dimensional rock discontinuity orientation from digital images of fracture traces. *Computers & Geosciences*, 29(1), 65–77.
- Kenefick, J. F. (1977). Applications of photogrammetry in shipbuilding. *Photogrammetric Engineering and Remote Sensing*, 43(9), 1169–1175.
- Kong, D., Wu, F., & Saroglou, C. (2020). Automatic identification and characterization of discontinuities in rock masses from 3D point clouds. *Engineering Geology*, 265, 105442. <https://doi.org/10.1016/J.ENGCEO.2019.105442>
- Kovács, G. (2019). An empirical comparison and evaluation of minority oversampling techniques on a large number of imbalanced datasets. *Applied Soft Computing*, 83, 105662. <https://doi.org/10.1016/J.ASOC.2019.105662>
- Kovin, O. (2011). Mapping of evaporite deformation in a potash mine using ground penetrating radar: Upper Kama deposit, Russia. *Journal of Applied Geophysics*, 74(2–3), 131–141.
- Kromer, R. A., Hutchinson, D. J., Lato, M. J., Gauthier, D., & Edwards, T. (2015). Identifying rock slope failure precursors using LiDAR for transportation corridor hazard management. *Engineering Geology*, 195, 93–103.
- Ku, C. Y., Hsu, S. M., Chiou, L. Bin, & Lin, G. F. (2009). An empirical model for estimating hydraulic conductivity of highly disturbed clastic sedimentary rocks in Taiwan. *Engineering Geology*, 109(3–4), 213–223. <https://doi.org/10.1016/J.ENGCEO.2009.08.008>
- Lague, D., Brodu, N., & Leroux, J. (2013a). Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 10–26. <https://doi.org/10.1016/j.isprsjprs.2013.04.009>
- Lague, D., Brodu, N., & Leroux, J. (2013b). Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 10–26. <https://doi.org/10.1016/J.ISPRSJPRS.2013.04.009>
- Lamarche, J., Lavenue, A. P. C., Gauthier, B. D. M., Guglielmi, Y., & Jayet, O. (2012). Relationships between fracture patterns, geodynamics and mechanical stratigraphy in Carbonates (South-East Basin, France). *Tectonophysics*, 581(December), 231–245. <https://doi.org/10.1016/j.tecto.2012.06.042>
- Landen, D. (1959). Impact of the Development of Photogrammetry upon Geology. *Journal of the Washington Academy of Sciences*, 49(7), 234–252.
- Lane, S. N., James, T. D., & Crowell, M. D. (2000). Application of digital photogrammetry to complex topography for geomorphological research. *The Photogrammetric Record*, 16(95), 793–821.
- Larsen, B., Grunnaleite, I., & Gudmundsson, A. (2010). How fracture systems affect permeability

- development in shallow-water carbonate rocks: An example from the Gargano Peninsula, Italy. *Journal of Structural Geology*, 32(9), 1212–1230. <https://doi.org/10.1016/j.jsg.2009.05.009>
- Lato, M. J., & Vöge, M. (2012). Automated mapping of rock discontinuities in 3D lidar and photogrammetry models. *International Journal of Rock Mechanics and Mining Sciences*, 54, 150–158. <https://doi.org/10.1016/j.ijrmms.2012.06.003>
- Laubach, S E, & Gale, J. F. W. (2006). OBTAINING FRACTURE INFORMATION FOR LOW-PERMEABILITY (TIGHT) GAS SANDSTONES FROM SIDEWALL CORES. *Journal of Petroleum Geology*, 29(2), 147–158. <https://doi.org/https://doi.org/10.1111/j.1747-5457.2006.00147.x>
- Laubach, Stephen E., Olson, J. E., & Cross, M. R. (2009). Mechanical and fracture stratigraphy. *AAPG Bulletin*, 93(11), 1413–1426. <https://doi.org/10.1306/07270909094>
- Lee, J., Kim, N., & Lee, J.-H. (2015). An over-sampling technique with rejection for imbalanced class learning. *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*, 1–6.
- Li, M., Han, S., Zhou, S., & Zhang, Y. (2018). An Improved Computing Method for 3D Mechanical Connectivity Rates Based on a Polyhedral Simulation Model of Discrete Fracture Network in Rock Masses. *Rock Mechanics and Rock Engineering*, 51(6), 1789–1800. <https://doi.org/10.1007/s00603-018-1423-3>
- Li, X., Chen, Z., Chen, J., & Zhu, H. (2019). Automatic characterization of rock mass discontinuities using 3D point clouds. *Engineering Geology*, 259, 105131.
- Loke, M. H., & Barker, R. D. (1996). Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method1. *Geophysical Prospecting*, 44(1), 131–152.
- Long, J. C. S., Remer, J. S., Wilson, C. R., & Witherspoon, P. A. (1982). Porous media equivalents for networks of discontinuous fractures. *Water Resources Research*, 18(3), 645–658. <https://doi.org/https://doi.org/10.1029/WR018i003p00645>
- Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828), 133–135.
- López-blanco, M. (2007). Long-term changes in extreme temperatures and precipitation in Spain. *Long-Term Changes in Extreme Temperatures and Precipitation in Spain*, 3(3), 331–342. <https://doi.org/10.2436/20.7010.01.1>
- López-Blanco, M., Gómez-Paccard, M., Costa, E., Garcés, M., Beamud, E., & Larrasoaña, J. C. (2010). Tectonic Versus Climatic Controls on Basin-margin Alluvialfan/fan Delta Complex Development, Montserrat, NE Spain. *72nd EAGE Conference and Exhibition Incorporating SPE EUROPEC 2010*, cp-161.
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2, 1150–1157.
- Manda, A. K., & Horsman, E. (2015). Fracturesis Jointitis: Causes, Symptoms, and Treatment in Groundwater Communities. *Groundwater*, 53(6), 836–840. <https://doi.org/https://doi.org/10.1111/gwat.12301>
- Marchesini, P., Grasmueck, M., Eberli, G. P., & Weger, R. J. (2013). 4D GPR for characterization of fluid Flow in carbonates: insights from structural-vs. stratigraphiccontrolled domains and comparison with eclipse dynamic modeling: American Association of Petroleum Geologists. *Search and Discovery Article*, 41179, 22.
- Marinos, P., Marinos, V., & Hoek, E. (2007). Geological Strength Index (GSI). A characterization tool for assessing engineering properties for rock masses. *Underground Works under Special Conditions - Proceedings of the Workshop (W1) on Underground Works under Special Conditions*, 13–21. <https://doi.org/10.1201/noe0415450287.ch2>
- Martinez, A., & Byrnes, A. P. (2001). Modeling dielectric-constant values of geologic materials: An aid to ground-penetrating radar data collection and interpretation. *Current Research in Earth Sciences*, 1–16.
- Massiot, C., Townend, J., Nicol, A., & McNamara, D. D. (2017). Statistical methods of fracture characterization using acoustic borehole televiewer log interpretation. *Journal of Geophysical Research: Solid Earth*, 122(8), 6836–6852. <https://doi.org/https://doi.org/10.1002/2017JB014115>
- Mauldon, M., & Mauldon, J. G. (1997). Fracture sampling on a cylinder: From scanlines to boreholes and tunnels. *Rock Mechanics and Rock Engineering*, 30(3), 129–144. <https://doi.org/10.1007/BF01047389>
- Mazzanti, P., Caporossi, P., Brunetti, A., Mohammadi, F. I., & Bozzano, F. (2021). Short-term

- geomorphological evolution of the Poggio Baldi landslide upper scarp via 3D change detection. *Landslides*, 18(7), 2367–2381.
- Mertz, J.-D. (1991). *Structures de porosité et propriétés de transport dans les grès* (Vol. 90, Issue 1). Persée-Portail des revues scientifiques en SHS.
- Micheletti, N., Foresti, L., Robert, S., Leuenberger, M., Pedrazzini, A., Jaboyedoff, M., & Kanevski, M. (2014). Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping. *Mathematical Geosciences*, 46(1), 33–57. <https://doi.org/10.1007/s11004-013-9511-0>
- Miller, R. P. (1965). *Engineering classification and index properties for intact rock*. University of Illinois at Urbana-Champaign.
- Mineo, S., Pappalardo, G., Mangiameli, M., Campolo, S., & Mussumeci, G. (2018). Rockfall analysis for preliminary hazard assessment of the cliff of Taormina Saracen Castle (Sicily). *Sustainability*, 10(2), 417.
- Monicard, R. (1982). Principal Methods of Evaluation. In *Drilling Mud and Cement Slurry Rheology Manual: Publication de la Chambre Syndicale de la Recherche et de la Production du Pétrole et du Gaz Naturel* (pp. 53–93). Springer Netherlands. https://doi.org/10.1007/978-94-010-9246-3_4
- Mook, W. G. (2002). *Isótopos ambientales en el ciclo hidrológico* (Vol. 1). IGME.
- Murtagh, F. (1991). Multilayer perceptrons for classification and regression. *Neurocomputing*, 2(5–6), 183–197.
- Nakamura, M., Kajiwara, Y., Otsuka, A., & Kimura, H. (2013). Lvq-smote-learning vector quantization based synthetic minority over-sampling technique for biomedical data. *BioData Mining*, 6(1), 1–10.
- Narr, W., & Suppe, J. (1991). Joint spacing in sedimentary rocks. *Journal of Structural Geology*, 13(9), 1037–1048. [https://doi.org/10.1016/0191-8141\(91\)90055-N](https://doi.org/10.1016/0191-8141(91)90055-N)
- Nascimento Da Silva, C. C., De Medeiros, W. E., De Sá, E. F. J., & Neto, P. X. (2004). Resistivity and ground-penetrating radar images of fractures in a crystalline aquifer: A case study in Caiçara farm - NE Brazil. *Journal of Applied Geophysics*, 56(4), 295–307. <https://doi.org/10.1016/j.jappgeo.2004.08.001>
- Neagu, T., & Cîrnaru, P. (2004). Genus Choffatella (SCHLUMBERGER), 1904 (Foraminifera) in the lower Aptian (Bedoulian) from Southern Dobrogea and SE part of the Romanian plain. *Acta Palaeontologica Romaniaae*, 4, 269–275.
- Neuman, S. P. (2005). Trends, prospects and challenges in quantifying flow and transport through fractured rocks. *Hydrogeology Journal*, 13(1), 124–147. <https://doi.org/10.1007/s10040-004-0397-2>
- Nicholson, G. A., & Bieniawski, Z. T. (1990). A nonlinear deformation modulus based on rock mass classification. *International Journal of Mining and Geological Engineering*, 8(3), 181–202. <https://doi.org/10.1007/BF01554041>
- Nilsson, N. J. (1965). *Learning machines*.
- Odling, N E, Gillespie, P., Bourguin, B., Castaing, C., Chiles, J. P., Christensen, N. P., Fillion, E., Genter, A., Olsen, C., Thrane, L., Trice, R., Aarseth, E., Walsh, J. J., & Watterson, J. (1999). Variations in fracture system geometry and their implications for fluid flow in fractures hydrocarbon reservoirs. *Petroleum Geoscience*, 5(4), 373 LP – 384. <https://doi.org/10.1144/petgeo.5.4.373>
- Odling, Noelle E. (1997). Scaling and connectivity of joint systems in sandstones from western Norway. *Journal of Structural Geology*, 19(10), 1257–1271. [https://doi.org/10.1016/S0191-8141\(97\)00041-2](https://doi.org/10.1016/S0191-8141(97)00041-2)
- Ogata, K., Storti, F., Balsamo, F., Tinterri, R., Bedogni, E., Fetter, M., Gomes, L., & Hatushika, R. (2017). Sedimentary facies control on mechanical and fracture stratigraphy in turbidites. *Bulletin of the Geological Society of America*, 129(1–2), 76–92. <https://doi.org/10.1130/B31517.1>
- Öge, I. F. (2017). Assessing Rock Mass Permeability Using Discontinuity Properties. *Procedia Engineering*, 191, 638–645. <https://doi.org/10.1016/j.proeng.2017.05.373>
- Olariu, M. I., Ferguson, J. F., Aiken, C. L. V., & Xu, X. (2008). Outcrop fracture characterization using terrestrial laser scanners: Deep-water Jackfork sandstone at big rock quarry, Arkansas. *Geosphere*, 4(1), 247–259. <https://doi.org/10.1130/GES00139.1>
- Olson, J. E. (2004). Predicting fracture swarms - The influence of subcritical crack growth and the crack-tip process on joint spacing in rock. *Geological Society Special Publication*, 231(May 2007), 73–87. <https://doi.org/10.1144/GSL.SP.2004.231.01.05>
- Olson, J. E., Laubach, S. E., & Lander, R. H. (2007). Combining diagenesis and mechanics to

- quantity fracture aperture distributions and fracture pattern permeability. *Geological Society Special Publication*, 270(June), 101–116. <https://doi.org/10.1144/GSL.SP.2007.270.01.08>
- Olson, J. E., Laubach, S. E., & Lander, R. H. (2009). Natural fracture characterization in tight gas sandstones: Integrating mechanics and diagenesis. *AAPG Bulletin*, 93(11), 1535–1549. <https://doi.org/10.1306/08110909100>
- Optech, T. (2014). ILRIS summary specification sheet. *Teledyne Optech Incorporated: Vaughan, ON, Canada*.
- Ortega, O. J., Marrett, R. A., & Laubach, S. E. (2006). A scale-independent approach to fracture intensity and average spacing measurement. *AAPG Bulletin*, 90(2), 193–208. <https://doi.org/10.1306/08250505059>
- Pahl, P. J. (1981). Estimating the mean length of discontinuity traces. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 18(3), 221–228. [https://doi.org/10.1016/0148-9062\(81\)90976-1](https://doi.org/10.1016/0148-9062(81)90976-1)
- PARKER III, J. M. (1942). Regional systematic jointing in slightly deformed sedimentary rocks. *Bulletin of the Geological Society of America*, 53(3), 381–408.
- Peacock, D. C. P., Harris, S. D., & Mauldon, M. (2003). Use of curved scanlines and boreholes to predict fracture frequencies. *Journal of Structural Geology*, 25(1), 109–119. [https://doi.org/10.1016/S0191-8141\(02\)00016-0](https://doi.org/10.1016/S0191-8141(02)00016-0)
- Peacock, D. C. P., & Mann, A. (2005). Evaluation of the controls on fracturing in reservoir rocks. *Journal of Petroleum Geology*, 28(4), 385–396. <https://doi.org/10.1111/j.1747-5457.2005.tb00089.x>
- Peacock, D. C. P., Nixon, C. W., Rotevatn, A., Sanderson, D. J., & Zuluaga, L. F. (2016). Glossary of fault and other fracture networks. *Journal of Structural Geology*, 92, 12–29. <https://doi.org/10.1016/j.jsg.2016.09.008>
- Peacock, D. C. P., & Sanderson, D. J. (2018). Structural analyses and fracture network characterisation: Seven pillars of wisdom. *Earth-Science Reviews*, 184(June), 13–28. <https://doi.org/10.1016/j.earscirev.2018.06.006>
- Pedregosa, F. (2011). Scikit-learn: Machine Learning in Python Fabian. *Journal Of Machine Learning Research* 12 (2011) 2825-2830, 125(9), 2825–2830. <https://doi.org/10.1289/EHP4713>
- Persson, P.-A. (1994). Voladura de Rocas e Ingeniería de los Explosivos (Rock Blasting and Explosives Engineering). *Estados Unidos de América*.
- Petit, J.-P., Massonat, G., Pueo, F., & Rawnsley, K. (1994). Rapport de forme des fractures de mode 1 dans les roches stratifiées: une Étude de cas dans le bassin permien de Lodève (France). *Bulletin Des Centres de Recherches Exploration-Production Elf-Aquitaine*, 18(1), 211–229.
- Picart, J., Cirés, J., Samsó, J., Salas, R., Segura, X., Péres, R., Solé, X., Galindo, J., Carmona, R., & Casanovas, J. (2016). *Vallirana, full 420-1-2 (71-32). Mapa geològic de Catalunya 1:25.000. 1ª Edició. Institut Cartogràfic i Geològic de Catalunya 1pp.*
- Pincus, H. J. (1951). Statistical Methods Applied to the Study of Rock Fractures: Quantitative Comparative Analysis of Fractures in Gneisses and Overlying Sedimentary Rocks of Northern New Jersey. *Geological Society of America Bulletin*, 62(2), 81–130.
- Pittet, B., Van Buchem, F. S. P., Hillgärtner, H., Razin, P., Grötsch, J., & Droste, H. (2002). Ecological succession, palaeoenvironmental change, and depositional sequences of Barremian–Aptian shallow-water carbonates in northern Oman. *Sedimentology*, 49(3), 555–581. <https://doi.org/https://doi.org/10.1046/j.1365-3091.2002.00460.x>
- POLLARD, D. D., & AYDIN, A. (1988). Progress in understanding jointing over the past century. *GSA Bulletin*, 100(8), 1181–1204. [https://doi.org/10.1130/0016-7606\(1988\)100<1181:PIUJOT>2.3.CO;2](https://doi.org/10.1130/0016-7606(1988)100<1181:PIUJOT>2.3.CO;2)
- Prakash, N., Manconi, A., & Loew, S. (2020). Mapping Landslides on EO Data: Performance of Deep Learning Models vs. Traditional Machine Learning Models. *Remote Sensing*, 12(3). <https://doi.org/10.3390/rs12030346>
- Price, N. J. (1966). Brittle fracture. *Fault and Joint Development in Brittle and Semi-Brittle Rock*, 2, 1–56.
- Priest, S. D. (1993). The Collection and Analysis of Discontinuity Orientation Data for Engineering Design, with Examples. *Comprehensive Rock Engineering*. Vol. 3, 167–192. <https://doi.org/10.1016/B978-0-08-042066-0.50015-X>
- Priest, S. D., & Hudson, J. A. (1976). Discontinuity spacings in rock. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 13(5), 135–148.

- [https://doi.org/10.1016/0148-9062\(76\)90818-4](https://doi.org/10.1016/0148-9062(76)90818-4)
- Priest, S. D., & Hudson, J. A. (1981). Estimation of discontinuity spacing and trace length using scanline surveys. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 18(3), 183–197. [https://doi.org/10.1016/0148-9062\(81\)90973-6](https://doi.org/10.1016/0148-9062(81)90973-6)
- Pruess, K., & Narasimhan, T. N. (1985). A practical method for modeling fluid and heat flow in fractured porous media. *Society of Petroleum Engineers Journal*, 25(01), 14–26.
- Qureshi, M. U. et al. (2014). *Qureshi, 2014.pdf*.
- Raggatt, H. G. (1954). Markings on joint surfaces in Anglesea member of Demon's Bluff formation, Anglesea, Victoria. *AAPG Bulletin*, 38(8), 1808–1810.
- Rasmussen, L. L., Cacciari, P. P., Futai, M. M., de Farias, M. M., & de Assis, A. P. (2019). Efficient 3D probabilistic stability analysis of rock tunnels using a Lattice Model and cloud computing. *Tunnelling and Underground Space Technology*, 85(December 2018), 282–293. <https://doi.org/10.1016/j.tust.2018.12.022>
- Rasolofosaon, P., Lucet, N., & Zinszner, B. (2008). Petroacoustics of carbonate reservoir rocks. *The Leading Edge*, 27(8), 1034–1039.
- Reid, T. R., & Harrison, J. P. (2000). A semi-automated methodology for discontinuity trace detection in digital images of rock mass exposures. *International Journal of Rock Mechanics and Mining Sciences*, 37(7), 1073–1089.
- Requena Mateu, G. (2010). *Caracterització de les zones de sortida de desprendiments de la paret de degotalls (Montserrat)*. https://upcommons.upc.edu/handle/2099.1/12578#YSYGA6_ga08.mendeley
- Robinson, P. C. (1983). Connectivity of fracture systems—a percolation theory approach. *Journal of Physics A: Mathematical and General*, 16(3), 605–614. <https://doi.org/10.1088/0305-4470/16/3/020>
- Rogers, S., Elmo, D., Webb, G., & Catalan, A. (2014). Volumetric Fracture Intensity Measurement for Improved Rock Mass Characterisation and Fragmentation Assessment in Block Caving Operations. *Rock Mechanics and Rock Engineering*, 48(2), 633–649. <https://doi.org/10.1007/s00603-014-0592-y>
- Rogers, Steve, Elmo, D., Dunphy, R., & Bearinger, D. (2010). *Understanding Hydraulic Fracture Geometry and Interactions in the Horn River Basin through DFN and Numerical Modeling*. <https://doi.org/10.2118/137488-MS>
- Romana, M., Tomás, R., & Serón, J. B. (2015). *Slope Mass Rating (SMR) Geomechanics Classification: Thirty Years Review. All Days*.
- Romana, Manuel. (1991). *SMR Classification. All Days*.
- Romano-Perez, C. A., & Diaz-Viera, M. A. (2012). A comparison of discrete fracture models for single phase flow in porous media by COMSOL Multiphysics® Software. *Proceedings of the 2015 COMSOL Conference in Boston, Boston, MA, USA*, 7–9.
- Roncella, R., Scaioni, M., & Forlani, G. (2004). Application of digital photogrammetry in geotechnics. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXth ISPRS Congress, Commission, 5*, 93–98.
- Rossen, W. R., Gu, Y., & Lake, L. W. (2000). *Connectivity and Permeability in Fracture Networks Obeying Power-Law Statistics. All Days*. <https://doi.org/10.2118/59720-MS>
- Rosser, N., Lim, M., Petley, D., Dunning, S., & Allison, R. (2007). Patterns of precursory rockfall prior to slope failure. *Journal of Geophysical Research: Earth Surface*, 112(F4). <https://doi.org/https://doi.org/10.1029/2006JF000642>
- Royan, M. (2015). *Rockfall Characterization and Prediction by Means of Terrestrial LiDAR*. Universitat de Barcelona Barcelona, Spain.
- Royán, M. J., Abellán, A., Jaboyedoff, M., Vilaplana, J. M., & Calvet, J. (2014). Spatio-temporal analysis of rockfall pre-failure deformation using Terrestrial LiDAR. *Landslides*, 11(4), 697–709. <https://doi.org/10.1007/s10346-013-0442-0>
- Royan, M. J., Vilaplana, J. M., Janeras, M., & Abellán, A. (2016). Detección e inventario de desprendimientos de rocas mediante el seguimiento con LiDAR Terrestre en la Montaña de Montserrat (Catalunya, España). *XIV Reunión Nacional de Geomorfología*, 5.
- Rushing, J. A., Newsham, K. E., Lasswell, P. M., Cox, J. C., & Blasingame, T. A. (2004). Klinkenberg-corrected permeability measurements in tight gas sands: steady-state versus unsteady-state techniques. *SPE Annual Technical Conference and Exhibition*.
- Sáez, J. A., Luengo, J., Stefanowski, J., & Herrera, F. (2015). SMOTE–IPF: Addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering. *Information Sciences*, 291(C), 184–203. <https://doi.org/10.1016/J.INS.2014.08.051>
- Saleem, Z. A. (1970). A Computer Method for Pumping-Test Analysis. *Groundwater*, 8(5), 21–24.

- <https://doi.org/https://doi.org/10.1111/j.1745-6584.1970.tb01318.x>
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>
- Sanchez. (2009). *Estudi Geològic de la cantera en explotació “La Fou” (no publicat)*.
- Sanderson, D. J., & Peacock, D. C. P. (2019). Line sampling of fracture swarms and corridors. *Journal of Structural Geology*, 122, 27–37. <https://doi.org/https://doi.org/10.1016/j.jsg.2019.02.006>
- Santana, D., Corominas, J., Mavrouli, O., & Garcia-Sellés, D. (2012). Magnitude-frequency relation for rockfall scars using a Terrestrial Laser Scanner. *Engineering Geology*, 144–145, 50–64. <https://doi.org/10.1016/j.enggeo.2012.07.001>
- Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197–227.
- Schovanec, H., Walton, G., Kromer, R., & Malsam, A. (2021). Development of improved semi-automated processing algorithms for the creation of rockfall databases. *Remote Sensing*, 13(8), 1479.
- Schroeder, R., van Buchem, F. S. P., Cherchi, A., Baghbani, D., Vincent, B., Immenhauser, A., & Granier, B. (2010). Revised orbitolinid biostratigraphic zonation for the Barremian--Aptian of the eastern Arabian Plate and implications for regional stratigraphic correlations. *GeoArabia Special Publication*, 4(1), 49–96.
- Seers, T. D., & Hodgetts, D. (2014). Comparison of digital outcrop and conventional data collection approaches for the characterization of naturally fractured reservoir analogues. *Geological Society, London, Special Publications*, 374(1), 51–77.
- Shackleton, J. R., Cooke, M. L., & Sussman, A. J. (2005). Evidence for temporally changing mechanical stratigraphy and effects on joint-network architecture. *Geology*, 33(2), 101–104. <https://doi.org/10.1130/G20930.1>
- Sharma, S., Bellinger, C., Krawczyk, B., Zaiane, O., & Japkowicz, N. (2018). Synthetic Oversampling with the Majority Class: A New Perspective on Handling Extreme Imbalance. *2018 IEEE International Conference on Data Mining (ICDM)*, 447–456. <https://doi.org/10.1109/ICDM.2018.00060>
- Sharp, G. C., Lee, S. W., & Wehe, D. K. (2002). ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1), 90–102. <https://doi.org/10.1109/34.982886>
- Sheldon, P. (1912). Some observations and experiments on joint planes. *The Journal of Geology*, 20(1), 53–79.
- Singhal, B. B. S., & Gupta, R. P. (2010). Applied hydrogeology of fractured rocks: Second edition. *Applied Hydrogeology of Fractured Rocks: Second Edition*, 1–408. <https://doi.org/10.1007/978-90-481-8799-7>
- Slob, S., Van Knapen, B., Hack, R., Turner, K., & Kemeny, J. (2005). Method for automated discontinuity analysis of rock slopes with three-dimensional laser scanning. *Transportation Research Record*, 1913, 187–194. <https://doi.org/10.3141/1913-18>
- Snavely, N., Seitz, S. M., & Szeliski, R. (2006). Photo tourism: exploring photo collections in 3D. In *ACM siggraph 2006 papers* (pp. 835–846).
- Snow, D. T. (1970). The frequency and apertures of fractures in rock. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 7(1), 23–40. [https://doi.org/10.1016/0148-9062\(70\)90025-2](https://doi.org/10.1016/0148-9062(70)90025-2)
- Sørensen, K. (1996). Pulled array continuous electrical profiling. *First Break*, 14(3).
- Stefanowski, J. (2013). *Overlapping, Rare Examples and Class Decomposition in Learning Classifiers from Imbalanced Data BT - Emerging Paradigms in Machine Learning* (S. Ramanna, L. C. Jain, & R. J. Howlett (Eds.); pp. 277–306). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28699-5_11
- Studio, V. (2019). URL: <https://visualstudio.microsoft.com/ru/vs>. Date Accessed, 6, 2021.
- Sturzenegger, M., & Stead, D. (2009). Close-range terrestrial digital photogrammetry and terrestrial laser scanning for discontinuity characterization on rock cuts. *Engineering Geology*, 106(3–4), 163–182. <https://doi.org/10.1016/J.ENGCEO.2009.03.004>
- Telling, J., Lyda, A., Hartzell, P., & Glennie, C. (2017). Review of Earth science research using terrestrial laser scanning. *Earth-Science Reviews*, 169(April), 35–68. <https://doi.org/10.1016/j.earscirev.2017.04.007>
- Terzaghi, R.D. (1965). *Sources of error in joiny surveys.pdf*.
- Terzaghi, Ruth D. (1965). Sources of Error in Joint Surveys. *Géotechnique*, 15(3), 287–304. <https://doi.org/10.1680/geot.1965.15.3.287>

- Theune, U., Schmitt, D., & Sacchi, M. (2005). *Looking inside Turtle Mountain : Mapping fractures with GPR. September*, 34–37.
- Tonini, M., & Abellan, A. (2014). Rockfall detection from terrestrial LiDAR point clouds: A clustering approach using R. *Journal of Spatial Information Science*, 8, 95–110.
- Torabi, S. R., Ataei, M., & Javanshir, M. (2011). Application of Schmidt rebound number for estimating rock strength under specific geological conditions. *Journal of Mining and Environment*, 2.
- Tsang, C.-F. (1999). LINKING THERMAL, HYDROLOGICAL, AND MECHANICAL PROCESSES IN FRACTURED ROCKS. *Annual Review of Earth and Planetary Sciences*, 27(1), 359–384. <https://doi.org/10.1146/annurev.earth.27.1.359>
- Tuia, D., Rattle, F., Pacifici, F., Kanevski, M. F., & Emery, W. J. (2009). Active Learning Methods for Remote Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7), 2218–2232. <https://doi.org/10.1109/TGRS.2008.2010404>
- Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 203(1153), 405–426.
- Umili, G., Ferrero, A., & Einstein, H. H. (2013). A new method for automatic discontinuity traces sampling on rock mass 3D model. *Computers & Geosciences*, 51, 182–192. <https://doi.org/10.1016/J.CAGEO.2012.07.026>
- Underwood, C. A., Cooke, M. L., Simo, J. A., Muldoon, M. A., Underwood, C. A., & Cooke, M. L. (2003). *Wisconsin*. 1(1), 121–142.
- van Veen, M., Hutchinson, D. J., Kromer, R., Lato, M., & Edwards, T. (2017). Effects of sampling interval on the frequency-magnitude relationship of rockfalls detected from terrestrial laser scanning using semi-automated methods. *Landslides*, 14(5), 1579–1592.
- Vatsavai, R. R., Shekhar, S., & Burk, T. E. (2005). A semi-supervised learning method for remote sensing data mining. *17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05)*, 5 pp. – 211. <https://doi.org/10.1109/ICTAI.2005.17>
- Vega, J. M., Solé, L., Obrador, A., & Marqués, M. A. (1973). *Hospitalet de Llobregat, hoja nº420. Mapa geológico de España 1:50.000. 2ª Serie. 1ª Edición. Servicio de Publicaciones, Ministerio de Industria y Energía, Madrid, 57pp.*
- Vilajosana, I., Suriñach, E., Abellán, A., Khazaradze, G., Garcia, D., & Llosa, J. (2008). Rockfall induced seismic signals: case study in Montserrat, Catalonia. *Natural Hazards and Earth System Sciences*, 8(4), 805–812. <https://doi.org/10.5194/nhess-8-805-2008>
- Vilas, L., Masse, J. P., & Arias, C. (1995). Orbitolina episodes in carbonate platform evolution: the early Aptian model from SE Spain. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 119(1–2), 35–45.
- Vlachopoulos, N., Vazaios, I., Forbes, B., & Carrapatoso, C. (2020). Rock Mass Structural Characterization Through DFN–LiDAR–DOS Methodology. *Geotechnical and Geological Engineering*, 38(6), 6231–6244. <https://doi.org/10.1007/s10706-020-01431-1>
- Voeckler, H., & Allen, D. M. (2012). Estimating regional-scale fractured bedrock hydraulic conductivity using discrete fracture network (DFN) modeling. *Hydrogeology Journal*, 20(6), 1081–1100. <https://doi.org/10.1007/s10040-012-0858-y>
- Volatili, T., Zambrano, M., Cilona, A., Huisman, B. A. H., Rustichelli, A., Giorgioni, M., Vittori, S., & Tondi, E. (2019). From fracture analysis to flow simulations in fractured carbonates: The case study of the Roman Valley Quarry (Majella Mountain, Italy). *Marine and Petroleum Geology*, 100(October 2018), 95–110. <https://doi.org/10.1016/j.marpetgeo.2018.10.040>
- Wang, Y., Xiao, J., Liu, L., & Wang, Y. (2021). Efficient rock mass point cloud registration based on local invariants. *Remote Sensing*, 13(8), 1540.
- Warren, J. E., & Root, P. J. (1963). The behavior of naturally fractured reservoirs. *Society of Petroleum Engineers Journal*, 3(03), 245–255.
- Watkins, H., Bond, C. E., Healy, D., & Butler, R. W. H. (2015). Appraisal of fracture sampling methods and a new workflow to characterise heterogeneous fracture networks at outcrop. *Journal of Structural Geology*, 72, 67–82. <https://doi.org/10.1016/J.JSG.2015.02.001>
- Wehr, A., & Lohr, U. (1999). Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2–3), 68–82.
- Weidner, L., Walton, G., & Kromer, R. (2019). Classification methods for point clouds in rock slope monitoring: A novel machine learning approach and comparative analysis. *Engineering Geology*, 263, 105326. <https://doi.org/10.1016/J.ENGGE.2019.105326>
- Weidner, L., Walton, G., & Kromer, R. (2020). *Automated Rock Slope Material Classification Using Machine Learning*.
- Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012).

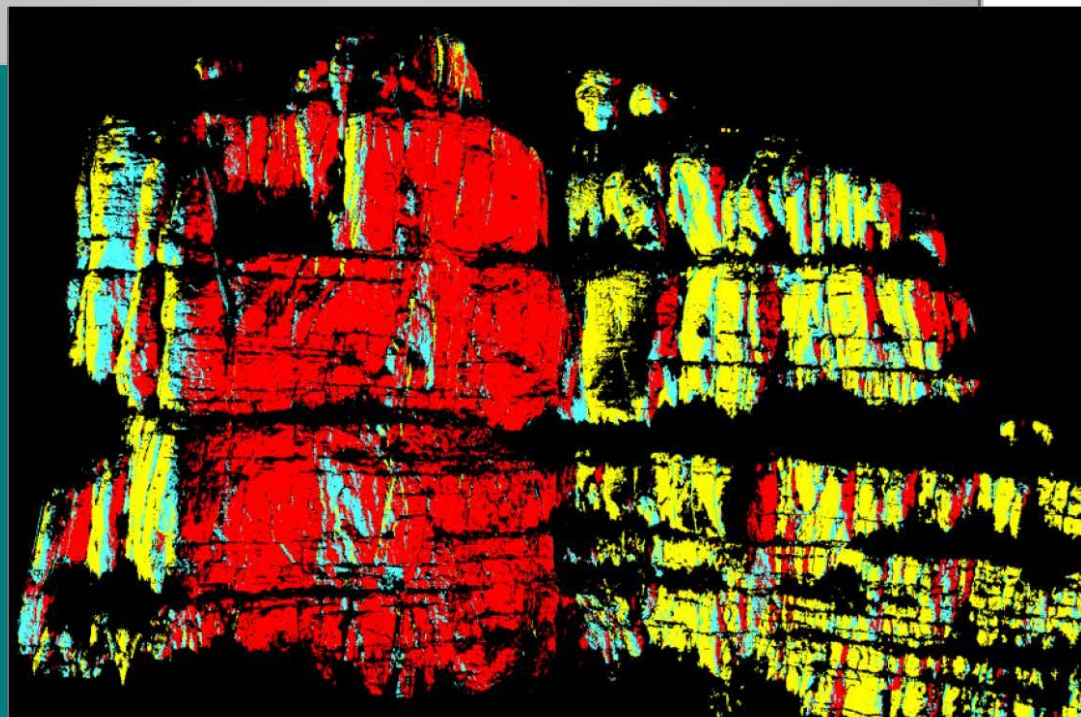
- 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314.
- Wilkinson, M. W., Jones, R. R., Woods, C. E., Gilment, S. R., McCaffrey, K. J. W., Kokkalas, S., & Long, J. J. (2016). A comparison of terrestrial laser scanning and structure-from-motion photogrammetry as methods for digital outcrop acquisition. *Geosphere*, 12(6), 1865–1880. <https://doi.org/10.1130/GES01342.1>
- Williams, J. G., Rosser, N. J., Hardy, R. J., Brain, M. J., & Afana, A. A. (2018). Optimising 4-D surface change detection: an approach for capturing rockfall magnitude–frequency. *Earth Surface Dynamics*, 6(1), 101–119. <https://doi.org/10.5194/esurf-6-101-2018>
- Williams, Jack G., Rosser, N. J., Hardy, R. J., & Brain, M. J. (2019). The importance of monitoring interval for rockfall magnitude-frequency estimation. *Journal of Geophysical Research: Earth Surface*, 124(12), 2841–2853.
- Wilson, C. E., Aydin, A., Karimi-Fard, M., Durllofsky, L. J., Sagy, A., Brodsky, E. E., Kreylos, O., & Kellogg, L. H. (2011). From outcrop to flow simulation: Constructing discrete fracture models from a LIDAR survey. *AAPG Bulletin*, 95(11), 1883–1906. <https://doi.org/10.1306/03241108148>
- Wines, D. R., & Lilly, P. A. (2002). Measurement and analysis of rock mass discontinuity spacing and frequency in part of the Fimiston Open Pit operation in Kalgoorlie, Western Australia: A case study. *International Journal of Rock Mechanics and Mining Sciences*, 39(5), 589–602. [https://doi.org/10.1016/S1365-1609\(02\)00003-5](https://doi.org/10.1016/S1365-1609(02)00003-5)
- Wolf, P. R. (1983). *Elements of photogrammetry: With air photo interpretation and remote sensing*.
- Wong, K. W. (1980). Basic mathematics of photogrammetry. *Manual of Photogrammetry*, 4, 37–101.
- Woodcock, N. H. (1977). Specification of fabric shapes using an eigenvalue method. *Geological Society of America Bulletin*, 88(9), 1231–1236.
- Woodworth, J. B. (1895). Some features of joints. *Science*, 903–904.
- Xiao, L., Zhang, Y., & Peng, G. (2018). Landslide Susceptibility Assessment Using Integrated Deep Learning Algorithm along the China-Nepal Highway. *Sensors*, 18(12). <https://doi.org/10.3390/s18124436>
- Yen, S. J., & Lee, Y. S. (2009). Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications*, 36(3), 5718–5727. <https://doi.org/10.1016/J.ESWA.2008.06.108>
- Zeeb, C., Gomez-Rivas, E., Bons, P. D., & Blum, P. (2013). Evaluation of sampling methods for fracture network characterization using outcrops. *AAPG Bulletin*, 97(9), 1545–1566. <https://doi.org/10.1306/02131312042>
- Zeeb, C., Gomez-Rivas, E., Bons, P. D., Virgo, S., & Blum, P. (2013). Fracture network evaluation program (FraNEP): A software for analyzing 2D fracture trace-line maps. *Computers & Geosciences*, 60, 11–22. <https://doi.org/https://doi.org/10.1016/j.cageo.2013.04.027>
- Zhang, H., Liu, S. guang, Han, Z., Zheng, L., Zhang, Y. bin, Wu, Y. qiang, Li, Y. ge, & Wang, W. (2016). A new algorithm to identify contact types between arbitrarily shaped polyhedral blocks for three-dimensional discontinuous deformation analysis. *Computers and Geotechnics*, 80, 1–15. <https://doi.org/10.1016/J.COMP GEO.2016.06.007>
- Zhang, P., Li, J., Yang, X., & Zhu, H. (2018). Semi-automatic extraction of rock discontinuities from point clouds using the ISODATA clustering algorithm and deviation from mean elevation. *International Journal of Rock Mechanics and Mining Sciences*, 110, 76–87.
- Zhou, B., Yang, C., Guo, H., & Hu, J. (2013). A quasi-linear SVM combined with assembled SMOTE for imbalanced data classification. *The 2013 International Joint Conference on Neural Networks (IJCNN)*, 1–7.
- Zhou, Q.-Y., Park, J., & Koltun, V. (2018). Open3D: A modern library for 3D data processing. *ArXiv Preprint ArXiv:1801.09847*.
- Zoumpikas, T., Puig, A., Salamó, M., García-Sellés, D., Blanco Nuñez, L., & Guinau, M. (2021). An intelligent framework for end-to-end rockfall detection. *International Journal of Intelligent Systems*, May. <https://doi.org/10.1002/int.22557>

APÈNDIXS

A.1. Propietats geomètriques de les fractures a l'àrea d'estudi La Fou

A.2. Article 1: Machine Learning-Based Rockfalls Detection with 3D Point Clouds

A.3. Article 2: An intelligent framework for end-to-end rockfall detection

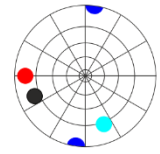
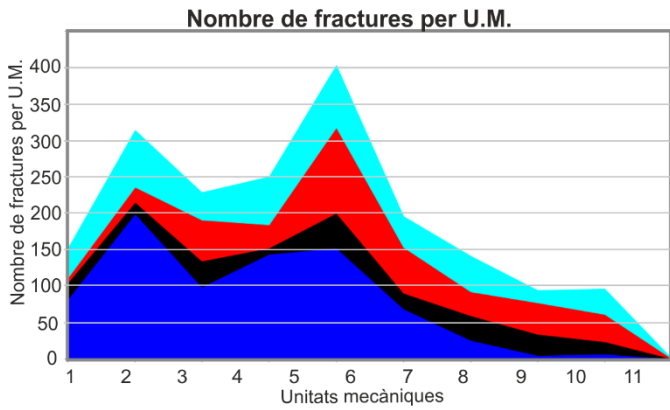


A.1. Propietats geomètriques de les fractures a l'àrea d'estudi La Fou

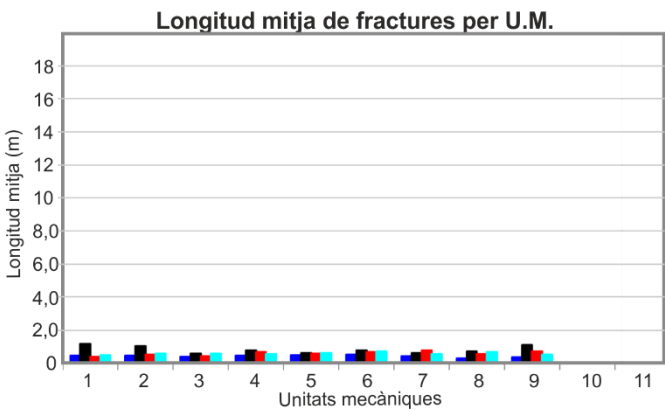
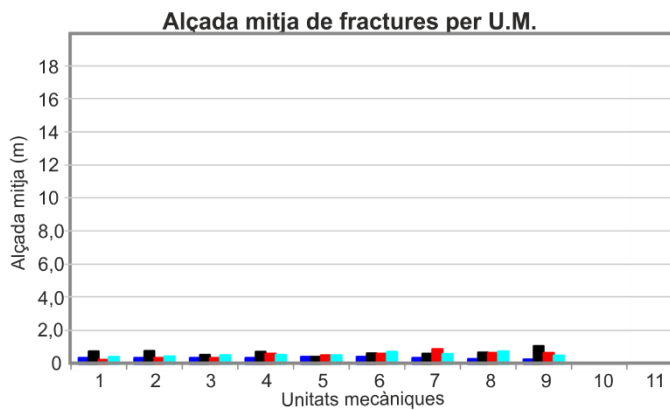
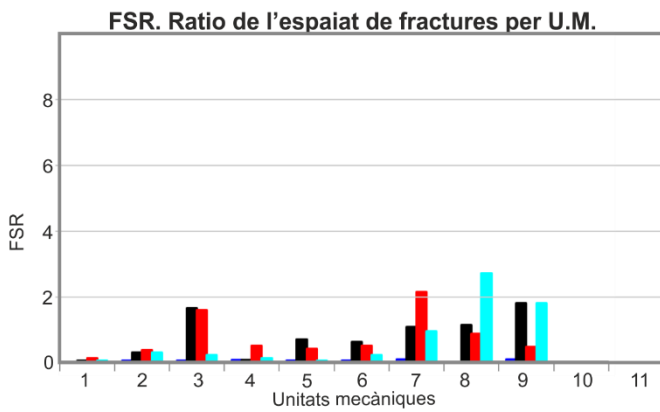
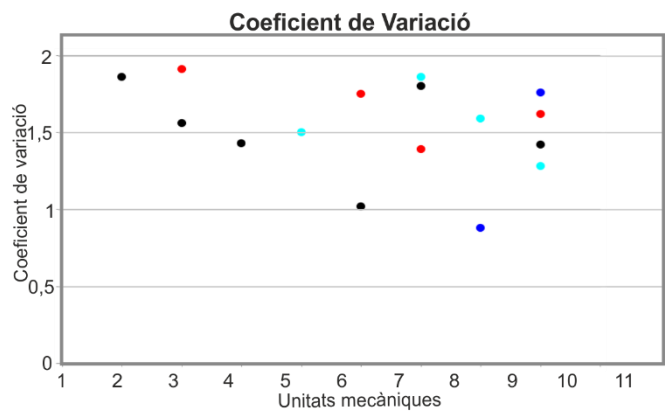
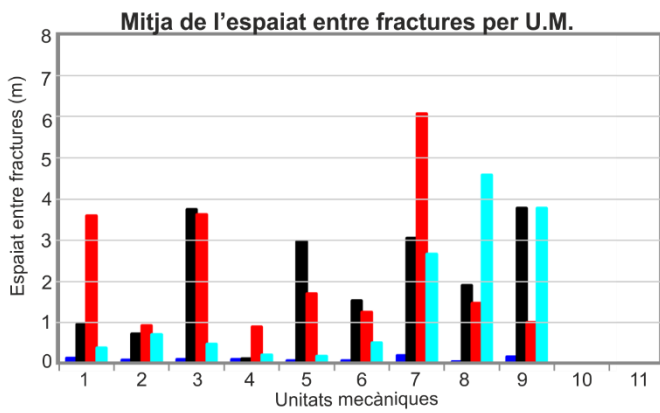
Afloraments fracturats digitalitzats.
Avaluació de les tècniques remotes en models *DFN* i aplicació de *Machine Learning*.

Gràfics de les propietats obtingut amb TLS

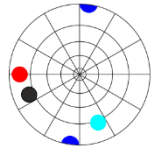
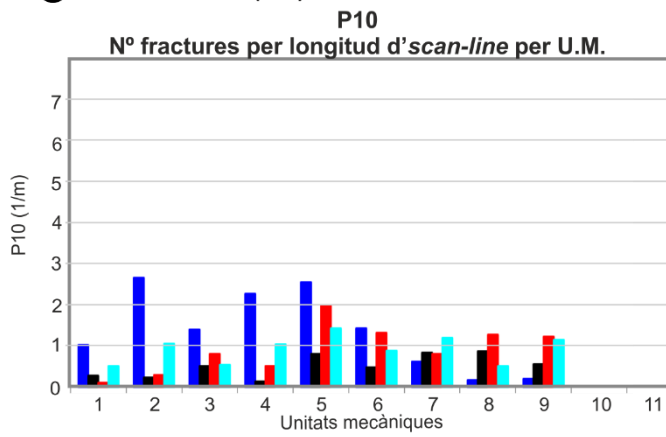
A TLS - Frontal 1 (EW)



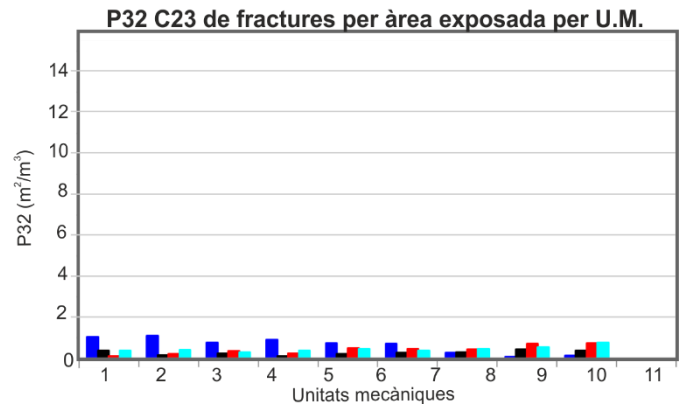
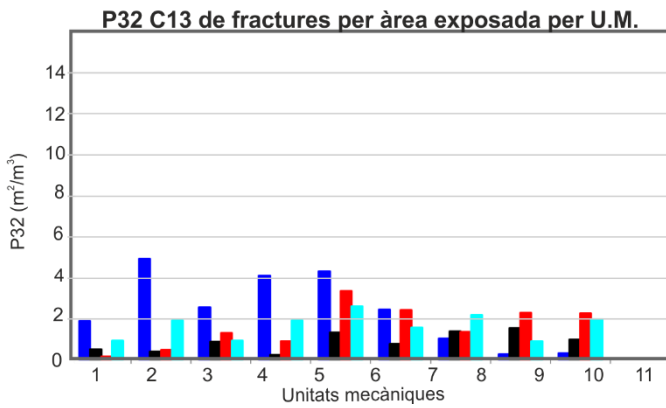
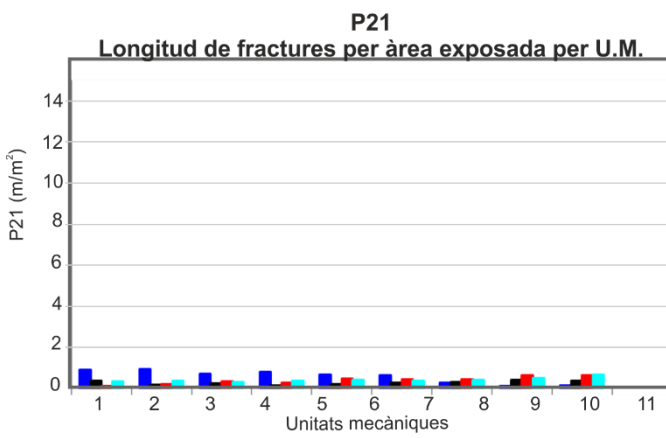
- Família I
- Família II
- Família III
- Família IV



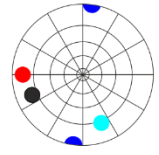
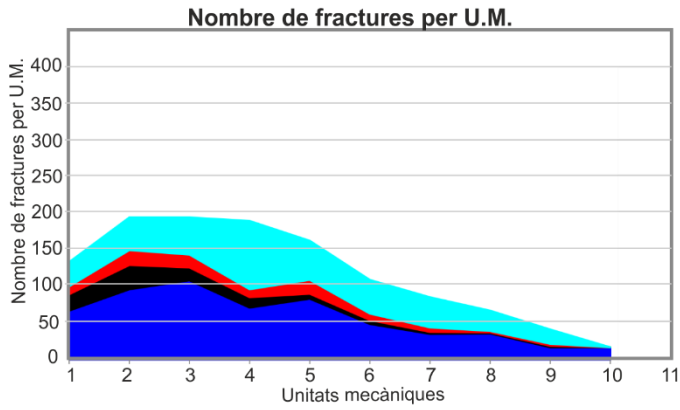
A TLS - Frontal 1 (EW)



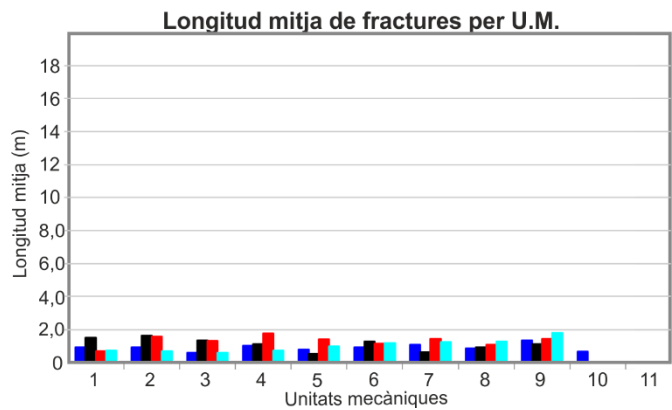
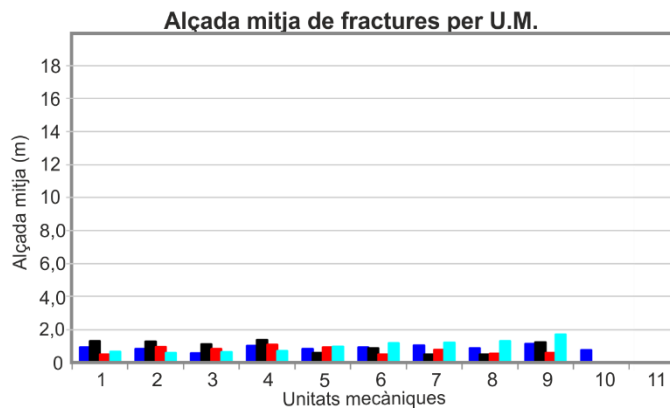
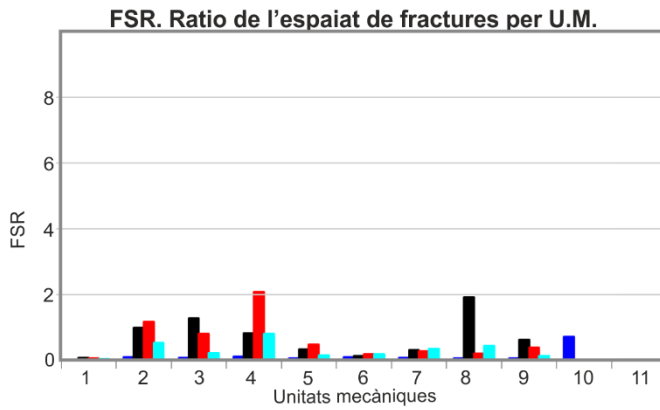
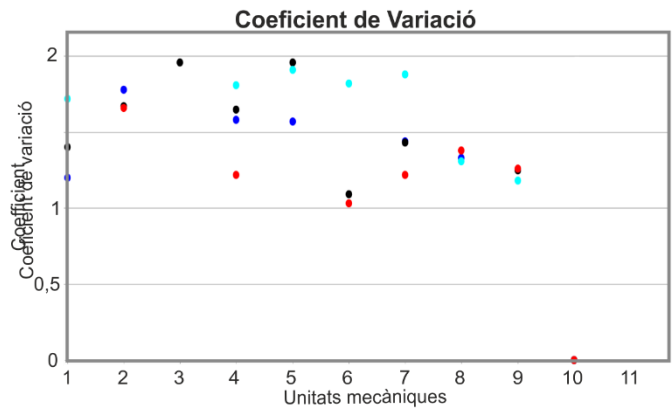
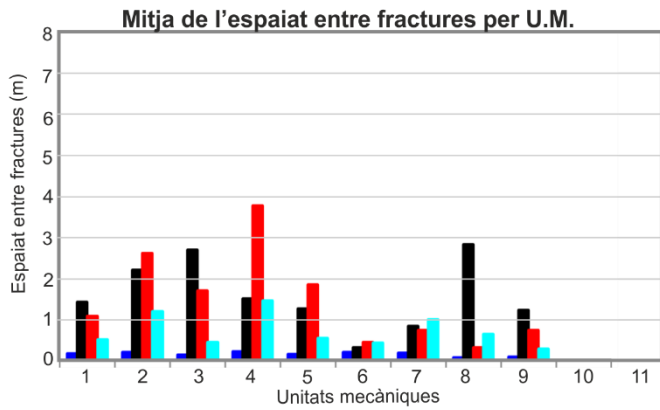
- Família I
- Família II
- Família III
- Família IV



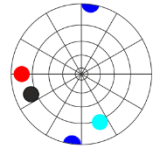
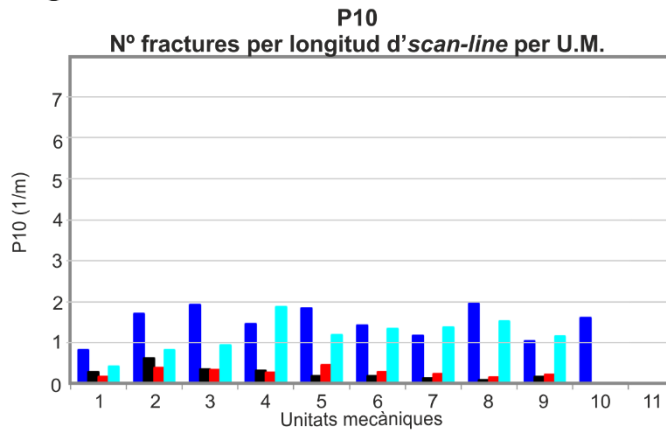
A TLS - Frontal 2 (EW)



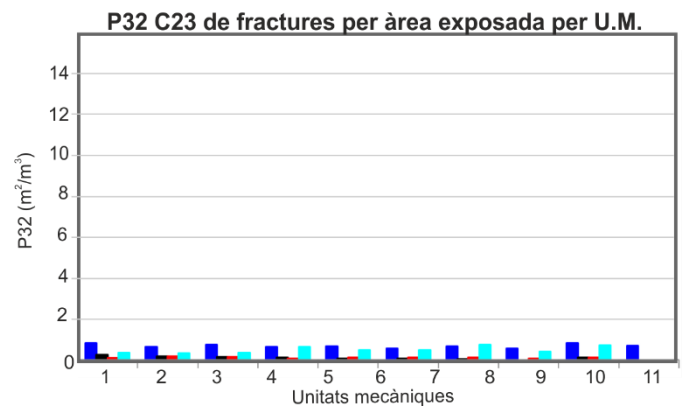
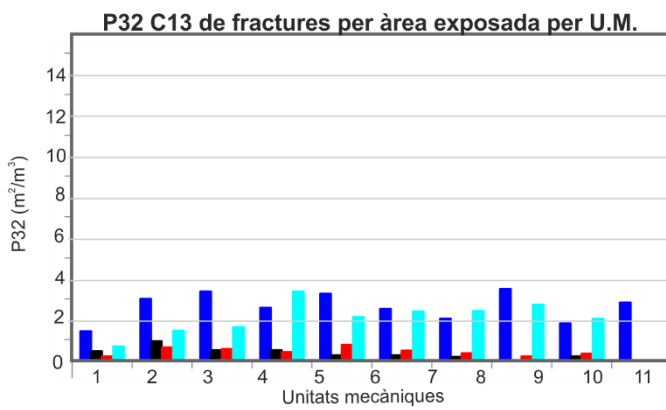
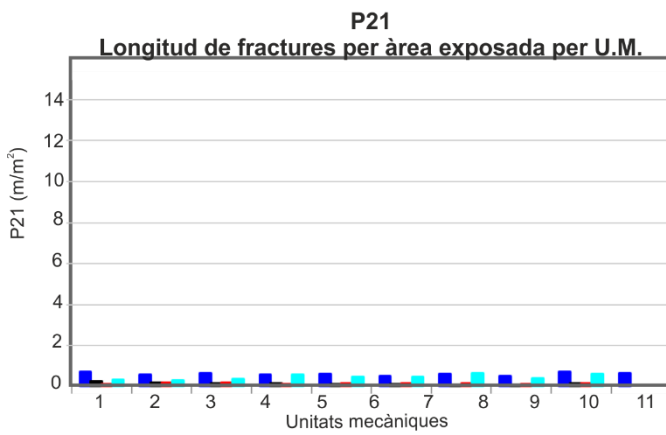
- Família I
- Família II
- Família III
- Família IV



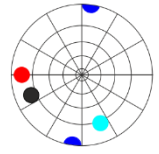
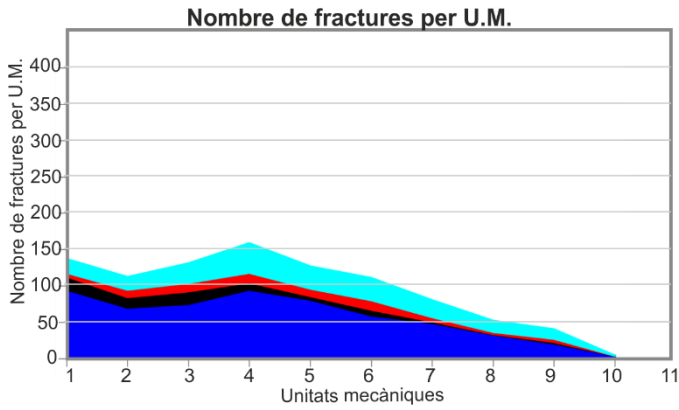
A TLS - Frontal 2 (EW)



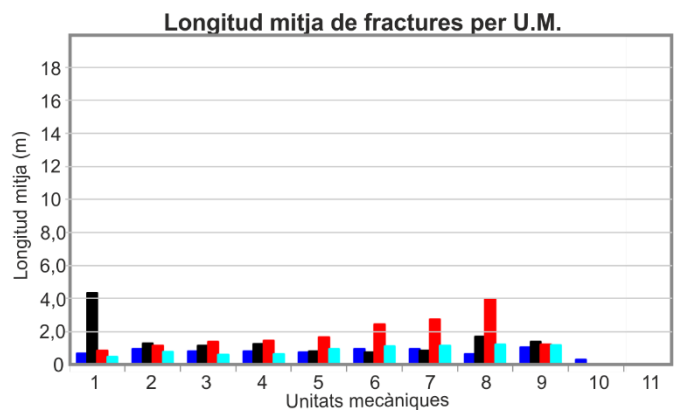
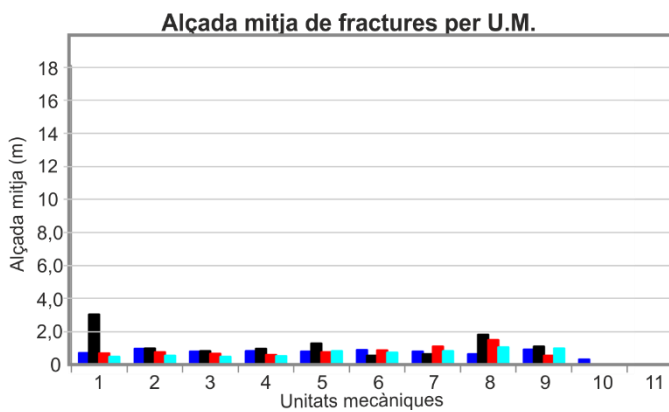
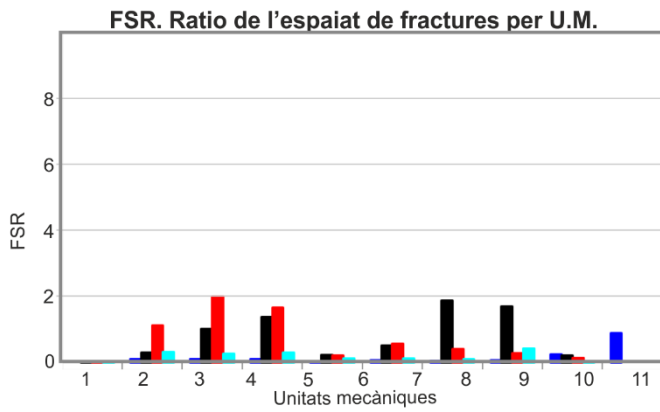
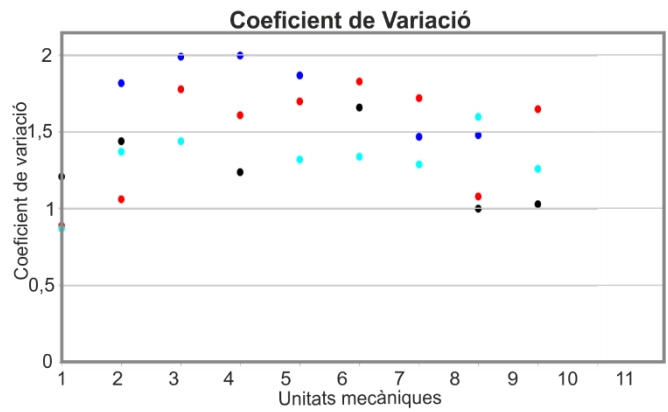
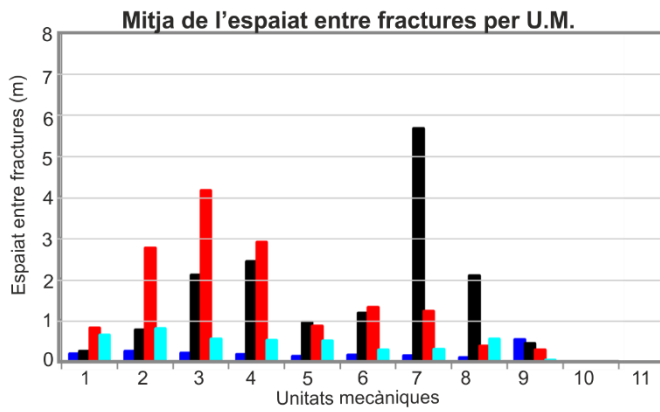
- Família I
- Família II
- Família III
- Família IV



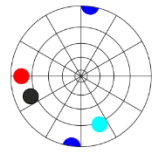
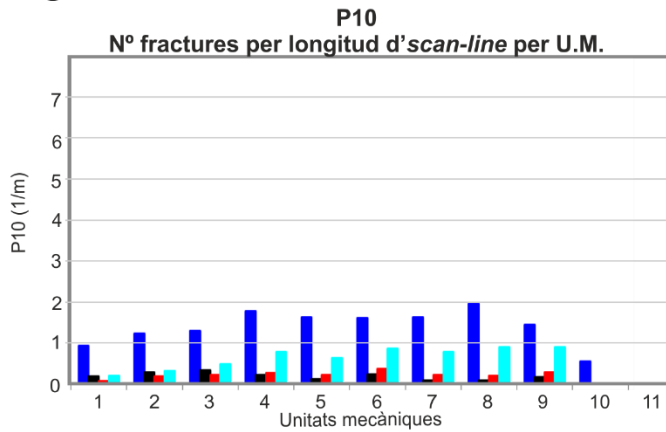
A TLS - Frontal 3 (EW)



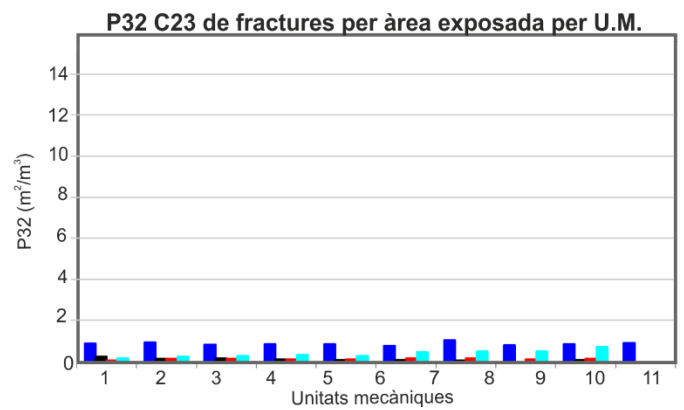
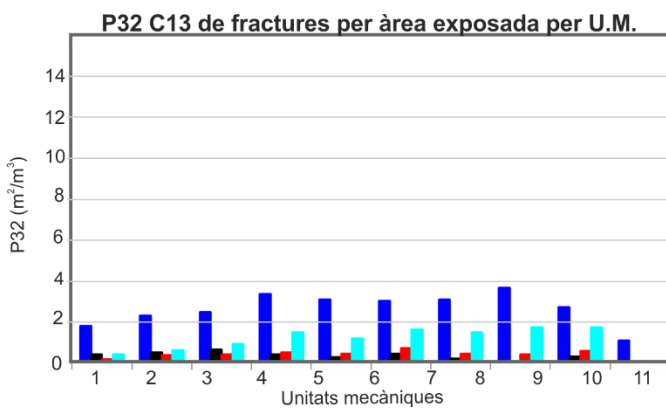
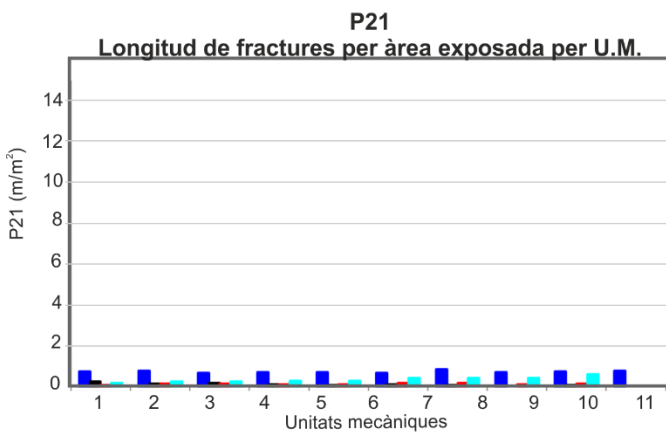
- Família I
- Família II
- Família III
- Família IV



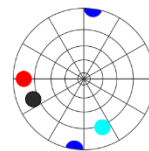
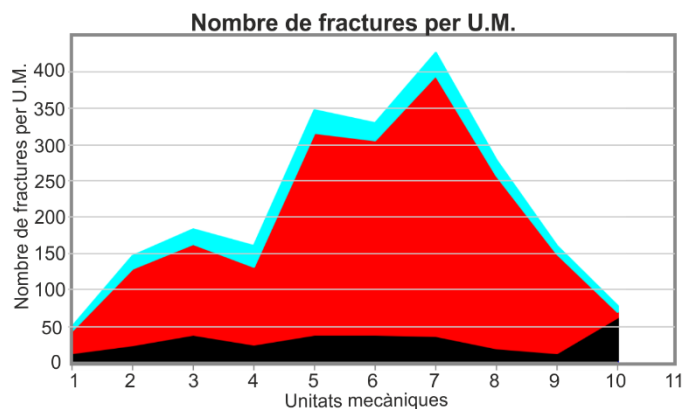
A TLS - Frontal 3 (EW)



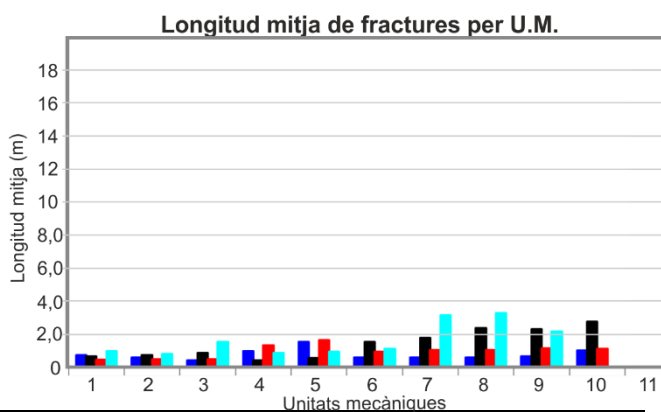
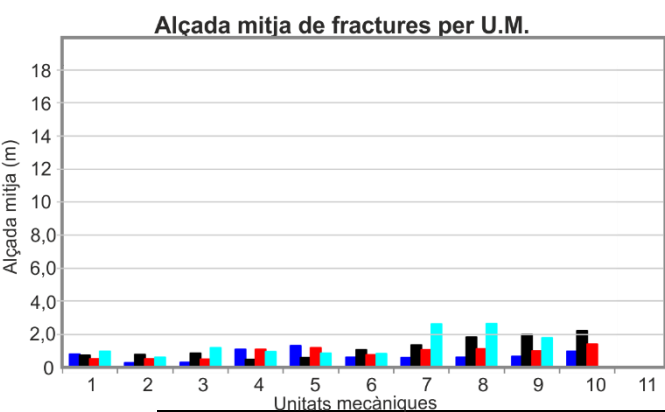
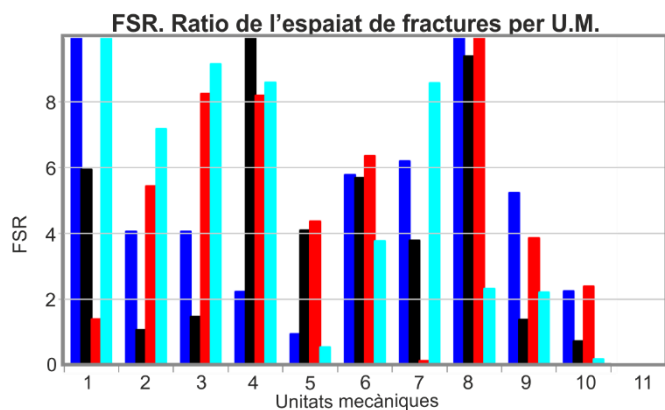
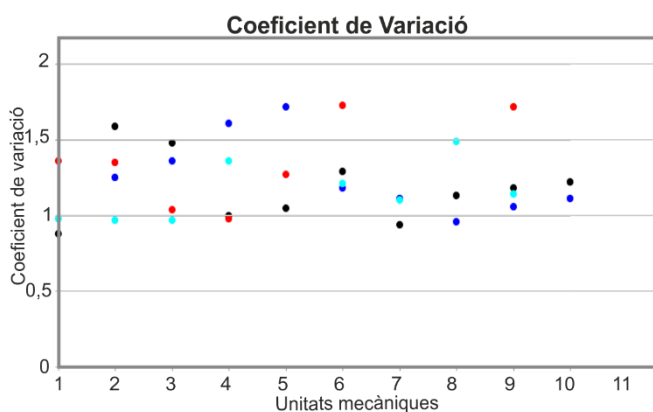
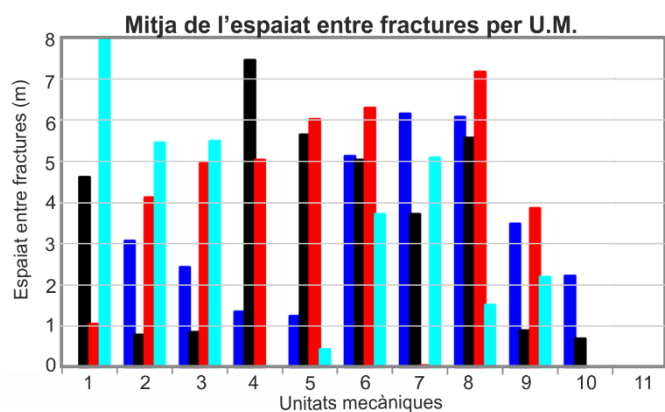
- Família I
- Família II
- Família III
- Família IV



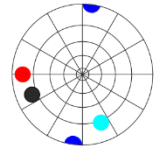
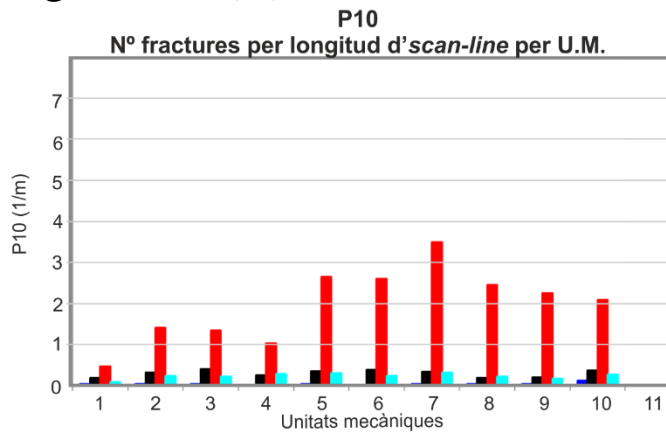
A TLS - Lateral (NS)



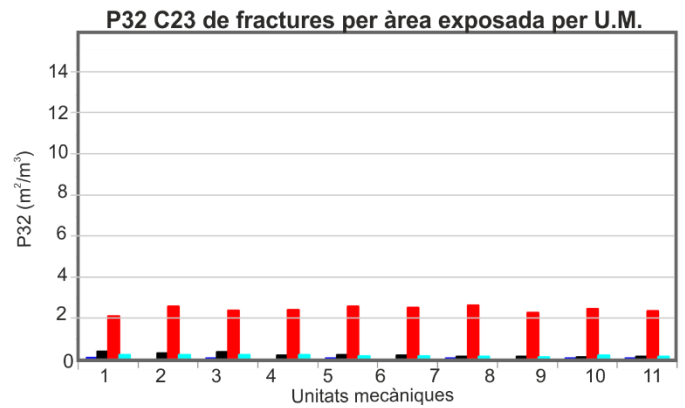
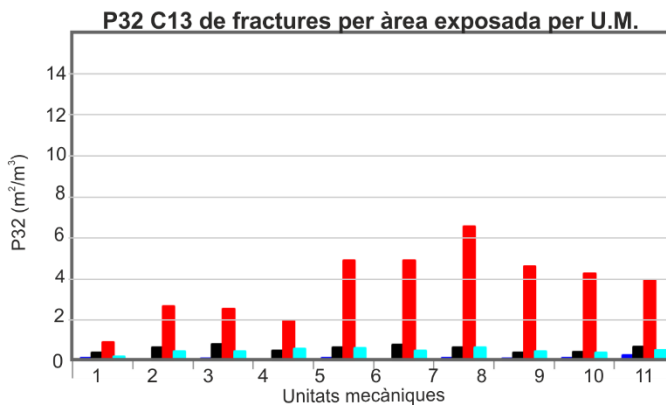
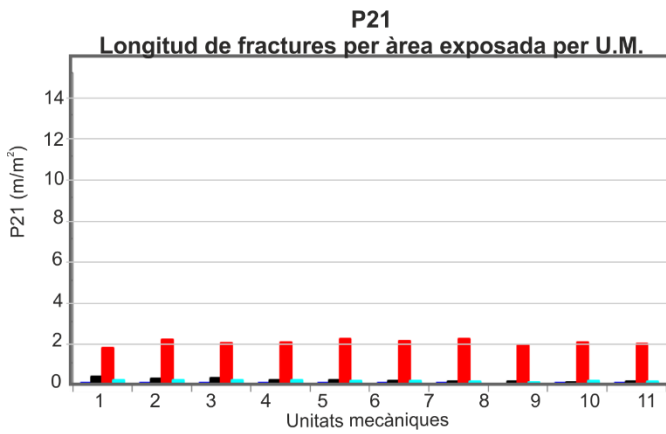
- Família I
- Família II
- Família III
- Família IV



A TLS - Lateral (NS)

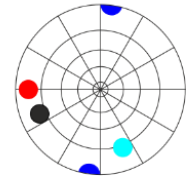
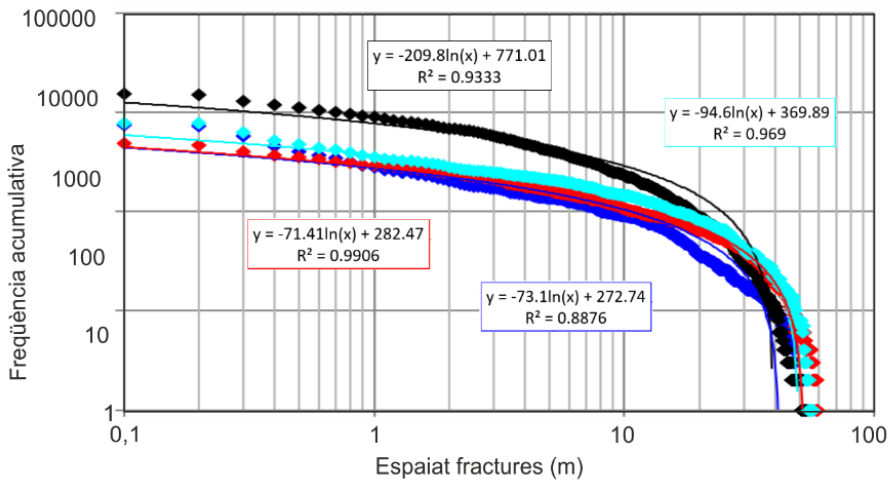


- Família I
- Família II
- Família III
- Família IV



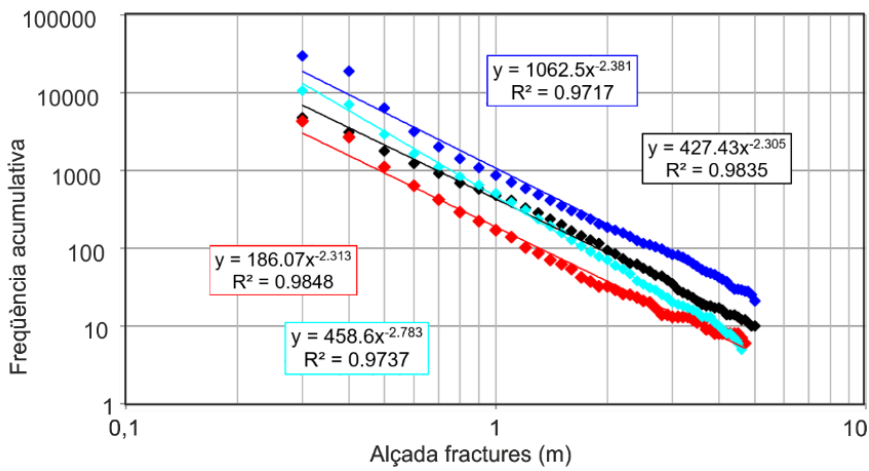
A TLS

Espaiat entre fractures (TLS)

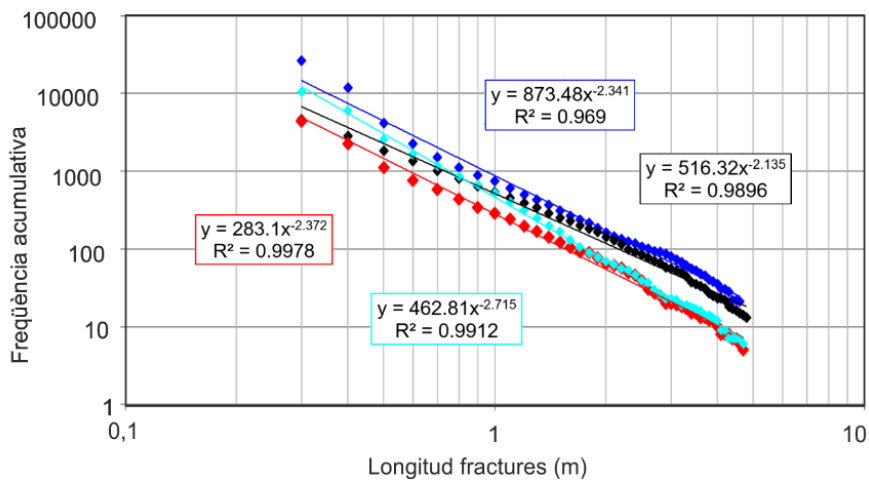


- Família I
- Família II
- Família III
- Família IV

Alçades de les fractures (TLS)



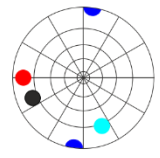
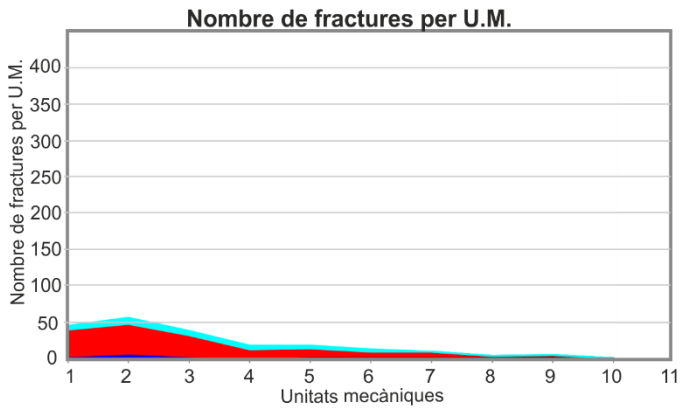
Longituds de les fractures (TLS)



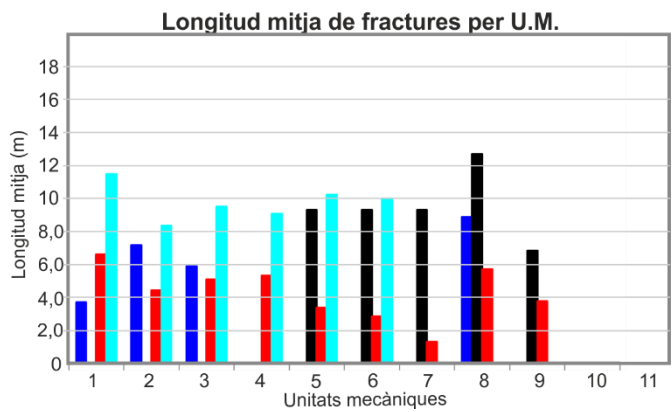
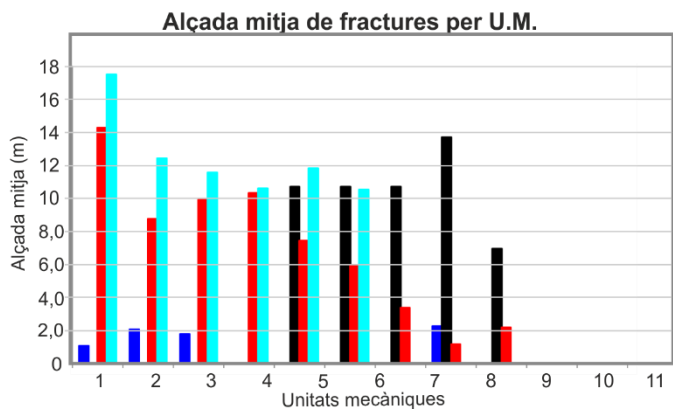
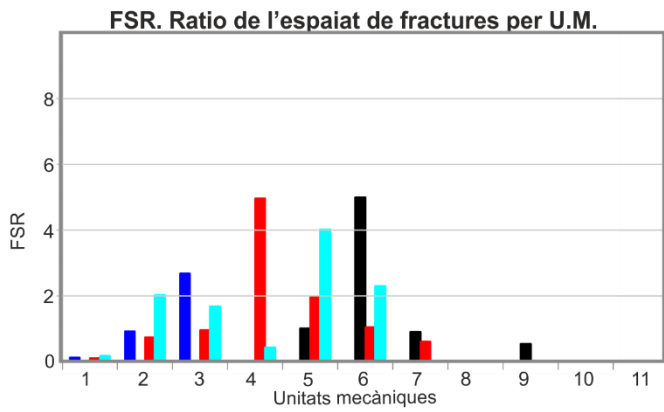
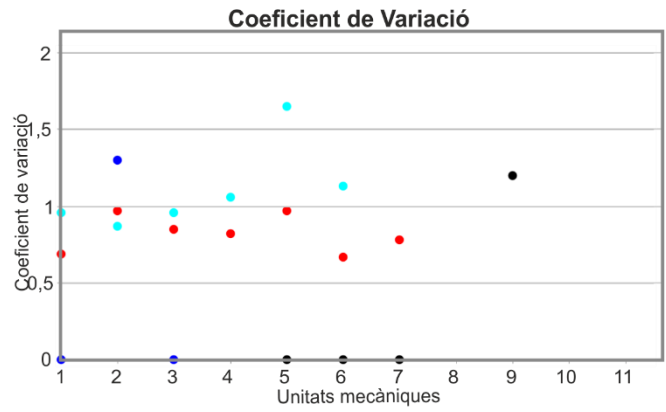
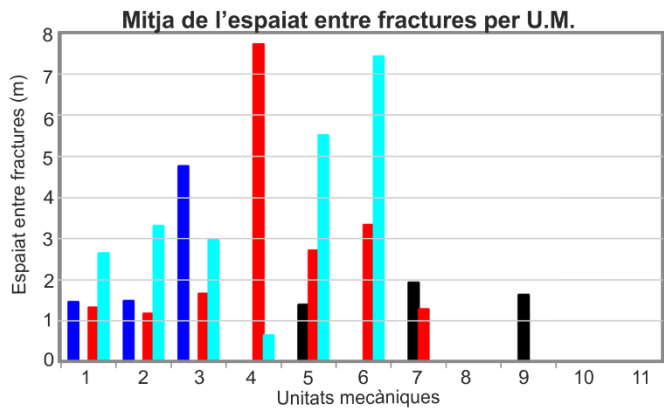
Afloraments fracturats digitalitzats.
Avaluació de les tècniques remotes en models *DFN* i aplicació de *Machine Learning*.

Gràfics de les propietats obtingut amb Fotogrametria

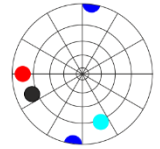
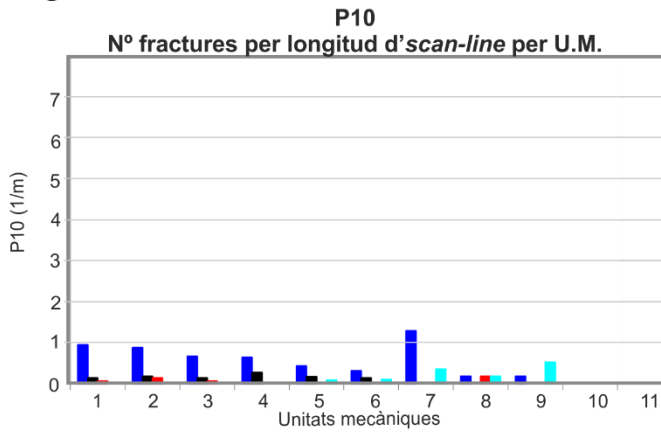
B Fotogrametria - Frontal 1 (EW)



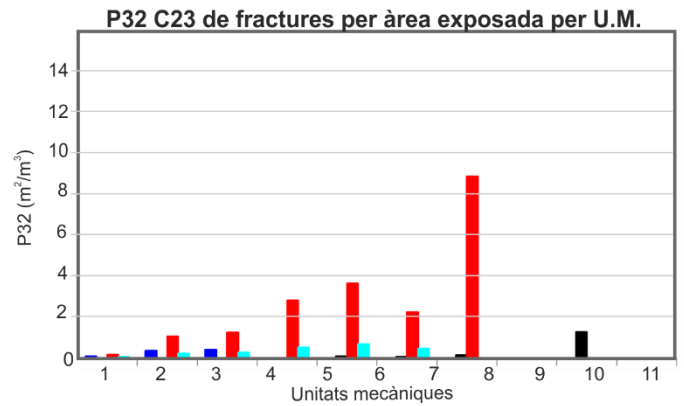
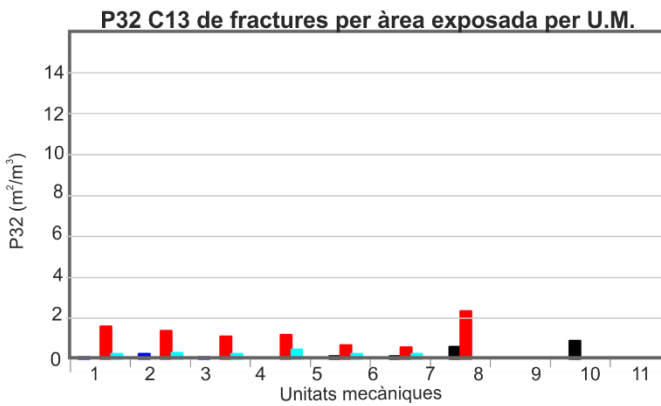
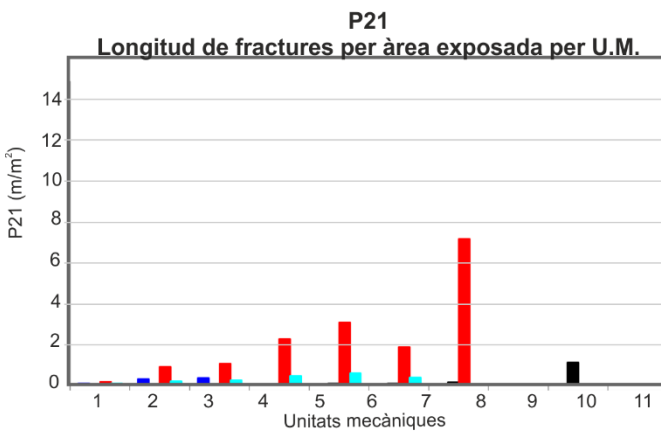
- Família I
- Família II
- Família III
- Família IV



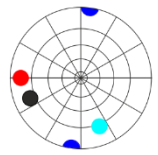
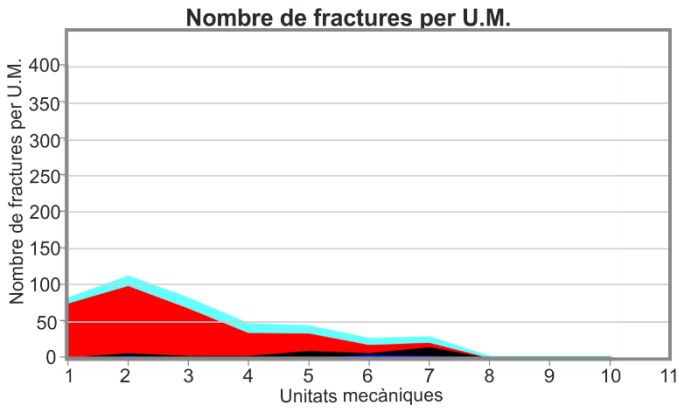
B Fotogrametria - Frontal 1 (EW)



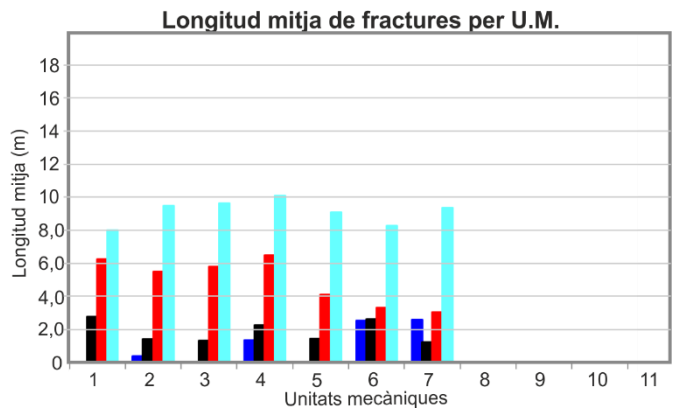
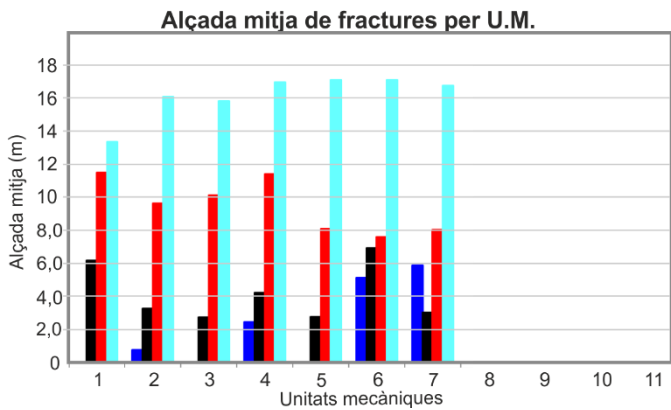
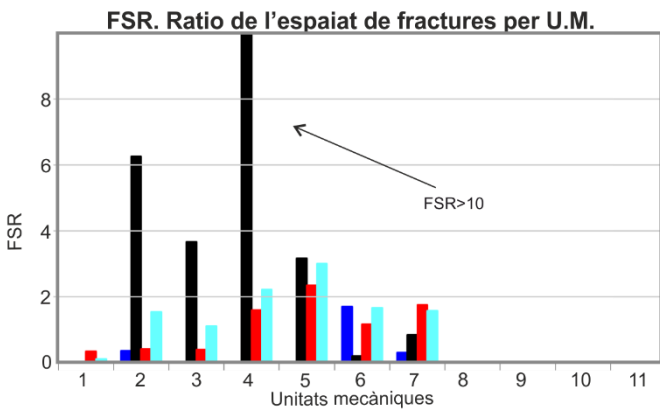
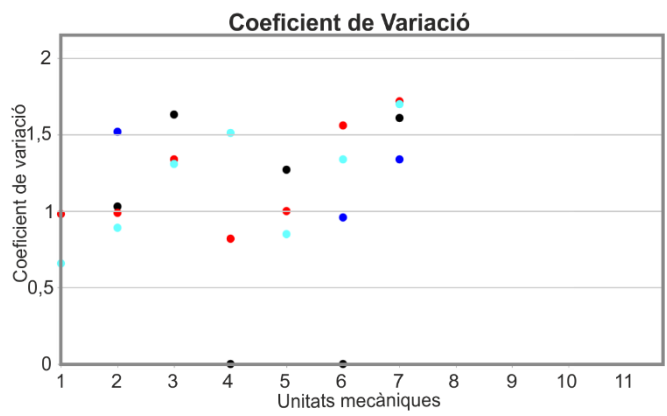
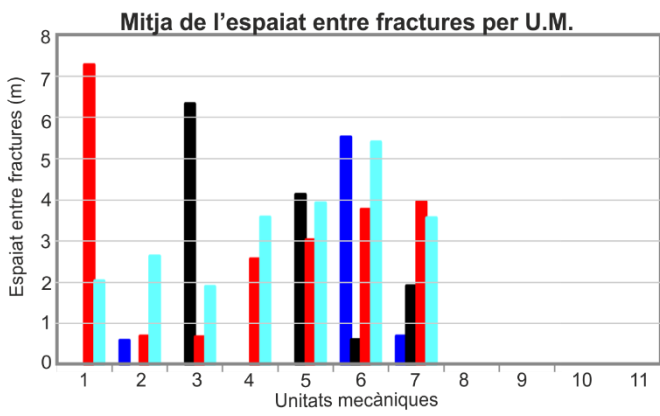
- Família I
- Família II
- Família III
- Família IV



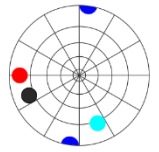
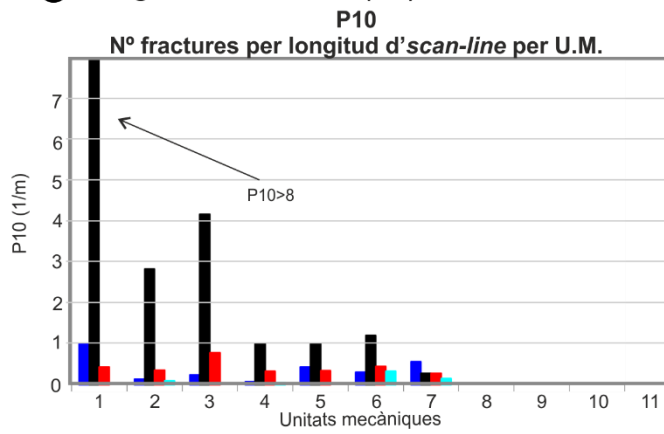
B Fotogrametria - Frontal 2 (EW)



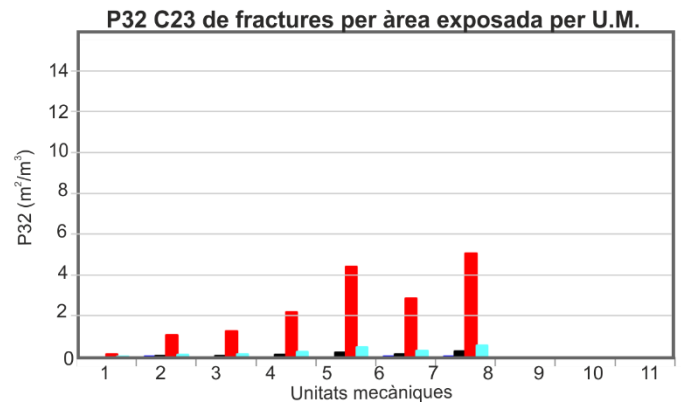
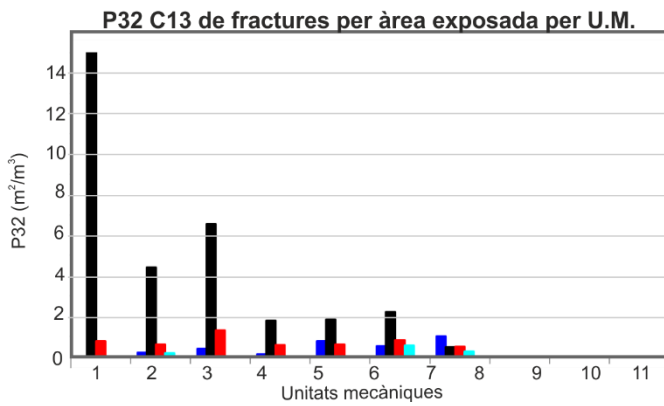
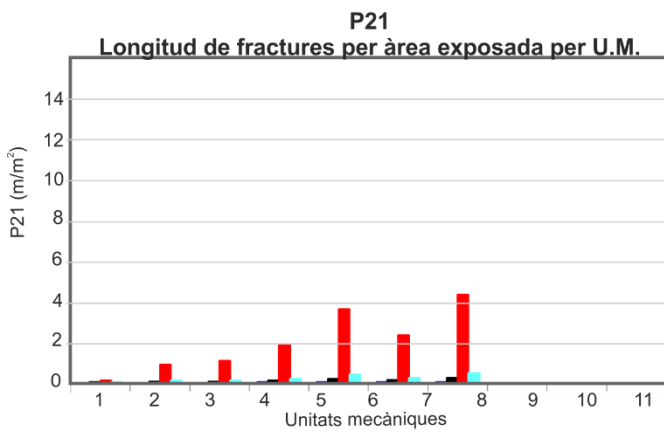
- Família I
- Família II
- Família III
- Família IV



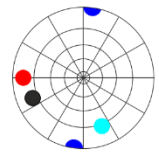
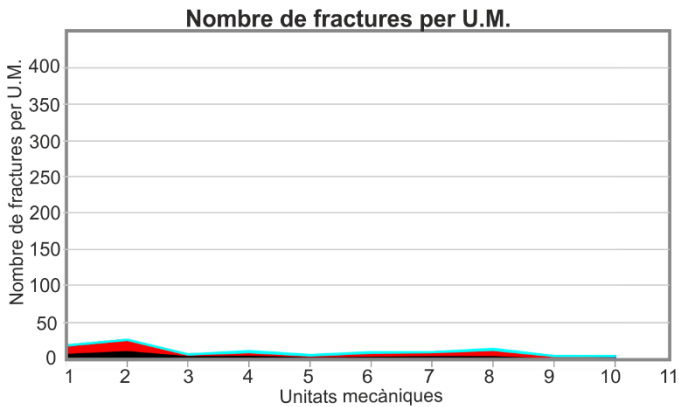
B Fotogrametria - Frontal 2 (EW)



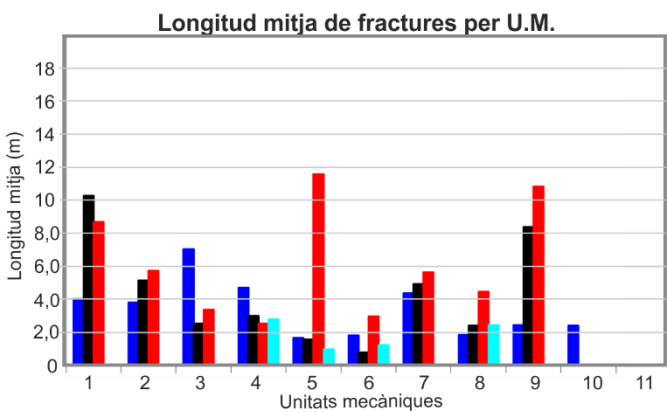
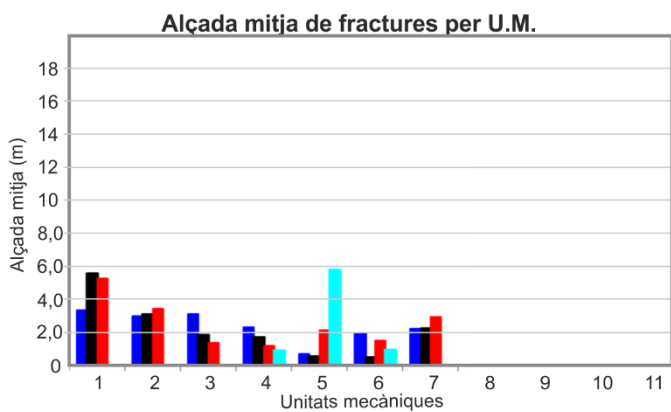
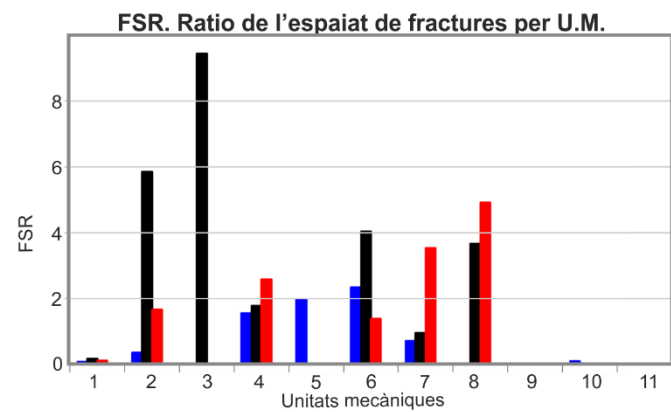
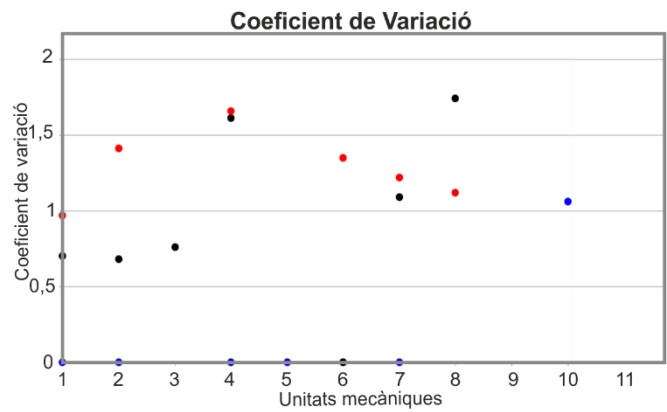
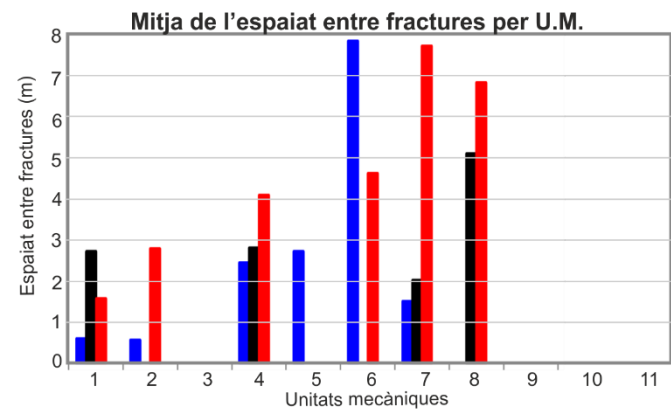
- Família I
- Família II
- Família III
- Família IV



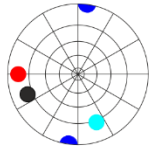
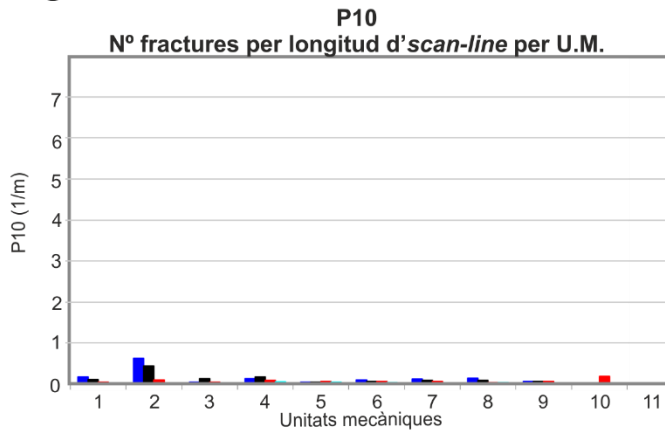
B Fotogrametria - Frontal 4 (EW)



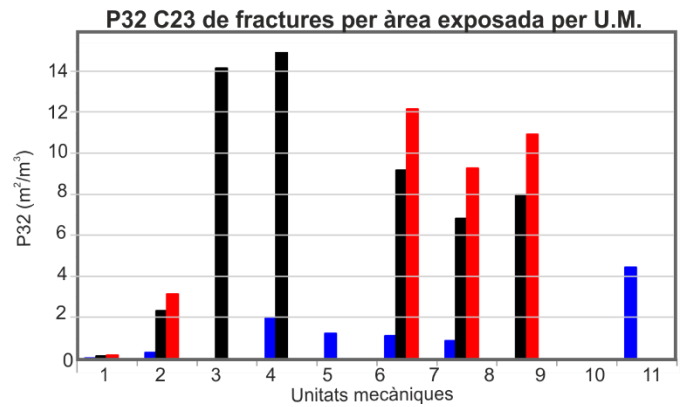
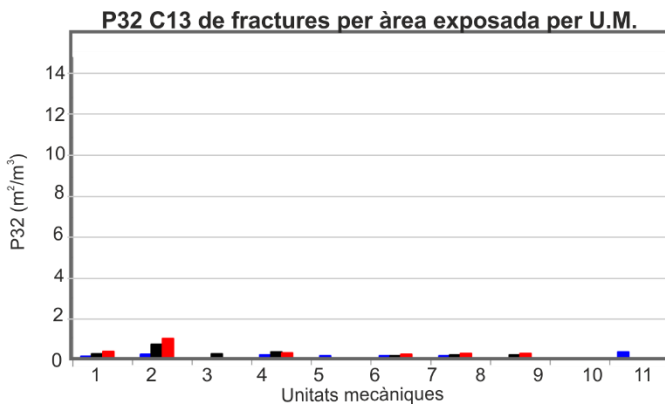
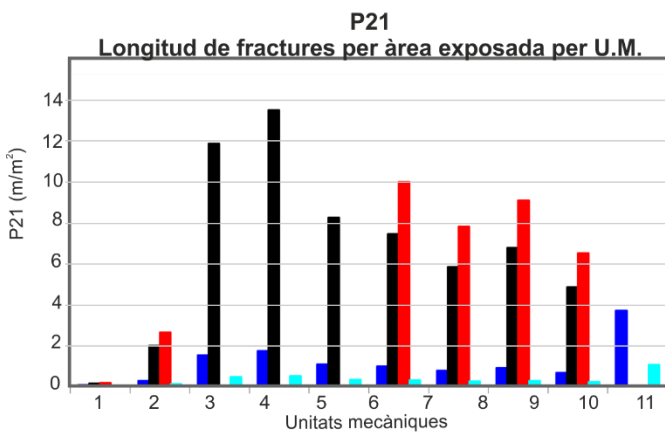
- Família I
- Família II
- Família III
- Família IV



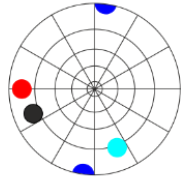
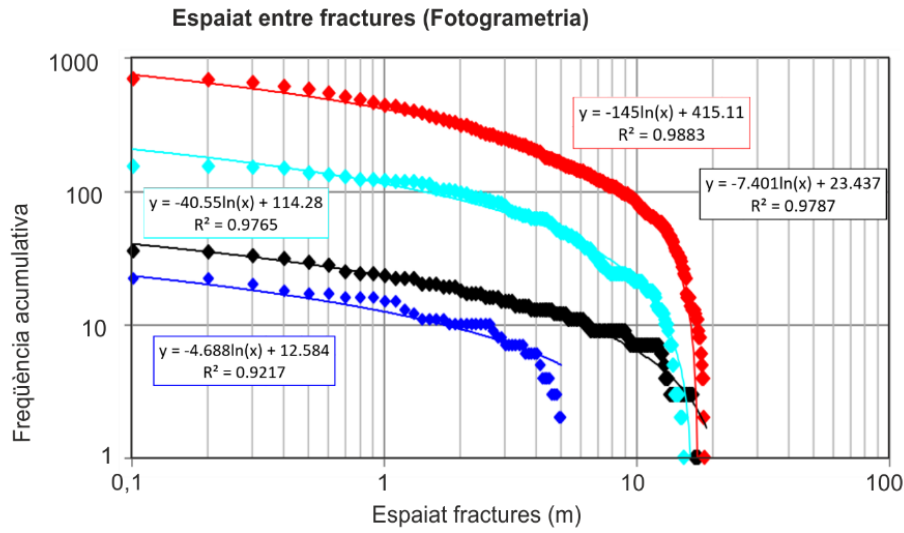
B Fotogrametria - Frontal 4 (EW)



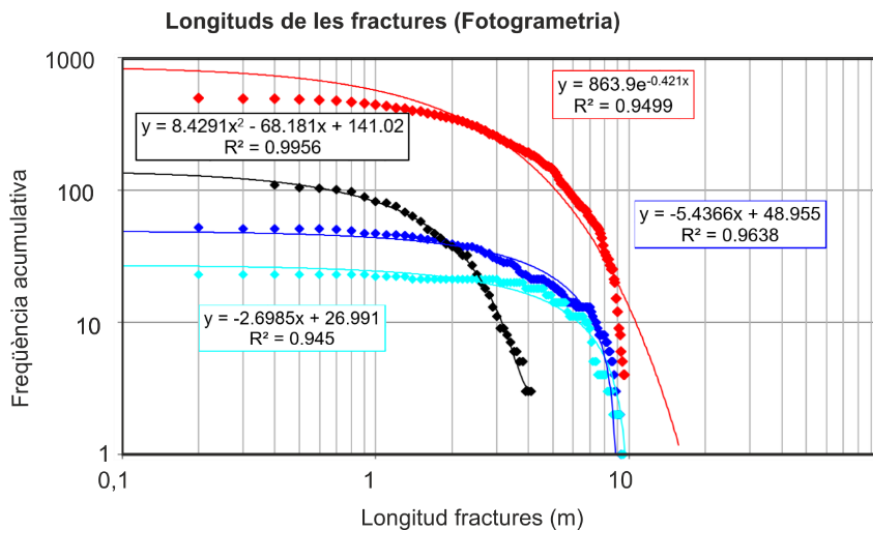
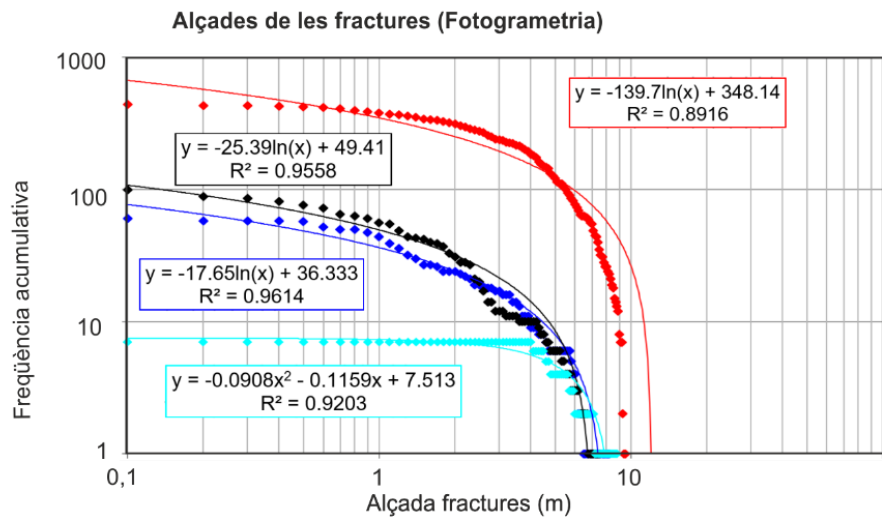
- Família I
- Família II
- Família III
- Família IV



B Fotogrametria



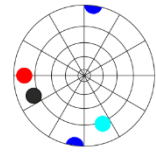
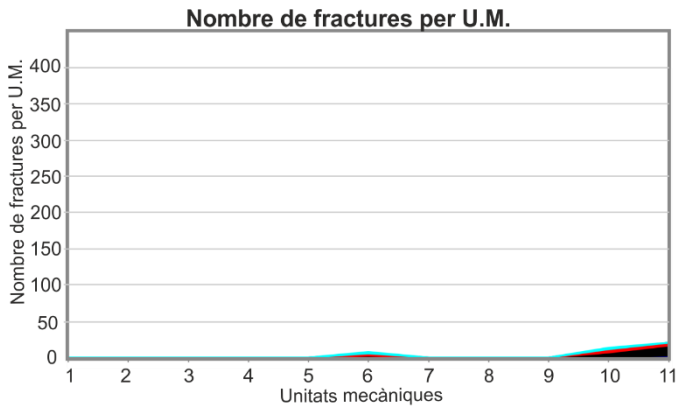
- Família I
- Família II
- Família III
- Família IV



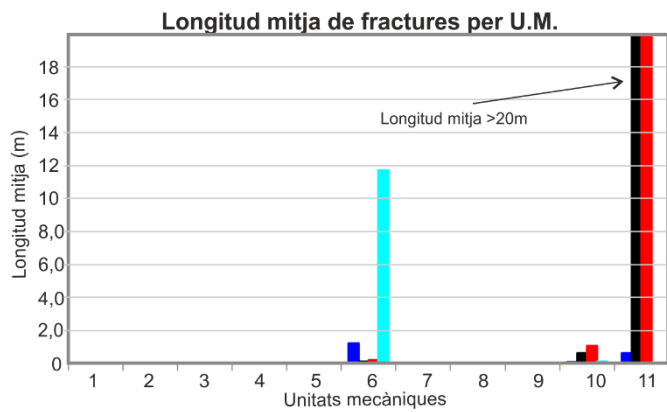
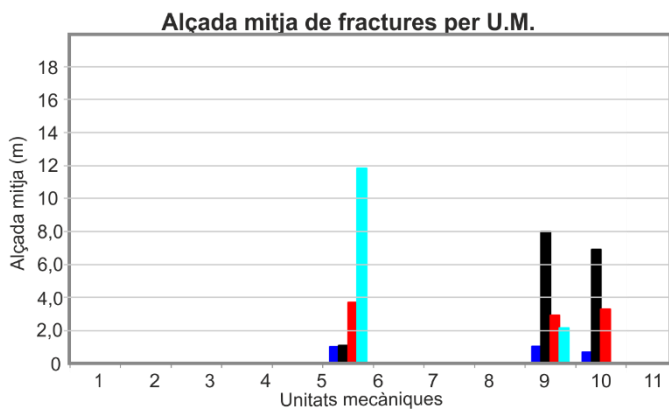
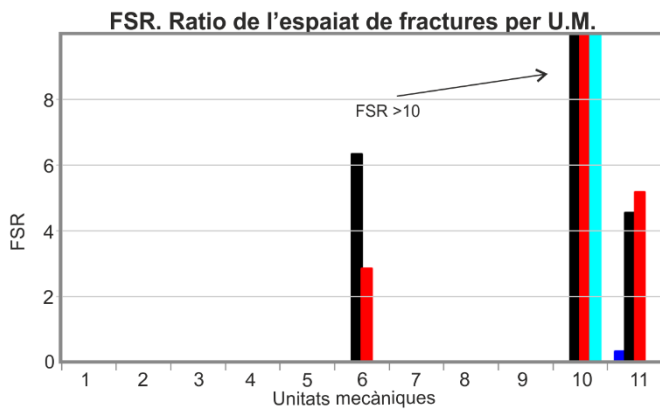
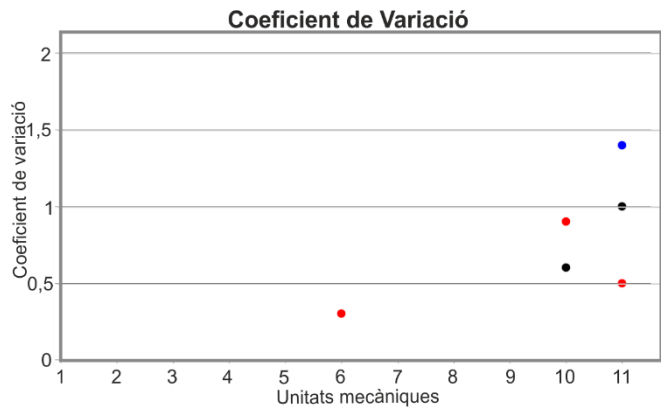
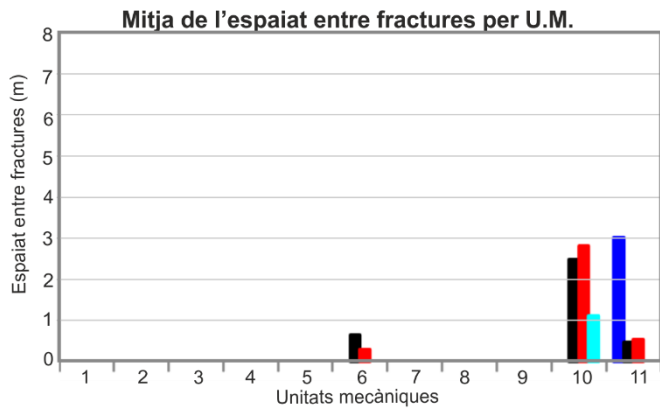
Afloraments fracturats digitalitzats.
Avaluació de les tècniques remotes en models *DFN* i aplicació de *Machine Learning*.

Gràfics de les propietats obtingut amb *scan-lines*

C Scan-lines - Frontal 1 (EW)



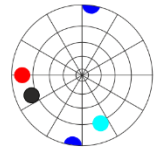
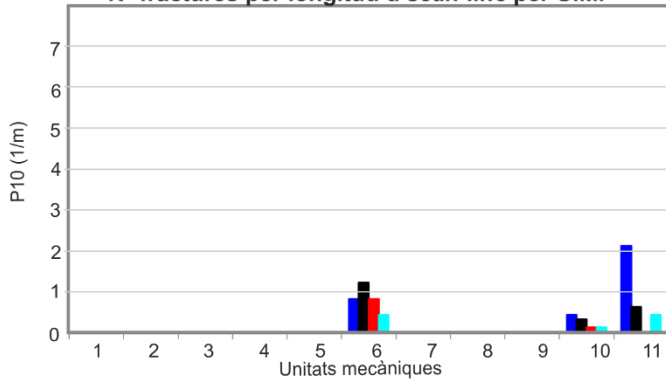
- Família I
- Família II
- Família III
- Família IV



C Scan-lines - Frontal 1 (EW)

P10

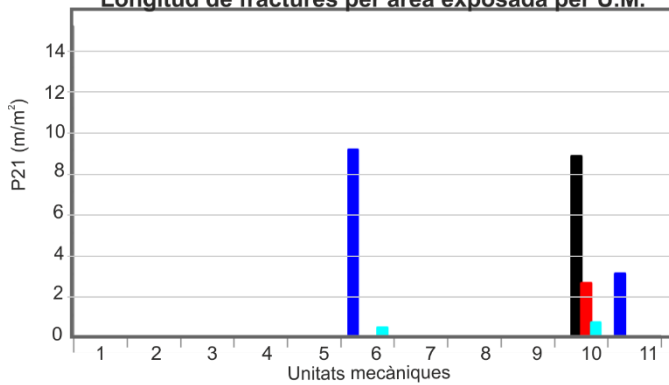
Nº fractures per longitud d' *scan-line* per U.M.



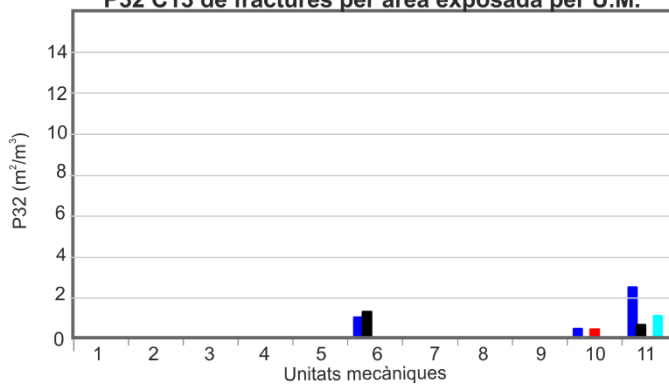
- Família I
- Família II
- Família III
- Família IV

P21

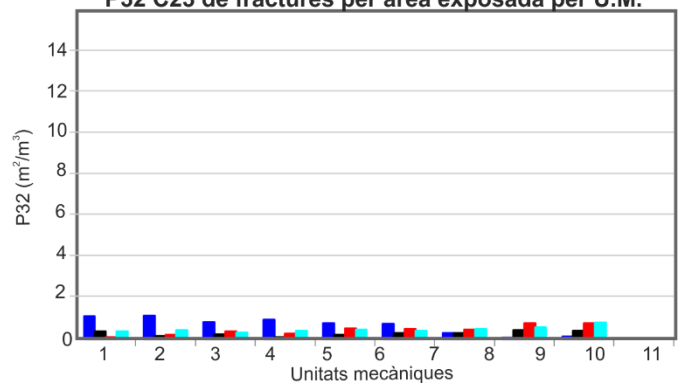
Longitud de fractures per àrea exposada per U.M.



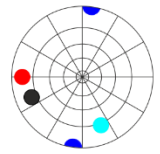
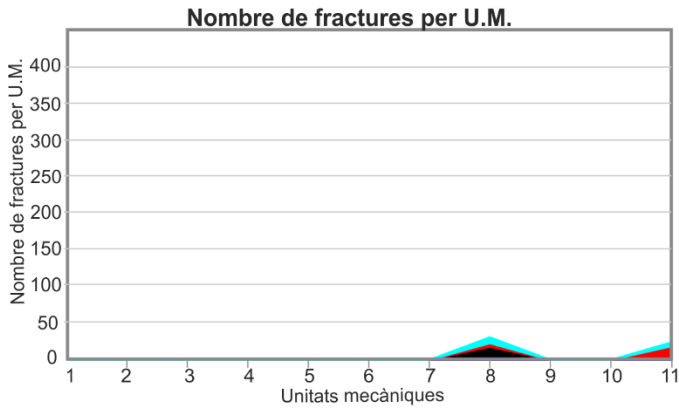
P32 C13 de fractures per àrea exposada per U.M.



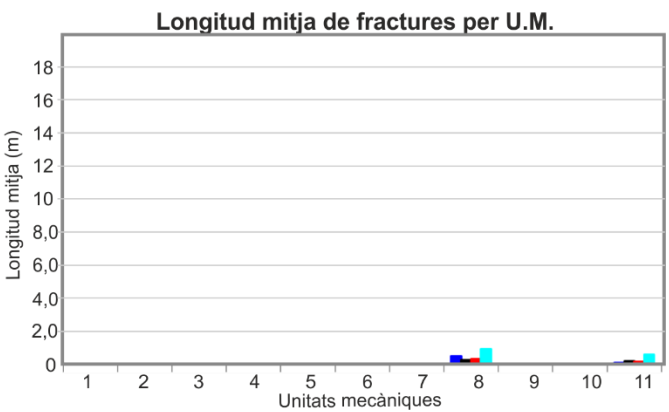
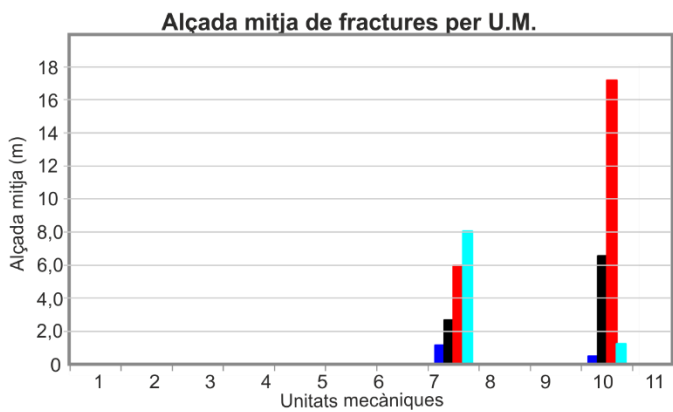
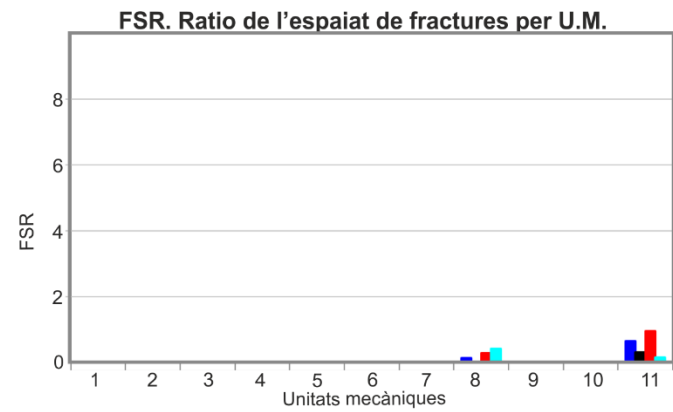
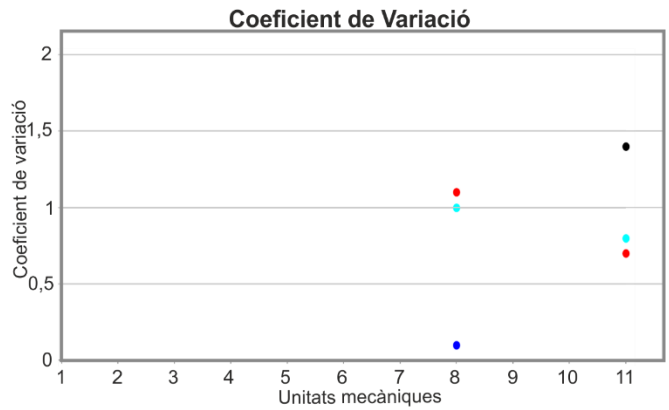
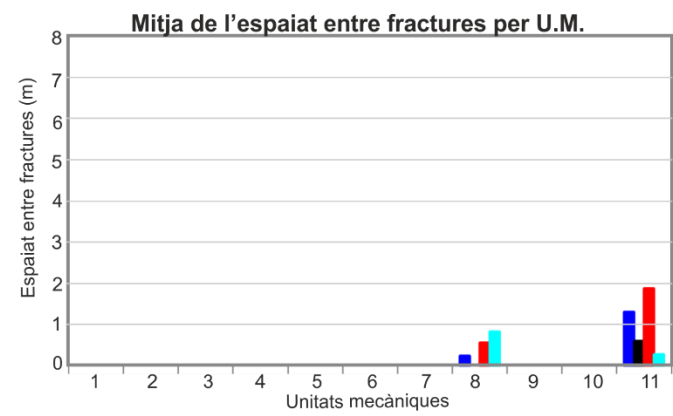
P32 C23 de fractures per àrea exposada per U.M.



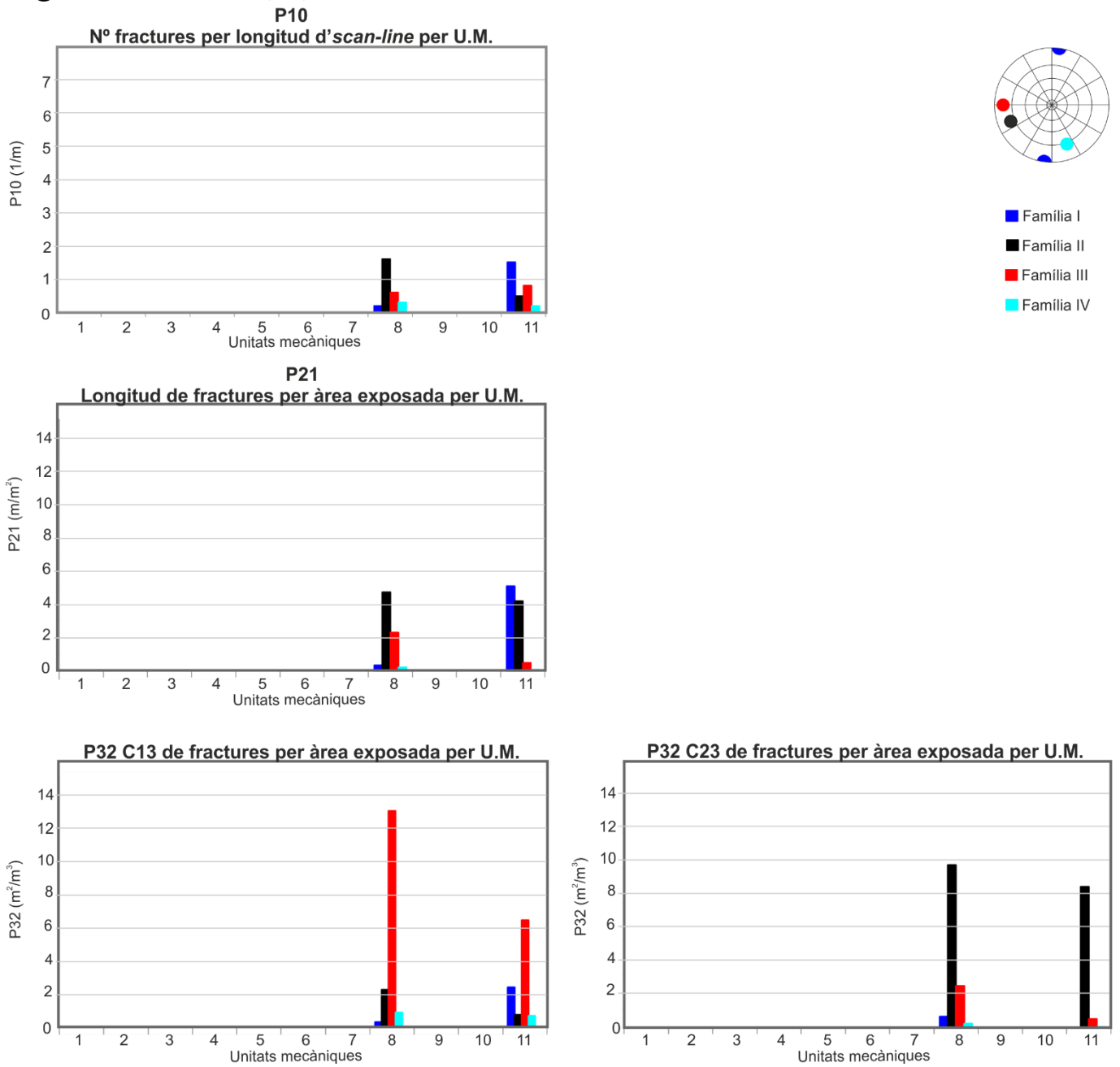
C Scan-lines - Frontal 2 (EW)



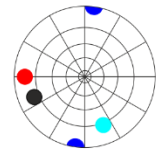
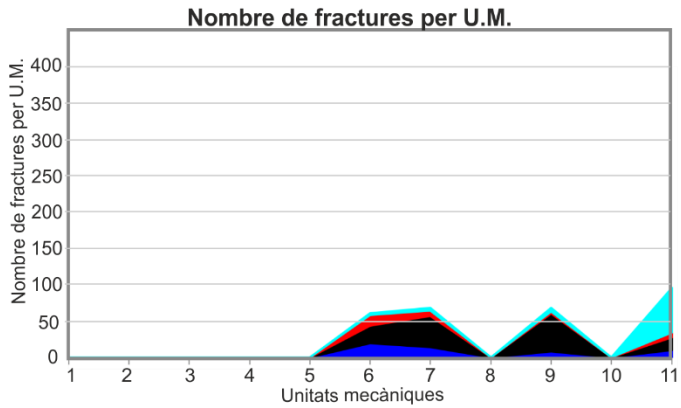
- Família I
- Família II
- Família III
- Família IV



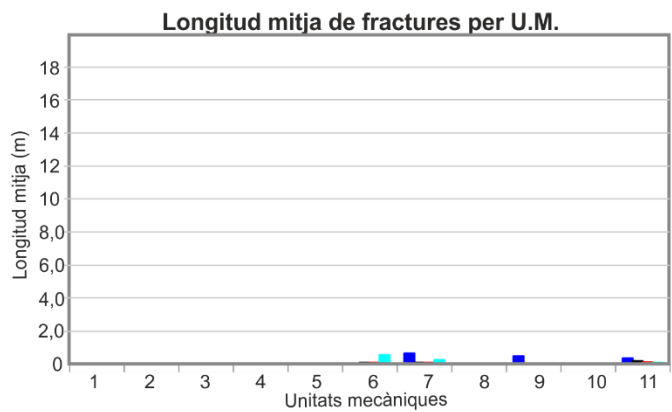
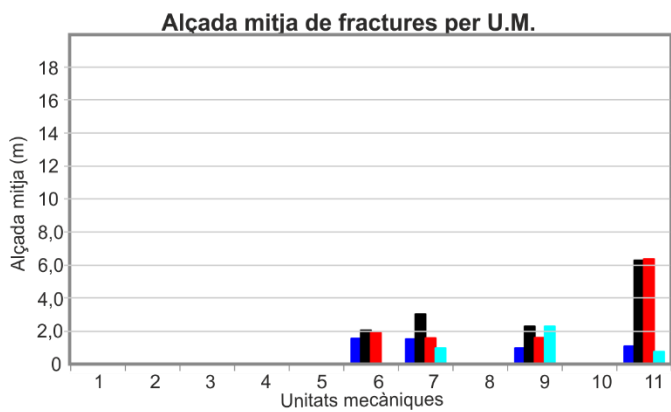
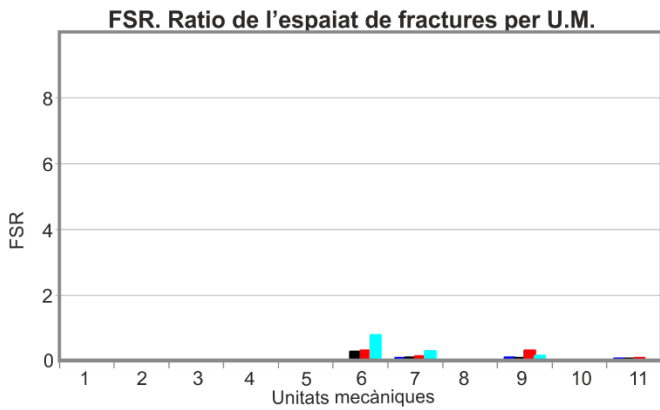
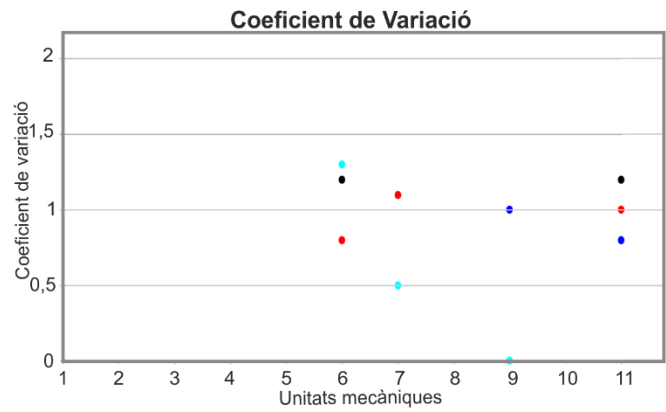
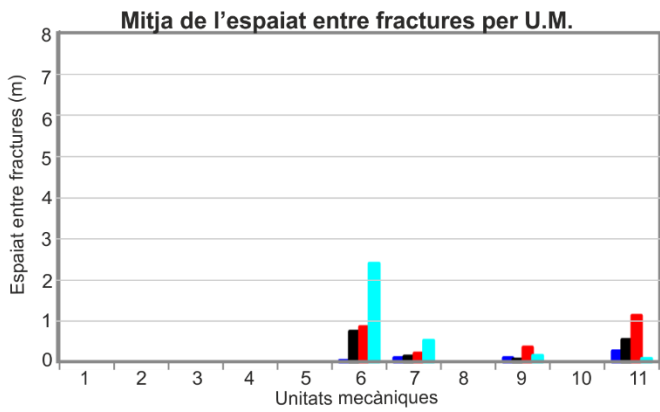
C Scan-lines - Frontal 2 (EW)



C Scan-lines - Frontal-4 (EW)



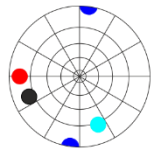
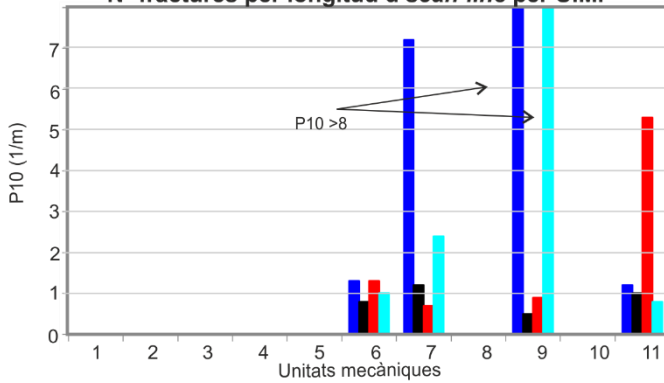
- Família I
- Família II
- Família III
- Família IV



C Scan-lines - Frontal 4 (EW)

P10

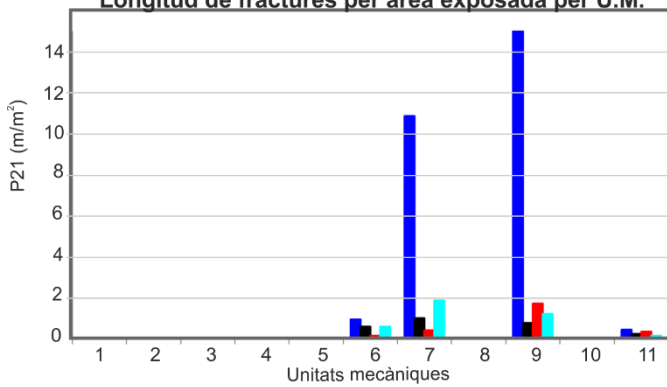
Nº fractures per longitud d'scan-line per U.M.



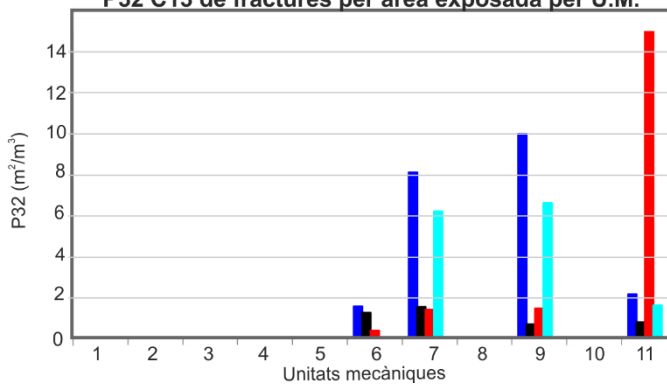
- Família I
- Família II
- Família III
- Família IV

P21

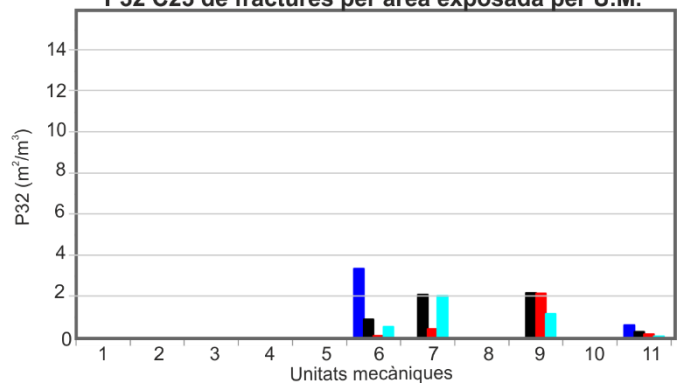
Longitud de fractures per àrea exposada per U.M.



P32 C13 de fractures per àrea exposada per U.M.



P32 C23 de fractures per àrea exposada per U.M.



A.2. Article 1: Machine learning-Based Rockfalls Detection with 3D Point Clouds

Afloraments fracturats digitalitzats.
Avaluació de les tècniques remotes en models *DFN* i aplicació de *Machine Learning*.



Article

Machine Learning-Based Rockfalls Detection with 3D Point Clouds, Example in the Montserrat Massif (Spain)

Laura Blanco ^{1,2}, David García-Sellés ^{3,*}, Marta Guinau ³, Thanasis Zoumpikas ⁴, Anna Puig ⁴, Maria Salamó ⁴, Oscar Gratacós ¹, Josep Anton Muñoz ¹, Marc Janeras ⁵ and Oriol Pedraza ⁵

- ¹ Departament de Dinàmica de la Terra i de l'Oceà, Grup de Geodinàmica i Anàlisi de Conques (GGAC), UB-Geomodels, Facultat de Ciències de la Terra, Universitat de Barcelona (UB), 08028 Barcelona, Spain
- ² Anufra, Soil & Water Consulting, 08028 Barcelona, Spain
- ³ Departament de Dinàmica de la Terra i de l'Oceà, GRC RISKINAT, UB-Geomodels, Facultat de Ciències de la Terra, Universitat de Barcelona (UB), 08028 Barcelona, Spain
- ⁴ WAI Research Group, Departament de Matemàtiques i Informàtica, Universitat de Barcelona (UB), 08007 Barcelona, Spain
- ⁵ Institut Cartogràfic i Geològic de Catalunya, Engineering Geology Unit, 08038 Barcelona, Spain
- * Correspondence: dgarcia@ub.edu

Abstract: Rock slope monitoring using 3D point cloud data allows the creation of rockfall inventories, provided that an efficient methodology is available to quantify the activity. However, monitoring with high temporal and spatial resolution entails the processing of a great volume of data, which can become a problem for the processing system. The standard methodology for monitoring includes the steps of data capture, point cloud alignment, the measure of differences, clustering differences, and identification of rockfalls. In this article, we propose a new methodology adapted from existing algorithms (multiscale model to model cloud comparison and density-based spatial clustering of applications with noise algorithm) and machine learning techniques to facilitate the identification of rockfalls from compared temporary 3D point clouds, possibly the step with most user interpretation. Point clouds are processed to generate 33 new features related to the rock cliff differences, predominant differences, or orientation for classification with 11 machine learning models, combined with 2 undersampling and 13 oversampling methods. The proposed methodology is divided into two software packages: point cloud monitoring and cluster classification. The prediction model applied in two study cases in the Montserrat conglomeratic massif (Barcelona, Spain) reveal that a reduction of 98% in the initial number of clusters is sufficient to identify the totality of rockfalls in the first case study. The second case study requires a 96% reduction to identify 90% of the rockfalls, suggesting that the homogeneity of the rockfall characteristics is a key factor for the correct prediction of the machine learning models.

Keywords: rockfall; TLS; point cloud; monitoring; machine learning



Citation: Blanco, L.; García-Sellés, D.; Guinau, M.; Zoumpikas, T.; Puig, A.; Salamó, M.; Gratacós, O.; Muñoz, J.A.; Janeras, M.; Pedraza, O. Machine Learning-Based Rockfalls Detection with 3D Point Clouds, Example in the Montserrat Massif (Spain). *Remote Sens.* **2022**, *14*, 4306. <https://doi.org/10.3390/rs14174306>

Academic Editors: Tommaso Carlà and Renato Macciotta

Received: 21 June 2022

Accepted: 21 August 2022

Published: 1 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rockfalls are one of the most frequent and dangerous phenomena in mountainous areas [1,2]. They consist of rock fragments that are detached from a steep rock face by descending rapidly while performing a free fall, rolling or bouncing [3]. They are essentially gravitational events triggered at high speed, which may cause severe damage to buildings, infrastructures, and lifelines due to their spatial and temporal frequency and their intensity (kinetic energy) [4,5]. Rockfall risk in mountainous areas is increasing as the population and economic activity increase [6]. Therefore, rigorous hazard and risk analyses are essential to provide the best possible protection measures [5]. The rockfall magnitude–frequency relationship is an important component of hazard and risk assessment for dangerous slopes, which can be evaluated using a rockfall inventory [7,8]. In addition to the magnitude–frequency relationship, information on the location of the source zone and the shape of the

block can be useful for understanding the failure processes that operate on the slope [7,9]. Thus, the identification and characterization of rockfall events on rock slopes constitute a necessary task for risk assessment and mitigation.

Nowadays, the development of remote sensing techniques such as terrestrial laser scanner (TLS) based on the lidar (light detection and ranging) technique and digital photogrammetry has been advantageous where steep and unstable slopes make conventional field data collection dangerous and impractical [10]. Such remote sensing techniques greatly facilitate the collection of a significant amount of high-quality data, i.e., 3D point clouds. As a result, the studies on the identification of changes in rock slopes and the acquisition of rockfall inventories are rapidly increasing [11–13].

1.1. Rockfall Source Analysis from Point Cloud Data

Recently, many applications of rockfall analysis employing 3D point cloud data produced from TLS or photogrammetry have been developed. In terms of rockfall source detection, Santana et al. [14] and Corominas et al. [15] identify rockfall scars and calculate their volumes, using a point cloud by means of identifying discontinuity surfaces and the minimum spacing between them. Other authors used single point clouds to perform rock cliff stability analysis such as Fanti et al. [16] and Mazzanti et al. [17]. However, the development of change detection algorithms, such as M3C2 [18], facilitates the identification of areas of loss on slopes (i.e., rockfalls) among subsequent 3D point clouds. The location, volume, and shape of rockfalls on the slope can be calculated and populated into a database. In this respect, Tonini and Abellan [19] detect and extract individual rockfall events that occurred during a time span by using clustering algorithms. Van Veen et al. [7] applied these methods to a hazardous slope that presents rockfall hazards to the CN Rail line in British Columbia, to build a database of rockfalls. Janeras et al. [20] quantified the frequency of small or even previously unnoticed rockfalls in Montserrat massif near Barcelona (NE Spain). Bonneau et al. [21] describe the rationale and consequences of measuring the dimensions of 3D rockfall objects obtained from successive point clouds, and introduce two unique algorithms to standardize the process. Bonneau and Hutchinson [22] use high-resolution photographs to validate changes described in TLS data analysis and to follow deposition patterns in a dynamic cliff talus system. Furthermore, in order to monitor progressive failures, change detection systems have been used to identify precursory indicators, such as cleft opening or pre-failure deformations [23–25].

The interval between point cloud acquisitions has a significant impact on both the identification of rockfall precursory indicators and the categorization of rockfalls with their shape and volume from change detection approaches [7,19,26]. In order to better understand progressive failures and to improve the magnitude–frequency relationship to characterize the rockfall activity on a rock cliff, fixed systems have been developed to acquire data with a high frequency or in quasi real-time [27–29]. Consequently, a large amount of high-quality data acquired by TLS and photogrammetry in high-frequency monitoring approaches has begun to overwhelm users [26,30]. Consequently, the process of change identification in rock cliffs from lidar or photogrammetric 3D point clouds has undergone a thorough development in the last years to address and improve the current time-consuming analytical methods.

The algorithms developed to automate the processes for the detection of change from multi-temporal 3D point cloud comparison usually follow these steps (see comparative in Figure 1): (a) point cloud classification to remove objects without interest (e.g., vegetation, or misplaced points due to moisture or edge effects); (b) point clouds alignment (i.e., placing the different point clouds in the same reference system); (c) computing differences between point clouds; (d) clustering neighboring points with significant differences, and (e) classification of clusters according to their nature (e.g., bedrock, vegetation, edge effects, soil, rockfall, or deformation) [24,25,31]. Numerous initiatives have been made to automate intelligent pipelines using machine learning techniques such as deep learning and neural networks, which have made tremendous progress, particularly in the task of identifying

rockfalls. However, in this proposal we try to automate not only the identification of rockfalls but also the evolution of the overall process, as data are collected over time until the final prediction is obtained.

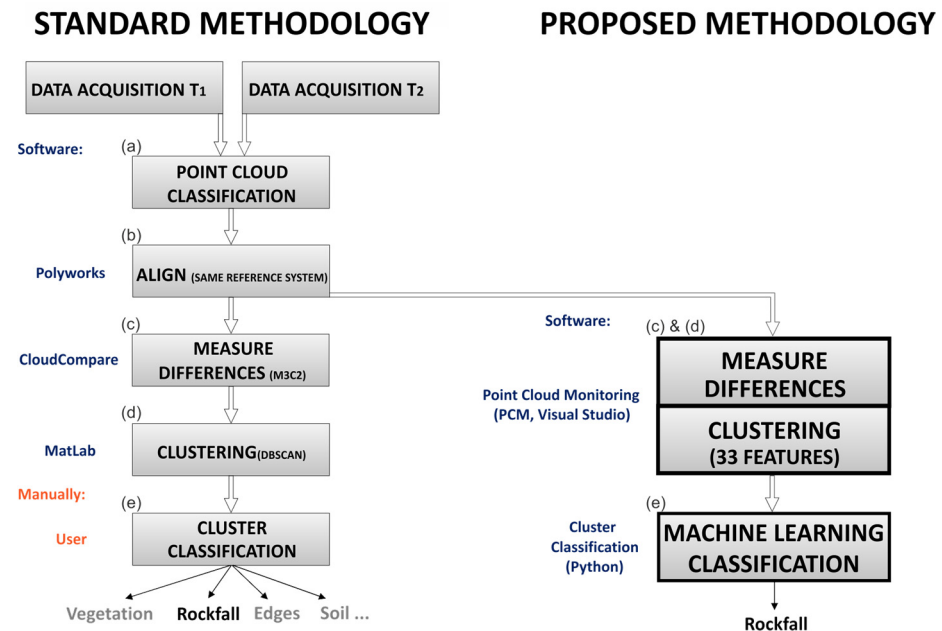


Figure 1. Comparative of the standard workflow with the method proposed in this study to identify rockfalls from point clouds. The proposed workflow uses 33 features and point cloud texture intensity for machine learning classification. Letters refer to text explanation in Section 1.1.

1.2. Improvements on Rockfall Detection from Point Cloud Comparison including Machine Learning Algorithms

Recent research studies utilize mainly machine learning algorithms for rock-slope and landslide monitoring and analysis ([32] and references therein). However, few studies address the automation of rockfall detection including machine learning processes in the different steps of the point cloud comparison.

Since differences in vegetation, soil, anthropic actions or rockfalls lead to inaccuracies in the point clouds alignment process, the first step for point clouds comparison is the filtering of unwanted areas of each point cloud, such as vegetation or edges, to infer optimal change detection and posterior rockfall classification [29]. It is sometimes difficult to separate bedrock from vegetation or another class of surface in terms of time consumption or supervision of the result. Different approaches consist of manually generating masks with the areas of interest in the point cloud scene [30,33]. Furthermore, based on the possibilities of analyzing the lidar return intensity, Williams et al. [29] remove points with certain criteria, thereby reducing uncertainty. Other approaches, such as CANUPO [34], classify regions of interest, training binary classifiers and combining them with some rules. Weidner et al. [33] present a random forest machine learning approach that improves the classification accuracy and efficiency compared to CANUPO. Other approaches, such as surface interpolation concepts with the cloth simulation filter (CSF, [35]), or multiscale curvature classification MCC [36], are also strategies for identifying relevant classes of surfaces (e.g., bedrock, soils, or vegetation). This step for removing unwanted points can be carried out, providing that the success of this operation is ensured. However, in cliffs with complex morphology and compounded of different materials (rock, soil, talus, vegetation, etc.), it is difficult to encompass a good automatic classification, requiring time-consuming manual filtering.

Other advances have focused on adapting and improving change detection algorithms. While the most commonly used algorithm is Multiscale Model-to-Model Cloud Comparison (M3C2) [18], other authors such as Williams et al. [29] and Kromer et al. [37] improved

the overall accuracy of change detection and streamlined the workflow when applied to large time series scan datasets. For the purpose of recognizing rockfall events, additional developments have been offered by DiFrancesco et al. [4], indicating enhancements to the performance of the change algorithm. According to these authors, the procedure of filtering away incorrect clusters (such as vegetation, edge faults, snow, or dampness) is often either disregarded or carried out manually once individual objects are recovered from the point cloud using clustering.

Several authors have developed different approaches to gain objectivity, efficiency, or productivity in cluster classification. Supervised or manual classification is one of the most commonly employed techniques, but identifying rockfalls according to several classes is a tedious and laborious undertaking that can sometimes become subjective and dependent on the criteria used by the expert classifiers [7,38,39]. More recent proposals are based on statistical model classifications and machine learning studies such as those by Schovanec et al. [30]. The aforementioned study proposes a way of filtering clusters using the random forest algorithm. However, they do not address the data imbalance issues (lower number of rockfall labeled clusters vs. higher number of no-rockfall clusters), caused by the sensitivity of the sensor and the complexity of the analyzed scene [32]. In this context, Zoumpakas et al. [32] focus their work on identifying 3D point cloud clusters of the rockfall class while dealing with an imbalanced classification task.

In this paper we propose a cluster classification method based on the developments by Zoumpakas et al. [32], using different machine learning models and resampling strategies in the classification of rockfall clusters in order to implement prediction models. In our opinion, machine learning algorithms are better trained by using cluster features than using 3D point characteristics. Consequently, we propose some adaptations of the already accepted algorithms for change detection (M3C2 by Lague et al. [18]) and clustering (Density-Based Spatial Clustering of Applications with Noise, DBSCAN by Ester et al. [40]) to compute geometric features. Thus, a total of 33 features are used to characterize clusters to improve the training stage for the machine learning classification. In this study, certain characteristics of the clusters are proposed to be able to feed the learning models. Although there could be more or others, in our proposal a global framework based on machine learning is presented, but the study and validation of the elicitation of the characteristics (features) used by the learning models is outside the scope of this proposal.

Zoumpakas et al. [32] concentrate on the machine learning training stage, analyzing the associated problems and proposing a series of solutions and tools (data normalization, balancing techniques, cross-validation, hyper-parameterization, classifier models, tools to identify the best classifier model). Our current proposal shows the adaptation of existing algorithms to better characterize clusters and incorporating the approaches proposed by Zoumpakas et al. [32] to solve the problems encountered in the training stage. The proposed intelligent framework is performed in two different rock cliffs of the Montserrat massif (NE Spain) to test the response of prediction models against rockfalls of different characteristics, in terms of volume and shape. This framework provides a novel ready-to-use point cloud machine learning software targeted at rockfall identification, which may be incorporated as a stand-alone component in a rockfall hazard decision support system. The main contributions of this work are the following:

- We propose an extension of the full end-to-end intelligent framework proposed by Zoumpakas et al. [32] for rockfall detection handling highly imbalanced data by reducing the number of clusters in our data. We further introduce geological properties to the framework itself.
- We implement the proposed intelligent system with real data from two different cliffs of Montserrat massif (NE Spain) to validate its efficacy and effectiveness.
- Our results show great performance and robustness, which is of paramount importance in rockfall detection.
- We provide a baseline methodology and a detection accuracy benchmark for future related experimental analyses.

- We have made fully accessible the applications developed in this work, the 3D point cloud data used, and an example of application in public repositories (see Section 2).

The paper is organized as follows. Section 2 exposes how these methods have been adapted to the proposed methodology, based on measuring new features that upgrade machine learning classifications. Section 3 shows the results of the study case of the Degotalls in the Montserrat massif (Barcelona, Spain), a rock cliff where the fracture pattern favors the rockfalls next to infrastructures [20].

2. Methods

In the first steps, the proposed methodology follows the standard processes of data capture, alignment, and a light manual point classification, which can be applied with the most commonly used processing software (e.g., CloudCompare [41] or Polyworks [42]). After that, the methodology is split into two steps. Firstly, the measurement of differences to create clusters with associated features, and secondly the classification of clusters to identify rockfalls (right side of Figure 1). Our contribution consists in implementing the new features necessary for characterizing clusters and classifying rockfall clusters with machine learning in order to increase the automation of the process. Thus, our methodology proposes the following four-fold contributions:

1. The adaptation of the M3C2 algorithm to measure differences point-to-point and to obtain the new associated features required for the machine learning processes. The main features are geometric such as difference between point clouds, reference and compared surface orientation, indexes of coplanarity and collinearity.
2. The development of a self-calibration method to automatically define the limit of detection (LoD) and differentiate real changes in the rock cliff from the system noise.
3. The adaptation of the DBSCAN algorithm for clustering point clouds and create new cluster features of predominance associated with the point differences (retreat or advance) in the cliff surface.
4. The analysis of different machine learning models to classify clusters of rockfalls.

The first three methods, measurements of differences, LoD and clustering, implemented in the “Point Cloud Monitoring” software (PCM), have been developed using Visual Studio 2019 [43] and the BASIC programming language. The fourth method related to the cluster classification is called “Cluster Classification” and it has been implemented using Python programming language. All the developed applications in this research are available at the GitHub repository of the group (https://github.com/Geomodels-UB/Risknat_Detection) (accessed on 12 June 2022). The 3D point clouds, and one example can be found in the UB repository: Point Cloud: <https://dataverse.csuc.cat/dataset.xhtml?persistentId=doi:10.34810/data201> (accessed on 20 July 2022). Example: <https://dataverse.csuc.cat/dataset.xhtml?persistentId=doi:10.34810/data199> (accessed on 20 July 2022).

The initial data format of the methodology is a 3D point cloud format (see Appendix A for detailed information), to which the computed new features are added (see the complete list in Appendix B). The cluster format (detailed in Appendix C) is a 3D point (the center of mass of used points) which is characterized with statistical and cluster features.

2.1. Adaptation of the M3C2 Algorithm

Different strategies have been developed to measure changes in rock slope surfaces over time. The most commonly employed techniques are: (a) point-to-point [41,44]; (b) point-to-model [19,45,46]; or (c) model-to-model [4,18,25]. The M3C2 method proposed by Lague et al. [18] and implemented in the software CloudCompare [47] is one of the most commonly employed in rockfall change detection [7,22,23,30]. In this case, the difference is measured along the normal vector calculated with a set of points from the reference point cloud and projected until intersection with a set of points of the compared 3D point cloud. In accordance with the objectives, the arrangement is configured with two parameters that define a cylinder, the diameter of search around an initial point and

the maximum depth [48]. The result of the point cloud comparison is a new point cloud including the feature difference (ϵ) as well as to the 3D coordinates and intensity or RGB.

In this work we propose an adaptation of the M3C2 method to obtain more attributes (features) that characterize rockfall clusters, which will feed the training of machine learning classifiers.

Essentially, the adaptation consists in compute features associated with the geometric relationship between reference point (initial 3D point cloud) and compared point (monitoring or second 3D point cloud). For this reason, the [49] algorithm M3C2 has been converted from model-to-model to point-to-point by changing the detection method to use real data points. Consequently, a normal (V_N in Figure 2) for each reference 3D point (see P_{REF} in Figure 2) is computed by means of the principal component analysis method [49] and the eigenvalues and eigenvectors method [50] with the set points selected by a radius search condition defined by the users (R_S in Figure 2). The compared point (see P_{COM} in Figure 2) is sought along this normal vector direction (V_N) with the closest point criterium. Thus, the feature difference is calculated with the closest point in an orthogonal direction to the reference 3D point. Likewise, the cylinder used by Lague et al. [18] is replaced by a double truncated cone or conical frustum. The user defines the maximal and the minimal horizontal distance and the maximal vertical distance (normal direction) of the search (see MaxHd, MinHd, MaxVd in Figure 2). This geometric figure allows a better fit when working with high densities of point clouds. The high densities of points reduce the distance between them, allowing to work point-to-point without the need for interpolation, especially in irregular scenarios. Monitoring always under the same conditions (TLS or camera position, point cloud density) it is strongly recommended. The process computes the normal vector as a feature at each point in both point clouds (reference and compared), decomposing the azimuth and slope values, with the aim of characterizing the surface of the rock cliff. Furthermore, the coplanarity and collinearity values are calculated as a result of applying the eigenvalues at each point of the point clouds [51].

The algorithm computes a total of 33 features (using Intensity texture, see the complete list in Appendix B, summarized in Table 1). All the new features are related to the geometric components of the distance (vertical and horizontal components), the angle between the two normal vectors (e.g., reference point cloud, and compared point cloud), and their direction (towards the TLS or from the TLS), and its geometric attributes are: collinearity and coplanarity.

Table 1. Summary of the new features computed with the adaptation of the M3C2 [18] algorithm implemented in PCM software. Comparison between the distances of the 3D point clouds is performed in the direction defined by the reference normal vectors. Results are associated with a new point cloud with these features.

Features	Significance
Distance	Distance between points (reference and compared)
Vertical Distance	Distance along the normal vector
Horizontal distance	Perpendicular distance to the normal vector
Angle between normal	Angularity between normal reference and compared
Direction	Direction of the normal vector with respect to the surface
Vector	Normal vector (i, j, k) for each point
Azimuth	Normal vector decomposed in orientation to North
Slope	Normal vector decomposed in orientation to horizontal
Collinearity	Distribution degree of neighboring points along a line
Coplanarity	Distribution degree of neighboring points along a plane

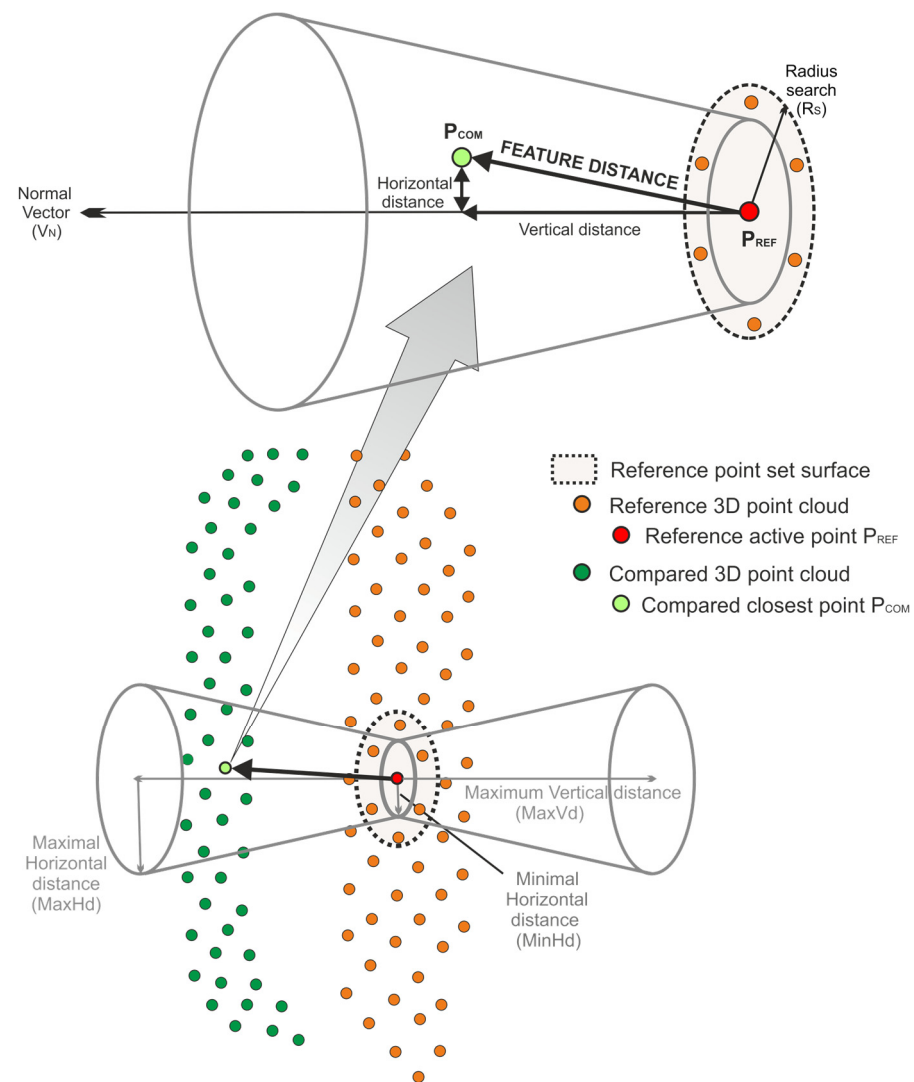


Figure 2. Depiction of the M3C2 adaption algorithm. From each point of the reference 3D point cloud, a normal vector (V_N) is computed with the neighboring points in a user defined search radius (R_s). The normal vector (V_N) defines the direction in which the closest point is sought (P_{COM}) in the compared 3D point cloud. The maximum and minimum horizontal distance and maximal vertical distances ($MaxHd$, $MinHd$, and $MaxVd$) are parameters assigned by the user to define a double truncated cone (for inward or outward searches).

2.2. Automatic Calibration

The LoD is the value at which the distance is considered representative of a real change or is assumed to be system noise. Thus, according to the LoD, points can be classified into three possible classes: (a) surface advance (i.e., precursory deformation); (b) surface retreat (i.e., rockfall); or (c) undetermined difference depending on the precision of the system (e.g., TLS precision, TLS–surface range, and software used). The undetermined type is not assigned to a true variation of the surfaces when the values of the difference are within the LoD limits. The value of the LoD is established by the knowledge of the user in the area or by experimental case studies [27,48]. Thus, it is necessary to quantify the LoD to distinguish between differences caused by the noise of the detection system and these corresponding to real changes in the slope. For this purpose, we propose to acquire two datasets with a time interval that is as short as possible, assuming that in this interval nothing has changed on the bedrock surface. The calculation of the differences between both point clouds (T_0 – T_1) are fitted to a Gaussian distribution model in which the function depends on the mean and the standard deviation establishing the precision of the system

(Figure 3a). Subsequently, when two monitoring point clouds (T_1 – T_2) are compared in order to determine the real changes, the values of the differences are also fitted to another Gaussian distribution (Figure 3b). On comparison of both probability density functions, the points of intersection indicate the limit value (LoD) to be considered as system error, and therefore indicate the probability of being assigned to a real change in the cliff (Figure 3c). In fact, the values of the feature difference between the upper LoD and the lower LoD are regarded as system noise. However, values greater than the upper LoD are regarded as probable advances in the bedrock, while difference values less than the lower LoD are regarded as probable mass losses due to possible rockfalls.

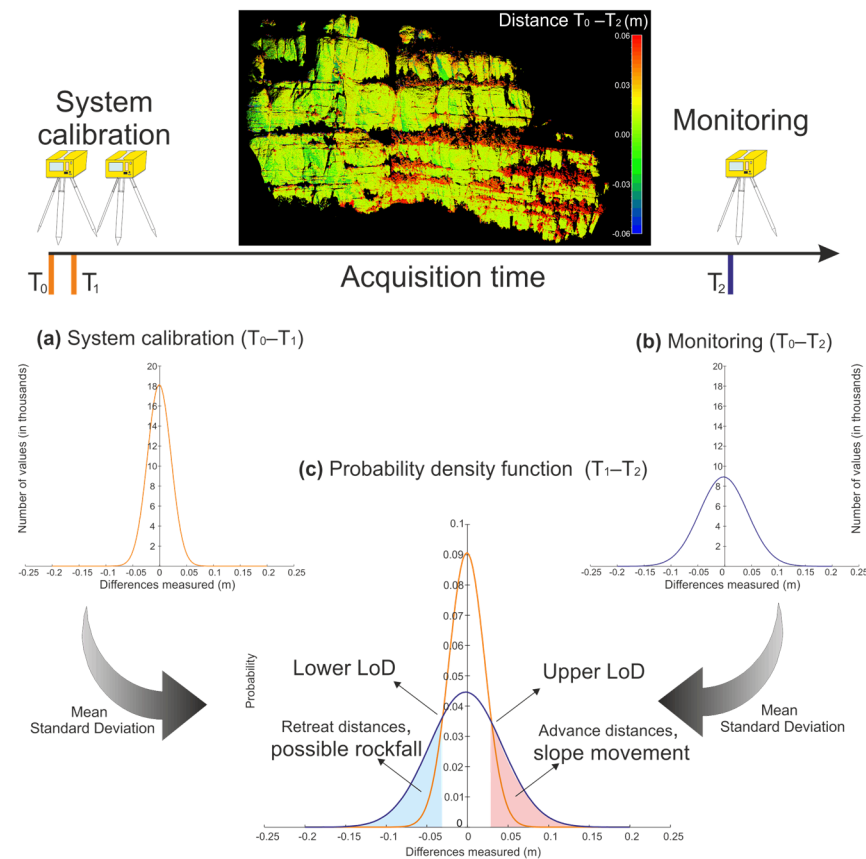


Figure 3. Depiction of the method for calculating the LoD and interpretation of the result. (a) Distribution of the differences between the initial data acquisition and the repetition in the minimum time interval (T_0 – T_1), with the aim of repeating the same conditions in order to calculate the systematic error. (b) Distribution of the feature differences calculated during the monitoring for a certain period of time (T_0 – T_2). (c) Superposition of distributions to calculate the LoD (intersection between orange and blue lines). The areas in red mark the values assigned to advanced processes, and those in blue assigned to decrease due to a loss of volume.

Factors that define the noise of the system are varied [25] and differ in each scan. During the scan, the angle of incidence of the emitted laser pulse and its return vary according to the decrease in the perpendicularity of the scene. Distance and incidence angle affect the precision of the difference values between point clouds. We avoid this loss of precision during the calibration by using only the points belonging to surfaces perpendicular to the TLS and at a representative mean distance from the outcrop. The calibration system values (defined by mean and standard deviation) have been introduced in PCM software to calculate the upper and lower LoD during monitoring.

2.3. DBSCAN Adaptation

After computing differences, it is necessary to cluster points in accordance with this feature in order to identify and reconstruct the rockfall shapes and volumes. DBSCAN algorithm [40] is widely used to cluster points [30,39,52–54]. The algorithm of clustering points conforms to the parameters of distance with respect to the other points, the values of the difference between point clouds, and a minimal number of points to define a cluster (parameters: ϵ , ϵ , and minPts). It is implemented in free software available in open-source libraries such as Open3d [55], Scikit-learn [56], or commercial software such as MATLAB.

This work proposes an adaptation of the DBSCAN algorithm that consists of analyzing the feature differences in the neighboring points during clustering in order to create four new features (Table 2); predominance, noise percentage, advance percentage, and retreat percentage. The feature predominance is defined for each reference point by the majority value classified as advance, retreat, or noise of its neighboring points according to the LoD (Figure 4). The remaining features quantify the percentage of each class in the neighboring points (advance, retreat, and noise).

Table 2. New features computed with the adaptation of the algorithm DBSCAN. The results are incorporated into the cluster event as cluster feature.

Feature	Significance
Predominance	Majority class (advance, retreat, or noise)
Noise percentage	Percentage of points classified as noise according to the LoD
Advance percentage	Percentage of points classified as advance according to the LoD
Retreat percentage	Percentage of points classified as retreat according to the LoD

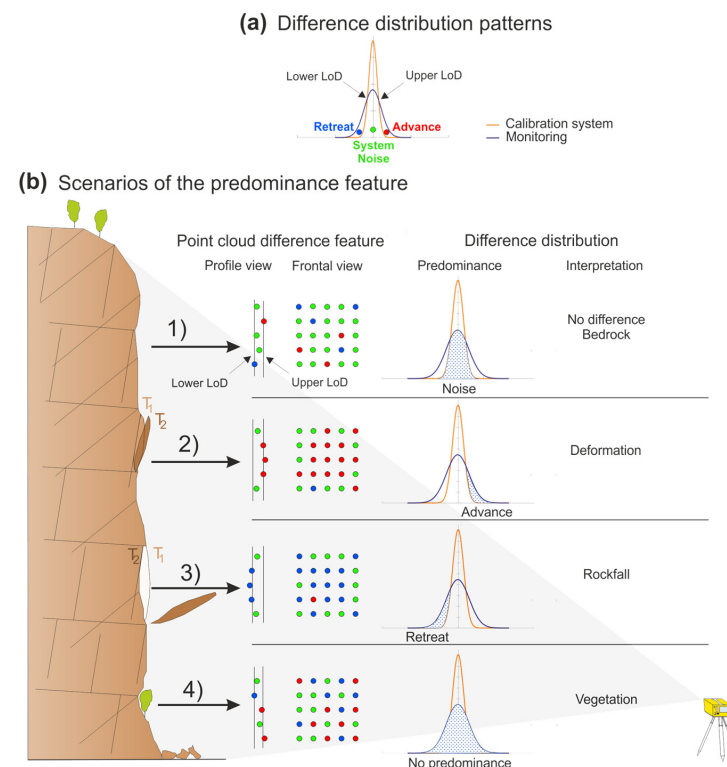


Figure 4. (a) Reference distribution of the differences between the calibration system and monitoring. (b) Possible scenarios during the monitoring of a rock cliff and its interpretation in the distribution of the differences: (1) System noise when the cluster distribution of the differences has a random distribution below the LoD. (2) Deformation when the values of differences are classified predominantly as advances. (3) Rockfall when the values of differences are predominantly classified as retreats. (4) This scenario can be interpreted as vegetation: e.g., when the advance and retreat classifications have a random distribution and exceed the LoD, and is mathematically similar to noise predominance.

Some of the features that characterize each cluster (detailed in Appendix C) have their origin in the features that characterize its points. The centroid of each cluster is represented by the center of mass of the point coordinates, computed with the principal component analysis method, and the quantifiable features are processed statistically (means and standard deviations). The specific features of each cluster, such as the volume, area or the number of points, are computed individually. The reference point clouds and the compared point clouds are triangulated separately with respect to a common plane base. The total volume corresponds to the sum of these volumes, and always preserves the positive and negative direction with respect to the TLS position.

2.4. Cluster Classification

On completion of the PCM software stage for the creation of clusters, the cluster classification step categorize clusters as rockfalls and not. This step involves different stages (see flowchart in Figure 5): (1) the training stage to train the models using a hand labeled dataset of clusters of rockfalls, and (2) the predictive (or testing) stage to identify non-classified clusters. As input of the system, we manually labeled clusters as “Candidate”, if they contain rockfalls validated with high-resolution images, and 0 otherwise (i.e., “Unknown”). We propose the categorization of “Candidate” to determine cluster that contain possible candidates to be a rockfall, and “Unknown” for clusters attributable to the rest of events. After that, we proceed with the data normalization to convert all the features in the range [0–1]. This process is necessary to avoid the classification models weigh more on some features than others.

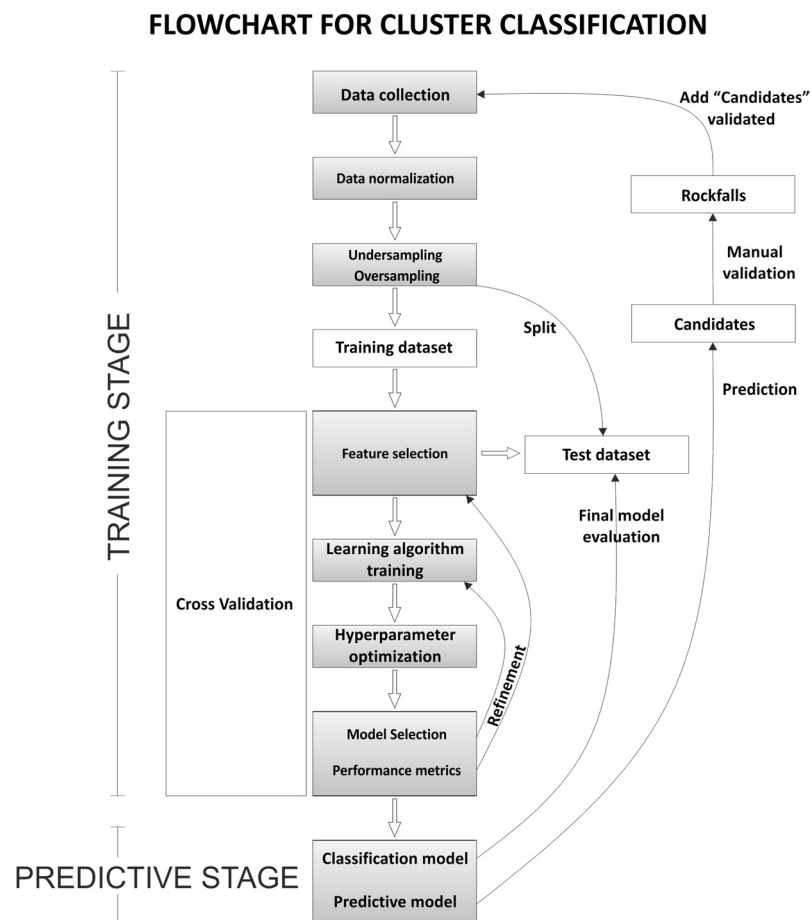


Figure 5. Flowchart of the supervised machine learning model to classifier clusters of rockfalls. Data collection corresponds to clusters of points created with PCM software.

In fact, this prior manual classification must guarantee a minimum number of clusters labeled as rockfalls. However, some scenarios in the training stage may present an imbalance between the number of clusters in each class due to the large number of “Unknown” clusters, such as in the case study of the Montserrat massif [20,57]. The number of items resulting from clustering differences, mostly consisting of the “Unknown” class (attributable to vegetation or edge effects), may range from tens to about one hundred times higher with respect to the validated “Candidate”. To correct this imbalance, Zoumpikas et al. [32] propose the implementation of resample strategies, either by reducing the majority classes (undersampling method) or synthetically increasing the number of minority classes (oversampling method). The resample methods implemented in the proposed methodology are shown in Table 3.

Table 3. Different resampling methods to correct the imbalance between classes.

Undersampling	Oversampling
Cluster Centroids	SMOTE [58] (<i>Synthetic Minority Oversampling Technique</i>)
Cluster Representatives [59]	ADASYN [60] (<i>Adaptive Synthetic Sampling</i>)
	SPIDER [61] (<i>Selective Pre-processing of Imbalanced Data</i>)
	SWIM [62] (<i>Sampling with the Majority</i>)
	Polynom-fit-SMOTE [63]
	ProWsyn [64] (<i>Proximity Weighted Synthetic</i>)
	SMOTE-IPF [65] (<i>SMOTE-Iterative Partitioning Filter</i>)
	LEE [66]
	SMOBD [67] (<i>Synthetic Minority Over-sampling Based on Samples Density</i>)
	G-SMOTE [68] (<i>Geometric-SMOTE</i>)
	LVQ-SMOTE [69] (<i>Learning Vector Quantization-SMOTE</i>)
	Assembled-SMOTE [70]
	SMOTE-TomekLinks [71]

The classification models used in this refinement process corresponds to two families, depending on the number of algorithms used in each model. The simplest “Single base learning”, i.e., models that just use one strategy to learn through a dataset and those that use multiple strategies, commonly named as “Ensemble learning” [72]. Table 4 shows the methods contemplated in this study.

Table 4. Different classifier models used to learn and define a pipeline with the nomenclature proposed by Ma et al. [72].

Single Base	Ensemble
Linear Discriminant Analysis [73]	AdaBoost Classifier [74]
Quadratic Discriminant Analysis [73]	Random Forest Classifier [73]
K-Nearest Neighbors Classifier [73]	Extra Trees Classifier [75]
Gaussian Naive Bayes [73]	XGBoost Classifier [76]
Decision Tree Classifier [73]	
Support Vector Classifier [77]	
Multi-Layer Perceptron Classifier [78]	

It should be noted that we use a stratified 10-fold cross-validation technique to assess the effectiveness of the parameters that define each classification model, evaluating and checking the models for independent datasets in a resampling procedure. The design of each model architecture is inferred by hyper-parameter tuning optimizing the scoring recall. Note that with the 10-fold cross-validation, we perform training on the 9 subsets, but we leave one subset for the evaluation of the trained model. Thus, we iterate 10 times with a different subset reserved for testing purpose each time.

In order to find out the classification model and its configuration (i.e., the hyper-parameters in Table 5 and the feature selection) that obtain the best “scoring recall”, we performed a set of experiments involving the analysis of different classification models summarized in Table 4. We symbolize this analysis as a “Refinement” in Figure 5.

Table 5. The table shown the configuration for the exhaustive search of the best parameter in each classifier model. The best parameter is implemented to obtain the best «scoring recall».

Classifier Models	Hyper-Parameters
Linear Discriminant Analysis	Solver: svd, lsqr, eigen
Quadratic Discriminant Analysis	Reg param: 0.1, 0.3, 0.5
K-Nearest Neighbors Classifier	Number of neighbors: 1, 17
Gaussian Naive Bayes	Var smoothing: logspace (0, −9, num = 100)
Decision Tree Classifier	Criterion: gini, entropy; Maximum depth: 3–15
Support Vector Classifier	C: 0.1, 1, 10; Gamma: 1, 0.01; Kernel: rbf
Multi-Layer Perceptron Classifier AdaBoost Classifier	Solver: lbfgs, SGD, ADAM; Activation: relu; Hidden layer sizes: 50, 100, 150 Number of estimators: 1–50; Learning rate: 0.2
Random Forest Classifier	Number of estimators: 1–20;
Extra Trees Classifier	Criterion: gini, entropy; Maximum depth: 3–15 Number of estimators: 1–20;
XGBoost Classifier	Criterion: gini, entropy; Maximum depth: 3–15 Nthread: 4; Booster: gblinear, gbtrees; Missing: −999 Learning rate: 0.1, 0.2, 0.3; Number of estimators: 50, 100, 500; Seed: 1337; Disable default metric: True

Once the classifier models are trained with the optimal score for recall results (fraction of rockfalls that are successfully identified), the predictor models are executed with a new monitoring collection dataset that contains unlabeled clusters. The results of the prediction models provide a list of clusters labeled as “Candidate”. The accuracy of the prediction models poses two challenges: (a) to achieve a low number of false positives (FP), i.e., clusters classified as “Candidate” but which do not correspond to real rockfalls; and (b) to avoid false negatives (FN), i.e., clusters of real rockfalls but classified as “Unknown”. The first challenge requires a great effort from the geologist to validate false cluster, while the second makes it impossible to validate candidate clusters of rockfalls. Therefore, these goals of making the real number of rockfalls is equal to TP are achieved, and the values of FP and FN are as close to zero or becomes zero. Thus, unlike with a classical methodology for selecting the best model using accuracies [32], we evaluate the results of the prediction by analyzing manually high-resolution images of the landscape, searching true positive, false positive, and false negative values.

This manual validation of the automatically classified clusters in relation to the observation of high-resolution images of the outcrop, allows to guarantee the validity of the labelling for future training and predictions. The clusters validated with this procedure as “Candidate” are added to the “Labeled Clusters” dataset to feed new trainings in future scenarios.

3. Study Sites and Processing

In this section, we detail the study sites and how the data has been collected and pre-processed.

Study Sites

The Montserrat massif is located about 40 km NW of Barcelona (Spain) with an extension of 35 km² with an elevation around 1000 m and a characteristic relief formed by rounded pinnacles and needles outlining many cliffs (see Figure 6a,b). The massif is a Natural Park and belongs to the Central Catalonia UNESCO Global Geopark as well as a tourist complex of cultural-religious heritage. The Montserrat sanctuary is located on the eastern part of the massif and attracts more than two million tourists and pilgrims per year. Thus, it requires large infrastructures, some of which are located next to cliffs with a high rockfall hazard. Precedents of rockfalls are numerous [20] all around the mountain, not excluding the sanctuary and the nearby Degotalls cliff (see Figure 6b,c).

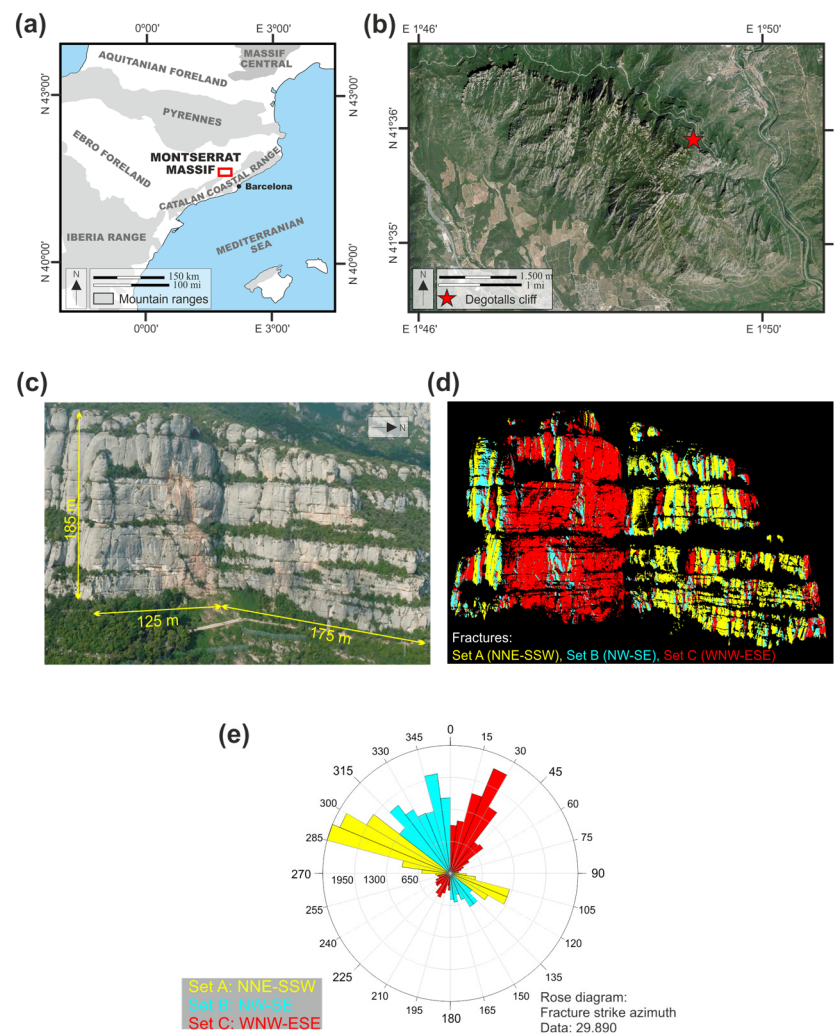


Figure 6. (a) Regional setting of the Montserrat massif, on the boundary of the Catalan Coastal Range and the Ebro Foreland Basin. (b) Degotalls Cliff location ($41^{\circ}35'54''N$, $1^{\circ}50'00''E$). Orthoimage of the area (source: Cartographic and Geological Institute of Catalonia), the Montserrat sanctuary is 600 m to the SW. (c) Degotalls study area. Left side, Degotalls N orientated E-W, and right side Degotalls E with N-S orientation. (d) Fracture orientations in Degotalls rock cliff. (e) Strike azimuth rose diagram of the fractures modeled with TLS.

Geologically, the massif consists of a succession of conglomerates (Montserrat conglomerates unit) more than 1000 m thick, interleaved by red siltstone and sandstone (La Salut and Artés Fm.) with sub-horizontal stratigraphic layers. The depositional system corresponds to a fan-delta complex accumulation in the Late-Eocene epoch along the southeastern margin of the Ebro Foreland Basin and adjacent to the Catalan Coastal Ranges [79–81].

A fracture pattern controls the morphology of the massif, with two orthogonal fracture sets oriented NNE-SSW (fracture set A) and WNW-ESE (fracture set C) [82] (as shown in Figure 6d,e). These sub-vertical, penetrative and high-frequency fracture sets cut the massif into blocks of decametric size, which together with the weathering action contribute to characterizing the peculiar landscape of the massif. The surface trace of the fractures can be followed for up to one kilometer on aerial photographs. Fracture set B with a NW-SE orientation has a lower frequency but contributes to the instability of the cliff, as well as its conjugate, with an NE-SW orientation and with a residual presence in the Degotalls area. The alphabetical order of the fracture sets defines their chronology from oldest to most recent [82]. Fractures represented in Figure 6d,e has been modeled with TLS data [83].

Due to the progressive event of rockfall failure that occurred in the North face of the Degotalls (hereafter Degotalls N) during the period 2001–2009 (see Figure 7) a risk mitigation plan was designed. In addition to protective measures, the plan triggered the monitoring of the Degotalls N and the orthogonal East face (hereafter Degotalls E) with TLS in 2007. The detachment of a high persistence fracture (fracture set C) and its fragmentation by the intersections with other fractures (fracture sets A and B) and stratigraphic layers resulted in rockfalls of decametric and metric dimensions.

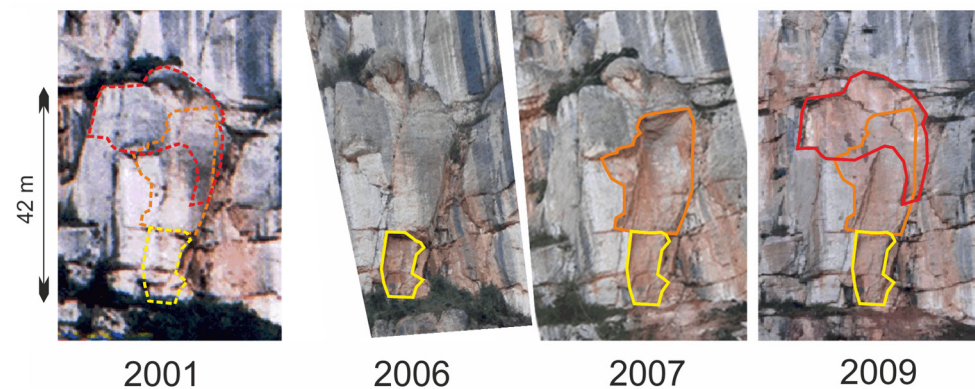


Figure 7. Depiction of the large blocks sequence detachment during the 2001–2009 period in Degotalls N cliff. The detachment was controlled by discontinuities produced by fractures and stratigraphic layers. In the 2001 image: Degotalls area before the rockfall (in dashed lines). In 2008, the surface of the cliff stabilized after different rockfalls episodes, resulting in a total rockfall volume higher than 1000 m^3 . Data collection and processing.

Rockfalls at the Degotalls cliff are classified into three categories according to the instability mechanisms and the volume [20]: (a) large blocks, which are detachments of great volumes measuring several cubic meters and controlled by mechanical discontinuities produced by fractures and stratigraphic layers; (b) pebbles or pebble aggregates of medium to small volume, generally less than 1 m^3 , caused by detachments of the matrix due to weathering or associated with small fractures; and (c) plates, corresponding to weathering flakes and thermal exfoliation in slabs with small volumes (cm^3 or dm^3).

The monitoring system consists of 13 point cloud acquisitions with TLS Optech Ilris-3D (accuracy of $\sigma = 0.7 \text{ cm}$ at 100 m) [84] over the last 14 years, as details in Figure 8a from two different TLS stations. Degotalls N requires one scanner image acquisition (see Figure 8b), while Degotalls E requires two scanner images from another TLS station, designated as the first section (North) and the second section (South), respectively, as shown in Figure 8c,d. The range of the stations is about 175 m , and the height of the cliff is 185 m . The density of the scans is approximately one point each 7 cm , and the returned intensity of 1535 nm (infrared region) is recorded as a feature. In addition, high-resolution images were acquired to validate the “Candidate” class of clusters classified by the machine learning models.

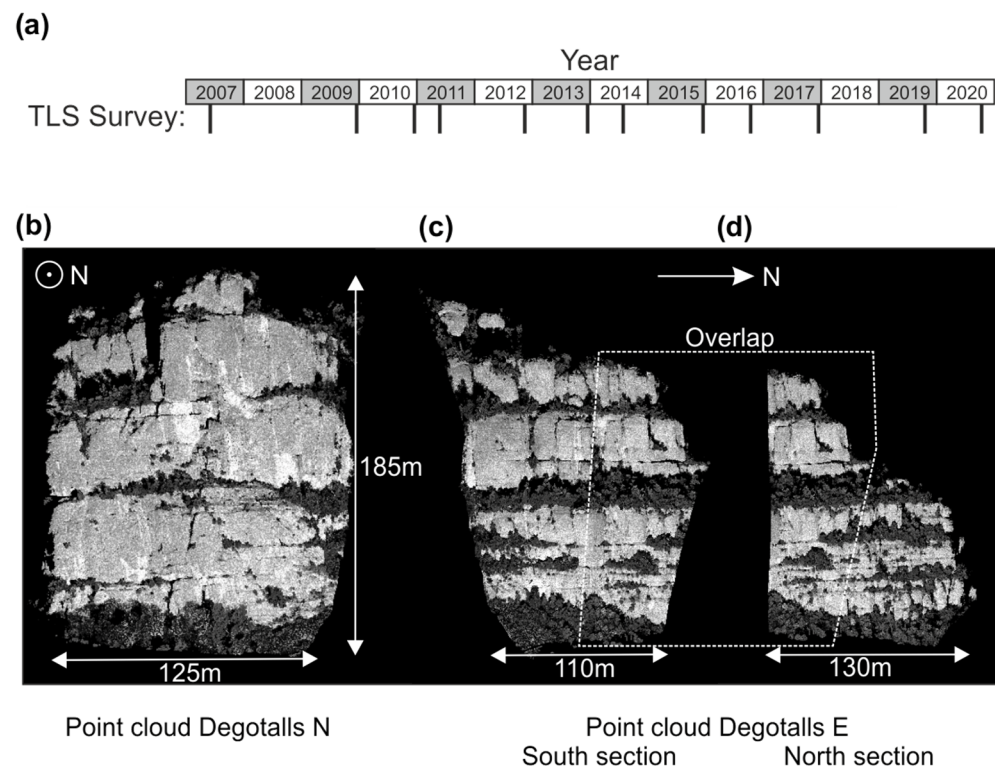


Figure 8. (a) Time series relating to the TLS surveys in the Degotalls area. All TLS acquisition were conducted with the same device from two stations (Degotalls N and Degotalls E). The survey also included high-resolution images to validate data. (b) Point cloud of the Degotalls N cliff: 2,370,000 points. Station 1 (c) South section point cloud of the Degotalls E cliff: 2,860,000 points. Station 2, orientation 1 (d) North section point cloud of the Degotalls E cliff: 2,060,000 points. Station 2, orientation 2 (re-orientation of the scanner). Point clouds and figures correspond to the texture intensity of the TLS returned signal (1530 nm).

A point cloud classification was executed to manually remove lush and easily identifiable vegetation, reducing uncertainty and facilitating the accuracy of the alignment process (see Figure 1). Degotalls scanner data were collected with point cloud format (Appendix A, with texture Intensity), measure differences (data output in Appendix B format) and clustering (data output in Appendix C format).

The data processing in the Degotalls cliffs was conducted with different strategies, the first objective being the identification of rockfalls and the second of previous deformation movements. The south section of Degotalls E cliff was compared throughout the monitoring period (2007–2020) with 12 consecutive comparisons, whereas the north section was performed with only one comparison for the same period. These different strategies of comparison are due to the reduced area of the north section and the need to identify previous deformation movements in the cliffs. Longer intervals are interpreted as sceneries more favorable to the identification of slow deformation movements. Otherwise, the Degotalls N cliff was compared in two batches, 2007–2017 and 2017–2020.

Throughout the monitoring period, point clouds were acquired from the same position, TLS device, and settings both for monitoring and calculation of the LoD of the system. Likewise, the differences between point clouds for this calibration were measured in areas with a perpendicular orientation to the TLS point position, without vegetation, and representative mean ranges to the cliff. In addition, the time interval between the calibration point clouds acquisition was 45 min. The dates of the point cloud acquisitions for monitoring are shown in Figure 8a.

4. Results

The parameters for the configuration of the double truncated cone and clustering are shown in Table 6a,b. The values were fitted according to the resolution of the point cloud and the expected cluster sizes, generally present in the most frequent antecedents. Thereby, the metric of the differences is evaluated for fitting to a Gaussian distribution, and the statistical parameters of the mean and the standard deviation are computed (Table 6c).

Table 6. Parameters to set the calibration and monitoring processes: (a) parameters used to define the geometry of the double truncated cone in the measure of the differences; (b) clustering parameters with the nomenclature equivalent to the DBSCAN algorithm; (c) mean and standard deviation to define the difference distribution for the TLS ILRIS-3D in both Degotalls areas for a mean range of 175 m.

(a) Distance Parameters	Degotalls Cliff (m)
Maximum vertical	0.5
Minimal horizontal	0.08
Maximal horizontal	0.10
(b) Clustering Settings	
Threshold distance between points (eps)	0.15
Minimum number of points (minPts)	10 points
(c) Degotalls TLS System Calibration Difference	
Mean	−0.000268
Standard deviation	0.019547

Once the differences have been computed and the LoD established (see Table 7), PCM software can process the clustering of points step. From each monitoring the number of clusters in Degotalls E oscillates around 5800 in the South section, around 2600 in the North section, and around 3700 clusters in Degotalls N.

Table 7. Statistical summary of the results of calculating the LoD in both Degotalls cliffs. This table summarizes the 13 comparisons in Degotalls E (12 in the South section and 1 in the North section) and 2 comparisons in Degotalls N.

Cliffs		LoD Mean (m)	LoD STD (m)	
Degotalls E: South section	Upper	0.03242	0.00336	
	Lower	−0.03189	0.00453	
	North section	Upper	0.03430	0.00590
		Lower	−0.03511	0.00385
Degotalls N:	Upper	0.03928	0.01685	
	Lower	−0.04026	0.01626	

The monitoring dataset of the period 2007–2009 in the South section of Degotalls E is constituted by 5957 clusters. The training dataset was manually analyzed to identify 10 real rockfalls, labeled as “Candidates”, and 1990 other clusters labeled as “Unknown”. Thus, the Cluster Classification pipeline (see flowchart in Figure 5) begins with the training stage with 1990 clusters from the “Unknown” class and 10 clusters from the “Candidate” class. The pipeline combines the 15 resampling techniques (shown in Table 3) with the 11 classification models (shown in Table 4) using 10-fold cross-validation procedure, after normalizing the dataset fitting 165 classification configurations (i.e., 15 resampling techniques times 11 classification models). Thereafter, trained models were used as predictive models with the rest of the 3957 unclassified clusters (test dataset with 5957 initial clusters, less 10 clusters “Candidate” and 1990 from the “Unknown” class) from the period 2007–2009 in order to identify new “Candidate” clusters.

The clusters labeled as “Candidate” by the 165 predictive configurations of the models were validated manually together with the 3957 clusters of the period 2007–2009, thereby identifying eight new rockfalls. Metric evaluation of the best predictive and the resampling model is shown in Tables 8 and 9 (Degotalls E, south section, 2007–2009).

Table 8. Summary of the predictive models results in the Degotalls where the disparity of the best methods is appreciated attending to the true positive (TP), false positive (FP), and false negative (FN) results. Degotalls N presents two solutions: ^a the best solution with false negatives = 0 or ^b the best solution, but accepting a reduced number of false negatives. * Initial manual classification. Clusters of “Rockfalls for training” are referred to in the “Candidate” class in the training stage. Real rockfall is referred to existent and known rockfalls on the bedrock.

Outcrop Period	Rockfalls for Training	Best Classifier Model	Best Resampling Method	Real Rockfalls	TP	FP	FN
Degotalls E							
South section							
2007–2009	10 *	Quadratic Discr.	Pol. Fit-SMOTE	8	8	91	0
2009–2010	18	Linear Discr. A.	Cluster Centr.	5	5	7	0
2010–2011	23	KNN C.	Cluster Centr.	4	4	139	0
2011–2012	27	XGBoost C.	S. TomekLinks	4	4	48	0
2012–2013	31	Extra Trees C.	Cluster Centr.	2	2	10	0
2013–2014	33	XGBoost C.	Cluster Centr.	1	1	1	0
2014–2015	34	SVC	Cluster Centr.	1	1	0	0
2015–2016	35	Linear Discr. A.	Stefanowsky	3	3	53	0
2016–2017	38	-	-	0			
2017–2019	38	Linear Discr. A.	Cluster Centr.	3	3	9	0
2019–2020	41	-	-	0			
2020–2020	41	Extra Trees C.	Stefanowsky	2	2	2	0
North section							
2007–2019	43	Quadratic Discr.	LVQ-SMOTE	22	22	97	0
Degotalls N							
2007–2017 ^a	10 *	Linear Discr. A.	Cluster Repres.	107	107	1211	0
2007–2017 ^b	10	Decision Tree C.	Cluster Repres.	107	104	296	3
2017–2019 ^a	117	Quadratic Discr.	SWIM	16	16	455	0
2017–2019 ^b	117	Quadratic Discr.	Pro WSyn	16	15	256	1

Table 9. Summary of the metric for the predictive models results with Recall and Accuracy parameters. Degotalls N presents two solutions: ^a the best solution with False Negatives = 0 or ^b the best solution, but accepting a reduced number of False Negatives.

Outcrop Period	Best Classifier Model	Best Resampling Method	Recall	Accuracy
Degotalls E				
South section				
2007–2009	Quadratic Discr.	Pol. Fit-SMOTE	1	0.979
2009–2010	Linear Discr. A.	Cluster Centr.	1	0.999
2010–2011	KNN C.	Cluster Centr.	1	0.979
2011–2012	XGBoost C.	S. TomekLinks	1	0.991
2012–2013	Extra Trees C.	Cluster Centr.	1	0.998
2013–2014	XGBoost C.	Cluster Centr.	1	0.999
2014–2015	SVC	Cluster Centr.	1	1
2015–2016	Linear Discr. A.	Stefanowsky	1	0.992
2016–2017	-	-		
2017–2019	Linear Discr. A.	Cluster Centr.	1	0.998
2019–2020	-	-		
2020–2020	Extra Trees C.	Stefanowsky	1	0.999
North section				
2007–2019	Quadratic Discr.	LVQ-SMOTE	1	0.968

Table 9. Cont.

Outcrop Period	Best Classifier Model	Best Resampling Method	Recall	Accuracy
Degotalls N				
2007–2017 ^a	Linear Discr. A.	Cluster Repres.	1	0.704
2007–2017 ^b	Decision Tree C.	Cluster Repres.	0.972	0.906
2017–2019 ^a	Quadratic Discr.	SWIM	1	0.891
2017–2019 ^b	Quadratic Discr.	Pro WSyn	0.937	0.935

Classification of the following monitoring period (2009–2010) uses the 18 validated candidates (10 + 8) and the totality of the unknown clusters from the 2007–2009 monitoring for the training stage. When it finalizes, 165 configurations of the models are fitted again to be used as predictive models with the 5100 “Unknown” clusters for the period 2009–2010. The new cluster “Candidate” proposed by the predictive models are validated again with high-resolution images and their metrics evaluated. This procedure is repeated until completion of the monitoring period and showed in Table 8, Degotalls E, South section.

The Degotalls E North section begins the learning stage with the 43 clusters labeled as “Candidate” from the South section in order to analyze one comparison monitoring of the period 2007–2019. The aim was to identify pre-deformation movements and rockfall clusters, although only 22 validated “Candidate” of rockfalls were identified (as shown in Table 8, Degotalls E, North section) and none of pre-deformation.

The Degotalls N pipeline was initialized with the manual identification of the new 10 rockfall clusters for the analysis of two comparison monitoring periods 2007–2017 and 2017–2019. The search for pre-deformation movement clusters was also negative. The results are shown in Tables 8 and 9, Degotalls N.

With the increase in the number of clusters labeled as “Candidate”, the Degotalls E South section dataset tends to decrease the number of false positives. Nevertheless, although the metric of the best classifier and resampling models reveal a high percentage of models with optimal results (Table 8). It is observed that the best classifier and resampling models are different for each comparison.

The manual cluster validation of the predictive results demonstrates the existence, especially in Degotalls E, of acceptable results in terms of true positives, false positives, and false negatives. It should be noted that the number of initial clusters for each monitoring is around 5000–6000. However, after the classification of the predictive models, there is a significant reduction in the number of clusters to be validated. Degotalls N presents an elevated number of false positives, especially when total identification of TP success is required for the results for the real number of rockfalls.

Table 10 shows the example when the first best predictive model of the first comparison in Degotalls E (i.e., quadratic discriminant analysis and polynom-fit-SMOTE) is used for the whole period. The results are not acceptable due to a large number of false positives and therefore, the large number of cluster candidates to be validated. Neither is observed a clear tendency to reduce FP.

However, there are correct identifications among the 165 configurations of the models (see Table 8), and above all the percentage of these configurations with TP is significant. It is difficult to select one model as the best one, the common practice in the machine learning field is to use a pipeline of models and perform cross-validation and to select the one or ones that obtain the performance (based on accuracy or recall, for example).

To solve this problem, we seek to validate only the clusters labeled as “Candidate” which were proposed by the totality of the predictive models to most improve the score. With this premise, the “Candidate” labeled clusters validated manually as rockfalls or true positives clusters in Degotalls E are always among the 115 clusters most predicted by the models. An initial average survey of 5800 clusters belonging to the “Unknown” class produced by each monitoring period, reveal a reduction of 98% in the number of clusters to be validated (true positives and false positives) with real rockfalls from the rock cliff (true

positives). In Degotalls N the reduction in “Candidate” clusters to be validated is 80.16% for complete identification of rockfalls, and in order to identify 96% of the real rockfalls, the reduction is 90% from a population of 3700 initial class of “Unknown” clusters.

Table 10. Quadratic discriminant analysis model and polynom-fit-SMOTE resampling results. This combination shows the best model for the 2007–2009 comparison, but not for the following comparison in true positive (TP), false positive (FP) and false negative (FN) results.

Outcrop Period	Real Rockfalls	TP	FP	FN
Degotalls E				
South section				
2007–2009	8	8	91	0
2009–2010	5	5	148	0
2010–2011	4	4	461	0
2011–2012	4	4	258	0
2012–2013	2	2	235	0
2013–2014	1	1	224	0
2014–2015	1	1	188	0
2015–2016	3	3	315	0
2016–2017	0	-	-	-
2017–2019	3	2	517	1
2019–2020	0	-	-	-
2020–2020	2	2	111	0

Figure 9 depicts one example of cluster validation, comparing the cluster images of one cluster before and after the monitoring, and visualizing the cluster features. The feature clusters are shown in Table 11.

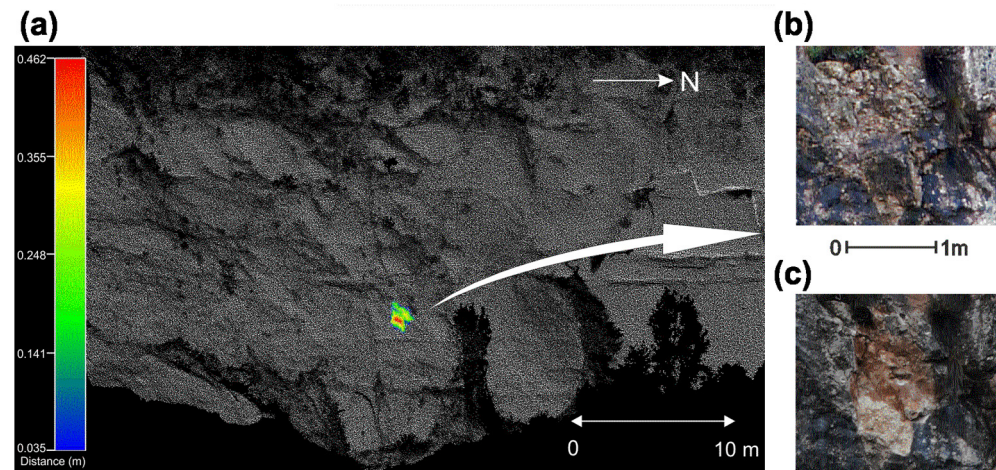
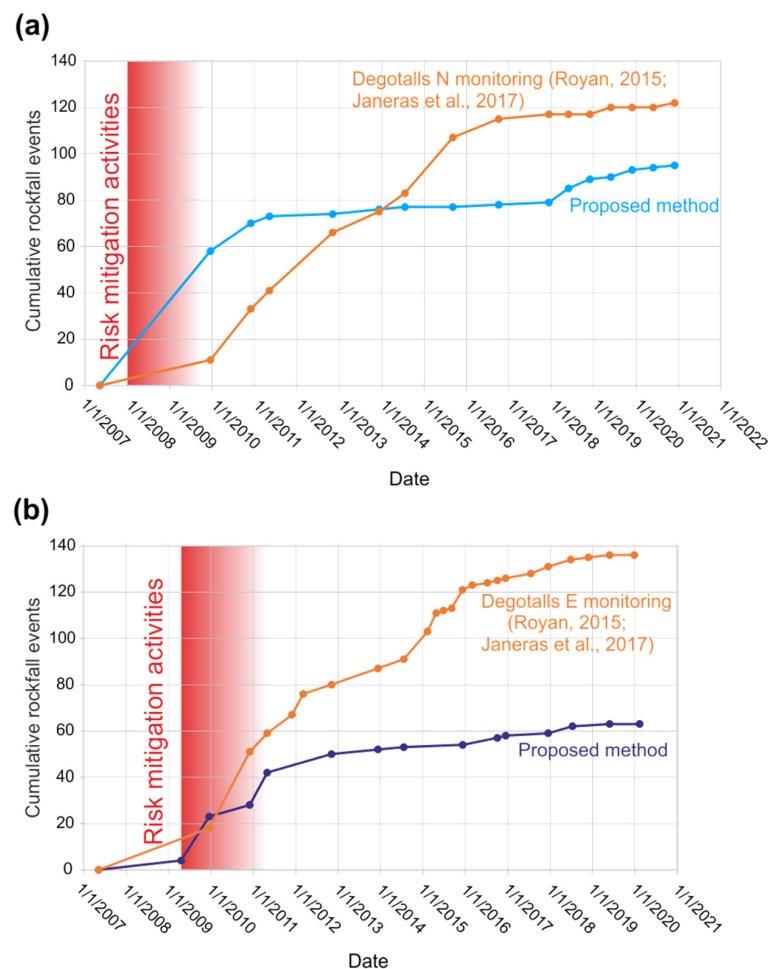


Figure 9. (a) Point cloud intensity in Degotalls E with the difference feature values of cluster #1326 (South section, period 2017–2019) shown in a multi-color scale. The dimensions in this example are 1.4 m (height) and 1 m (wide). (b) Cluster image before the rockfall. (c) Image post rockfall where the wedge and the surface of the fractures that control the detachment are visible (Fracture set A and the conjugate fracture set B with orientation NE-SW). This rockfall was classified as a large block.

The total number of clusters labeled as “Candidate” and validated as rockfall events at Degotalls E during the monitoring period was 65. At Degotalls N the number of candidates validated was 133, but 40 clusters of them correspond in fact, to only two large events occurred in December of 2008, therefore we assess 95 rockfalls (see Figure 10). This re-count is due to the configuration of the monitoring process focused on preventing the loss of small rockfalls. The number of events registered using the standard methodology [20,57] is also shown in Figure 10.

Table 11. Summary of features of the cluster shown in Figure 9. Cluster #1326 (Degotalls E, south section, period 2017–2019).

Cluster Feature	Value	Cluster Feature	Value
Cluster Number	1326	Points %: Noise	10.48%
Centroid Coord. X	−23.169 m	Advance	0.21%
Coord. Y	192.847 m	Retreat	89.31%
Coord. Z	−10.063 m	Intensity Ref.	210.32
Number of points	428	Intensity Comp.	228.43
Positive Volume	0.27164 m ³	Azimuth Ref.	165.10°
Area	1.45 m ²	Azimuth Comp.	166.20°
Predominance: Mean	2 (Retreat)	Slope Ref.	71.11°
Standard deviation	0.07	Slope Comp.	70.91°

**Figure 10.** Rockfall events in (a) Degotalls N and (b) Degotalls E Orange lines represent the number of events registered with the methodology used to date with the standard methodology for monitoring point clouds [20,57]. Cyan and Purple lines represent the results of the proposed methodology in this study. Mitigation activities are marked from start to finish in red.

5. Discussion

In this study, an inventory of rockfalls is constructed from point clouds with two new methods (PCM and Cluster Classification). PCM software is implemented to characterize and identify clusters of differences during monitoring processes, and Cluster Classification software classifies the nature of the clusters. Specifically, it is trained to classify clusters of rockfalls with machine learning techniques. The inventory was created covering the period 2007–2020 for the Degotalls cliff area (Montserrat massif, Spain). The cliff, divided into

Degotalls E and Degotalls N, accounts for 65 and 95 rockfalls, respectively, and these results are adjusted to the expected and known values of the previous studies in the area [20,57].

PCM software calculates the LoD (as shown in Table 7), and the order of magnitude of the results are similar to those used previously in the monitoring of the Degotalls and recommended in previous works [20,57]. In general, values of the LoD in Degotalls N are higher (20%) than those in Degotalls E, but this is attributable to the greatest height of the Degotalls N cliff, and therefore, with a somewhat greater TLS–cliff distance.

Measure differences and clustering points are integrated processes in PCM software and offer similar results to previous studies conducted in the area with the M3C2 and DBSCAN algorithms by Royán et al. [57] and Janeras et al. [20]. PCM software complements results with features that characterize firstly the points and subsequently, the clusters (e.g., predominance, coplanarity, or normal vector) to feed the machine learning classification process.

The results of cluster classification provide differences in the number of rockfalls counted with respect to previous works, but this fact is attributable to the different methods of validation. In this work, only those clusters labeled as “Candidate” with a clear validation in high-resolution images have been accepted. The disparate results of the predictive models for both scenarios have been interpreted in Degotalls E as promising, which opens up the possibility of studying in greater detail the importance of each feature in the contribution of the models. Moreover, this can be extended to the study of the efficiency of each resample technique and each classification model to evolve the identification of deformation clusters. At Degotalls E, the manual validation of the first 16 clusters labeled as “Candidate” most frequently proposed by the 165 predictive models identifies 65% of the rockfalls. If the manual validation is extended to the first 115 cluster candidates, the identification reaches 100% of the rockfalls. This implies a significant reduction in the initial clusters to validate them as rockfalls. The efficiency of predictive models also tends to increase when the rockfall database increases in number with new identifications in future monitoring.

Otherwise, the results from Degotalls N, with a higher percentage of clusters to be validated, the 370 clusters labeled as “Candidate” to identify 90% of the rockfalls show the potential for further improvement focused on reducing this percentage.

The analysis of these validated clusters reveals a relationship between the feature volume and the ratio of real TP identifications of the predictive models. The predictive models have a lower percentage of real TP identifications with large volume rockfalls (see Figure 11a), while in Degotalls E (shown in Figure 11b) large blocks are not presents. Degotalls N presents the clusters labeled as “Candidate” least predicted by the models belonging mainly to the category of large blocks, with volumes greater than 0.1 m³. On the other hand, the clusters corresponding to the category of plates with smaller volumes obtain the highest levels of real TP identifications in the prediction because they present more homogeneous characteristics and, therefore, facilitate the training.

Clusters of rockfalls in Degotalls N can be regarded heterogeneous (in terms of features: e.g., volume, orientation, intensity), and therefore are more difficult to learn in the training stage since the characteristics of the clusters are not as polarized as in Degotalls E. The two largest-volume clusters in Degotalls N have the lowest percentage of model predictions because they have a high degree of singularity, which does not contribute to defining a homogeneous class for the model learning. An increase in rockfalls in the dataset collection may correct this problem, but it is difficult to increase ratio between the number of “Candidate” and “Unknown” along the time with the current scenarios.

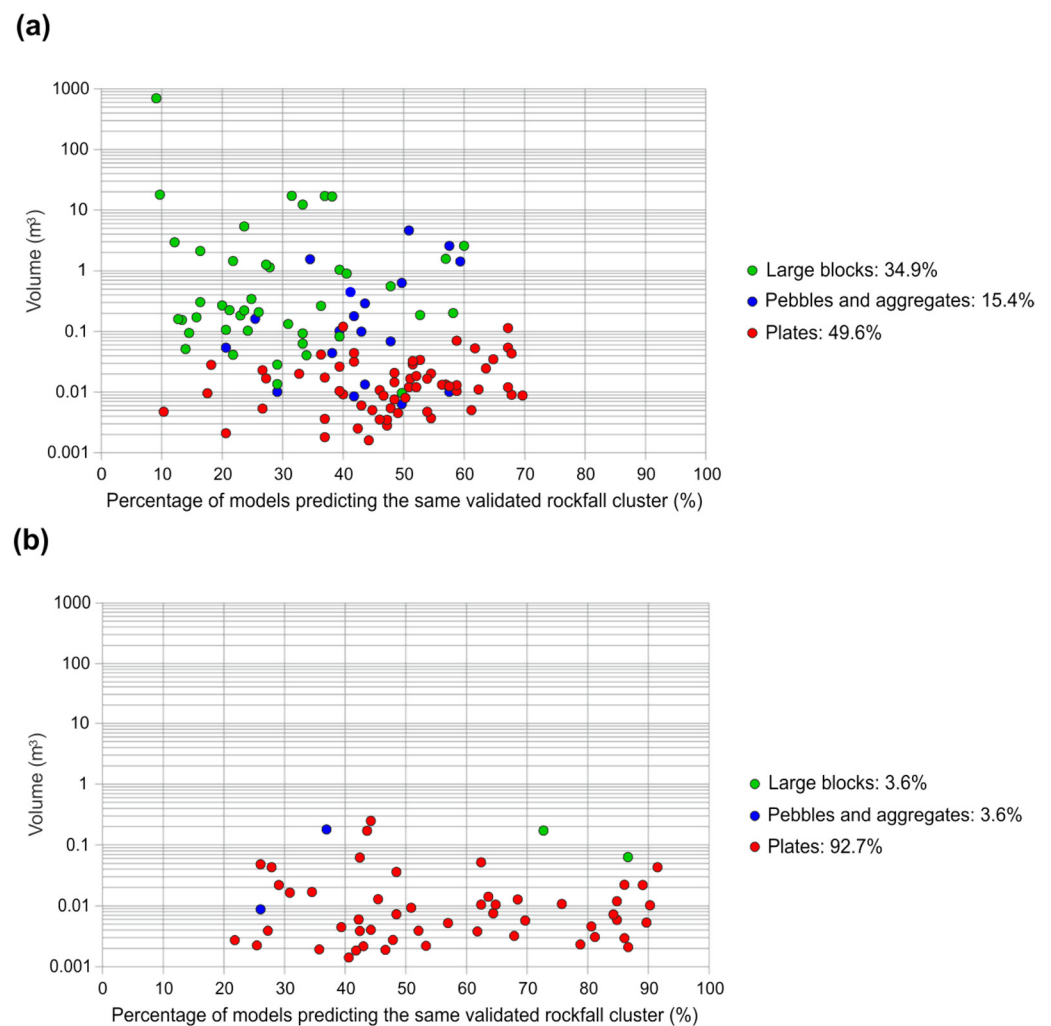


Figure 11. Relationship between the volume of the rockfall clusters, classes of rockfalls, and the percentage of models predicting the same validated rockfall clusters at the Degotalls. (a) Rockfall classes of Degotalls N are majority plates, usually associated with weathering processes with small volumes, and large blocks, due to the large detachment during the 2007–2009 period. Both classes define a heterogeneous scenario that gives rise to more difficulties in the training stage of predictive models. (b) Degotalls E presents a homogeneous class with small volumes that facilitate the identification of predictive models.

The large blocks ($>1 \text{ m}^3$) category was predominant in Degotalls N during the first two years of monitoring, and were associated with the large event controlled by fractures in 2001 and their subsequent risk mitigation activities, as it can be seen in Figure 7. In the second stage, both rock cliffs enter a stability period in which a reduced and constant number of rockfalls belonging to the category plates predominated, linked to the lingering weathering process. In Degotalls E, we observed some detachments of small volumes with some events controlled by fractures around the cubic meter of volumes. The increase in registered events in both rock cliffs (look at Figure 10) since 2018 corresponds to the category plates with small volumes (see Figure 11) of less than half a cubic meter.

The cumulative distribution of volumes registered in the Degotalls presents slightly different power law exponents in both cliffs, despite the fact that the rock mass structural conditions are the same, except rock face orientation. In the case of Degotalls N, the distribution covers a wider range of orders of magnitude due to the 2007–2009 large rockfalls event (see Figure 12). In consequence, the common sample in the range of 0.01 to 10 m^3 is the most representative of the rockfall activity that we are clearly detecting with

TLS in the Degotalls area. These results should determine the scenarios to be considered in further hazard assessment for the sanctuary parking area [85].

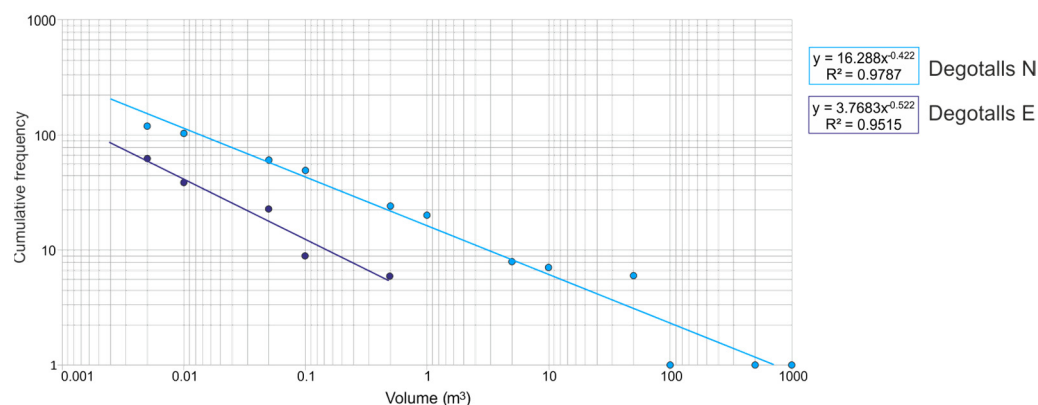


Figure 12. Cumulative frequency–rockfalls volume in both Degotalls cliffs. The power law functions are depicted on the upper left for each scenario. The values of volume are grouped into intervals.

A possible new strategy to optimize the results in Degotalls N is to feed the training stage with each different category (large block, pebble, and plate), but always considering a significant number of “Candidate” labeled clusters.

On comparison with the results from the standard methodology in the Degotalls [20,57], a variation is found to exist in the amount and temporal distribution of the events. This can undoubtedly be attributed to the reduction in the number of clusters to be validated allowing an increase in the quality of the validation that reduces doubtful cases. Furthermore, the objective assessment of each cluster as a candidate for a rockfall improves the process, and contributes to a better interpretation of the rock cliff evolution and the valuation of risk mitigation activities.

Analysis of the nature of the clusters has been unable to identify any pre-deformation process with which to create a new class. The study of the area in years prior to the rockfall has been unproductive. The reason for this non-identification may be due to using a temporal resolution that is too low to record this process. Possibly, the modification of the temporal and spatial resolution during the monitoring allows the identification of pre-deformed clusters, and therefore enables their classification with our methodology.

The time required to process the measure differences process may constitute a limitation of the process, even though it is an automatic process. Several factors are involved in controlling this step, and it may be convenient to minimize this time.

6. Conclusions

Monitoring rock cliffs to identify rockfalls with point clouds requires dataset processing methods, many of which are developed by the research community. Periodic digital capture of cliff surfaces with instruments such as TLS does not directly provide an inventory of rockfalls. For this aim, it is necessary to process the captured data and configure the temporal and spatial resolutions of the TLS in accordance with the dynamics of the rockfalls from the cliff. The capture of point clouds, their alignment, the measure of the differences, and the clustering are more highly evolved aspects. However, classification of the clusters is a factor that is not so frequently addressed. The present work proposes a solution to this issue based on machine learning and predictive models.

In this paper, we propose two developments, PCM and Cluster Classification applications, to automatically classify clusters with machine learning by taking advantage of the fact that clusters that contain rockfalls have similar characteristics, a fact that facilitates the learning stage. As observed in Degotalls N, singular rockfalls (those covering large volumes) are the most complicated to learn. The proposed modifications of the already existing algorithms that deal with the creation of clusters have been very useful for measure 33 features especially significant during the classification. However, the discrimination

power of each feature introduced in the learning process has not been tested in this work, an issue that should be addressed in future studies.

The machine learning process implies the development of 165 prediction models based on the 11 classification models, combined with 15 resampling methods (13 undersampling and 2 oversampling) to balance the unbalances between the number of clusters in the rockfall class of the non-rockfall class. A total of 11 classification models have used the balanced data to classify the rockfalls class using 10-fold cross-validation and hyper-parameterization techniques.

In summary, the following conclusions are derived from this work:

- Monitoring rockfalls in rock cliffs with point cloud is a difficult task that can benefit from machine learning strategies, provided that both techniques are appropriately combined. We validate this assumption with the attempt to identify rockfalls in the rock cliff of the Montserrat massif (Spain).
- We have observed the difficulty of correlating classification models, trained with clusters of rockfalls, with the best prediction model. For this reason, we use all the combinations of prediction models to validate the most proposed candidates.
- The success of the rockfall prediction models depends on the homogeneity/heterogeneity of the features that characterize the different categories of the rockfall clusters (large blocks, pebbles and plates) used to train the classification models.
- Rockfalls in the Degotalls (Montserrat, Spain) are currently in a phase of stabilization, and those that occur are of small volume and attributable to plates associated with weathering processes. However, since 2018 a slight increase in cases has been observed.

Author Contributions: L.B. and D.G.-S. conceptualized, created the code, co-acquired data, analyzed, writing, and editing the manuscript; methodology, writing—review and editing M.G.; algorithm and writing—review and editing, T.Z., A.P. and M.S.; writing—review, O.G., J.A.M., M.J. and O.P.; validation and writing—review. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Spanish MINEICO with the projects SABREM (PID2020-117598GB-100, funded by MCIN/AEI/10.13039/501100011033), PROMONTEC (CGL2017-84720-R AEI/FEDER, EU) and SALTEC CGL2017-85532-P (AEI/FEDER, EU) and AGAUR (Agència de Gestió d’Ajuts Universitaris i de Recerca) project 2016 DI 069, and the European Union’s Horizon 2020 research and innovation program under the grant agreement Marie Skłodowska-Curie No 860843. Data from the Montserrat massif were funded by the Institut Cartogràfic i Geològic de Catalunya (ICGC). Anna Puig and Maria Salamó also thank the Generalitat de Catalunya for its support under project 2017-SGR-341.



Marie Skłodowska-Curie
Actions

Data Availability Statement: Data are available in public repositories indicated in Section 2. Upon request, more data may be available.

Acknowledgments: Xavier Blanch and Manuel J. Royán for their contributions in the early processing of the Montserrat dataset. We would like to thank Nicolás Pascual González for his contribution to this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Format point cloud.

Point coordinates
1. Coordinate X
2. Coordinate Y
3. Coordinate Z
With intensity
4. Intensity
Or RGB texture
5. Red
6. Green
7. Blue
Or intensity and texture
8. Intensity
9. Red
10. Green
11. Blue

Appendix B

Table A2. Features associated with the point cloud after calculating differences. * No computed as features. ¹ Only computer with RGB format is used. ² Only computed with Intensity format is used.

1. n *	Reference point index
2. m	Compared point index
3. Coordinate X	Point reference coordinates
4. Coordinate Y	
5. Coordinate Z	
6. Code_n *	Reference index texture (0 n/a, 1 Intensity, 2 RGB, 3 RGB + Int)
7. R ¹	Texture reference points RGB format
8. G ¹	
9. B ¹	
10. Intensity ²	Reference intensity texture
11. Vector_i	Reference vector normal vector components
12. Vector_j	
13. Vector_k	
14. Orientation	Reference strike azimuth (degree)
15. Dip	Reference strike slope (degree)
16. Collinearity	Reference point index of collinearity
17. Coplanarity	Reference point index of coplanarity
18. Selected	Number of points to calculate the normal vector
19. Distance	Distance selected between closest and average
20. Vertical Distance	Vertical distance along vector with direction (+ or −)
21. Horizontal Distance	Horizontal distance component between points
22. Distance closest	Shorter distance between Refer. and Comp. point
23. Coordinate X	Point compared coordinates
24. Coordinate Y	
25. Coordinate Z	
26. Code_m *	Compared index texture (0 n/a, 1 Intensity, 2 RGB, 3 RGB + Int)
27. R ¹	Texture compared points RGB format
28. G ¹	
29. B ¹	
30. Intensity ²	Compared intensity texture
31. vector_i	Compared vector normal vector components
32. vector_j	
33. vector_k	

Table A2. *Cont.*

34. Orientation	Compared strike azimuth (degree)
35. Dip	Compared strike slope (degree)
36. Collinearity	Compared point index of collinearity
37. Coplanarity	Compared point index of coplanarity
38. Selected	Number of points to calculate the normal vector
39. Angle	Angle between Ref. and Comp. normal vectors
40. Angle_Direction	Angle with direction
41. Minimal_distance	Shortest distance between those inscribed in the geometric figure
42. Average_distance	Average distance between those inscribed in the geometric figure
43. Maxima_distance	Longest distance between those inscribed in the geometric figure
44. Dev. Stand_distance	Dev.Stand distance between those inscribed in the geometric figure
45. Selected points	Number of points inscribed in the geometric figure

Appendix C

Table A3. Features associated to the Clusters, format order. * No used in ML classification. ¹ Only computer with RGB format is used. ² Only computed with Intensity format is used.

1. Cluster identification *	
2. Coordinate X *	
3. Coordinate Y *	Coordinates of cluster centroid
4. Coordinate Z *	
5. Item_number *	Cluster differences number
6. Points_number	Number of points
7. TotalVolume	Cluster total volume
8. PositiveVolume	Volume behind the reference surface and TLS
9. NegativeVolume	Volume in front the reference surface and TLS
10. Area	Planimetric cluster area 2D, perpendicular to TLS
11. Code *	Cluster classification (Unknown, Candidate)
12. Confidence *	Confidence index
13. Predominance_Mean	Mean predominance (noise 0, advance 1, retreat 2)
14. Predominance_Sigma	STD predominance classification (0, 1, 2)
15. Percentage_1_Mean	Mean advance predominance (1)
16. Percentage_1_Sigma	STD advance predominance (1) classification
17. Percentage_0_Mean	Mean noise predominance (0)
18. Percentage_0_Sigma	STD noise predominance (0) classification
19. Percentage_2_Mean	Mean retreat predominance (2)
20. Percentage_2_Sigma	STD retreat predominance (2) classification
21. OrientationSetsRef	Reference cluster strike azimuth
22. OrientationSetsCom	Compared cluster strike azimuth
23. IndexTextureRef *	Reference texture index (0, 1 Int, 2 RGB, 3 RGB + Int)
24. R_Mean_Ref ¹	
25. R_Sigma_Ref ¹	
26. G_mean_Ref ¹	
27. G_Sigma_Ref ¹	
28. B_mean_Ref ¹	Texture. Mean & Std of reference clusters.
29. B_Sigma_Ref ¹	
30. I_Mean_Ref ²	
31. I_Sigma_Ref ²	
32. IndexTextureCom *	Compared texture index (0, 1 Int, 2 RGB, 3 RGB + Int)
33. R_mean_Com ¹	
34. R_Sigma_Com ¹	
35. G_mean_Com ¹	
36. G_Sigma_Com ¹	
37. B_mean_Com ¹	Texture. Mean & Std dev. of compared clusters
38. B_Sigma_Com ¹	
39. I_Mean_Com ²	

Table A3. Cont.

40. I_Sigma_Com ²	
41. AziRef_Mean	Mean strike azimuth of Reference points
42. SloRef_Mean	Mean strike slope of Reference points
43. AziCom_Mean	Mean strike azimuth of Compared points
44. SloCom_Mean	Mean strike slope of Compared points
45. CopRef_Mean	Mean coplanarity of Reference points
46. CopRef_Sigma	STD coplanarity of Reference points
47. ColRef_Mean	Mean collinearity of Reference points
48. ColRef_Sigma	STD collinearity of Reference points
49. CopCom_Mean	Mean coplanarity of Compared points
50. CopCom_Sigma	STD coplanarity of Compared points
51. ColCom_Mean	Mean collinearity of Compared points
52. ColCom_Sigma	STD collinearity of Compared points
53. ang_Mean	Mean angularity between normal vectors
54. ang_Sigma	STD angularity between normal vectors
55. Reference File *	String
56. Compared File *	String

References

1. Erisman, T.H.; Abele, G. *Dynamics of Rockslides and Rockfalls*; Springer: Berlin/Heidelberg, Germany, 2001; ISBN 978-3-642-08653-3.
2. Whalley, W.B. Rockfalls. In *Slope Instability*; Brunson, D., Prior, D.B., Eds.; Wiley: Chichester, UK, 1984; pp. 217–256.
3. Hung, O.; Leroueil, S.; Picarelli, L. The Varnes Classification of Landslide Types, an Update. *Landslides* **2014**, *11*, 167–194. [[CrossRef](#)]
4. DiFrancesco, P.-M.; Bonneau, D.; Hutchinson, D.J. The Implications of M3C2 Projection Diameter on 3D Semi-Automated Rockfall Extraction from Sequential Terrestrial Laser Scanning Point Clouds. *Remote Sens.* **2020**, *12*, 1885. [[CrossRef](#)]
5. Volkwein, A.; Schellenberg, K.; Labiouse, V.; Agliardi, F.; Berger, F.; Bourrier, F.; Dorren, L.K.A.; Gerber, W.; Jaboyedoff, M. Rockfall Characterisation and Structural Protection—A Review. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 2617–2651. [[CrossRef](#)]
6. Corominas, J.; Copons, R.; Moya, J.; Vilaplana, J.M.; Altimir, J.; Amigó, J. Quantitative Assessment of the Residual Risk in a Rockfall Protected Area. *Landslides* **2005**, *2*, 343–357. [[CrossRef](#)]
7. van Veen, M.; Hutchinson, D.J.; Kromer, R.; Lato, M.; Edwards, T. Effects of Sampling Interval on the Frequency—Magnitude Relationship of Rockfalls Detected from Terrestrial Laser Scanning Using Semi-Automated Methods. *Landslides* **2017**, *14*, 1579–1592. [[CrossRef](#)]
8. Williams, J.G.; Rosser, N.J.; Hardy, R.J.; Brain, M.J. The Importance of Monitoring Interval for Rockfall Magnitude-Frequency Estimation. *J. Geophys. Res. Earth Surf.* **2019**, *124*, 2841–2853. [[CrossRef](#)]
9. Ritchie, A.M. Evaluation of Rockfall and Its Control. *Highw. Res. Rec.* **1963**, *17*, 13–28.
10. Sturzenegger, M.; Stead, D. Quantifying Discontinuity Orientation and Persistence on High Mountain Rock Slopes and Large Landslides Using Terrestrial Remote Sensing Techniques. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 267–287. [[CrossRef](#)]
11. Abellán, A.; Oppikofer, T.; Jaboyedoff, M.; Rosser, N.J.; Lim, M.; Lato, M.J. Terrestrial Laser Scanning of Rock Slope Instabilities. *Earth Surf. Process. Landf.* **2014**, *39*, 80–97. [[CrossRef](#)]
12. Abellán, A.; Derron, M.-H.; Jaboyedoff, M. “Use of 3D Point Clouds in Geohazards” Special Issue: Current Challenges and Future Trends. *Remote Sens.* **2016**, *8*, 130. [[CrossRef](#)]
13. Telling, J.; Lyda, A.; Hartzell, P.; Glennie, C. Review of Earth Science Research Using Terrestrial Laser Scanning. *Earth Sci. Rev.* **2017**, *169*, 35–68. [[CrossRef](#)]
14. Santana, D.; Corominas, J.; Mavrouli, O.; Garcia-Sellés, D. Magnitude–Frequency Relation for Rockfall Scars Using a Terrestrial Laser Scanner. *Eng. Geol.* **2012**, *145–146*, 50–64. [[CrossRef](#)]
15. Corominas, J.; Mavrouli, O.; Ruiz-Carulla, R. Rockfall Occurrence and Fragmentation. In *Advancing Culture of Living with Landslides*; Springer International Publishing: Cham, Switzerland, 2017; pp. 75–97. [[CrossRef](#)]
16. Fanti, R.; Gigli, G.; Lombardi, L.; Tapete, D.; Canuti, P. Terrestrial Laser Scanning for Rockfall Stability Analysis in the Cultural Heritage Site of Pitigliano (Italy). *Landslides* **2013**, *10*, 409–420. [[CrossRef](#)]
17. Mazzanti, P.; Schilirò, L.; Martino, S.; Antonielli, B.; Brizi, E.; Brunetti, A.; Margottini, C.; Scarascia Mugnozza, G. The Contribution of Terrestrial Laser Scanning to the Analysis of Cliff Slope Stability in Sugano (Central Italy). *Remote Sens.* **2018**, *10*, 1475. [[CrossRef](#)]
18. Lague, D.; Brodu, N.; Leroux, J. Accurate 3D Comparison of Complex Topography with Terrestrial Laser Scanner: Application to the Rangitikei Canyon (N-Z). *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 10–26. [[CrossRef](#)]
19. Tonini, M.; Abellán, A. Rockfall Detection from Terrestrial Lidar Point Clouds: A clustering approach using R. *J. Spat. Inf. Sci.* **2013**, *8*, 95–110. [[CrossRef](#)]
20. Janeras, M.; Jara, J.-A.; Royán, M.J.; Vilaplana, J.-M.; Aguasca, A.; Fàbregas, X.; Gili, J.A.; Buxó, P. Multi-technique Approach to Rockfall Monitoring in the Montserrat Massif (Catalonia, NE Spain). *Eng. Geol.* **2017**, *219*, 4–20. [[CrossRef](#)]

21. Bonneau, D.; DiFrancesco, P.M.; Jean Hutchinson, D. Surface Reconstruction for Three-Dimensional Rockfall Volumetric Analysis. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 548. [[CrossRef](#)]
22. Bonneau, D.A.; Hutchinson, D.J. The Use of Terrestrial Laser Scanning for the Characterization of a Cliff-Talus System in the Thompson River Valley, British Columbia, Canada. *Geomorphology* **2019**, *327*, 598–609. [[CrossRef](#)]
23. Hendrickx, H.; Le Roy, G.; Helmstetter, A.; Pointner, E.; Larose, E.; Braillard, L.; Nyssen, J.; Delaloye, R.; Amaury, F. Timing, Volume and Precursory Indicators of Rock and Cliff Fall on a Permafrost Mountain Ridge (Mattertal, Switzerland). *Earth Surf. Process Landf.* **2022**, *47*, 1532–1549. [[CrossRef](#)]
24. Rosser, N.; Lim, M.; Petley, D.; Dunning, S.; Allison, R. Patterns of Precursory Rockfall Prior to Slope Failure. *J. Geophys. Res.* **2007**, *112*, 148–227. [[CrossRef](#)]
25. Kromer, R.; Hutchinson, D.; Lato, M.; Gauthier, D.; Edwards, T. Identifying Rock Slope Failure Precursors Using LiDAR for Transportation Corridor Hazard Management. *Eng. Geol.* **2015**, *195*, 93–103. [[CrossRef](#)]
26. Carrea, D.; Abellan, A.; Derron, M.H.; Jaboyedoff, M. Automatic Rockfalls Volume Estimation Based on Terrestrial Laser Scanning Data. In *Engineering Geology for Society and Territory—Volume 2: Landslide Processes*; Springer International Publishing: Cham, Switzerland, 2015; pp. 425–428, ISBN 9783319090573.
27. Blanch, X.; Eltner, A.; Guinau, M.; Abellan, A. Multi-Epoch and Multi-Imagery (MEMI) Photogrammetric Workflow for Enhanced Change Detection Using Time-Lapse Cameras. *Remote Sens.* **2021**, *13*, 1460. [[CrossRef](#)]
28. Kromer, R.; Walton, G.; Gray, B.; Lato, M.; Group, R. Development and Optimization of an Automated Fixed-Location Time Lapse Photogrammetric Rock Slope Monitoring System. *Remote Sens.* **2019**, *11*, 1890. [[CrossRef](#)]
29. Williams, J.; Rosser, N.J.; Hardy, R.; Brain, M.; Afana, A. Optimising 4-D Surface Change Detection: An Approach for Capturing Rockfall Magnitude–Frequency. *Earth Surf. Dyn.* **2018**, *6*, 101–119. [[CrossRef](#)]
30. Schovanec, H.; Walton, G.; Kromer, R.; Malsam, A. Development of Improved Semi-Automated Processing Algorithms for the Creation of Rockfall Databases. *Remote Sens.* **2021**, *13*, 1479. [[CrossRef](#)]
31. Eberhardt, E.; Stead, D.; Coggan, J.S. Numerical Analysis of Initiation and Progressive Failure in Natural Rock Slopes—the 1991 Randa Rockslide. *Int. J. Rock Mech. Min. Sci.* **2004**, *41*, 69–87. [[CrossRef](#)]
32. Zoumpikas, T.; Puig, A.; Salamó, M.; García-Sellés, D.; Blanco-Nuñez, L.; Guinau, M. An Intelligent framework for End-to-End Rockfall Detection. *Int. J. Intell. Syst.* **2021**, *36*, 6471–6502. [[CrossRef](#)]
33. Weidner, L.; Walton, G.; Kromer, R. Classification Methods for Point Clouds in Rock Slope Monitoring: A Novel Machine Learning Approach and Comparative Analysis. *Eng. Geol.* **2019**, *263*, 105326. [[CrossRef](#)]
34. Brodu, N.; Lague, D. 3D Terrestrial Lidar Data Classification of Complex Natural Scenes Using a Multi-Scale Dimensionality Criterion: Applications in Geomorphology. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 121–134. [[CrossRef](#)]
35. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [[CrossRef](#)]
36. Evans, J.S.; Hudak, A.T. A Multiscale Curvature Algorithm For Classifying Discrete Return LiDAR in Forested Environments. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1029–1038. [[CrossRef](#)]
37. Kromer, R.; Lato, M.; Hutchinson, D.J.; Gauthier, D.; Edwards, T. Managing Rockfall Risk through Baseline Monitoring of Precursors Using a Terrestrial Laser Scanner. *Can. Geotech. J.* **2017**, *54*, 953–967. [[CrossRef](#)]
38. Mazzanti, P.; Caporossi, P.; Brunetti, A.; Mohammadi, F.I.; Bozzano, F. Short-Term Geomorphological Evolution of the Poggio Baldi Landslide Upper Scarp via 3D Change Detection. *Landslides* **2021**, *18*, 2367–2381. [[CrossRef](#)]
39. Royán, M.J.; Abellán, A.; Jaboyedoff, M.; Vilaplana, J.M.; Calvet, J. Spatio-Temporal Analysis of Rockfall Pre-Failure Deformation Using Terrestrial LiDAR. *Landslides* **2014**, *11*, 697–709. [[CrossRef](#)]
40. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland, OR, USA, 2–4 August 1996; Simoudis, E., Fayyad, U., Han, J., Eds.; AAAI Press: Menlo Park, CA, USA, 1996; pp. 226–231.
41. Girardeau-Montaut, D.; Roux, M.; Marc, R.; Thibault, G. Change Detection on Points Cloud Data Acquired with a Ground Laser scanner. In *Proceedings of the ISPRS WG III/3, III/4, V/3 Workshop “Laser Scanning 2005”*, Enschede, The Netherlands, 12–14 September 2005; Vosselman, G., Brenner, C., Eds.; 2005; pp. 30–35. Available online: <https://www.isprs.org/proceedings/xxxvi/3-w19/> (accessed on 21 June 2022).
42. Innovmetric. Polyworks. Quebec City. 2022. Available online: <https://www.innovmetric.com> (accessed on 18 May 2022).
43. Visual Studio 2019. Microsoft. Available online: <https://visualstudio.microsoft.com> (accessed on 18 May 2022).
44. Barnhart, T.B.; Crosby, B.T. Comparing Two Methods of Surface Change Detection on an Evolving Thermokarst Using High-Temporal-Frequency Terrestrial Laser Scanning, Selawik River, Alaska. *Remote Sens.* **2013**, *5*, 2813–2837. [[CrossRef](#)]
45. Cignoni, P.; Rocchini, C.; Scopigno, R. Metro: Measuring Error on Simplified Surfaces. *Comput. Graph. Forum* **1998**, *17*, 167–174. [[CrossRef](#)]
46. Kazhdan, M.; Bolitho, M.; Hoppe, H. Poisson Surface Reconstruction. In *Eurographics Symposium on Geometry Processing*; Sheffer, A., Pothier, K., Eds.; The Eurographics Association, 2006. Available online: <http://diglib.org.org/handle/10.2312/SGP.SGP06.061-070> (accessed on 18 May 2022).
47. Girardeau-Montaut, D. CloudCompare, Version 2.12.1 Alpha. Available online: <http://www.cloudcompare.org/> (accessed on 18 May 2022).

48. Abellán, A.; Jaboyedoff, M.; Oppikofer, T.; Vilaplana, J.M. Detection of Millimetric Deformation Using a Terrestrial Laser Scanner: Experiment and Application to a Rockfall Event. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 365–372. [CrossRef]
49. Jolliffe, I. Principal Component Analysis. In *International Encyclopedia of Statistical Science*; Springer: Berlin/Heidelberg, Germany, 2011. [CrossRef]
50. Woodcock, N. Specification of fabric shapes using an Eigenvalue method. *Geol. Soc. Am. Bull.* **1977**, *88*, 1231–1236. [CrossRef]
51. García-Sellés, D.; Falivene, O.; Arbués, P.; Gratacós, O.; Tavani, S.; Muñoz, J.A. Supervised Identification and Reconstruction of Near-Planar Geological Surfaces from Terrestrial Laser Scanning. *Comput. Geosci.* **2011**, *37*, 1584–1594. [CrossRef]
52. Benjamin, J.; Rosser, N.J.; Brain, M.J. Emergent Characteristics of Rockfall Inventories Captured at a Regional Scale. *Earth Surf. Process Landf.* **2020**, *45*, 2773–2787. [CrossRef]
53. Carrea, D.; Abellan, A.; Derron, M.-H.; Gauvin, N.; Jaboyedoff, M. MATLAB Virtual Toolbox for Retrospective Rockfall Source Detection and Volume Estimation Using 3D Point Clouds: A Case Study of a Subalpine Molasse Cliff. *Geosciences* **2021**, *11*, 75. [CrossRef]
54. Wang, Y.; Xiao, J.; Liu, L.; Wang, Y. Efficient Rock Mass Point Cloud Registration Based on Local Invariants. *Remote Sens.* **2021**, *13*, 1540. [CrossRef]
55. Zhou, Q.-Y.; Park, J.; Koltun, V. Open3D: A Modern Library for 3D Data Processing. *arXiv* **2018**, arXiv:1801.09847. [CrossRef]
56. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. [CrossRef]
57. Royan, M. Rockfall Characterization and Prediction by Means of Terrestrial LiDAR. Ph.D. Thesis, Universitat de Barcelona, Barcelona, Spain, September 2015. Available online: <http://hdl.handle.net/10803/334400> (accessed on 12 June 2022).
58. Yen, S.J.; Lee, Y.S. Cluster-Based Under-Sampling Approaches for Imbalanced Data Distributions. *Expert Syst. Appl.* **2008**, *36*, 5718–5727. [CrossRef]
59. Chawla, N.; Bowyer, K.; Hall, L.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
60. He, H.; Bai, Y.; Garcia, E.A.; Li, S. ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–8 June 2008; pp. 1322–1328. [CrossRef]
61. Stefanowski, J.; Wilk, S. Selective Pre-processing of Imbalanced Data for Improving Classification Performance. In *Data Warehousing and Knowledge Discovery*; Song, I.Y., Eder, J., Nguyen, T.M., Eds.; DaWaK 2008; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5182, pp. 283–292. [CrossRef]
62. Sharma, S.; Bellinger, C.; Krawczyk, B.; Zaiane, O.; Japkowicz, N. Synthetic Oversampling with the Majority Class: A New Perspective on Handling Extreme Imbalance. In Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM), Singapore, 17–20 November 2018; pp. 447–456. [CrossRef]
63. Gazzah, S.; Amara, N.E. New Oversampling Approaches Based on Polynomial Fitting for Imbalanced Data Sets. In Proceedings of the Eighth IAPR International Workshop on Document Analysis Systems, Nara, Japan, 16–19 September 2008; pp. 677–684. [CrossRef]
64. Barua, S.; Islam, M.; Murase, K. ProWSyn: Proximity Weighted Synthetic Oversampling Technique for Imbalanced Data Set Learning. In *Advances in Knowledge Discovery and Data Mining*; Pei, J., Tseng, V.S., Cao, L., Motoda, H., Xu, G., Eds.; Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); PAKDD Springer: Berlin/Heidelberg, Germany, 2013; Volume 7819, pp. 317–328. [CrossRef]
65. Sáez, J.A.; Luengo, J.; Stefanowski, J.; Herrera, F. SMOTE-IPF: Addressing the Noisy and Borderline Examples Problem in Imbalanced Classification by a re-Sampling Method with Filtering. *Inf. Sci.* **2015**, *291*, 184–203. [CrossRef]
66. Lee, J.; Kim, N.; Lee, J.-H. An Over-Sampling Technique with Rejection for Imbalanced Class Learning. In Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication, Bali, Indonesia, 8–10 January 2015; ACM: New York, NY, USA; pp. 1–6. [CrossRef]
67. Cao, Q.; Wang, S. Applying Over-Sampling Technique Based on Data Density and Cost-Sensitive SVM to Imbalanced Learning. In Proceedings of the 4th International Conference on Information Management, Innovation Management and Industrial Engineering, Shenzhen, China, 26–27 November 2011; Volume 2, pp. 543–548. [CrossRef]
68. Douzas, G.; Bação, F. Geometric SMOTE: Effective Oversampling for Imbalanced Learning Through a Geometric Extension of SMOTE. *arXiv* **2017**, arXiv:1709.07377. [CrossRef]
69. Nakamura, M.; Kajiwara, Y.; Otsuka, A.; Kimura, H. LVQ-SMOTE—Learning Vector Quantization based Synthetic Minority Over-sampling Technique for biomedical data. *BioData Min.* **2013**, *6*, 16. [CrossRef]
70. Zhou, B.; Yang, C.; Guo, H.; Hu, J. A Quasi-Linear SVM Combined with Assembled SMOTE for Imbalanced Data Classification. In Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA, 4–9 August 2013; pp. 1–7. [CrossRef]
71. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *ACM SIGKDD Explor. Newsl.* **2004**, *6*, 20–29. [CrossRef]
72. Ma, Z.; Mei, G.; Piccialli, F. Machine Learning for Landslides Prevention: A Survey. *Neural Comput. Appl.* **2021**, *33*, 10881–10907. [CrossRef]
73. Hastie, T.; Friedman, J.; Tibshirani, R. *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2001. [CrossRef]

74. Awad, M.; Khanna, R. Support Vector Machines for Classification. In *Efficient Learning Machines*; Apress: Berkeley, CA, USA, 2015; pp. 39–66. [[CrossRef](#)]
75. Murtagh, F. Multilayer Perceptrons for Classification and Regression. *Neurocomputing* **1991**, *2*, 183–197. [[CrossRef](#)]
76. Zhu, J.; Zou, H.; Rosset, S.; Hastie, T. Multi-class AdaBoost. *Stat. Interface* **2009**, *2*, 349–360. [[CrossRef](#)]
77. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely Randomized Trees. *Mach. Learn.* **2006**, *63*, 3–42. [[CrossRef](#)]
78. Chen, T.; Guestrin, C. XGBoost. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; ACM: New York, NY, USA; pp. 785–794. [[CrossRef](#)]
79. Anadón, P.; Marzo, M.; Puigdefàbregas, C. The Eocene fan-delta of Montserrat (Southeastern Ebro Basin, Spain). In *6th European Meeting Excursion Guidebook*; Milà, M.D., Rosell, J., Eds.; IAS/Institut d’Estudis Ilerdencs: Lleida, Spain, 1985; pp. 109–146.
80. López-Blanco, M.; Marzo, M.; Burbank, D.W.; Vergés, J.; Roca, E.; Anadón, P.; Piña, J. Tectonic and Climatic Controls on the Development of Foreland Fan Deltas: Montserrat and Sant Llorenç Del Munt Systems (Middle Eocene, Ebro Basin, NE Spain). *Sediment. Geol.* **2000**, *138*, 17–39. [[CrossRef](#)]
81. Gómez-Paccard, M.; López-Blanco, M.; Costa, E.; Garcés, M.; Beamud, E.; Larrasoana, J.C. Tectonic and Climatic Controls on the Sequential Arrangement of an Alluvial Fan/Fan-Delta Complex (Montserrat, Eocene, Ebro Basin, NE Spain). *Basin Res.* **2012**, *24*, 437–455. [[CrossRef](#)]
82. Alsaker, E.; Gabrielsen, R.H.; Roca, E. The Significance of the Fracture Pattern of the Late-Eocene Montserrat Fan-Delta, Catalan Coastal Ranges (NE Spain). *Tectonophysics* **1996**, *266*, 465–491. [[CrossRef](#)]
83. García-Sellés, D.; Sarmiento, S.; Gratacós, O.; Granada, P.; Carrera, N.; Lakshminantha, M.R.; Cordova, J.C.; Muñoz, J.A. Fracture analog of the sub-Andean Devonian of southern Bolivia: Lidar applied to Abra Del Condor. In *Petroleum Basins and Hydrocarbon Potential of the Andes of Peru and Bolivia*; Zamora, G., McClay, K.M., Ramos, V., Eds.; AAPG Memoir, 2018; Volume 117, pp. 577–612. Available online: <https://pubs.geoscienceworld.org/books/book/2153/chapter-abstract/120760614/Fracture-Analog-of-the-Sub-Andean-Devonian-of?redirectedFrom=fulltext> (accessed on 18 May 2022).
84. Teledyne Optech. *ILRIS Summary Specification Sheet*; Teledyne Optech Incorporated: Vaughan, ON, Canada, 2014.
85. Mineo, S.; Pappalardo, G.; Mangiameli, M.; Campolo, S.; Mussumeci, G. Rockfall Analysis for Preliminary Hazard Assessment of the Cliff of Taormina Saracen Castle (Sicily). *Sustainability* **2018**, *10*, 417. [[CrossRef](#)]

A.3. Article 2: An intelligent framework for end-to-end rockfall detection

An intelligent framework for end-to-end rockfall detection

Thanasis Zoumpikas¹  | Anna Puig¹  | Maria Salamó¹  |
David García-Sellés²  | Laura Blanco Nuñez^{3,4} | Marta Guinau² 

¹Department of Mathematics and Computer Science, WAI Research Group, IMUB and UBICS Institutes, University of Barcelona, Barcelona, Spain

²Department of Earth and Ocean Dynamics, RISKINAT Research Group, Geomodels Institute, University of Barcelona, Barcelona, Spain

³Department of Earth and Ocean Dynamics, GGAC Research Group, Geomodels Institute, University of Barcelona, Barcelona, Spain

⁴Anufra—Soil and Water Consulting, Barcelona, Spain

Correspondence

Thanasis Zoumpikas, Department of Mathematics and Computer Science, WAI Research Group, IMUB and UBICS Institutes, University of Barcelona, Barcelona 08007, Spain.
Email: thanasis.zoumpikas@ub.edu

Funding information

Marie Skłodowska-Curie Actions, Grant/Award Number: 860843

Abstract

Rockfall detection is a crucial procedure in the field of geology, which helps to reduce the associated risks. Currently, geologists identify rockfall events almost manually utilizing point cloud and imagery data obtained from different capture devices such as Terrestrial Laser Scanner (TLS) or digital cameras. Multitemporal comparison of the point clouds obtained with these techniques requires a tedious visual inspection to identify rockfall events which implies inaccuracies that depend on several factors such as human expertise and the sensibility of the sensors. This paper addresses this issue and provides an intelligent framework for rockfall event detection for any individual working in the intersection of the geology domain and decision support systems. The development of such an analysis framework presents major research challenges and justifies exhaustive experimental analysis. In particular, we propose an intelligent system that utilizes multiple machine learning algorithms to detect rockfall clusters of point cloud data. Due to the extremely imbalanced nature of the problem, a plethora of state-of-the-art resampling techniques accompanied by multiple models and feature selection procedures are being investigated. Various machine

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *International Journal of Intelligent Systems* published by Wiley Periodicals LLC

learning pipeline combinations have been examined and benchmarked applying well-known metrics to be incorporated into our system. Specifically, we developed machine learning techniques and applied them to analyze point cloud data extracted from TLS in two distinct case studies, involving different geological contexts: the basaltic cliff of Castellfolit de la Roca and the conglomerate Montserrat Massif, both located in Spain. Our experimental results indicate that some of the above-mentioned machine learning pipelines can be utilized to detect rockfall incidents on mountain walls, with experimentally validated accuracy.

KEYWORDS

geology, imbalanced classification, intelligent systems, machine learning, rockfall monitoring

1 | INTRODUCTION

In the field of geology, one crucial task is rockfall detection, which helps to reduce the risk of future hazards.¹ Recently, technologies that are able to characterize the geometrical properties of rock slopes and cliffs have emerged.² As a result, the possibility of detecting changes in a cliff with high precision has been increased greatly.

Currently, geoscientists utilize specific methodologies to detect changes in rock slopes mainly by comparing measurements at different points in time.³ However, the problem with such techniques is that they rely greatly on the sensitivity of the sensor that captures the data or even the measurement tool. Consequently, due to potential measurement errors, the users need to examine case by case analyzing the nature of the change in the cliff, that is, vegetation, edge effect, noise or random objects. For instance, it is often difficult to distinguish rock detachments from other changes optically in point cloud or imagery data.

The issues presented above justify the need for a new automated intelligent system designed to detect rockfalls. Therefore, in this paper, we propose an intelligent framework for rockfall detection. At the time of writing, a thorough search for such a system has failed to yield results. In addition, although there are recent studies such as Reference [4] that develop rockfall identification frameworks using point cloud data, they do not address the data imbalance issues, which are significant in the geology field. We consider our study capable of elucidating the field towards the implementation of a general machine learning classification framework able to handle cases with extremely imbalanced data. The level of data imbalance in this study is significant. This is justified in the later sections of this article since there is a low number of rockfall labeled instances. Depending on the nature of the scenery, the vegetation or edge effects can be numerous compared to rockfall events.

In this article, we examine several of the intelligent methods dealing with rockfall and landslide detection, implement the relative machine learning models accompanied by various

resampling strategies to handle the imbalanced nature of the study and apply them on point cloud data to identify rockfall events. Two well differentiated geological environments have been selected as case studies. The first one is located at the Montserrat Massif (Barcelona, Spain), which corresponds to a fractured conglomerate cliff, called Degotalls. The second case study involves data from the basaltic lava flow cliff of Castellfollit de la Roca (Girona, Spain). The acquired data are preprocessed and used as input in the learning stage of our intelligent framework. We systematically compare various models to select the most accurate for each case study. Then, our prototype system utilizes the aforementioned selected models to perform effective rockfall event identifications.

Our framework provides a novel intelligent solution for the geoscientists, which may be incorporated as a stand-alone component in a geological decision support system targeted to rockfall monitoring. Its novelty is also derived from the fact that it is a ready-to-use geological point cloud machine learning software able to deal with imbalanced data, which is the only published work providing such a stand-alone rockfall detection system at the time of writing. Also, our results portray great robustness and generalizability performance, which is of utmost importance in rockfall detection.

Finally, the main contributions of this article are the following:

- We propose a full intelligent framework for rockfall detection handling highly imbalanced data.
- We conduct experiments with real data from two distinct case studies to validate the efficacy and effectiveness of our proposed intelligent system.
- We develop a web-based rockfall detection system.
- We provide a baseline methodology and a detection accuracy benchmark for future related experimental analyses.

The paper is organized as follows. Section 2 contains a synopsis of related efforts and Section 3 summarizes the design specifications and the implementation of the developed framework, while also provides the background information needed to understand the whole procedure of detecting rockfall events. In Section 4, we present the experimental analysis of the examined intelligent process, concerning common performance metrics for imbalanced classification, and the prototype web-based Rockfall Detection System. Section 5 includes the final observations of our experiments. Finally, Section 6 includes our concluding comments and potential issues for further investigation.

2 | RELATED WORK

Various research studies have been conducted on the intersection of the geology and machine learning domains. Over recent years, several machine learning applications have emerged in the geoscience field and there is a growing enthusiasm for intelligent methods.⁵ However, the geoscience domain presents new and special challenges for machine learning algorithms and methodologies, because of the combinations of geoscience properties encountered in each specific case. In addition, there is an open need for novel machine learning research and automatic intelligent analysis, especially in the geosciences field as highlighted by Reference [6]. The automation of intelligent pipelines performing specific, special analysis will bring significant advances in both the geoscience and machine learning domains, especially in the

task of detecting objects and events, estimating and long-term predictions for geoscience variables and in extracting knowledge from geoscience data.

The detection of deformations and rockfall events have been studied in the literature for a couple of decades.^{7,8} The number of scientific articles on the use of modern and advanced sensors such as Terrestrial Laser Scanner (TLS) in rockfall and landslide studies has escalated considerably in the last years. However, further research into the development of new methodologies using such sensors is required to address and improve the current time consuming analytical methods due to the high volume of data that these sensors produce.⁹⁻¹³ Research studies on the *identification* of landslide and rockfall events and on general classification tasks in the geology domain present variations regarding their input data, feature inventories, methodologies, and in general their specific objectives.

Regarding the input data, we could categorize the related research studies into three broad categories: (i) the imagery data studies; (ii) point cloud studies; and (iii) other sensory data studies. Considering imagery data studies, remote-sensing techniques are considered to be among the most important for landslide event *detection* and *monitoring*. Multiple efforts by researchers working in the remote sensing field applied in geology appear in the literature, utilizing intelligent methods on high-resolution Earth Observation imagery data.¹⁴ Imagery data seem to be a popular data source utilized in numerous studies. Recently published articles, such as References [15-21] use images as their main data source. In addition, researchers involved in Reference [22] constructed 3D images using a multiview stereo algorithm and consumer-grade cameras.

On the other hand, the use of TLS devices to extract point cloud data for the categorization of landslide kinematics is significant. Also, research studies show that it seems to be necessary to fuse point cloud data from various sources, namely TLS combined with Airborne Laser Scanner (ALS) or Structure from Motion (SfM) methods combined with TLS, to overcome the limitations of each individual technique.²³ Point clouds are heavily utilized by various researchers, mainly because they carry detailed, high quality information. Specifically, recent studies, including References [4,24-30] use point clouds as their main data source for their analysis. Moreover, advanced research studies use both images and point clouds portraying a speciality and differentiate themselves from the majority of the related studies.³¹⁻³³ Furthermore, Reference [34] use images and TLS point cloud data to model slopes. Last but not least, Reference [35] utilize data from wireless sensor networks and machine learning algorithms to monitor and forecast landslides in real time.

Several researchers have compiled a brief selection of the features used for landslide and rockfall identification. These features can be categorized in five groups, namely morphological, hydrological, geological, land cover features and features obtained from other sources, like rainfall intensity, according to Reference [14]. Pure geological features like the lithology of a studied field or geo-structural information appear in older studies such as Reference [36]. Recently, various scientists have utilized explanatory variables derived from basic cartographic operations on geological records. These variables are commonly used to form land cover features based on distances to faults, roads, and rivers, which are measured using simple spatial operations in GIS software.³⁷ In addition, Reference [20] use combined morphological predictors, namely elevation, aspect, curvature, slope, with hydrological ones like wetness index and rainfall intensity data. Similarly, Reference [38] utilize morphological and hydrological features adding multiple curvature measurements, such as plan and surface curvature to explain the mechanisms of landslides and to confirm the state of vegetation, roads and, in general, geometric deformations.

Furthermore, there are variations regarding the methodologies utilized in rockfall and landslide *identification* and *forecasting*, ranging from methods using traditional geology and statistics to more intelligent ones using machine learning approaches such as deep learning and neural networks. Multiple efforts have been presented introducing analysis frameworks able to deal with landslide and rockfall events.

Traditional geology is mainly utilized in research articles published more than 10 years ago, such as Reference [39] although recent papers still investigate the application of such methods.² Lague et al.¹¹ proposed a technique, which compares and combines two different sets of point clouds obtained from a TLS device. According to their methodology, two different time frames of point cloud data are combined and later compared to identify whether the input data form deformation or rockfall event. The aforementioned method uses traditional statistics and lacks generalization performance.

Machine learning-based methodological workflows are the most used techniques, as there are various research articles utilizing them. Recent research studies such as References [25-30,32,33,35] utilize mainly machine learning algorithms for rock-slope and landslide *monitoring* and *analysis*. More elaborate machine learning methods outperform the majority, examples include Reference [40], who developed a landslide monitoring approach for TLS point cloud data, which integrates a specialized machine learning classification involving topological rules in an object-based framework of analysis.

Although deep learning and neural networks are highly popular in other domains, such as the biomedical domain, achieving significant results in classification tasks,^{41,42} are currently utilized in fewer studies in the geology domain, such as References [43,44]. However, advanced neural network architectures using specialized loss functions are being employed to tackle specific issues in the field such as data imbalance in landslide analysis, showing important advances in terms of performance.¹⁴ Also, Xiao et al.¹⁵ present a landslide susceptibility assessment framework based on deep learning algorithms using multisource imagery data. A wide assortment of recently published research studies combines and compares machine learning, neural networks, and ensemble-based techniques to achieve higher performance in terms of accuracy and generalizability.^{4,15-21,31,45}

The data imbalanced issue has been addressed in multiple fields, for instance, in classification approaches in the biomedical domain.^{46,47} Ijaz et al.⁴⁶ used Synthetic Minority Over-Sampling Technique (SMOTE) combined with Random Forest (RF) classifier to deal with imbalanced data and model diseases. In the geoscience domain, the landslide and rockfall *identification* are two tasks that present significant imbalanced data issues. Prakash et al.¹⁴ approached the imbalanced learning task by utilizing sophisticated loss functions in the training phase and data augmentation techniques. Stumpf et al.⁴⁸ proposed a machine learning-based analysis framework with a repetitive strategy to handle class imbalance. Zhao et al.⁴⁹ tackled minor data imbalance issues in predicting landslide susceptibility by utilizing a voting system and the random processing of samples with a random forest algorithm.

In a more recent framework,⁴ proposed combining ALS and TLS data with GIS and introducing a hybrid ensemble model and a 3D kinematic rockfall forecasting model able to deal with rockfall hazard assessment, achieving promising accuracy. However, they did not address the imbalanced nature of the study or deal with imbalanced data. Our proposal considers a novel machine learning approach that also deals with the problem of imbalanced data. Specifically, we propose an intelligent analysis framework and rockfall detection decision support software that incorporate various processing stages, which are considered to be essential for such problems, namely clustering, resampling, model parameterization, and feature selection.

Regarding the input data sources, we consider only point cloud data acquired by a TLS device from two distinct geological contexts, with features derived mostly from the morphology of the terrain.

3 | END-TO-END MACHINE LEARNING FRAMEWORK: DESIGN AND IMPLEMENTATION

This section introduces our developed end-to-end machine learning framework for detecting rockfall events. First, we explain the basic procedure of detecting rockfalls. Then, we analytically present our proposed framework.

3.1 | Background on detecting rockfalls

One of the most common processes in the detection of rockfall events on mountain cliffs and slopes is explained in Reference [3] and depicted in Figure 1, specifically in Steps (A)–(E). The methodology implies capturing periodical measurements of the same cliff face at different time frames with a TLS, denoted as Step (A) in Figure 1. TLS is a measuring device which offers the ability to collect dense point-clouds of objects. It also provides high-precision and high-accuracy data and is widely used in the geology domain.⁵⁰ After the capture, the procedure continues with the detection of changes at the surface of the cliff from point cloud data comparison, which is mostly performed with a technique called *m3c2*,¹¹ denoted as Step (B). As a result, a new point cloud is obtained containing the metric distances between the compared dense clouds. A clustering algorithm is then applied in Step (C) following the change detection step outputting a set of clusters, which are aggregations of points with a significant distance value. These clusters allow for the management of subsets of point clouds with specific topological properties. A statistical analysis then generates a set of features (Step (D)) to be evaluated manually, through a visual inspection with expert criteria, to determine whether they are rockfalls or random noise (Step (E)). In most cases, samples, that is, the clusters, happen to be random noise or even measurement errors by the scanner itself.

The current semimanual classification task of the TLS point cloud data for the detection of rockfall events presents two main challenges. The first concerns the sensibility of the sensor because in some cases the detection of movements is smaller than the device's margin of error. The second concerns the process of distinguishing rock movement events from other kinds of events, such as the movement of the sensor between measurements, the appearance of

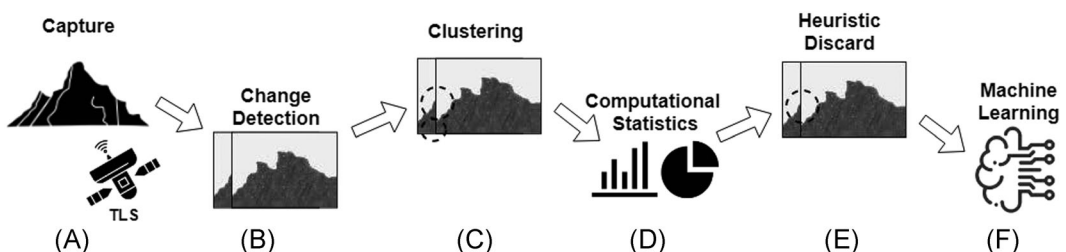


FIGURE 1 Procedure for detecting rockfalls. TLS, terrestrial laser scanner

vegetation or even random noise. The use of clustering techniques helps to mitigate the aforementioned issues.

Machine learning algorithms present an interesting and currently widely accepted solution for the automatic classification of TLS point cloud data for the detection of rockfall events. However, rockfall detection is considered to be a highly imbalanced classification task, due to the rarity of a rockfall event in a relative data set. Moreover, clustered point cloud data present patterns that are not easily distinguishable, while having high dimensionality, due to the considerable number of features. In addition, automating the above-mentioned process seems to be an interesting solution for an individual working in the geology domain. There is an open need for a consistent and concise intelligent rockfall identification framework, as detailed by Reference [6]. In Section 3.2, we propose a developed machine learning framework, which represents Step (F) in Figure 1.

3.2 | Framework overview

An outline of our proposal can be seen in Figure 3, where we synthesize our framework by providing its methodology workflow, and in Figure 2, where we present its component-based system architecture.

First, we collect point cloud data in different time frames (Figure 2A) using the TLS device and store it in a database (Figure 2B). Next, we preprocess it (Figure 2C) to be fed in our analysis framework. Then, we statistically analyse the input data, obtaining descriptive statistics to elucidate the field regarding the data. The obtained data set is then stored and normalized. Subsequently, to prevent issues of imbalance, we define our resampling module,

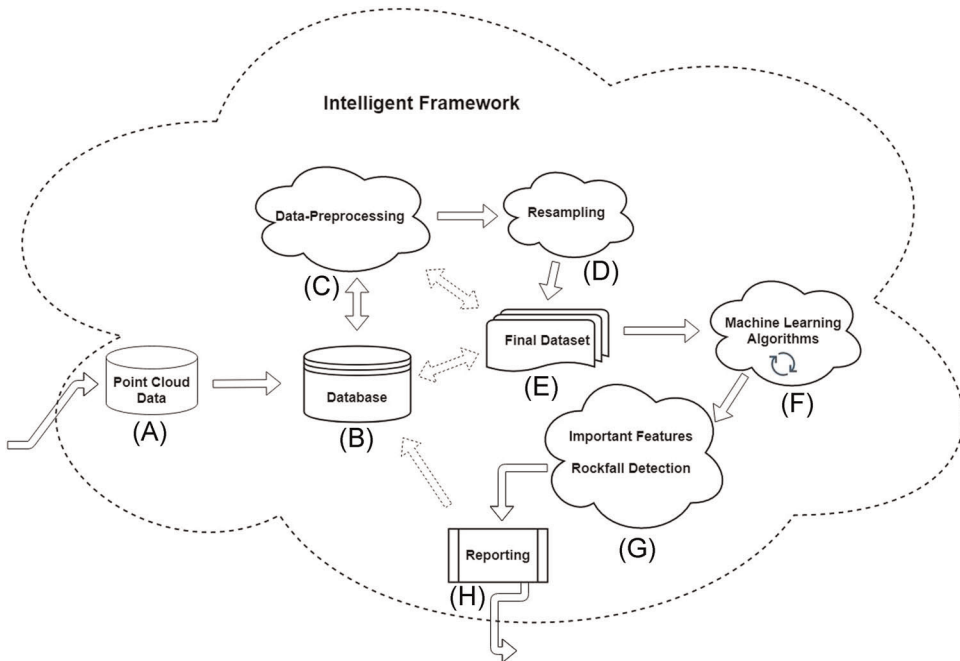


FIGURE 2 Intelligent framework—methodology workflow [Color figure can be viewed at wileyonlinelibrary.com]

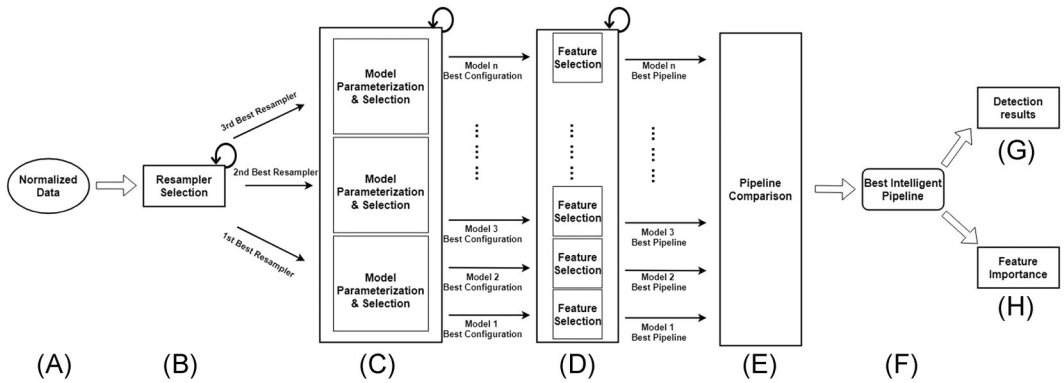


FIGURE 3 Component-based system architecture

portrayed in Figure 2D, which balances the normalized data (Figure 2E) to be used as input for machine learning algorithms (Figure 2F). Several intelligent machine learning pipelines are utilized to detect rockfall events and to identify the most significant features (Figure 2G). Finally, a reporting component is used to display the results in a convenient and illustrative manner, as depicted in Figure 2H.

Regarding our methodology, we initially feed normalized point cloud data in our intelligent framework as depicted in Figure 3A. Then, in the resampling process (Figure 3B), an assortment of several resampling strategies accompanied by multiple machine learning models are investigated to handle this imbalanced classification task. The three best resamplers are then selected and paired with a variety of models in the model selection and parametrization stage (Figure 3C). The completion of the hyper-parameterization stage follows the feature selection phase (Figure 3D), in which the best parameterized models of each model variant accompanied by each one of the three best resamplers are experimentally evaluated using different numbers of features. The output of the aforementioned stage consists of multiple properly parameterized machine learning pipelines (Figure 3E). These pipelines are then statistically compared to decide on the best intelligent pipeline to be used in our prototype system (Figure 3F), which is able to produce detection results (Figure 3G) and rank feature importance (Figure 3H) with promising accuracy. In the following sub-sections, we analytically present our framework and methodological approach.

3.3 | Data collection

The data used in this study has been freely provided by the RISKMAT* research group and the GEOMODELS** Research Institute, which belong to the University of Barcelona.

In this article, we utilize two distinct case studies. Regarding the first case study, we use point cloud data measured from the Degotalls cliff in Montserrat Massif, located in Barcelona, Spain. The data set consists of clustered TLS point cloud data aggregated from temporal point cloud measurements, from 2007 to 2020 in eight nearly regular time steps. The second case study includes clustered TLS point cloud data measured from the cliff of Castellfollit de la Roca, located in Girona, Spain, from 2008 to 2012. We would like to highlight that the two landscapes are very different in terms of their geological and imagery properties.

Actually, in each case study, the data is composed of a set of statistics computed from the point cloud data, using a technique based on the *m3c2* methodology,¹¹ which creates clusters of points, as described in Section 3.1, including several numerical features that are used in the proposed workflow to identify rockfalls. Then, the aforementioned statistical set is fed in our analysis framework. The Degotalls data consist of 6004 instances, that is, clusters, and has 37 distinct numerical features, while the Castellfollit data consist of 10,371 instances and has 31 numerical features. The aforementioned features include the coordinates of the point clouds and other statistically computed values. A list of these variables with a brief explanation can be found in Appendix S1.

3.4 | Data analysis

A data exploration analysis of the point cloud data is provided in this section. This analysis is critical because it elaborates the entire data collection process. Particularly, in Tables 4 and 5 in Appendix S2, we display the summary statistics for all the data features used in the Degotalls and Castellfollit case studies, respectively, supplying the foundation for our subsequent analysis. For clarification purposes, we chose to use an abbreviation for each variable of the data. The complete variable names followed by a brief explanation can be found in Appendix S1.

The classification labels of the clusters are determined by the event that causes the change at the surface of the cliff. Rockfall events are denoted as “Candidate.” Also, we call “Precursor” the events in which the rock presents a small movement before rockfall. In addition, the vegetation of the cliffs is denoted as “Vegetation” and the unknown or human-based events as “Unknow.” The artifacts due to the edge effects or the TLS noise are denoted as “Limit_effect.”

Observing the graphs in Figure 4, it is clear that the data in each case study are considered highly imbalanced. In the Degotalls case study (Figure 4A), we have only 65 rockfall candidate samples compared to the total of 6004 samples. On the other hand, the Castellfollit case study (Figure 4C) includes only 38 compared to the total of 10,371 samples. Please note, that the labeling process is done by expert geoscientists with a visual inspection of the 6004 and the 10,371 clusters.

3.5 | Data preprocessing

Section 3.4 reveals that the initial data contains a wide range of values of significant size, indicating that normalization is essential to facilitate the faster convergence of the optimization algorithms and the machine learning models to achieve the best performance.⁵¹ Thus, we first normalize the data using z-score normalization technique as proposed in Reference [51], which is defined in Equation (1).

$$X_{norm} = \frac{X - \text{mean}(X)}{\text{std}(X)}, \quad (1)$$

where X denotes the feature array and X_{norm} is the normalized X as resulting from the subtraction of its mean ($\text{mean}(X)$) and division by its SD ($\text{std}(X)$).

The normalized clustered data of point clouds are used for the training of our machine learning models. Then, a vector (v) is constructed that models the target label naively as follows. If the cluster is considered a rockfall event, which can also be denoted as rockfall

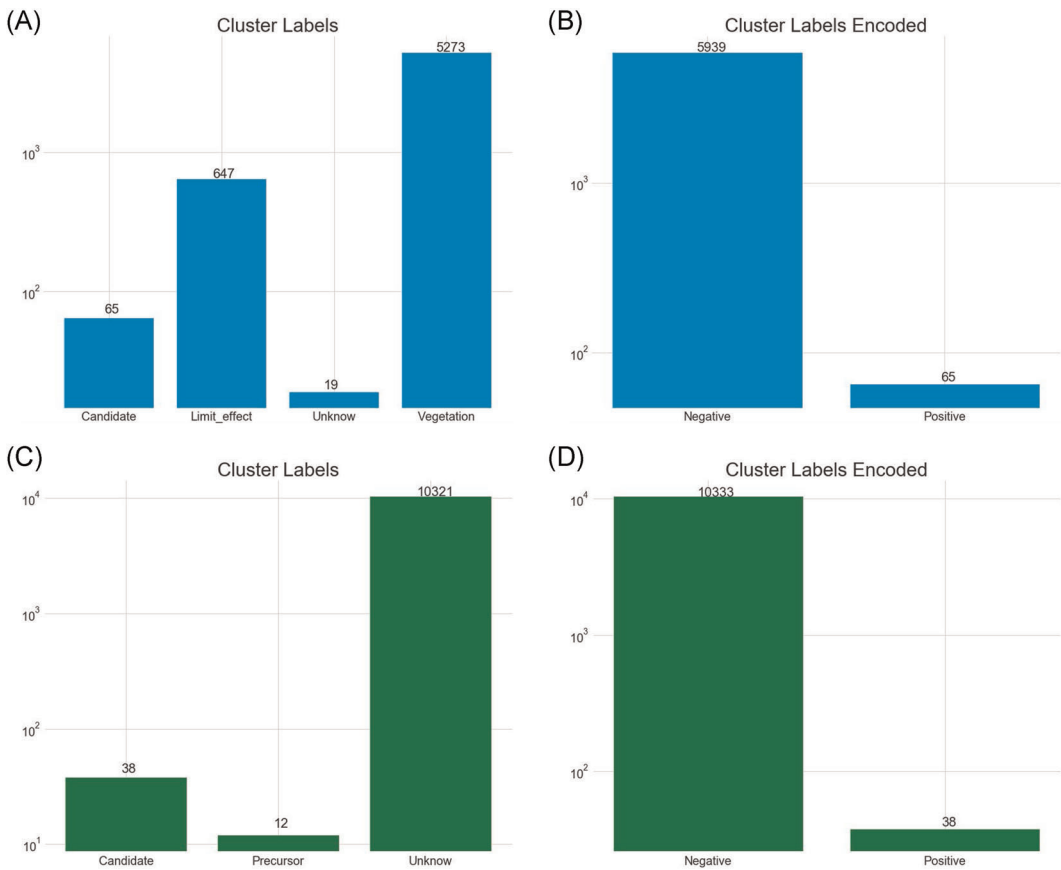


FIGURE 4 Target labels summary. Y-axis is in log-scale for comparison purposes. The actual values are included as text above each bar (A) Cluster labels—Degotalls case study, (B) encoded cluster labels for classification task—Degotalls case study, (C) cluster labels—Castellfollit case study, (D) encoded cluster labels for classification task—Castellfollit case study [Color figure can be viewed at wileyonlinelibrary.com]

candidate event, then we set a positive label ($v_i = 1$), and if not, we assign a negative label ($v_i = 0$). A summary of cluster labels is depicted in Figure 4, where subfigures (A) and (C) show the distribution of cluster labels before and subfigures (B) and (D) after encoding in the Degotalls and Castellfollit case studies, respectively.

Please note that the rockfall candidates belong in the class where $v_i = 1$ and the various “Limit_effect,” “Unknow,” “Vegetation,” “Precursor” cases belong in the class where $v_i = 0$.

3.6 | Resampling techniques

Rockfall events are characterized by their rarity and thus our task is considered to be an imbalanced classification task. This is also justified by Figure 4, where the candidate clusters for rockfall represent roughly 1% and 0.4% of the total cases in Degotalls and Castellfollit, respectively. There are two well-known approaches to balance the data instances, namely *undersampling* and *oversampling* methods.⁵² Undersampling removes elements from the majority class, while oversampling creates synthetic samples from the minority class. There are

also ensemble methods that oversample the minority class and undersample the majority one simultaneously.⁵³ Below, we briefly describe each one of the resamplers examined in this study. Finally, we conclude by utilizing the three best approaches in our subsequent analysis phases, according to their performance.

3.6.1 | Undersampling methods

The simplest way to perform undersampling is to do it at random, that is, by eliminating individual samples from the majority class without using any heuristics. More elaborate techniques employ some kind of heuristic processing to exclude elements that are informationally irrelevant to the overall data set. Some of the aforementioned methods, aggregate distinct samples to undersample the data using well-known clustering techniques, such as References [54].

In this study, we utilize two undersampling methodologies based on the K-means clustering algorithm. Specifically, we use two of the methods proposed by Reference [54] utilizing the K-means with the number of clusters proportional to the number of elements belonging to minority class. The first methodology swaps a cluster of samples belonging to the majority class with the cluster centroid obtained by the K-means algorithm (denoted as Cluster Centroids in this writing). The second one utilizes the same approach but uses the cluster center of the K-means instead of the cluster centroid (denoted as Cluster Representatives). Similar undersampling techniques utilizing clustering appear in recent studies such as Reference [55].

3.6.2 | Oversampling methods

Oversampling techniques perform the balancing of the data set by focusing on the minority classes. Various oversampling approaches are described in the literature.⁵⁶ SMOTE, one of the most widely used oversampling techniques, creates synthetic samples by combining features of the minority class's nearest neighbors.⁵⁷ Numerous variants of the aforementioned algorithm appear in the literature.⁵⁸ Kovacs et al.⁵⁶ presented a detailed empirical comparison of multiple variants of minority oversampling methods involving a wide assortment of imbalanced data sets. In this study, we focus on some of them, specifically the ones that we utilized for the purpose of this study.

The ADASYN oversampler determines the number of artificial samples to be produced for a given point by using a weighted density distribution for distinct minority class elements based on their level of difficulty in learning.⁵⁹ In contrast, SMOTE weights uniformly all minority points. The major difference between SMOTE and ADASYN lays in the generation of synthetic sample points for minority data points.

Barua et al.⁶⁰ introduced a “proximity weighted synthetic oversampling technique” (ProWSyn), which calculates weight values for the minority data samples based on their proximity information. Proximity, in this case, is portrayed as the distance between the sample and the boundary.⁶⁰ Another variant of SMOTE is the SMOTE-IPF, which is an extension through a new component, an “iterative ensemble-based noise filter called Iterative-Partitioning Filter” (IPF). SMOTE-IPF is considered to be a solution to the problems caused by noisy and borderline examples in unbalanced datasets as presented and explained in Reference [61].

SMOTE algorithm lacks distribution and density information of the data, both of which are critical for correctly synthesizing minority examples. The SMOBD algorithm addresses the above issues by effectively removing the effects of noise.⁶² In addition, the Assembled-SMOTE approach implements oversampling by taking data distribution information into account to prevent the overlapping between classes.⁶³

Other methods, such as the Lee algorithm, produce synthetic samples and determine whether to reject or accept them based on their location, that is, a synthetic sample's nearest neighbors.⁶⁴ Batista et al.⁶⁵ proposed an oversampling method call SMOTE-TomekLinks, which is essentially a SMOTE algorithm with the additional application of Tomek links to the over-sampled training set as a data cleaning technique. Thus, the aforementioned technique improves the oversampling accuracy by removing examples from both classes. Tomek links are introduced in Reference [66] and can be defined as follows:

“Given two examples E_i and E_j belonging to different classes, and $d(E_i, E_j)$ is the distance between E_i and E_j . A (E_i, E_j) pair is called a Tomek link if there is not an example E_l , such that $d(E_i, E_l) < d(E_i, E_j)$ or $d(E_j, E_l) < d(E_i, E_j)$.”⁶⁵

Furthermore, to effectively identify minority class instances, the CCR algorithm integrates the cleaning of the decision boundary around minority elements and controlled synthetic oversampling.⁶⁷ Moreover, the G-SMOTE approach pioneered hybrid oversampling that is driven by previously unknown patterns derived from the class with minor samples and randomization. The authors utilize concurrent oversampling and undersampling to deal with heavily skewed data distributions.⁶⁸ Additionally, LVQ-SMOTE integrates the SMOTE oversampler with feature codebooks learned by vector quantization to create synthetic samples that use more feature space than the other SMOTE variants.⁶⁹ Also, a comparable performance can be obtained using the Polynom-fit SMOTE, which oversamples the minority class using polynomial fitting functions.⁷⁰

Moreover, the research study of Reference [71] proposed a method for selective pre-processing, called SPIDER, that combines filtering and oversampling of imbalanced data. Also, the SWIM algorithm creates synthetic elements with the same Mahalanobis distance from the majority class as currently available minority data samples.⁷²

In this study, we examine and experiment with the 16 resamplers explained above, namely Cluster Centroids, Cluster Representatives, SMOTE, ProWSyn, SMOTE-IPF, SWIM, SMOBD, Lee, ADASYN, Assembled-SMOTE, SMOTE-TomekLinks, CCR, G-SMOTE, LVQ-SMOTE, Polynom-fit-SMOTE, and SPIDER.

3.7 | Models

This section presents the machine learning models trained and evaluated in the rockfall detection task. In addition, we provide basic information regarding the models utilized. However, a complete explanation and analysis of the algorithmic aspects of each model is out of the scope of this paper.

Our aim is to deal with rockfall detection with models belonging to two distinct and broad families of models, namely “single base learning” and “ensemble learning” algorithms, which are utilized in similar landslide classification tasks, as presented in Reference [73]. Single base learners are simpler models and learn by using a single algorithm, such as a decision tree

algorithm, while ensemble learners are comprised usually of multiple single base learners.⁷⁴ In this case, and following the notation of Ma et al.,⁷³ we use the following “single base learning” algorithms: **Linear Discriminant Analysis (LDA)**, **Quadratic Discriminant Analysis (QDA)**, **K-Nearest Neighbors Classifier (KNN)**, **Gaussian Naive Bayes (GNB)**, **Decision Tree Classifier (DT)**, **Support Vector Classifier (SVC)** and **Multilayer Perceptron (MLP) Classifier**, and the following “ensemble learning” algorithms: **AdaBoost Classifier (AdaB)**, **Random Forest Classifier (RF)**, **Extra Trees Classifier (ET)**, **XGBoost Classifier (XGB)**. Following, there is a list containing all the models examined including their hyper-parameters.

Regarding the utilized “single base learning” approaches, we explain the models below:

Linear Discriminant Analysis (LDA) is a linear decision boundary classifier. The boundary of this classifier is created by fitting class conditional densities to the input data. In this model, each class is fitted with a Gaussian density function. The covariance matrix for all classes is considered to be the same.⁷⁵ We experiment with solvers such as singular value decomposition (SVD), eigenvalue decomposition (EIGEN) and least squares solution (LSQR), and with automatic or no shrinkage parameters.

Quadratic Discriminant Analysis (QDA) is a classifier that uses a quadratic decision boundary created by fitting class conditional densities to the input data to classify objects. The covariance matrix for all classes is assumed to be the same.⁷⁵ We utilize a regularization parameter of 0, that is, we do not regularize the per-class covariance approximations.

K-Nearest Neighbors Classifier (KNN) is a memory-based algorithm. Naively, given a query point x_0 , we find the k training points $x_{(r)}$, $r = 1, \dots, k$ nearest in distance to x_0 , and then classify it using a majority vote policy among the k neighbors.⁷⁵ We experiment using various numbers of nearest neighbors, such as 1, 3, 5 and 9 neighbors.

Gaussian Naive Bayes (GNB) supposes that given a class $C = j$, the k number of features X_k are independent: $f_j(X) = \prod_{k=1}^p f_{jk}(X_k)$.⁷⁵ GNB considers continuous valued features and models them all as Gaussian (normal) distributions. To maintain measurement stability, a value of $1e - 9$ of the highest variance of all features is appended to variances.

Decision Tree Classifier (DT) divides the feature space into rectangles, and then fits a simple model to each of them. Theoretically, tree-based classifiers are simple but efficient and accurate.⁷⁵ For this particular classifier, we experiment with two strategies for splitting each node, namely the best split and random split. Regarding the criterion for splitting or alternatively the function to calculate the quality of a split, we experiment with Gini impurity⁷⁶ and information gain metrics.

Support Vector Classifier (SVC) algorithm creates a hyperplane or set of hyperplanes in a high-dimensional space that can be used for classification. The margins of the hyperplanes are described by support vectors, which are discovered after an optimization process involving an objective function regularized by an error term and a constraint. SVMs may use kernel functions to create linear or nonlinear decision boundaries. Different kernel functions can be specified for the decision function.⁷⁷ In this particular case, we experiment with common kernels such as polynomial with multiple distinct degree values ranging from 2 to 8, a radial-basis function, and sigmoid. Regarding the regularization parameter, we have used multiple values, such as 0.1, 1, 10, and 100.

Multi-Layer Perceptron (MLP) Classifier is a class of feed-forward artificial neural networks.⁷⁸ Alternatively, the MLP classifier can be considered to be a shallow deep neural network. The investigated hyper-parameters are hidden layer sizes of 50, 100, 150,

and 200 and activation functions for the hidden layers, namely rectified linear unit (ReLU) $f(x) = \max(0, x)$, hyperbolic tangent (TanH) $f(x) = \tanh(x)$, and logistic sigmoid $f(x) = \frac{1}{1 + \exp(-x)}$. Besides, three optimizers are tested namely LBFGS, which is an optimizer belonging to the group of quasi-Newton methods, the stochastic gradient descent (SGD) and Adam, which is a stochastic gradient-based optimizer.

Following, we provide information on the utilized “ensemble learning” models:

AdaBoost Classifier (AdaB) is a meta-estimator and an ensemble machine learning algorithm that initially fits a classifier on the original input data and then fits additional copies of the classifier on the same data but conditionally. It fits the additional copies where the weights of wrongly classified instances are changed such that subsequent classifiers concentrate more on difficult cases.^{79,80} Multiple numbers of estimators are investigated, such as 10, 50, 100, and 500. The base estimator is chosen to be a simple decision tree classifier with a max tree depth of 5.

Random Forest Classifier (RF) is a meta-estimator belonging to the ensemble group of algorithms that uses averaging methods to improve predictive accuracy. Also, it manages overfitting by fitting multiple decision tree classifiers on various subsamples of the original input data.⁷⁵ Multiple numbers of trees are investigated, such as 10, 100, 500, and 1000. The quality measurement of a split of a tree-node is done by applying the Gini impurity and the information gain strategy. Both approaches are commonly used in the decision tree algorithms.⁷⁶

Extra Trees Classifier (ET) is a meta estimator and an ensemble machine learning algorithm that uses averaging methods to boost predictive accuracy. Also, similarly to RF classifier, it manages overfitting by fitting multiple randomized decision trees on different sub-samples of the original input data.⁸¹ Different numbers of estimators are investigated, namely 10, 100, 500, and 1000. Also, the function to calculate the quality of a split of a tree-node is set to be the Gini impurity⁷⁶ or the information gain.

XGBoost Classifier (XGB) is a scalable end-to-end ensemble machine learning framework. It is actually a version of gradient boosted decision trees that has been engineered to produce cutting-edge results while staying fast and effective.⁸² Different values of boosting learning rate are examined, ranging from 10^{-1} to 10^{-4} . In addition, various numbers of gradient boosted trees are used such as 10, 50, and 100, together with a variety of boosters such as *gbtree*, which is a version of a regression tree as a weak learner, *gblinear*, which uses generalized linear regression with L1 and L2 shrinkage, and *dart*, which drops trees in such way as to reduce the overfitting issue.

Even though by using the “single base learning” algorithms, many pitfalls may be observed, such as overfitting and unreliable results,⁷³ in reality they are not always weak, as the correct utilization of them often leads to better outcomes.⁷⁴ For instance, DT is able to mine and identify significant relationships and key features, but it achieves lower accuracy values than its competitors.⁸³ Additionally, in simple rockfall and landslide classification applications, SVC algorithm achieves marginally better accuracy than the other methods, but the quality of its results is highly depended on the quality of the input data.⁸⁴ Besides, MLP achieved reasonably accurate results in similar studies.⁸⁵

On the other hand, although “ensemble learning” algorithms, in general, are able to increase the robustness and generalizability of the “single base learning” ones,⁷³ they still need to be properly constructed and selected.⁷⁴ Essentially, in similar identification tasks an

RF model is considered to outperform other models and the stability of such ensemble models is ensured by the numerous base learners, that is, sets of decision trees.⁸⁶ Hong et al.⁸⁷ highlight that tree-based ensemble models could greatly improve the accuracy in landslide classification tasks.

Theoretically, GNB or SVC perform well in circumstances where limited and imbalanced data is available⁸⁸ although, practically, this phenomenon is not always witnessed.⁸⁹ LDA and QDA algorithms are utilized in classification studies, that involve geological features and phenomena.⁹⁰ While LDA can easily distinguish the data between two classes by applying linear combinations to input features, QDA enhances this ability and can also deal with classes having different covariance values, that is, imbalanced data.

Intuitively, facing the complex and challenging task of rockfall detection, which involves imbalanced point cloud data having a certain amount of features to handle, and examining studies such as Reference [91], where certain single base learners outperform some ensemble learners, lead us to use this selection of classifier models starting from simpler models like GNB, KNN, and so on, to more complex and elaborated ensemble methods like AdaB, ET, and XGB, to get strong performance proofs and to experimentally compare the ensemble methods with baseline methods. In addition, to further provide performance evidence of the utilized machine learning models, we use *10-fold cross-validation* procedures, as explained in the Section 4.2, which clearly facilitate the model selection by providing significant insights on the accuracy and robustness of each model.

3.8 | Feature selection

In this phase of the analysis framework, see Figure 3D, the best selected and properly parameterized models of each model variant accompanied with each one of the three best resamplers are used as inputs in a grid search algorithm aiming to select the exact number of features that achieve the highest performance. The criterion used for the univariate feature selection is mutual information between two random variables. Mutual information is a non-negative value that measures the dependency between the variables.⁹² According to the aforementioned methodology, we keep only the k highest scoring features, with k ranging from 5 to 37 (all features) in the Degotalls case study and from 5 to 31 (all features) in the Castellfolit case study.

3.9 | Pipeline selection

The input of this stage consists of the multiple machine learning pipelines, each properly parameterized with the best number of features, as shown in Figure 3E. These pipelines are statistically compared to conclude with the best intelligent pipeline to be used in our prototype system. The statistical comparison of the various pipelines is conducted by applying the Friedman and Nemenyi test, a nonparametric statistical test and a post-hoc test respectively, as described in Reference [93]. We run every pipeline $n = 10$ times using *10-fold cross-validation*. In each run, we rank the pipelines from 1 to n_{models} . By averaging the ranks table to get the averaged ranks for each pipeline, we can infer that two models are significantly different if their average ranks vary by at least the critical difference, as explained in Reference [94]. The critical difference is computed with the following formula⁹³:

$$CD = q_{\alpha} \sqrt{\frac{n_{models} * (n_{models} - 1)}{6 * n}}, \quad (2)$$

where q_{α} denotes the critical value of the two tailed Nemenyi test, which depends on the chosen α level of statistical significance and the number of models or in this case the pipelines, n_{models} .

3.10 | Feature importance

In this phase of our framework, see Figure 3H, we calculate and visualize the feature importance of the input data set. This step provides significant insights in terms of the interpretability and explainability of the outcome produced by the utilized algorithm. Thus, we utilize a generalized feature importance measurement, the permutation feature importance as introduced and described in Reference [95]. The computation of permutation feature importance is done by first calculating a baseline metric, defined by a performance scoring function, and evaluated on a data set X . Next, a feature from the validation set is permuted and the metric is evaluated again. Finally, the permutation importance is defined as the difference between the baseline metric and the metric obtained from permutating the aforementioned feature.

In our study, we utilize an extensive permutation feature importance test to get robust results and insights regarding the most significant features utilized by the classifier. Specifically, we use the most statistical significant classifier as concluded from the computations in the Pipeline Selection stage of our framework, depicted in Figure 3E, with a *10-fold cross validation* technique on input data and 100 feature permutations in each fold. Then, we average all the values obtained and express them in percentage values for interpretability purposes. In addition, we use K-means clustering algorithm to group the feature importance obtained to get a better view of the most important features and their neighbors.

4 | EXPERIMENTAL ANALYSIS

This section presents the experimental analysis of our designed intelligent framework, based on well-known performance metrics. First, we explain the accuracy metric and then we present evaluation tables for each stage of the framework. Additionally, an ablation study is conducted to justify the addition of each component to the final prototype implementation.

4.1 | Performance metrics

Our aim is to design our models for a classification task, in which a false negative is usually more disastrous than a false-positive for preliminary rockfall detection. The nature of our problem is imbalanced as discussed in Section 3.1. For this reason, we utilize a special metric called balanced accuracy, as introduced and explained in Reference [96]. The authors of the aforementioned study define balanced accuracy as a performance metric for imbalanced classification tasks. The average accuracy obtained in either class can be naively characterized as balanced accuracy. The balanced accuracy is given by

$$Acc_b = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right), \quad (3)$$

where FP, TP, FN, and TN, are the false-positives, true-positives, false-negatives, and true-negatives, respectively, according to the confusion matrix, all of them clearly defined in Reference [96].

4.2 | Experimental evaluation

In this section, we evaluate the various machine learning models, which we proposed and concisely discussed in Section 3.7. We first exhibit the performance evaluation of the various resamplers in Table 1 and then a thorough evaluation of all the models used in the model selection and parameterization and feature selection phases in Table 2. The entire evaluation process is based on the balanced accuracy metric, denoted in Equation (3), using stratified *10-fold cross-validation*, which provides robust accuracy results.

4.2.1 | Resamplers

Table 1 displays the average performance of each resampler in the Degotalls (A) and Castellfollit (B) case studies. The depicted values are the average performances of all machine learning models evaluated with a stratified *10-fold cross validation* procedure, without prior parameterization, paired with each resampler. We average the balanced accuracy metric values resampler-wise, to select the three best resamplers to proceed to the model selection and parameterization stage. In both case studies, the Cluster Centroids resampler achieves the highest balanced accuracy while being the most robust method. It achieves a \overline{Acc}_b of 0.89 with 3.83% error and 0.82 with 7.78% error in the Degotalls and Castellfollit case studies respectively. Observing the two subtables in Table 1, the results obtained in both cases do not seem to differ much in terms of ranking.

4.2.2 | Model selection

Table 2 shows the performance evaluation of the model selection and parameterization phase in the Degotalls (A) and Castellfollit (B) case studies. Specifically, it displays only the six best methods for each case study. For clarification purposes, the full tables are included in Tables 6 and 7, respectively in Appendix S2. For the selection of the best hyper-parameters of each model, a grid search algorithm is utilized using a stratified *10-fold cross-validation* procedure for each available combination of parameters. Please note that the depicted values in the above-mentioned table are the balanced accuracy measurements of another stratified *10-fold cross-validation* procedure on the data set utilizing only the best properly configured models that resulted from the aforementioned grid search. It is clear that in the Degotalls case study, the XGBoost classifier parameterized the best, accompanied by the SMOTE-IPF oversampler, noted as XGB-SMOTE_IPF in Table 2A, and performed better than all the other models while remaining robust. It achieves a \overline{Acc}_b of 0.94 with a 3.83% error score. Regarding the best hyper-parameters of the XGBoost

TABLE 1 Acc_b metric summary for each resampling method. $\overline{Acc_b}$ denotes the average value of Acc_b of the 10-fold cross validation

(a) Degotalls case study.

Resampling Method	$\overline{Acc_b}$	Error (%)
Cluster Centroids	0.89	3.83
ProWSyn	0.87	5.34
SMOTE-IPF	0.87	5.56
SWIM	0.86	5.65
SMOBD	0.86	6.29
Lee	0.86	5.48
ADASYN	0.86	6.28
Assembled-SMOTE	0.86	6.16
SMOTE	0.86	5.85
SMOTE-TomekLinks	0.86	6.56
CCR	0.85	7.91
G-SMOTE	0.85	7.15
LVQ-SMOTE	0.84	6.20
Polynom-fit-SMOTE	0.83	7.56
SPIDER	0.80	6.29
Cluster Representatives	0.78	7.67

(b) Castellfollit case study.

Resampling Method	$\overline{Acc_b}$	Error (%)
Cluster Centroids	0.82	7.78
ProWSyn	0.71	16.94
CCR	0.70	15.88
SWIM	0.70	14.90
Lee	0.68	17.52
SMOTE	0.68	18.13
LVQ-SMOTE	0.68	18.33
ADASYN	0.68	18.50
SMOBD	0.68	18.86
SMOTE-IPF	0.68	17.76
SMOTE-TomekLinks	0.68	17.33
Assembled-SMOTE	0.67	18.08
G_SMOTE	0.66	19.83
Polynom-fit-SMOTE	0.65	18.00
SPIDER	0.58	17.29
Cluster Representatives	0.56	18.28

Note: With dark grey we denote the best, with lighter grey the second best and with pale grey the third best.

classifier, a linear booster, namely gblinear, with a learning rate of 0.1 with 50 estimators, was chosen. On the other hand, in the Castellfollit case study, the Linear Discriminant Analysis classifier paired with the ProWSyn oversampler, noted as LDA-ProWSyn in Table 2B, appears to be the most robust technique in terms of balanced accuracy, achieving a $\overline{Acc_b}$ of 0.93 with 0.36% of error, using Eigenvalue Decomposition as the solver and with an automatic shrinkage parameter.

4.2.3 | Feature selection

In addition, Table 2 also displays the performance results of the feature selection phase in the Degotalls (A) and Castellfollit (B) case studies, in which each properly parameterized model

TABLE 2 Acc_b metric summary. $\overline{Acc_b}$ denotes the average value of Acc_b of the 10-fold cross validation**(a) Degotalls case study.**

Method	Model Parameterization		Feature Selection		
	$\overline{Acc_b}$	Error (%)	$\overline{Acc_b}$	Error (%)	Features
XGB-SMOTE_IPF	0.94	3.83	0.95	3.78	35
MLP-SMOTE_IPF	0.94	5.54	0.95	5.64	35
KNN-ClusterCentroids	0.94	4.29	0.95	4.09	36
XGB-ProWSyn	0.94	3.84	0.95	3.68	35
SVC-ProWSyn	0.94	5.68	0.95	4.14	17
LDA-SMOTE_IPF	0.94	4.13	0.95	4.07	31

(b) Castellfolit case study.

Method	Model Parameterization		Feature Selection		
	$\overline{Acc_b}$	Error (%)	$\overline{Acc_b}$	Error (%)	Features
XGB-ProWSyn	0.93	8.11	0.94	6.92	30
LDA-ProWSyn	0.93	0.36	0.93	3.98	30
MLP-CCR	0.93	8.21	0.93	8.18	30
LDA-CCR	0.93	0.54	0.93	4.06	29
XGB-CCR	0.92	8.11	0.92	8.17	30
SVC-ClusterCentroids	0.91	8.18	0.91	8.26	30

Note: With dark grey we denote the best model in terms of accuracy and then robustness and with lighter grey the second best for each phase. "Features" column displays the number of features utilized by each algorithm to achieve this score.

accompanied by a resampler is evaluated using the best number of features over a stratified 10-fold cross-validation process. In the first case study, the XGBoost classifier accompanied by the SMOTE-IPF oversampler with 35 features seems to be the best performing model, while in the second best model is the XGBoost classifier paired with ProWSyn with 30 features, as displayed in Table 2.

Generalizing, we could say that in both case studies the top performing methods are approximately the same utilizing around the same percentage of features compared to the total available number of features in each case.

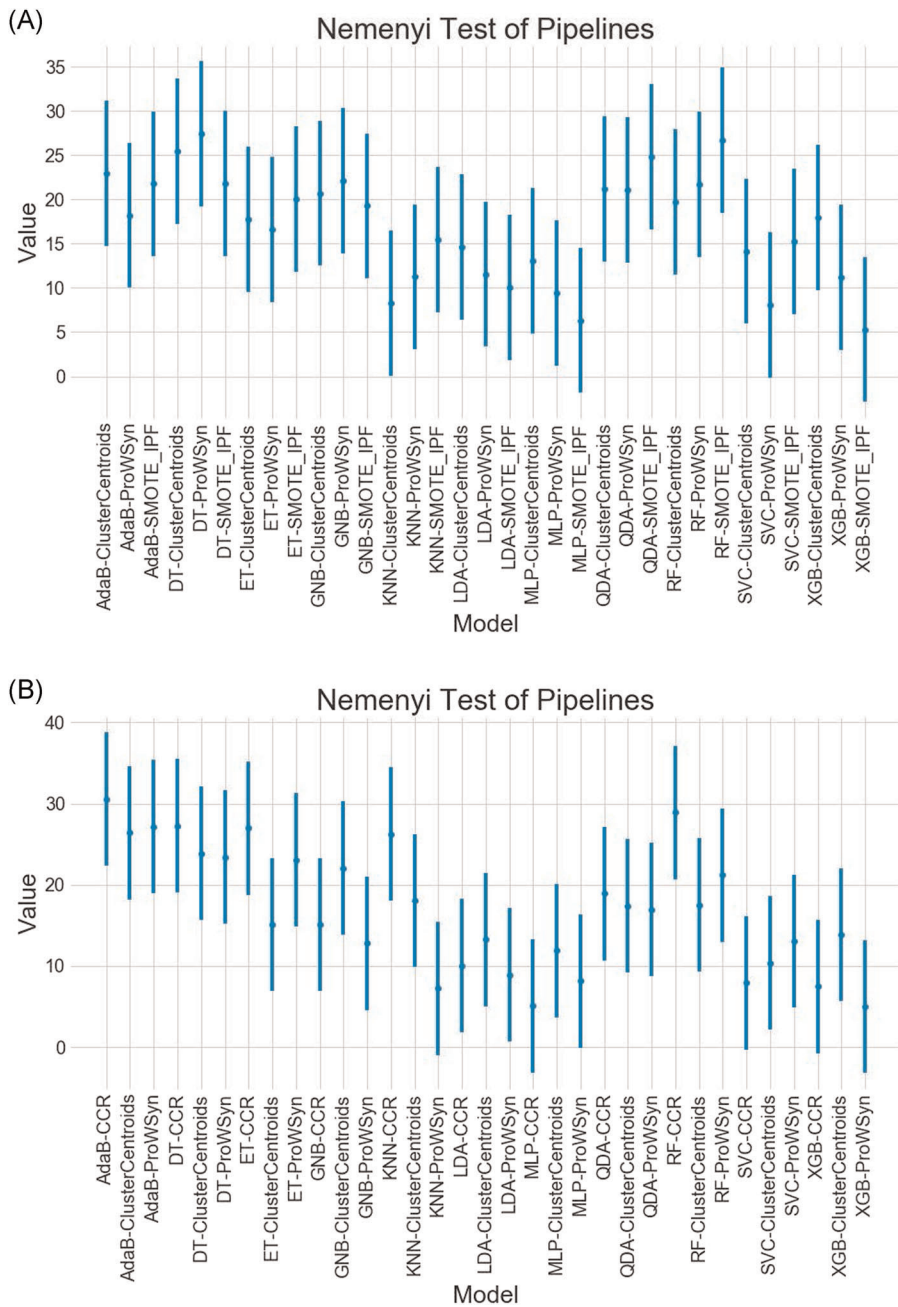


FIGURE 5 Nemenyi test—scores of model pipelines (A) Degotalls case study, (B) Castellfollit case study [Color figure can be viewed at wileyonlinelibrary.com]

4.2.4 | Pipeline selection

In Figure 5, there is an illustration of the Friedman and Nemenyi techniques, which are computed as described in Section 3.9, for the Degotalls (A) and Castellfollit (B) case studies. This process provides enough evidence to identify which of the methods are more statistically

significant than the others. We use a critical distance corresponding to 95% statistical significance. Please note that between two algorithmic approaches there is a statistical significance when the lines do not overlap. The best algorithm is the one with the minimum rank value.

Figure 5A shows that the XGBoost classifier properly parameterized and paired with the SMOTE-IPF oversampler with 35 features performs significantly better than the majority of the machine learning pipelines examined. In addition, the MLP classifier paired with the SMOTE-IPF undersampler with 35 features could also be an acceptable solution.

Regarding the second case study in Figure 5B, the XGBoost classifier properly parameterized and paired with the ProWSyn oversampler with 30 features is the best machine learning pipeline by a significant margin. Furthermore, the MLP classifier accompanied by the CCR oversampler with 30 features could also be an alternative acceptable solution.

4.2.5 | Feature importance

In Figure 6, there is a visualization of permutation feature importance technique, as described and explained in Section 3.10. Please note that the colorization refers to the distinct clusters acquired from the K-means algorithm. For the calculation of the importance of feature in the Degotalls case study, we use the XGB classifier paired with the SMOTE-IPF oversampler with 35 features, which we concluded performed better than the majority of all other methods tested, while for the Castellfolit case study, we utilize the XGB classifier paired with the ProWSyn oversampler with 30 features.

Observing Figure 6A,B, we can conclude that the most important features for each case study appear to be the same. Moreover, the highly important features for rockfall event detection are the coordinates of the point clouds.

4.3 | Ablation study

An ablation study is crucial for the development of an intelligent system because it helps to understand the contribution of each component to the overall system. Table 3 portrays how the balanced accuracy metric is increased with the addition of each proposed component in our analysis framework, in the Degotalls (A) and Castellfolit (B) case studies. It shows the increment of the aforementioned metric across all models ($\overline{Acc_b}$) and the best performing model (Acc_b^{best}) in each phase. The baseline is considered a state in which all the models are trained and evaluated without prior resampling of data, parameterization and feature selection. It is clear in both case studies that the addition of each component provides an interesting performance increment, improving the accuracy and robustness of the utilized method.

4.4 | Prototype implementation

The implementation of a web-based system for rockfall detection is presented in this section. Informatively, the end system incorporates the best intelligent pipeline according to our performance evaluation. To showcase the effectiveness of our intelligent system, we have designed a simple application that detects whether the input point cloud data are considered candidates for a rockfall event. Figure 7 illustrates the architecture of the developed prototype system for

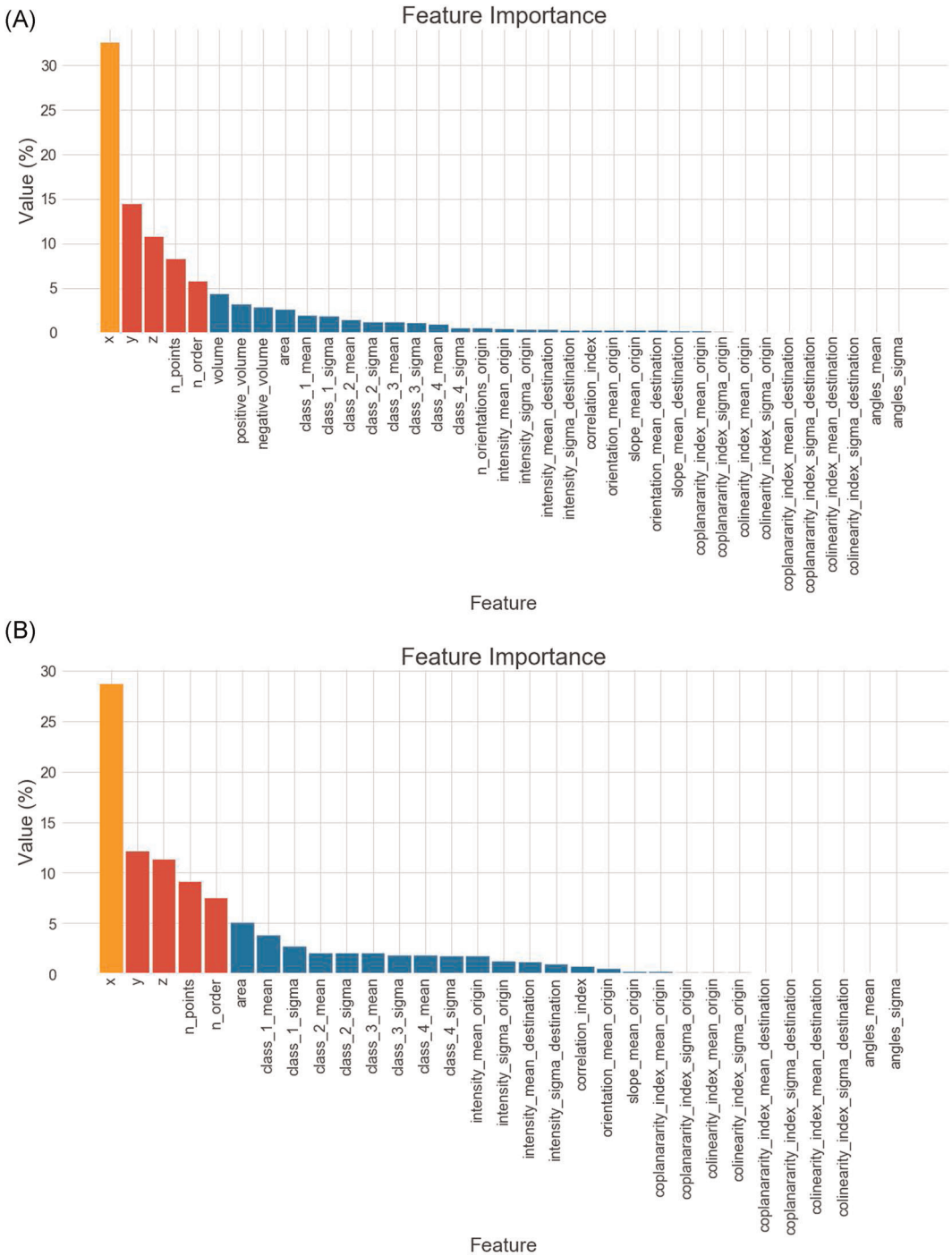


FIGURE 6 Feature Importance. Orange color denotes high importance, blue and red denote mid and low importance respectively (A) Degotalls case study and (B) Castellfollit case study [Color figure can be viewed at wileyonlinelibrary.com]

the visualization of the predictions. A basic but well-designed table is used to source input feature values in the front-end, as shown in Figure 8, and a button that triggers the initialization of the detection procedure. In the back-end, we normalize the data entered by the users, if needed, and generate a prediction using our intelligent framework, which is the pretrained and loaded machine learning pipeline. Finally, the results are displayed to the user in a fast and accurate manner.

Our web application is placed at this [GitHub repository](#). In addition, we used the Python programming language to develop our models and run our experiments. Particularly, for the development of the models, we mainly use the Scikit-learn Python library.⁹⁷ The Matplotlib library is used for visualization.⁹⁸ Also, for developing the web application, we utilize a microweb framework, which is called Flask.⁹⁹

5 | FINAL OBSERVATIONS

In this section, we provide the final observations of our experiments and analysis. In the particular case studies, we conclude that some of the resampling methods and machine learning models examined could be used to form intelligent pipelines to detect rockfall events. In the case of the Degotalls mountain wall TLS data, we come to the conclusion that the XGBoost classifier using a linear booster, namely *gblinear*, with a learning rate of 0.1 and 50 estimators utilizing 35 features of the input data set accompanied by the SMOTE-IPF resampler, achieved the best outcome, with a robust balanced accuracy score of 95%. Furthermore, in the Castellfollit case study, we conclude that the XGBoost classifier with a linear booster, a learning rate of 0.1 and 100 estimators utilizing 30 features accompanied by the ProWSyn resampler achieved the best balanced accuracy score of 94%.

Intuitively, we could say that the second-order gradients and the advanced regularization of the XGBoost algorithm during the learning phase, help to identify better the data relations. In addition, using the *gblinear* booster, the algorithm builds multiple regularized

TABLE 3 Ablation study—results based on Acc_b metric

(A) Degotalls case study			
Method	\overline{Acc}_b	Error (%)	Acc_b^{best}
Baseline	0.78	7.65	0.84
+Resampling	0.85	3.25	0.89
+Model parameterization	0.89	4.65	0.94
+Feature selection	0.91	4.45	0.95
(B) Castellfollit case study			
Method	\overline{Acc}_b	Error (%)	Acc_b^{best}
Baseline	0.56	18.04	0.80
+Resampling	0.68	8.21	0.82
+Model parameterization	0.79	15.93	0.93
+Feature selection	0.82	11.75	0.94

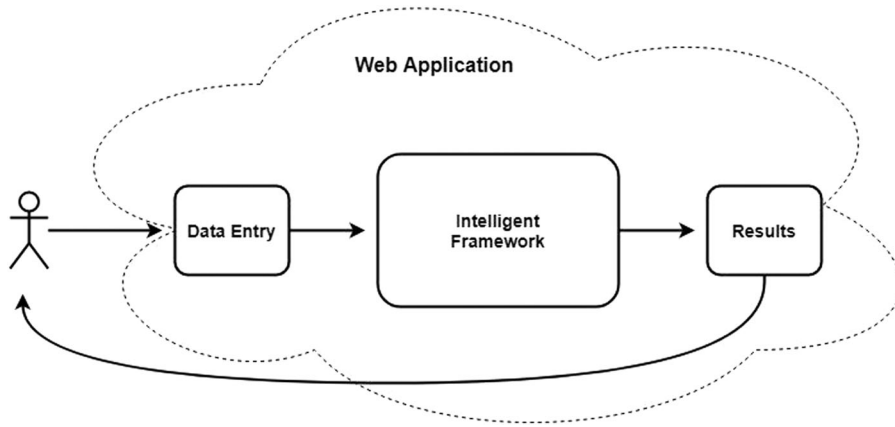


FIGURE 7 Rockfall detection—web application architecture [Color figure can be viewed at wileyonlinelibrary.com]

Rockfall Detection System				
x	y	z	number of points	number of order
volume	positive volume	negative volume	area	class 1 mean
class 1 sigma	class 2 mean	class 2 sigma	class 3 mean	class 3 sigma
class 4 mean	class 4 sigma	n orientations origin	intensity mean origin	intensity sigma_origin
intensity mean destination	intensity sigma destination	correlation index	orientation mean origin	slope mean origin
orientation mean destination	slope mean destination	coplanarity index mean	coplanarity index sigma	colinearity index mean o
colinearity index sigma o	coplanarity index mean	coplanarity index sigma	colinearity index mean d	colinearity index sigma d
angles mean	angles sigma			

Candidate for Rockfall: True

FIGURE 8 Rockfall detection—web application [Color figure can be viewed at wileyonlinelibrary.com]

linear models to later incorporate them in a generalized linear model with advanced regularization. It seems that learning inner linear models and then additively producing a generalized linear model facilitates the achievement of a higher accuracy than the other algorithms in both case studies while remaining fast. Additionally, the Castellfolit case study appears to be more difficult than the Degotalls case study in terms of model learning, because of its data imbalance of 0.4% compared to 1% in the Degotalls study. This can also be justified by the fact that the XGBoost algorithm utilized twice as many estimators in Castellfolit than Degotalls.

Moreover, the ablation study, portrayed in Table 3 justifies the need for each component of our system. The increment in balanced accuracy metric seems to be significant with the addition of the special framework stages, namely resampling, model selection and parameterization and feature selection. Specifically, the most significant addition appears to be the resampling module of our system, displaying an increment in accuracy from the baseline of approximately 9% and 21% in the Degotalls and Castellfolit case studies, respectively. Additionally, in general, an ablation study provides an interesting view of how the developed

intelligent system behaves by adding additional components. This is a good practice to identify the performance gains of a developed system.

To provide further evidence for our developed framework, we visualize the predicted point cloud data along with the original data of the Degotalls case study, as displayed in Figure 9. We chose to display the former visualization because it includes four incorrectly classified rockfall events representing roughly 6% of the total rockfall events in the data set, as depicted in Figure 4A. For clarification purposes, in the Castellfollit case study, all rockfall events in the initial data set were correctly classified by the XGBoost classifier using 30 features paired with the ProWSyn resampler, which is the best selected pipeline, as discussed above.

5.1 | Practical implications

Point cloud data of a 3D scan may contain millions of points with highly detailed information, but they also contain scattered, disjointed information and in most cases with a lot of noise.¹⁰⁰ Geoscientists, to properly deal with such huge data and correctly identify rockfall candidates, go through a semimanual repetitive process, as explained in Section 3.1. The proposed intelligent system will be a key tool in the hands of geology engineers, scientists, and any individual working at the intersection of the geology domain and decision support systems. It aims to provide meaningful and accurate insights into the task of rockfall detection in distinct geological contexts. Our learning methodology provides a robust and accurate way to identify rockfalls using point clouds. Indeed, our automatic intelligent system advances the accuracy and efficiency of the geoscientists' work, saving a lot of time from their work-routines on identifying such events.

Moreover, the final intelligent system could easily be incorporated as a stand-alone component in a rockfall monitoring system providing support and warning alarms. The web application that incorporates our intelligent system could automate the whole rockfall detection process by providing a user-friendly front-end, explained in Section 4.4 and shown in Figure 8.

Particular practical implications of our proposed intelligent system, which enhance the accuracy and efficiency of the geoscientists' work, are further explained below with examples. For instance, having identified the best intelligent pipeline for a specific mountain cliff could aid the rockfall detection in other similar geological contexts, that is, similar mountain cliff types. This is especially relevant for geologists as they regularly capture point clouds from the same mountains to observe changes in their surface structure. Therefore, having a pretrained model on such mountains is essential to observe and detect the evolution of changes automatically. This means, that in a run of our web-based application, using the pretrained best intelligent pipeline derived by analysing a mountain cliff, an individual could identify rockfall candidates in a similar mountain. In addition, taking a closer look at the outcome of our intelligent system, that is, a predicted rockfall candidate cluster of points, geoscientists could later focus on this specific region of the mountain cliff and take additional 3D scans of only this region to be further processed and analysed. To sum up, our system accelerates both the process and the analysis of identifying rockfall events.

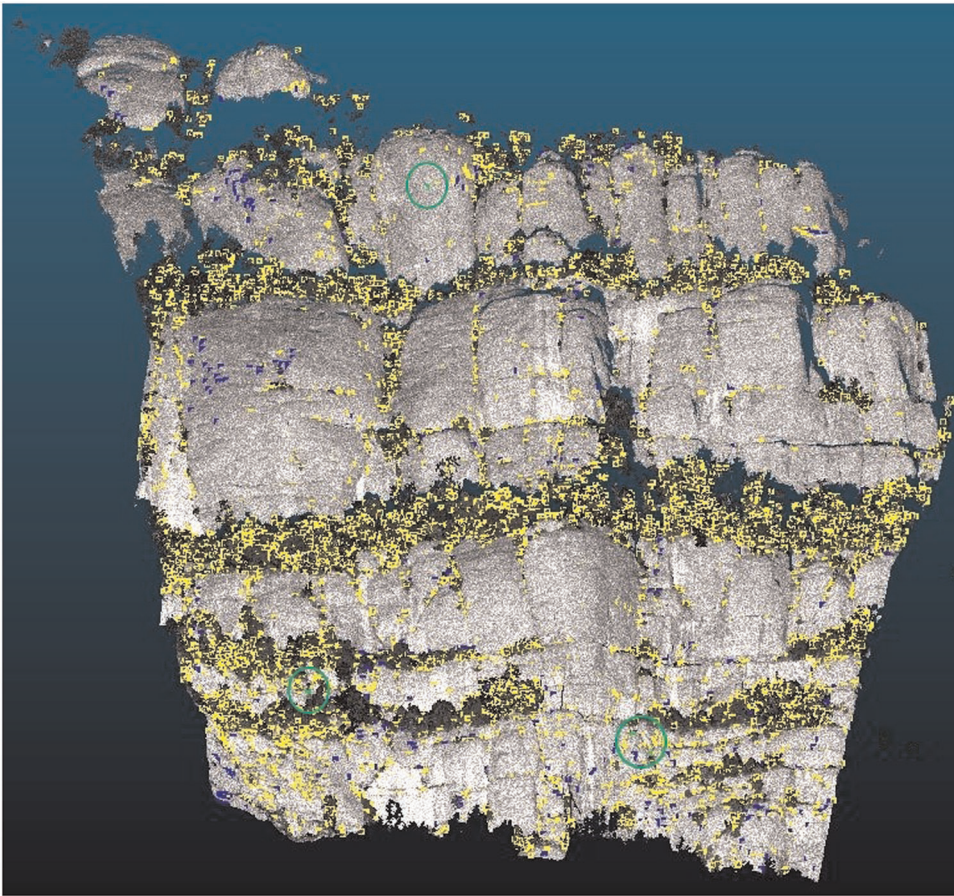


FIGURE 9 Point clouds visualization—Degotalls case study. We denote the correctly and incorrectly classified instances with yellow and blue colors, respectively. We further highlight them in green and circle, the events that are initially rockfall events but were incorrectly classified as not rockfall [Color figure can be viewed at wileyonlinelibrary.com]

5.2 | Limitations

Currently, the proposed intelligent system is analyzed and applied in two distinct geological contexts, that is, two types of mountain cliffs. The potential next steps are to enrich our database with more point cloud data coming from various geological contexts. Additionally, the use of clustering techniques on geological point cloud data increases resource efficiency by decreasing the processing time and memory footprint of the data but in some cases it eliminates some detailed geometrical and topological properties of them. Thus, this can be considered another limitation of the current study, which, however, could be lifted by processing and analysing clustered with raw point cloud data simultaneously.

Although deep learning models are popular in a vast majority of challenging engineering tasks, they are not analysed in this study. On the other hand, even though advanced neural network architectures may produce higher accuracy with lower error rates in a great variety of engineering tasks, they are not the panacea for every specific task they are employed to. The

case of analyzing imbalanced geological point cloud data with deep learning needs to be carefully investigated to properly select the right methods and algorithms to deal with the challenging task of rockfall and landslide detection.

6 | CONCLUSION AND FUTURE WORK

Rockfall events are considered to be a great hazard in multiple regions across the world. People working in the geology domain put a lot of effort into identifying such events. However, rockfall event detection utilizing data from remote sensors such as LiDAR or commercial cameras still need specialized analysis tools that can improve the accuracy and efficiency of the distinction of such events from other changes triggered by edge effects, data noise, vegetation, and so on. We believe that an intelligent decision assisting tool would facilitate geologists' work. However, the development of such a tool seems to be a complex process, due to the inherent nature of the data.

This study presents a thorough examination of several computational approaches for detecting rockfall events. In particular, we focus on the development of an end-to-end machine learning framework able to analyze clustered point cloud data from a TLS device. In the presented framework, due to the imbalanced nature of the rockfall event detection, we implement various resampling methodologies, parameterize several machine learning models and create intelligent pipelines to tackle the abovementioned task.

We experimentally evaluate our framework on two case studies, involving data from TLS measurements of a cliff at the Degotalls, Montserrat Massif in Barcelona and the cliff at the Castellfollit de la Roca in Girona, both located in Spain. We introduce and examine our design analytically and validate it experimentally. These actions clarify the components of the overall process required to set up the pillars of a reliable and intelligent rockfall detection system. We demonstrate the feasibility of creating a practical system for detecting rockfalls.

Our study provides an opportunity to clarify some key issues and provide concrete scientific and technical solutions at the intersection of geology and machine learning. There is a major potential for further additional research on the identification of rockfalls. Additionally, raw point cloud data carry an enormous amount of information and utilizing such information effectively and efficiently remains a challenging task. For this, a system utilizing raw point cloud data as input is being developed, in which the proposed methodology is used in conjunction with various added machine learning and neural network based methodologies to detect rockfall events directly from temporal point cloud data, showing good potential. Analyzing raw point cloud data without any grouping techniques using sophisticated deep learning architectures may bring advances in rockfall detection, because of the complete utilization of the inner geometrical and topological features of the data.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 860843. In addition, the authors want to acknowledge the support from PROMONTEC (CGL2017-84720-R AEI/FEDER, UE) and SALTEC CGL2017-85532-P (AEI/FEDER, UE) projects, founded by the Spanish MINEICO, and AGAUR project 2016 DI 069 (Agència de Gestió d'Ajuts Universitaris i de Recerca). Data from Castellfollit de la Roca were acquired with the support of the Spanish Ministry of Science and Education (predoctoral grant 2004-1852) and funded by the Natural

Park of the Garrotxa Volcanic Field (PNZVG) and the following projects: MEC CGL2006-06596 (DALMASA), TopoIberia CSD2006-0004/Consolider-Ingenio2010, MEC CGL2010-18609 (NUTESA). Data from Montserrat Massif were funded by the Institut Cartogràfic i Geològic de Catalunya (ICGC). Anna Puig and Maria Salamó also acknowledge Generalitat de Catalunya, for its support under project 2017-SGR-341. Moreover, we would like to thank Nicolas Pascual González for his contribution in this study.

ENDNOTES

*<http://www.ub.edu/risknat/>

**<http://www.ub.edu/geomodels/>

ORCID

Thanasis Zoumpikas  <http://orcid.org/0000-0002-3736-1155>

Anna Puig  <https://orcid.org/0000-0002-2184-2800>

Maria Salamó  <https://orcid.org/0000-0003-1939-8963>

David García-Sellés  <https://orcid.org/0000-0002-5995-5712>

Marta Guinau  <https://orcid.org/0000-0002-2898-4625>

REFERENCES

1. Robbins BA, Stephens IJ, Marcuson WF. Geotechnical engineering. In: *Encyclopedia of Geology*. Elsevier; 2021:377-392.
2. Robiati E, Vanneschi F, Venn C. Application of remote sensing data for evaluation of rockfall potential within a quarry slope. *ISPRS Int J Geo-Inform*. 2019;8:367.
3. DiFrancesco PM, Bonneau D, Hutchinson DJ. The implications of M3C2 projection diameter on 3D semi-automated rockfall extraction from sequential terrestrial laser scanning point clouds. *Remote Sensing*. 2020;12:1885.
4. Fanos AM, Pradhan B, Alamri A, Lee C-W. Machine learning-based and 3D kinematic models for rockfall hazard assessment using LiDAR data and GIS. *Remote Sensing*. 2020;12:1755.
5. Dramsch JS. 70 years of machine learning in geoscience in review. In: *Advances in Geophysics*. Vol 61. Academic Press Inc.; 2020:1-55.
6. Karpatne A, Ebert-Uphoff I, Ravela S, Babaie HA, Kumar V. Machine learning for the geosciences: challenges and opportunities. *IEEE Trans Knowl Data Eng*. 2019;31:1544-1554.
7. Crosta GB, Agliardi F. Failure forecast for large rock slides by surface displacement measurements. *Canadian Geotech J*. 2003;40:176-191.
8. Carrara A. Multivariate models for landslide hazard evaluation. *J Int Assoc Math Geol*. 1983;15:403-426.
9. Abellán A, Calvet J, Vilaplana JM, Blanchard J. Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring. *Geomorphology*. 2010;119:162-171.
10. Abellán A, Oppikofer T, Jaboyedoff M, Rosser NJ, Lim M, Lato MJ. Terrestrial laser scanning of rock slope instabilities. *Earth Surface Process Landforms*. 2014;39:80-97.
11. Lague D, Brodu N, Leroux J. Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z). *ISPRS J Photogrammetry Remote Sensing*. 2013;82:10-26.
12. Oppikofer T, Jaboyedoff M, Blikra L, Derron M-H, Metzger R. Characterization and monitoring of the Åknes rockslide using terrestrial laser scanning. *Natural Hazards Earth Syst Sci*. 2009;9:1003-1019.
13. Jaboyedoff M, Oppikofer T, Abellán A, et al. Use of LIDAR in landslide investigations: a review. *Natural Hazards*. 2012;61:5-28.
14. Prakash N, Manconi A, Loew S. Mapping landslides on EO data: performance of deep learning models vs. traditional machine learning models. *Remote Sensing*. 2020;12:346.

15. Xiao L, Zhang Y, Peng G. Landslide susceptibility assessment using integrated deep learning algorithm along the China-Nepal highway. *Sensors*. 2018;18.
16. Fanos AM, Pradhan B, Mansor S, Yusoff ZMd, Abdullah AFbin. A hybrid model using machine learning methods and GIS for potential rockfall source identification from airborne laser scanning data. In: *Landslides*. Vol 15, 2018:1833-1850.
17. Pham BT, Shirzadi A, Tien BD, Prakash I, Dholakia MB. A hybrid machine learning ensemble approach based on a Radial Basis Function neural network and Rotation Forest for landslide susceptibility modeling: a case study in the Himalayan area, India. *Int J Sediment Res*. 2018;33:157-170.
18. Bandura L, Halpert AD, Zhang Z. Machine learning in the interpreter's toolbox: unsupervised, supervised, and deep learning applications. In: *2018 SEG International Exposition and Annual Meeting (SEG 2018)*. Society of Exploration Geophysicists; 2019:4633-4637.
19. Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena SR, Tiede D, Aryal J. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*. 2019;11:196.
20. Wang H, Zhang L, Yin K, Luo H, Li J. Landslide identification using machine learning. *Geosci Front*. 2020.
21. Bui DT, Tsangaratos P, Nguyen VT, Liem NV, Trinh PT. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena*. 2020;188:104426.
22. James MR, Robson S. Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. *J Geophys Res: Earth Surface*. 2012;117.
23. Lissak C, Bartsch A, De Michele M, et al. Remote sensing for assessing landslides and associated hazards. *Survey Geophys*. 2020;41:1391-1435.
24. Mayr A, Rutzinger M, Geitner C. Multitemporal analysis of objects in 3D point clouds for landslide monitoring. *Int Arch Photogrammetry Remote Sensing Spatial Inform Sci—ISPRS Archives*. 2018;42:691-697.
25. Weidner L, Walton G, Kromer R. Classification methods for point clouds in rock slope monitoring: a novel machine learning approach and comparative analysis. *Eng Geol*. 2019;263.
26. Weidner L, Walton G, Kromer R. Generalization considerations and solutions for point cloud hillslope classifiers. *Geomorphology*. 2020;354:107039.
27. Kong D, Wu F, Saroglou C. Automatic identification and characterization of discontinuities in rock masses from 3D point clouds. *Eng Geol*. 2020;265:105442.
28. Huu PNT. Lidar point cloud classification using expectation maximization algorithm. *Int J Comput Sci Inform Technol*. 2020;12:1-13.
29. Weidner L, Walton G, Kromer R. Automated rock slope material classification using machine learning. In: *54th U.S. Rock Mechanics/Geomechanics Symposium*. American Rock Mechanics Association; 2020.
30. Wang Y, Huang J, Tang H. Automatic identification of the critical slip surface of slopes. *Eng Geol*. 2020; 273:105672.
31. Fanos AM, Pradhan B. A novel hybrid machine learning-based model for rockfall source identification in presence of other landslide types using LiDAR and GIS. *Earth Syst Environ*. 2019;3: 491-506.
32. Bernsteiner H, BroÁová N, Eischeid I, et al. Machine learning for classification of an eroding scarp surface using terrestrial photogrammetry with nir and rgb imagery. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol 5. Copernicus GmbH; 2020:431-437.
33. Loghini AM, Pfeifer N, Otepka-Schremmer J. Supervised classification and its repeatability for point clouds from dense Vhr tri-stereo satellite image matching using machine learning. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol 5, 2020:525-532.
34. Salvini R, Francioni M, Riccucci S, Bonciani F, Callegari I. Photogrammetry and laser scanning for analyzing slope stability and rock fall runoff along the Domodossola-Iselle railway, the Italian Alps. *Geomorphology*. 2013;185:110-122.
35. Hemalatha T, Ramesh MV, Rangan VP. Effective and accelerated forewarning of landslides using wireless sensor networks and machine learning. *IEEE Sensors J*. 2019;19:9964-9975.
36. Agliardi F, Crosta G, Zanchi A. Structural constraints on deep-seated slope deformation kinematics. *Eng Geol*. 2001;59:83-102.

37. Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F. A review of statistically-based landslide susceptibility models. *Earth-Sci Rev.* 2018;180:60-91.
38. Micheletti N, Foresti L, Robert S, et al. Machine learning feature selection methods for landslide susceptibility mapping. *Mathe Geosci.* 2014;46:33-57.
39. Abellán A, Jaboyedoff M, Oppikofer T, Vilaplana JM. Detection of millimetric deformation using a terrestrial laser scanner: experiment and application to a rockfall event. *Natural Hazards Earth Syst Sci.* 2009;9:365-372.
40. Mayr A, Rutzinger M, Bremer M, Oude ES, Stumpf F, Geitner C. Object-based classification of terrestrial laser scanning point clouds for landslide monitoring. *Photogramm Record.* 2017;32:377-397.
41. Alfian G, Syafrudin M, Ijaz MF, Syaekhoni MA, Fitriyani NL, Rhee J. A personalized healthcare monitoring system for diabetic patients by utilizing BLE-based sensors and real-time data processing. *Sensors.* 2018;18.
42. Srinivasu PN, SivaSai JG, Ijaz MF, Bhoi AK, Kim W, Kang JJ. Classification of skin disease using deep learning neural networks with mobilenet V2 and LSTM. *Sensors.* 2021;21:2852.
43. Shi M, Sun W, Zhang T, Liu Y, Wang S, Song X. Geology prediction based on operation data of tbm: comparison between deep neural network and soft computing methods. In: *1st International Conference on Industrial Artificial Intelligence, IAI 2019.* Institute of Electrical and Electronics Engineers Inc.; 2019.
44. Chen J, Zeng Z, Jiang P, Tang H. Deformation prediction of landslide based on functional network. *Neurocomputing.* 2015;149:151-157.
45. Wang Z, Jia K. Frustum ConvNet: sliding frustums to aggregate local point-wise features for amodal 3D object detection. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* IEEE; 2019:1742-1749.
46. Ijaz MF, Alfian G, Syafrudin M, Rhee J. Hybrid prediction model for type 2 diabetes and hypertension using DBSCAN-based outlier detection, synthetic minority over sampling technique (SMOTE), and random forest. *Appl Sci.* 2018;8.
47. Ijaz MF, Attique M, Son Y. Data-driven cervical cancer prediction model with outlier detection and over-sampling methods. *Sensors.* 2020;20:2809.
48. Stumpf A, Kerle N. Object-oriented mapping of landslides using random forests. *Remote Sens Environ.* 2011;115:2564-2577.
49. Zhao L, Wu X, Niu R, Wang Y, Zhang K. Using the rotation and random forest models of ensemble learning to predict landslide susceptibility. *Geomatics Nat Hazards Risk.* 2020;11:1542-1564.
50. Williams R, Brasington J, Vericat D, Hicks M, Labrosse F, Neal M. Monitoring braided river change using terrestrial laser scanning and optical bathymetric mapping. In: *Developments in Earth Surface Processes.* Vol 15. Elsevier B.V.; 2011:507-532.
51. Singh D, Singh B. Investigating the impact of data normalization on classification performance. *Appl Soft Comput J.* 2019;97:105524.
52. Hernandez J, Carrasco-Ochoa JA, Martínez-Trinidad JF. An empirical study of oversampling and undersampling for instance selection methods on imbalance datasets. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* Vol 8258. Berlin, Heidelberg: Springer; 2013:262-269.
53. Junsomboon N, Phienthrakul T. Combining over-sampling and under-sampling techniques for imbalance dataset. In: *ACM International Conference Proceeding Series; Part F1283.* New York, NY: Association for Computing Machinery; 2017:243-247.
54. Yen SJ, Lee YS. Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Syst Appl.* 2009;36:5718-5727.
55. Wang N, Zhao S, Cui S, Fan W. A hybrid ensemble learning method for the identification of gang-related arson cases. *Knowl-Based Syst.* 2021;218:106875.
56. Kovács G. An empirical comparison and evaluation of minority oversampling techniques on a large number of imbalanced datasets. *Appl Soft Comput J.* 2019;83:105662.
57. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res.* 2002;16:321-357.
58. Fernandez A, Garcia S, Herrera F, Chawla NV. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *J Artif Intell Res.* 2018;61:863-905.

59. He H, Bai Y, Garcia EA, Li S. ADASYN: adaptive synthetic sampling approach for imbalanced learning. In: *Proceedings of the International Joint Conference on Neural Networks*. 2008:1322-1328.
60. Barua S, Islam MdM, Murase K. ProWSyn: proximity weighted synthetic oversampling technique for imbalanced data set learning. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol 7819. Berlin, Heidelberg: Springer; 2013: 317-328.
61. Sáez JA, Luengo J, Stefanowski J, Herrera F. SMOTE-IPF: addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering. *Inform Sci*. 2015;291:184-203.
62. Cao Q, Wang S. Applying over-sampling technique based on data density and cost-sensitive SVM to imbalanced learning. In: *Proceedings—2011 4th International Conference on Information Management, Innovation Management and Industrial Engineering, ICIII 2011*. Vol 2; 2011:543-548.
63. Zhou B, Yang C, Guo H, Hu J. A quasi-linear SVM combined with assembled SMOTE for imbalanced data classification. In: *Proceedings of the International Joint Conference on Neural Networks*; 2013.
64. Lee J, Kim NR, Lee JH. An over-sampling technique with rejection for imbalanced class learning. In: *ACM ICMCOM 2015—Proceedings*. Vol 2015. New York, NY: Association for Computing Machinery, Inc, 1-6.
65. Batista GEAPA, Prati RC, Monard MC. MC. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explor Newsletter*. 2004;6:20-29.
66. Tomek I. Two modifications of CNN. *IEEE Trans Syst Man Cybernet*. 1976;SMC-6:769-772.
67. Koziarski M, Wozniak M. CCR: a combined cleaning and resampling algorithm for imbalanced data classification. *Int J Appl Math Comput Sci*. 2017;27:727-736.
68. Sandhan T, Choi JY. Handling imbalanced datasets by partially guided hybrid sampling for pattern recognition. In: *Proceedings—International Conference on Pattern Recognition*. Institute of Electrical and Electronics Engineers Inc.; 2014:1449-1453.
69. Nakamura M, Kajiwara Y, Otsuka A, Kimura H. LVQ-SMOTE - Learning vector quantization based synthetic minority over-sampling technique for biomedical data. *BioData Mining*. 2013;6:16.
70. Gazzah S, Amara NEB. New oversampling approaches based on polynomial fitting for imbalanced data sets. In: *DAS 2008—Proceedings of the 8th LAPR International Workshop on Document Analysis Systems*. 2008:677-684.
71. Stefanowski J, Wilk S. Selective pre-processing of imbalanced data for improving classification performance. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol 5182. Berlin, Heidelberg: Springer; 2008:283-292.
72. Sharma S, Bellinger C, Krawczyk B, Zaiane O, Japkowicz N. Synthetic oversampling with the majority class: a new perspective on handling extreme imbalance. In: *Proceedings—IEEE International Conference on Data Mining, ICDM*. Vol 2018-Novem. Institute of Electrical and Electronics Engineers Inc.; 2018: 447-456.
73. Ma Z, Mei G, Piccialli F. Machine learning for landslides prevention: a survey. *Neural Comput Appl*. 2020.
74. Zhou Z-H. Ensemble learning. In: *Encyclopedia of Biometrics*. Boston, MA: Springer US; 2009:270-273.
75. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY: Springer; 2009.
76. D'Ambrosio A, Tutore VA. Conditional classification trees by weighting the gini impurity measure. In: *Studies in Classification, Data Analysis, and Knowledge Organization*. Berlin, Heidelberg: Springer; 2011: 273-280.
77. Awad M, Khanna R, Awad M, Khanna R. Support vector machines for classification. In: *Efficient Learning Machines*. Apress; 2015:39-66.
78. Murtagh F. Multilayer perceptrons for classification and regression. *Neurocomputing*. 1991;2:183-197.
79. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci*. 1997;55:119-139.
80. Zhu J, Zou H, Rosset S, Hastie T. *Multi-class AdaBoost*, Technical Report; 2009.
81. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63:3-42.
82. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. August 13–17. Association for Computing Machinery; 2016:785-794.

83. Hwang S, Guevarra IF, Yu B. Slope failure prediction using a decision tree: a case of engineered slopes in South Korea. *Eng Geol.* 2009;104:126-134.
84. Marjanović M, Bajat B, Abolmasov B, Kovačević M. Machine learning and landslide assessment in a GIS environment. In: *Advances in Geographic Information Science*. Cham: Springer; 2018:191-213.
85. Lee S, Ryu J-H, Won J-S, Park H-J. Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Eng Geol.* 2004;71:289-302.
86. Chen W, Xie X, Wang J, et al. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *CATENA.* 2017;151:147-160.
87. Hong H, Liu J, Bui DT, et al. Landslide susceptibility mapping using J48 decision tree with AdaBoost, bagging and rotation forest ensembles in the Guangchang area (China). *CATENA.* 2018;163:399-413.
88. Forman G, Cohen I. Learning from little: comparison of classifiers given little training. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol 3202. Springer Verlag; 2004:161-172.
89. Klein D, Manning CD. Conditional structure versus conditional estimation in NLP models. In: *Proceedings of the ACL-02 conference on Empirical methods in natural language processing—EMNLP '02*. Vol 10. Morristown, NJ: Association for Computational Linguistics; 2002:9-16.
90. Hong H, Naghibi SA, Moradi DM, Pourghasemi HR, Chen W. A comparative assessment between linear and quadratic discriminant analyses (LDA-QDA) with frequency ratio and weights-of-evidence models for forest fire susceptibility mapping in China. *Arabian J Geosci.* 2017;10:167.
91. Li Y, Chen W. A comparative performance assessment of ensemble learning for credit scoring. *Mathematics.* 2020;8:1756.
92. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E—Stat Phys Plasmas Fluids Related Interdiscip Topics.* 2004;69:16.
93. Demšar J. *Statistical Comparisons of Classifiers over Multiple Data Sets*, Technical Report; 2006.
94. Mozetič I, Torgo L, Cerqueira V, Smailović J. *How to Evaluate Sentiment Classifiers for Twitter Time-Ordered Data?*; 2018.
95. Breiman L. Random forests. *Mach Learn.* 2001;45:5-32.
96. Brodersen KH, Ong CS, Stephan KE, Buhmann JM. The balanced accuracy and its posterior distribution. In: *Proceedings—International Conference on Pattern Recognition*. 2010:3121-3124.
97. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825-2830.
98. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng.* 2007;9:90-95.
99. The Pallets Projects. *Flask|The Pallets Projects*; 2010.
100. Bello SA, Yu S, Wang C, Adam JM, Li J. Review: deep learning on 3D point clouds. *Remote Sensing.* 2020; 12:1729.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Zoumpikas T, Puig A, Salamó M, García-Sellés D, Blanco Nuñez L, Guinau M. An intelligent framework for end-to-end rockfall detection. *Int J Intell Syst.* 2021;1-32. <https://doi.org/10.1002/int.22557>



UNIVERSITAT DE
BARCELONA

Anufra
Soil & Water Consulting



Institut de Recerca
Geomodels

GGAC
Grup de Geodinàmica i Anàlisi de Conques